Running-head: Gender Representation in Bilinguals

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

Sayaka Sato and Pascal M. Gygax
University of Fribourg, Switzerland

Ute Gabriel
Norwegian University of Science and Technology, Trondheim, Norway

In Press in *Bilingualism: Language & Cognition*

Address for correspondence:
Sayaka Sato
Department of Psychology
University of Fribourg
Rue de Faucigny 2
1700 Fribourg, Switzerland
E-mail: sayaka.sato@unifr.ch

Abstract

We 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 to 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; high proficient L2 participants resembling closer to native speakers of the L2 and less proficient L2 participants being influenced more by their native language.

Keywords: gender representation, grammatical gender, gender stereotypes, masculine form, generic interpretation

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

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'[1] world knowledge. Proponents of constructionist accounts (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 processes, studies have employed occupational role nouns associated to 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, Oakhill, & Reynolds, 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 taken on by women than men). Thusly 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, Sarrasin, 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* (i.e., *the chair*: French, feminine) or *der Bleistift* (i.e., *the pen*: German, masculine) but in cases of personal nouns, the grammatical and biological gender of the person typically correspond, as in *une étudiante* (i.e., *a female student*: French, feminine) and *un étudiant* (i.e., *a male* student: French, masculine). This rule of gender categorization in grammatical gender languages is fundamental, as verbs and adjectives are modified and 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, 1981; 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, Reyes, Hernandez, 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, Wicha et al. (2005) in a series of experiments employing a sentence-picture semantic judgment paradigm 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; Gabriel, Gygax, Sarrasin, Garnham & Oakhill, 2008; to some extent Norwegian: Gabriel & Gygax, 2008)[2].

Consequently, the use of the masculine form as generic may have serious implications when put into a real-world context: Vervecken and Hannover (2012) for example 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 occupational role nouns among French middle school pupils. The pupils were generally more confident to uptake jobs that were stereotypically congruent to their own sex. However, for occupations that were stereotypically incongruent, the usage of female-inclusive forms (i.e., feminine form: *Mathématicien(ne)*, 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 off-line studies in which participants are given the opportunity to reflect on their beliefs and perceptions, rather than from on-line task studies where clearer interpretations of the type of information being immediately activated can be made. Gygax and colleagues (2008) adapted an on-line 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, 2011; 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 worthwhile to mention the differentiating notion of CONCEPTUAL and SEMANTIC levels of representation. The conceptual level of representation refers to the non-linguistic and multi-model 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 L2 competence and proficiency have not reached a sufficient level, the access to a concept in L2 is most likely controlled by a 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, Krajciova, & 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 a 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. Secondarily, 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 few studies have been conducted on the way bilinguals represent gender, 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, Kozlowska-Macgregor, & Leung, 2004) in the target language. Still, some of the studies on bilingualism are highly relevant for 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 to 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 for 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 found (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. Second, it showed that even rather minor morphosyntactic features could influence bilinguals' comprehension. Third and finally, 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. (2008)'s

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 anticipate that it will be particularly apparent in the female stereotype condition, again, as French and English have been shown to generate two opposite representations in this condition.

Method

Participants

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; range = 18-33; mean start age of L2 acquisition: 12; range = 3-12; mean years of L2 study: 7; range: 4-22; 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.

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; range = 18-29; mean start age of L2 acquisition: 10; range =3-19; mean years of L2 study: 10; range: 4-22; 50 female, 11 male). Five participants were removed from the analyses as their L2 proficiency were too low (less than a third correct on the C-Test). Participants were either granted course credits for experiment participation or paid £5.

Materials and Design

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 either a FEMALE (e.g., *social workers*, *assistants sociaux*), MALE (e.g., *surgeon*, *chirurgiens*) or NEUTRAL (e.g., *musicians, musiciens*) stereotypical role noun in its masculine plural form in the French version and in the plural form in the English version as its main subject referent (see example 1a)[3]. 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 example 1b & 2b) or *women* (see examples 1c & 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.

(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* and *women* 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 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 whilst 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 example 3a & 3b), a second version where there was a mismatch in sex of the role noun mentioned in the first and second sentences (see example 4a & 4b below), and finally pairs in which there was a semantic incoherence (see examples 5a & 5b below).

(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.

(5b) Since sunny weather was forecast some of the men weren't wearing a coat.

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 A 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.

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.

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).

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 by which the second sentence is mapped onto the representation of the first one. The instructions stressed that they should read at normal speed as they would normally read a book, and to 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 "*no*" button. Participants were asked to keep their pointer of their dominant hand on the "*yes*" button and 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.

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 also that we do not present separate results for the two language orders (i.e., L1 or L2 first), as they did not show any 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. Less proficient French-English speakers on the other hand 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 high 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-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 towards 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, high 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) X 2 (Language: L1 vs. L2) X 3 (Stereotype: Female vs. Male vs. Neutral) X 2 (Continuation: Men vs. Women) mixed ANOVA, with Proficiency as a between-participant factor and Language, Stereotype and Continuation as within-participant factors. Second, as we expected possible shift and proficiency effects to be seen most clearly for female stereotyped role nouns, we ran the ANOVAs for female stereotyped role nouns only.

Proficiency Measures

C-test scores were taken into account to establish L2 proficiency[4]. 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.

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 ($W_s = 325$, $z = -6.26$, $p < .001$).

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).

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 X Continuation effect, $F(1, 58) = 6.80$, $p < .01$ (see Figure 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).

<Insert Figure 1 about here>

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.

Finally, there was a trend towards a significant four-way interaction of Language X Stereotype X Continuation X Proficiency, $F(2, 116) = 2.44$, $p = .09$, qualified, as expected (Hypothesis 3), by a significant Language X Stereotype X 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) X 2 (Language: English vs. French) X 2 (Continuation: Men vs. Women) ANOVA only for responses to female stereotypical role nouns. As predicted, there was a significant Proficiency by Language by 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 for *men* and .49 for *women*) and L2 (.64 for *men* and .52 for *women*) (see Figure 2).

<Insert Figure 2 about here>

Positive Response Times

There was only a significant Language by 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* than to *men* continuations (*ns*). This result also supports Hypotheses 1 and 2. No other main or interaction effects were significant.

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 ($W_s = 231$, $z = -6.38$, $p < .001$).

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 by 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 by Stereotype by Continuation interaction effect (see Figure 1), $F(2, 118) = 4.05$, $p < .05$, which confirmed that in L1 English, there was a Stereotype by 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 towards 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) X 2 (Language: English vs. French) X 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).

<Insert Figure 3 about here>

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 wo*men* continuations, which was further qualified by a significant Language by 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.

Again, we tested our specific hypothesis on the effects of Proficiency on female stereotypical role nouns. Though the analyses revealed a significant Language by Continuation interaction $F(1, 59) = 4.47, p < .05$ revealing that in L1 English, participants responded 235 milliseconds faster for wo*men* 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*), there was no effect of proficiency.

Discussion

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 and colleagues (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, 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 to 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, our study hence addresses issues associated to 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: 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, 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; Oxford University Press), 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 their 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 seem 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 a 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*), 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 at least for 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 recognition (Costa, Caramazza, & Sebastian-Galles, 2000; Dijkstra, et al., 1999). For example, the English noun *golfers* shares orthographic (and phonologic) 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 to *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 reran 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 (Garnham, Gabriel, Sarrasin, Gygax, & Oakhill, 2012; Gabriel & Gygax, 2008; Gygax et al., 2008). The advantage of this task is that it is effective in addressing the ease by which certain types of information are integrated into on-going mental models, and thusly 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 onto 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 (1996, 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 to focus their attention on one aspect and speakers of another language onto another. This process constitutes one of the essential components of what

he 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 (1996, 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 towards 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.

References

Athanasopoulos, P. (2009). Cognitive Representation of Colour in Bilinguals: The Case of Greek Blues. *Bilingualism: Language and Cognition, 12,* 83–95. doi: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. doi: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. doi:10.1017/S0142716408080053

Banaji, M., & Hardin, C. (1996). Automatic stereotyping. *Psychological Science, 7,* 136–141. doi: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. doi:10.1162/0898929052880101

Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian-German bilingual children. *International Journal of Bilingualism, 11,* 251 –273. doi: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. 357–384. Hove, U.K.: Psychology Press.

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. doi:10.1111/j.1559-1816.1973.tb01290.x

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.

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. doi:10.4074/S0003503308002030

Braun, F., Sczesny, S., & Stahlberg, D. (2005). Cognitive effects of masculine generics in German: An overview of empirical findings. *Communications*, *30*, 1–21. doi: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. doi:10.1017/S0272263108080327

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. doi: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*, *28*, 277–293. doi:10.1017/S0142716407070142

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. doi:10.1080/027249896392531

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. doi:10.3406/psy.2005.29694

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.

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.

Cook, V. J. & Bassetti, B. (Eds.). (2011). *Language and bilingual cognition.* New York and Hove: Psychology 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. doi:10.1037//0278-7393.26.5.1283

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 A, 53,* 397–428. doi:10.1080/027249800390547

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. doi:10.1017/S136672890100044X

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. doi:10.1006/jmla.1999.2654

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. doi:10.1017/S1366728998000121

Duffy, S., & Keir, J. (2004). Violating stereotypes: Eye movements and comprehension

    processes when text conflicts with world knowledge. *Memory & Cognition, 32,*

    551–559. doi:10.3758/BF03195846

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests.

    Language Testing, 23, 290–325. doi:10.1191/0265532206lt330oa

Flaherty, M. (2001). How a Language Gender System Creeps into Perception. *Journal*

    of Cross-Cultural Psychology, 32, 18–31. doi: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.

    doi:10.1177/026765830101700301

Gabriel, U., & Gygax, P. (2008). Can societal language amendments change gender

    representation? The case of Norway. *Scandinavian Journal of Psychology, 49,* 451–

    457. doi:10.1111/j.1467-9450.2008.00650.x

Gabriel, U., Gygax, P., Sarrasin, 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.

    doi:10.3758/BRM.40.1.206

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. 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. doi: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., & Reynolds, D. (2002). Are inferences from stereotyped role name to characters' gender made elaboratively? *Memory & Cognition, 30,* 439–446. doi:10.3758/BF03194944

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395. doi:10.1037//0033-295X.101.3.371

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*(3), 143–151. doi:10.1024/1421-0185.67.3.143

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. doi:10.1080/01690960701702035

Hosoda, M., & Stone, D. L. (2000). Current gender stereotypes and their evaluative content. *Perceptual and Motor Skills, 90,* 1283–1294. doi:10.2466/PMS.90.3.1283-1294

Irmen, L. (2007). What's in a (Role) Name? Formal and Conceptual Aspects of Comprehending Personal Nouns. *Journal of Psycholinguistic Research, 36,* 431–456. doi:10.1007/s10936-007-9053-z

Irmen, L., & Köhncke, A. (1996). Zur Psychologie des "generischen" Maskulinums [On the psychology of the "generic" masculine]. *Sprache & Kognition*, *15*, 152–166. doi:10.1515/zfgl.33.2-3.212

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. doi:10.1177/0261927X04266810

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. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Karmiloff-Smith, A. (1981). *A Functional Approach to Child Language: A Study of Determiners and Reference*. Cambridge University Press.

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.

Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test1. *Language Testing, 1*, 134 –146. doi:10.1177/026553228400100202

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. doi: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. doi: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. doi:10.1191/0267658302sr201oa

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2002). Language at Work: Children's Gendered Interpretations of Occupational Titles. *Child Development, 73,* 810–28.

Lyster, R. (2007). *Learning and teaching languages through content: a counterbalanced approach*. John Benjamins Publishing Company. doi:10.1111/1467-8624.00440

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. doi:10.1017/S1366728903001020

Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition, 33,* 972–983. doi:10.3758/BF03193206

Paradis, M. (1997). The cognitive neuropsychology of bilingualism. In A. De Groot & J. Kroll (eds.), *Tutorials in bilingualism. Psycholinguistic perspectives,* pp.15-27. Québec: CIRAL.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition, 3,* 209–230. doi:10.1017/S1366728999000322

Pavlenko, A. (2011). *Thinking and Speaking in Two Languages*. Bristol: Multilingual Matters.

Quick Placement Test. (2001). Oxford: Oxford University Press.

Rahimi, M., & Saadat, M. (2005). A verbal protocol analysis of a C-test. *Iranian Journal of Applied Linguistics, 8,* 55–85.

Reynolds, D. J., Garnham, A., & Oakhill, J. (2006). Evidence of immediate activation of gender information from a social role name. *Quarterly Journal of Experimental Psychology (2006), 59,* 886–903. doi:10.1080/02724980543000088

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. doi:10.1191/0267658306sr259oa

Scheutz, M., & Eberhard, K. M. (2004). Effects of morphosyntactic gender features in bilingual language processing. *Cognitive Science, 28,* 559–588. doi:10.1207/s15516709cog2804_3

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. doi:10.1037//0096-3445.131.3.377

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. doi:10.1016/j.jml.2004.05.003

Slobin, D. I. (1996). 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. (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.

Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name Your Favorite Musician. *Journal of Language and Social Psychology, 20*, 464 –469. doi:10.1177/0261927X01020004004

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. doi:10.1007/BF00291751

Vervecken, D. & Hannover, B. (2012). *Changing (s)expectations: The influence of teachers' language use on pupils' gendered representations of occupations, vocational aspirations, and self-efficacy*. Poster presented at the 13th Annual Meeting of the Society for Personality and Social Psychology, San Diego, CA, USA.

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. doi:10.1162/jocn.1996.8.3.231

White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Leung, Y.-K. I. (2004). Gender and Number Agreement in Nonnative Spanish. *Applied Psycholinguistics, 25*, 105–133. doi:10.1017/S0142716404001067

Wicha, N. Y. Y., Orozco-Figueroa, A., Reyes, I., Hernandez, A., Gavaldón de Barreto, L. & Bates, E. A., (2005). When zebras become painted donkeys: Grammatical gender and semantic priming interact during picture integration in a spoken Spanish sentence. *Language & Cognitive Processes, 20*, 553-587. doi:10.1080/01690960444000241

Zwaan, R., & Radvansky, G. (1998). Situation models in language comprehension and

memory. *Psychological Bulletin, 123,* 162–185. doi:10.1037//0033-2909.123.2.162

Footnotes

[1] 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.

[2] 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.

[3] To 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.

[4] Self assessment scores of L2 proficiency was significantly correlated to the performance of C-tests scores, $r = .78$, $p < .001$.

Appendix A

Example passage of English C-test used to evaluate L2 proficiency.

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.

Figure 1. Proportion of positive responses of each native group in their respective L1 (error
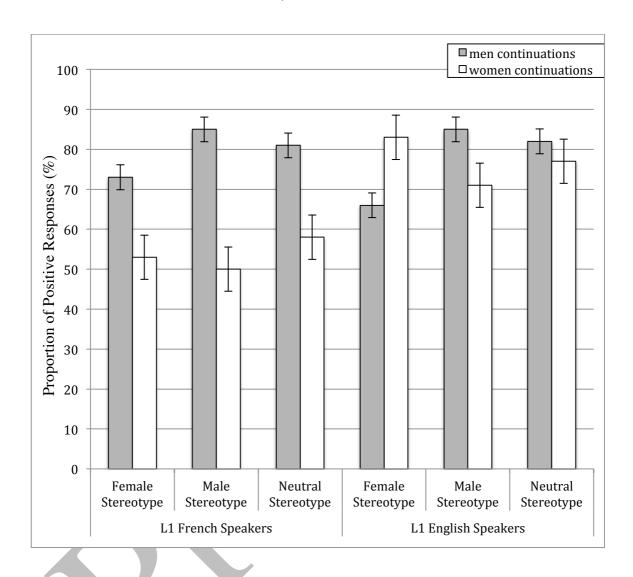
bars indicate standard error of the mean).

Figure 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).
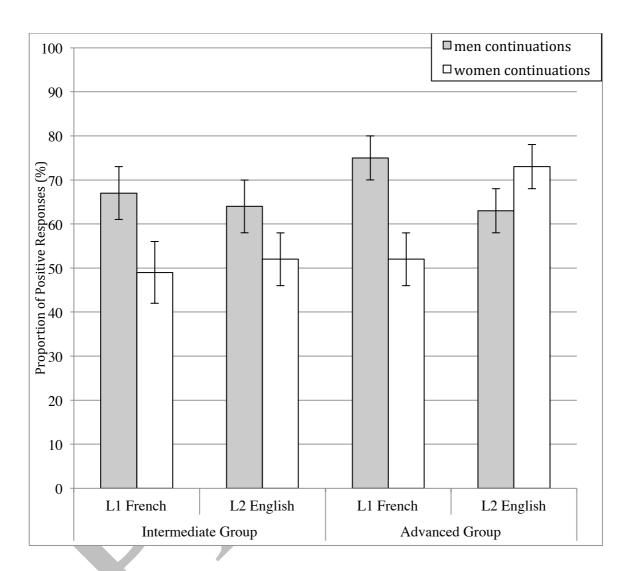
Figure 3. L1 English speakers' proportion of positive responses of each continuation only in the *female stereotype* condition. Proficiency is divided into *intermediate* and *advanced* groups (error bars indicate standard error of the mean).