The Dissertation Committee for Alisa Baron Certifies that this is the approved version of the following dissertation:

Predictive Use of Matched and Mismatched Gender-Marked Articles in Spanish-English Bilinguals

Committee:

Maya Henry, Supervisor

Lisa M. Bedore, Co-Supervisor

Elizabeth Peña

Zenzi Griffin

Predictive Use of Matched and Mismatched Gender-Marked Articles in Spanish-English Bilinguals

by

Alisa Baron

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2018

Dedication

I would like to dedicate my dissertation to my family for their unwavering support, motivation, love and understanding throughout my studies, which has culminated in this dissertation.

Acknowledgements

I would like to thank the undergraduate assistants and research assistant for their support throughout the process. Thank you to Karla Sanchez, Yarely Ramirez, Ivette Villarreal, Kathy Regalado, and Alice Regalado. Thank you to those who participated in naming agreement: Roberto Cofresí, Yarely Ramirez, Carlos Nye, Ivette Villareal, Maritza Jacobs, Leah Joseph, Stephanie Grasso, Scott Prath, Maximiliano Ramirez, Ellen Kester, Jissel Anaya, Katherine Johnson, Maria Mitidieri, Daniela Melo, Ricardo Nieto, Leandro Moreno Bello. Thank you for those of you who donated towards my HornRaiser campaign so that I could recruit children and adults to participate in my study. I would also like to thank Dr. Marcus Johnson from SR Research, for his continuous support during the programming and analysis stages of the eye-tracking task. Lastly, I would like to thank my dissertation committee Drs. Lisa Bedore, Zenzi Griffin, Maya Henry, and Elizabeth Peña for their informative feedback and support throughout the dissertation process.

Predictive use of matched and mismatched gender-marked articles in Spanish-English bilinguals

Alisa Baron, PhD

The University of Texas at Austin, 2018

Supervisor: Maya Henry

Co-Supervisor: Lisa M. Bedore

Abstract: Most behavioral research involving typically-developing children has been devoted to understanding language production processes, but there is limited information on language comprehension; thus, part of the developmental picture is incomplete. Although there is a growing body of literature focusing on production of grammatical forms for bilingual children, there is a critical need to understand how children's comprehension skills develop in conjunction with production skills by understanding their language knowledge and language experience. With regard to grammatical class, articles are especially important because they precede nouns in most contexts in Spanish and therefore are used with a high frequency in all aspects of language. Articles should be studied in the language processing of elementary age children to understand what children attend to during language comprehension. In this study, the visual world paradigm was used to examine gendered articles using phonological competitors in trials with informative (different-gender grammatical trials), uninformative (same-gender grammatical trials), or incorrect (ungrammatical trials) articles in bilingual children and adults. Participants named common nouns and completed an eye-tracking task, a grammaticality judgment task, and a standardized

vocabulary test in both English and Spanish. Bilingual children ages 5-6 and 8-9 did not show gender sensitivity in informative vs. uninformative trials but were significantly slower on ungrammatical trials. Bilingual adults showed sensitivity to gender and were significantly faster on informative trials relative to uninformative trials, which in turn were significantly faster than ungrammatical trials regardless of participants' profile of current Spanish language input. Children may be merging their representation of articles in the two languages and not find the gender cue in Spanish to be necessary. Spanish, a gendered language and English, a non-gendered language, may be in competition during this developmental period. Bilingual adults are able to quickly and accurately process the incoming gendered information, and are therefore able to demonstrate gender sensitivity.

Table of Contents

List of Tables

List of Figures

Chapter 1: Review of the Literature

Most behavioral research involving typically-developing children has been devoted to understanding language production and comprehension in relation to each other, and especially to early word learning and vocabulary development. There is limited research regarding language comprehension processes, especially in the morphosyntactic domain. Within the literature on adult language processing, there are more studies on both language comprehension and production, but there is a focus on monolingual speakers and adult second language learners (Brouwer, Sprenger, & Unsworth, 2017; Kroff, Dussias, Gerfen, Perrotti, & Bajo, 2016; Lew-Williams & Fernald, 2007, 2010; Montrul, Foote, & Peripiñán, 2008; Morales, Paolieri, Dussias, Kroff, Gerfen, & Bajo, 2016; White, Valenzuela, Kozlowska-Macgregor, & Leung, 2004). Studying children and adults who have been exposed to two languages since early childhood can help us understand the development and maintenance of their languages over time.

GRAMMATICAL GENDER

Spanish has a rich system of inflectional morphology in relation to English. Nouns in English and Spanish are marked for number (singular or plural) but in Spanish, nouns are also classified according to gender. Articles are especially important because they precede nouns in most contexts in Romance languages such as Spanish, and therefore are used with a high frequency in all aspects of language. Grammatical gender is binary in Spanish where all nouns are either masculine or feminine. Gender is an inherent property of nouns where animate nouns are assigned a gender that is semantically motivated while inanimate nouns are assigned a grammatical, semantically arbitrary gender (Corbett, 1991). Noun gender and morphological marking in Spanish co-

occur in a predictable way where 99.8% of nouns ending in *–o* are masculine and 96.3% of nouns ending in *–a* are feminine, while nouns ending in *–e* or consonants can be either masculine or feminine (Teschner & Russell, 1984).

Gender is learned and understood both on lexical and syntactic levels. Grammatical knowledge of gender assignment affects both lexical access (Bölte & Connine, 2004; Spinelli, Meunier, & Seigneuric, 2006) and syntactic processing (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Foucart & Frenck-Mestre, 2011; Gunter, Friederici, & Schriefers, 2000). At the lexical level, one must learn the meaning of the noun with the gender feature (gender assignment) and establish agreement between the noun and other elements in the sentence (gender agreement) at the syntactic level. Grammatical gender is an important morphosyntactic cue to identify words and build syntactic representations in real time (i.e. Foucart & Frenck-Mestre, 2011; Hopp, 2013; Wicha, Moreno, & Kutas, 2004). Prior gender information provided by an article can reduce the search space in the lexicon to only those elements with a particular gender (Friederici & Jacobsen, 1999). Thus, gender cues help listeners keep track of the references in a sentence (Bates et al., 1996) and facilitate the interpretation of speech.

Article use begins early in development but takes time to acquire. Articles are frequently produced in error or are omitted by toddlers learning language until they acquire articles around age 4 (Bedore & Leonard, 2005) or even later, between ages 5 and 7 (e.g. Castilla-Earls et al., 2016; Jackson-Maldonado & Maldonado, 2017). Children with language impairment continue to produce article gender and number errors over an extended period of time (e.g. Bedore & Leonard, 2001). Children may omit articles, as they are weakly stressed syllables (Gerken, 1991). Without articles, the sentence is typically ungrammatical even though articles do not provide key information for comprehension of the main idea. The gender of inanimate nouns is arbitrary, lacks

semantic value, and is encoded in multiple constituents. Gender is encoded in the article and noun, as well as the adjective (if present) (ex. la pelota roja [the.FEM red.FEM ball.FEM), thus appearing redundant. By contrast, in English, grammatical gender is not instantiated.

USING EYE-TRACKING TO STUDY LANGUAGE PROCESSING

People process speech incrementally, by making use of what they have heard so far to interpret words and sentences that are not yet complete (e.g. Bates et al.,1996; Fernald, Swingley, & Pinto, 2001; Friederici & Jacobsen, 1999; Marslen-Wilson, 1987). Children as young as two years old can identify the referent of a familiar word before the speaker has completed it by attending to an appropriate picture as a speech signal unfolds (Fernald & Hurtado, 2006; Fernald et al., 2001). In contrast, behavioral tasks are offline measures, which evaluate comprehension by assessing complex behaviors that children make in response to language input after this input ends, and not while they are listening to it and trying to make sense of it. There are limited tools for testing comprehension, as there is a high probability of children responding at chance. For example, when a child is given directions or a question, the response to these directions is seen by a child pointing to the correct pictures, or responding accurately to the question. A response is scored as correct or incorrect but what they attended to in order to execute the direction or respond to a question is unknown. Consequently, these measures might be missing some of the subtle, real-time processes underlying comprehension.

Very little is known about the real-time behavior related to bilingual language processing of children and adults. While behavioral tasks reflect a “look-back” approach from the point of view of the tester, eye gaze measures of receptive language reflect an online record of language processing. This provides the ability to examine processing

demands at the sentence and word level with minimal response demand. To show comprehension, the child simply needs to look at the image that is being named. In contrast, in offline measures, children's comprehension is judged based on complex non-verbal responses to words like pointing or following a direction. Eye-tracking has been helpful in uncovering the types of information people are and are not able to use when making syntactic decisions, providing insight into the organization of the language processing system and its developmental trajectory (Dussias, 2010). Eye movements to a visual scene are closely time locked with auditory input and are related to underlying activation levels of word candidates. Also, because eye movements are recorded while participants are presented with continuous spoken language, an eye-tracking task provides a sensitive, implicit measure of processing without interrupting the flow of speech (e.g. Tanenhaus, Magnuson, Dahan, & Chambers, 2000; Tanenhaus & Trueswell, 2006). Eye gaze seems to capture the changes in comprehension of the language input moment-to-moment (Borovksy, Elman, & Fernald, 2012). Therefore, eye-tracking is informative in this line of research as it is an online processing tool and may offer insights about the mechanisms underlying receptive language deficits.

GENDER-MARKED LANGUAGE IN MONOLINGUAL CHILDREN

Two and three-year-olds learning Spanish as their first language use gender-marked articles to more rapidly identify visual referents when the gender-marked article was informative. Lew-Williams and Fernald (2007) found that monolingual Spanish-speaking children take advantage of gender-marked words in real-time to interpret spoken sentences rapidly. Children looked at paired pictures of objects with names of either the same (e.g., la pelota [the.FEM ball], la galleta [the.FEM cookie]) or different grammatical gender (la pelota vs. el zapato [the.MASC shoe], as they heard a Spanish

sentence referring to one of the objects. Children oriented faster to the correct referent on different-gender trials, when the article was potentially informative, than on same-gender trials, when the article revealed nothing about the following noun. This ability to exploit morphosyntactic information in the process of establishing reference reveals how a young child learning a richly inflected language makes progress toward becoming a native listener.

Vocabulary size has been found to be strongly associated with speed of comprehension in looking tasks (Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008). Fernald et al. (2006) conducted a longitudinal study with toddlers from 15 to 24 months of age and noted that the online processing measures were relatively stable during this time period and were correlated with vocabulary growth during this time. Six years later, a follow-up study was conducted and processing efficiency while the children were toddlers was strongly linked to scores on standardized tests in elementary school (Marchman & Fernald, 2008). Although Lew-Williams and Fernald (2007) found a correlation between gender processing and vocabulary, Brouwer, et al. (2017) noted that receptive vocabulary scores were not determining factors for gender processing in Dutch.

Grammaticality judgment, an offline comprehension measure, has been included in a battery of tasks in conjunction with eye-tracking tasks. In Dutch, a language with 3 genders (masculine, feminine, and neuter forms), a gender grammaticality judgment task was completed with 4 and 6-year-olds using puppets (Unsworth & Hulk, 2010). Children listened to prerecorded sentences from 2 puppets (one grammatical and the other ungrammatical) and decided which puppet was correct. On average, the children scored around 70% accuracy and some performed at chance.

In summary, children learning a language with gender marking who are as young as 2 can orient faster to a target when a gendered article is informative than when it is uninformative. A majority of the evidence suggests that vocabulary size is associated with gender processing. A grammaticality judgment measure shows evidence of metalinguistic awareness of grammaticality and is a method frequently utilized with an eye-tracking task as they both target gender.

EFFECTS OF GENDER MISMATCH ON PERFORMANCE IN MONOLINGUAL SPEAKERS

Gender mismatches yield inhibitory effects in Romance languages. Monolingual French and Spanish-speaking children and adults experience these effects in online and offline tasks (e.g., Bates et al., 1996; Colé & Segui, 1994; Faussart et al., 1999; Jacobsen, 1999; Jakubowicz & Faussart, 1998; Lew-Williams & Fernald, 2007; Wicha et al., 2005). When the gender of an article and an adjacent noun are mismatched (e.g., el (the.MASC) pelota (ball.FEM), noun recognition is significantly slowed. In Italian, in three separate tasks (word-repetition, gender-monitoring, and grammaticality judgment) performance was also slower when the preceding adjective was gender mismatched with the noun and faster when the adjective was matched with the noun (Bates et al., 1996). In a study by van Heugten and Shi (2009), 25-month-old French-learners were presented with picture pairs that referred to nouns with matched or mismatched genders. They calculated proportion of looking time for three successive time windows of 500 ms each starting with 500 ms after article onset. Children recognized nouns more efficiently and spent more time looking at the target when correct and informative articles (different-gender grammatical trials) were used rather than when incorrect (ungrammatical trials) or uninformative (same-gender grammatical trials) articles were used. While target processing in same-gender grammatical trials recovered in the following time window, ungrammatical articles continued to affect processing efficiency until much later in the

trial. Therefore, ungrammatical articles impeded comprehension for a longer time than uninformative grammatical articles. Jakubowicz and Faussart (1998) conducted a study on French-speaking adults and found that mismatched gendered article-noun pairs were associated with slower lexical decisions than matched pairs. When the gender of an article or adjective is matched, or congruent with the following noun, noun recognition is facilitated relative to a neutral baseline and is slowed when it is mismatched or incongruent (Friederici & Jacobsen, 1999; Dahan, Swingley, Tanenhaus, & Magnuson, 2000). To summarize, when an article-noun pair has mismatched gender, noun recognition, processing efficiency, and lexical decisions are slower than when an article-noun pair is matched in gender.

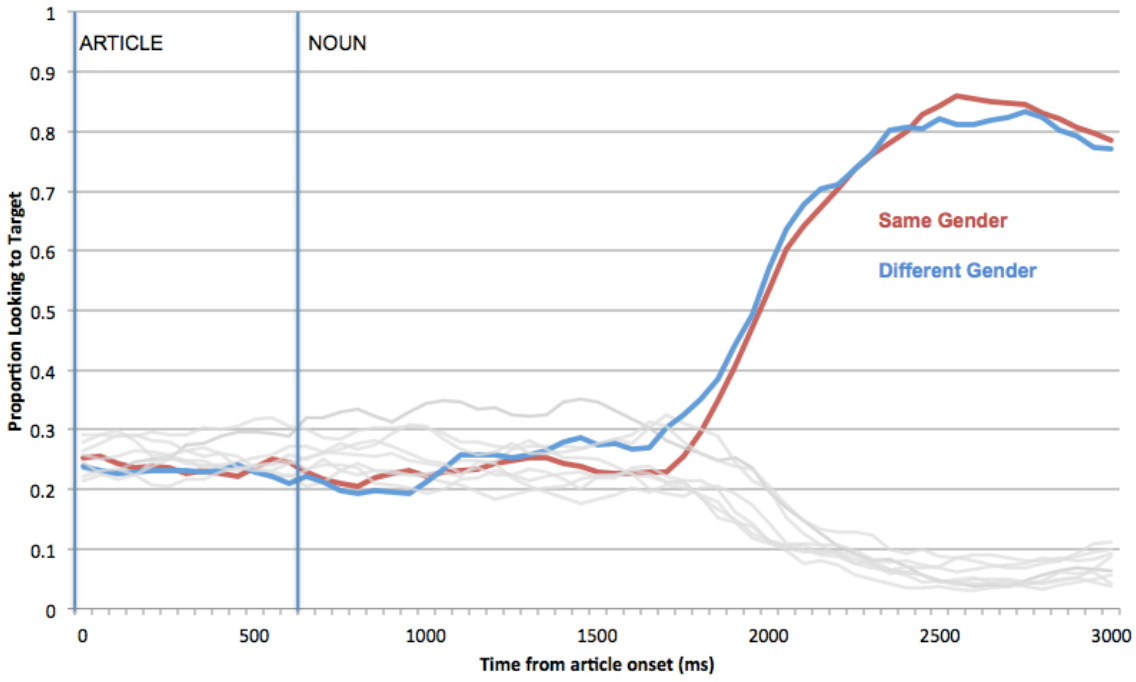
GENDER-MARKED LANGUAGE IN BILINGUAL CHILDREN

The gender-marking effect has been studied primarily with monolingual children. More than half of the world's population is bilingual. Around 20% of people in the U.S. are bilingual and Spanish-English bilinguals represent about half of all bilinguals in the U.S. This leads to a question of whether bilinguals will show the same effect and whether it depends on when they acquired and started using the gender-marked language and/or how much time they currently spend listening and speaking in each language. Given that gender marking can help with tracking of referents and disambiguation of referential constructions (Cacciari, Carreiras, & Barbolini Cionini, 1997; Corbett, 1991), one might predict that bilinguals should show the gender-marking effect. Adult second language acquisition research seems to show that early bilinguals produce no, or very few, gender production errors, whereas late bilinguals make a substantial number of gender errors (e.g., Carroll, 1989). Bearing this in mind, the presence of a gender-marking effect in bilinguals may depend on the age at which they started acquiring and using the

language(s) that have gender marking. Alternately, this may mean that the longer an individual has used a gendered language, the more accurate they become.

The gender-marking effect has only recently been investigated in bilingual children. Following the Lew-Williams and Fernald (2007) framework, a study with 36 Spanish-English bilingual children aged 5;6-8;6 was conducted (Baron & Griffin, in preparation). Children saw uninformative (same gender) and informative (different-gender) article conditions with four objects (1 target and 3 distractors). A time course analysis of orienting to the same-gender (Figure 1; red line) or different-gender (Figure 1; blue line) target among distractors (light gray lines) from that study is included below.

Figure 1. Time course of children's looking to the target object and distractors in same-gender and different-gender conditions.



The bilingual children in this study oriented 43 ms faster to the different-gender trials, when the article was potentially informative; however, in contrast to Lew-Williams &

Fernald (2007), this was not a statistically significant difference (p = .16). Age of first exposure to English, current Spanish language use, and number of different words produced in a Spanish narrative sample were marginally significant predictors of whether a child was sensitive to gender. Although age was not a significant predictor, it was included in the best fit model and contributed to the variance explained by the model.

To date, there has only been one study that has focused on gender-marking in Spanish-English bilingual children using an eye-tracking paradigm. The bilingual children in this study did not show a gender-marking effect. Literature addressing adult second language learners would suggest that age of acquisition and proficiency affect a person's ability to be sensitive to gender. Thus, the gender-marking effect in bilingual children needs to be studied in more depth by taking age of acquisition and proficiency into consideration as mitigating factors.

PROCESSING OF GENDER-MARKED LANGUAGE IN MONO- AND BILINGUAL ADULTS

Researchers using eye-tracking paradigms with adults have revealed that adults whose native language is gendered, identify the referent more rapidly when the gender-marked article was informative than uninformative (Brouwer et al., 2017; Kroff, et al., 2016; Lew-Williams & Fernald, 2007; Morales et al., 2016). Adults started to use the gender cue approximately 500-1000 ms after article onset (Brouwer et al., 2017). Sequential bilingual adults who were born in Latin America, with a mean age of English acquisition of 9.2, and were relatively balanced bilinguals (with slightly more Spanish use), had similar results for the feminine conditions (informative and uninformative) as monolinguals, but there was no significant effect for the masculine condition for informative versus uninformative articles (Kroff et al., 2016). The authors explain that since *el*, the masculine article, is used extensively as the default article in code-switching,

this may lead bilinguals to ignore it as a cue when followed by a noun during comprehension. They also posit that this gender sensitivity difference may be due to attrition, which manifested first in the masculine article, or perhaps due to exposure to other bilinguals who may overproduce the masculine article and may not reliably use the masculine article with masculine nouns. Other researchers have found similar patterns and attributed this to interrupted or incomplete knowledge of gender in heritage speakers (early bilinguals) (Montrul et al., 2008).

Several experiments show that experience-related factors may influence how adults use grammatical gender in real-time processing. English-speaking adults, who learned Spanish as a second language (L2), did not show earlier looks to informative than uninformative gendered articles and were slower at identifying the target noun compared to Spanish monolingual peers (Lew-Williams & Fernald, 2010). Further, L2 learners of Spanish persisted in producing gender agreement errors even at higher levels of proficiency (Montrul et al., 2008). They tended to overgeneralize the masculine form, but were more accurate with increased exposure and proficiency (i.e. White et al., 2004). In spoken language comprehension and reading experiments, English-speaking adults who learned Spanish revealed sensitivity to grammatical gender similar to native speakers, but this sensitivity was affected by the level of proficiency (Dussias et al., 2013; Gillon-Dowens et al., 2010; Keating, 2009). Dussias et al. (2013) noted that high-proficiency English-speaking learners of Spanish processed grammatical gender agreement in article-noun pairs in a manner that was qualitatively equivalent, but slower, than native Spanish speakers. In contrast, low-proficiency late learners did not exhibit the ability to process gender like their monolingual peers. Additionally, Guillelmon and Grosjean (2001) conducted a study with English-French bilinguals and monolingual French speakers who heard sentences with a masculine, feminine or neutral article. Both correct (or congruent

gender) trials and incorrect (or incongruent gender) trials were compared to neutral gender trials. Early English-French bilinguals behaved like monolinguals in their sensitivity to gender congruency and incongruency conditions, which means they were faster on congruent trials than neutral trials and slower on incongruent trials than neutral trials. Late bilinguals, on the other hand, did not show a congruency or incongruency effect when controlling for the speed of response and gender-production skills.

L2 learners can attain native-like processing of grammatical gender when syntactic structures are similar across the two languages, but not when they differ (e.g. Foucart & Frenck-Mestre, 2011; Sabourin & Haverkort, 2003). In an ERP study for example, Dussias et al. (2013) found that low proficiency Italian learners of Spanish showed gender sensitivity in feminine target nouns but not for masculine nouns. Italian, like Spanish, has a two-gender system where Italian words that end in *–o* are generally masculine and words that end in *–a* are generally feminine. In this study, about 84% of the words used, shared the same gender between the two languages. Dussias and colleagues explain that about 17% of the words had an opaque gender, where the noun ended in a vowel other than *–o* and *–a* or a consonant, and of this subset of opaque words, 80% were masculine.

In sum, monolingual adult speakers of gendered languages (Spanish, French, Dutch) orient to targets faster when the gendered article is informative than uninformative. Gender sensitivity (both in informative vs. uninformative and matched vs. mismatched trials) in adults learning Spanish as a second language, is mediated by age of acquisition and proficiency. However, gender sensitivity is more pronounced in feminine target nouns rather than masculine target nouns. Researchers posit that this may be due to the amount of experience with both languages as well as overgeneralization or default use of the masculine gendered article.

PROCESSING OF PHONOLOGICAL COMPETITORS IN EYE-TRACKING PARADIGMS

During the recognition of spoken words, multiple word candidates are simultaneously activated and compete against each other (e.g., Marslen-Wilson & Welsh, 1978; McQueen, Norris, & Cutler, 1994). The cohort model, first proposed by Marslen-Wilson & Welsh (1978), explains how this process occurs. The first few phonemes of a spoken word activate a set or cohort of word candidates that are consistent with that target word. Activated word candidates compete for recognition until they no longer match incoming segmental information. As more acoustic information is analyzed, candidates that are no longer consistent with the target word drop out of the set. This process continues until only one word candidate matches the target word. For example, listeners will look at 'candy' when instructed to "pick up the candle" (Allopenna, Magnuson, & Tanenhaus, 1998). Within the visual world paradigm, while participants hear the name of a target picture, they look more often to pictures with names that are similar in onset with the target than to pictures with phonologically unrelated names. The phonological competitor activation can help determine if the gender-marked article preceding the noun can influence which lexical candidate enters the competitor set and therefore allows participants more time to process and recognize the correct noun (Dahan et al., 2000).

Phonological competitors have been found to delay fixations to targets (Allopenna et al., 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1996) but the delay is modulated by the presence of grammatical gender. Phonological competition between target and competitor items in different gender trials was attenuated by the presence of grammatical gender, allowing French-speaking participants to circumvent the effect of the non-target phonological competitor item on the bases of the preceding gender cue (Dahan et al., 2000). Dahan et al. (2000) tested the activation of competitors that matched

the initial sounds of a target noun but mismatched the gender marking on the article in French. They found that the presence of a gender-marked definite article could prevent early activation of competitors inconsistent with that gender. In other words, when phonological competitors matched in initial sounds with the target noun but mismatched in gender marking on the preceding article, early competitor activation was eliminated. Weber and Paris (2004) conducted a study with French participants who knew German. When a target and a phonological competitor were the same gender in both German and French, French participants fixated on the phonological competitors more than the distractors. However, when the target and phonological competitor were the same gender in German but different gender in French, early fixations to the phonological competitor picture were reduced. Phonological competitor activation in the non-native language (German) appeared to be constrained by native gender information as the French listeners used native French gender information to constrain competitor activation in German. Native German listeners did not exhibit this difference. Thus, phonological competitors can help explain how people narrow down the competitor set when listening to a stimulus and looking at multiple images.

WHAT DO REACTION TIMES TELL US?

Reaction times in word recognition/word processing and reading tasks provide some evidence for automatic language competence and processing (Jiang, 2007; Segalowitz, 2003). Quick and effortless responses reflect automatic processing, as they do not require additional attention resources or conscious control (Segalowitz, 2003). Domínguez, Cuetos, and Segui (1999) showed faster reaction times to masculine nouns than feminine nouns for a lexical-decision experiment in L1 Spanish speakers. Some studies have shown better performance on gender matched vs. mismatched contexts, but

Alarcón (2009) found that both beginner and advanced L2 learners of Spanish exhibited faster response times when the article and noun were mismatched in a sentence completion task while advanced learners were significantly faster at establishing agreement when the noun was feminine rather than masculine. In regards to gender reaction time research overall, conclusions are inconsistent and generalizations cannot yet be made as there is significant variation across experimental design and participant proficiency (Alarcón, 2009). Therefore, more research is needed in order to conclude what reaction times contribute to our understanding of gender processing.

THEORIES ADDRESSING GENDER SENSITIVITY

To account for the observations noted above in regards to gender sensitivity, multiple theories make predictions about the ways that listeners will respond to gender cues. The Failed Features Hypothesis (e.g. Hawkins & Franceschina), Full Transfer Full Access Model (e.g. Schwartz & Sprouse, 1996), and Competition Model (MacWhinney, 1987) are the theories that are considered. The Failed Features Hypothesis would predict that grammatical gender sensitivity is not possible for learners whose native language does not possess this feature as L2 adults are unlikely to take advantage of gendered article information (e.g. Hawkins & Franceschina, 2004). This is to say that learners of a gendered language can learn the gender of nouns, but learning is typically based on noun phonology. They do not appear to have a syntactic reflex of gender agreement and therefore gender processing does not become an integral part of their system. Thus, we would expect only native speakers of a gendered language to have a strong representation of this grammatical feature and to accurately produce article-name pairs.

In contrast, the Full Transfer Full Access Model (Schwartz & Sprouse, 1996; White, 2003) states that as one begins to learn the L2, the representation of grammatical

features is based on the L1, but as people have 'full access' to universal grammar, the new grammatical features can be acquired in the L2. Researchers have noted that L2 learners of a gendered language are sensitive to gender agreement violations as the structure is unique to the L2. This converges with the assumptions of the 'full access' model (Foucart & Frenck-Mestre, 2012; Tokowicz & MacWhinney, 2005). As grammatical gender agreement does not exist in English, it does not lead to cross-linguistic competition; but if both languages are gendered, speakers transfer the agreement rules from their L1 to their L2 (Foucart & Frenck-Mestre, 2011; Sabourin & Stowe, 2008). Therefore, in eye-tracking paradigms we expect all L2 learners to show gender sensitivity, regardless of the native language, as all speakers can acquire new grammatical features. Although grammatical gender does not exist in a language like English, this model does not posit any competition between the two languages so gender sensitivity can be acquired and maintained without interference from another language.

One can also consider gender-marked articles in terms of the Competition Model, in which the utility of a cue's strength varies as a result of learning and processing (MacWhinney, 1987). Forms are initially transferred on the basis of their ability to apply to new cases. However, if this transfer leads to an error or is unnecessary, the strength of the transfer is weakened. As English does not have gendered articles, the strength of the cue may be weakened. Four patterns emerge when considering cue strength: (1) transfer of L1 onto L2, (2) abandonment of L1 for L2, (3) merger of L1 and L2, and (4) partial attainment of separate L1 and L2 systems. The merger of L1 and L2 or partial attainment of separate L1 and L2 systems are the patterns that are most likely to affect gender processing. Early on in learning, the concept being expressed (gender sensitivity) would be more strongly associated with the form consistent with the L1 contingencies (ex. *el gato* [the.MASC cat]), than with the form consistent with the L2 contingencies (ex. *the*

cat) (Trenkic, 2009). With more experience of an L2 where gender is not marked, though, the strength of the association with no gender (*the cat*) is likely to increase. Another factor determining the outcome of the competition is the cognitive architecture and mechanisms involved in language processing, and the capacity-limited nature of working memory (MacWhinney, Bates, & Kliegl, 1984). The two competing referential forms differ in terms of their complexity, with *el gato* being structurally more complex than *the cat* due to the additional gender component. Other concurrent processes and representations may restrict the resources available for referential processing. The more demand other processes make on the limited resources, the more likely it is that they will encroach on the space needed for gender processing/sensitivity. This, in turn, will cause the processing of the more complex expression (*el gato*) to become increasingly unaffordable or unnecessary, and leave the simpler form (*the cat*) the winner. As more cognitive resources become available with increased proficiency, there will be fewer instances in which resources are exceeded to the extent where they will completely preclude the processing of the more complex expression (*el gato*). Thus, proficiency/current language use may lead to predictable patterns of gender sensitivity. If representation of articles in the L1 and L2 merge, within this process, we would expect to see changing levels of gender processing and accuracy.

The Competition Model posits that the knowledge and acquisition of another language will affect grammatical features in both languages. Within the Full Access Full Transfer Model, however, grammatical features in the first language are firmly rooted and thus should not be affected by ones learning of another language.

CURRENT STUDY

In order to investigate how Spanish-English bilinguals process and attend to gendered articles in Spanish, a visual world paradigm was used to examine gender-marked articles in multiple contexts: informative articles (different-gender grammatical trials), incorrect articles (ungrammatical trials), or uninformative articles (same-gender grammatical trials) in bilingual children and adults. The following research questions were addressed:

1a. Do bilingual children (ages 5-6 and 8-9) and adults take advantage of informative grammatical gender marking on articles (rather than uninformative) when listening to Spanish, as defined by looking on target objects earlier when they are the item in a display with a particular grammatical gender? If they do take advantage, do they do so to a similar extent?

1b. Do bilingual children (ages 5-6 and 8-9) and adults take advantage of matched grammatical gender marking on articles (rather than mismatched) when listening to Spanish, as defined by looking on target objects earlier when they are the item in a display with a particular grammatical gender? If they do take advantage, do they do so to a similar extent?

2. Do experiential (age of first exposure to English, cumulative exposure to English, current Spanish language input and output) and Spanish language measures (a vocabulary measure, article-noun naming accuracy, and grammaticality judgment accuracy) influence bilinguals' sensitivity to gender?

HYPOTHESES

Research Question 1

In previous research, when adults were compared to children on the same task, adults were significantly faster than the children and both groups showed sensitivity to gender where they oriented faster on the informative trials than the uninformative trials (Lew-Williams & Fernald, 2007). Accordingly, for Research Question 1a, The Full Access Full Transfer Model would predict that older groups, who have more language experience, should be successively faster than younger groups in orienting to the target noun in the informative trials than the uninformative trials. Further, this model would predict that all participants (regardless of age) would show some degree of gender sensitivity, as Spanish is the L1 and grammatical features are based on the L1. By contrast, according to the Competition Model framework, we expect to see sensitivity to gender in young children with less experience and exposure to English (5 to 6-year-olds), no sensitivity to gender in older children as they have more experience and exposure to English (8 to 9-year-olds), and then again see gender sensitivity in adults who have extended experience and exposure to both languages. When speaking exclusively (or mostly) Spanish at home, younger children are considered less proficient in English and therefore the strength of gender sensitivity as a cue should be strong and not weakened by English. However, as children's interactions increase in English speaking environments in school and community at large, they will be expected to use more English, which will lead to a higher proficiency in English, while still acquiring and maintaining their Spanish. Competition between the languages may weaken the gender sensitivity cue. As adults have more cognitive resources available and have increased language proficiency and cumulative exposure, they will be faster at orienting to the target noun on the informative trials than the uninformative trials. The bulk of the adult second language

literature points to the finding that gender sensitivity is impacted by age of acquisition and proficiency (Dussias et al., 2013; Gillon-Dowens et al., 2010; Keating, 2009; Lew-Williams & Fernald, 2010). Even adults who grew up speaking a gendered L1 and learn a new gendered L2 later, show gender sensitivity in their L2 to those nouns that share the same gender as their L1 (Dussias et al., 2013). Thus, we expect that the Competition Model will best explain gender sensitivity across these three age groups and will result in an interaction of age group and trial type.

For Research Question 1b, we expect that all participants will be faster at orienting to the target noun on the uninformative than the ungrammatical trials. This again supports the Full Access Full Transfer Model as gender is a grammatical feature that exists in all participants' L1. Additionally, Gillelmon and Grosjean (2001) noted that early bilinguals showed sensitivity to matched (uninformative) and mismatched (ungrammatical) trials and van Heugten and Shi (2009) saw that the ungrammatical nature of the articles continued to affect processing efficiency until later in the trial. As this is a matter of grammaticality (is the article-noun pair grammatically correct or not), the Competition Model also predicts that all participants will be faster at the uninformative trials than the ungrammatical trials as the grammaticality of a sentence is considered in both English and Spanish and does not lead to competition or weakening of a cue. In sum, both models have the same prediction for all age groups in response to matched and mismatched gender sensitivity.

Research Question 2

Research Question 2 focuses on the influence of experiential and language measures on an individual's ability to show gender sensitivity. As other researchers have found that gender sensitivity is influenced by proficiency of the gendered language,

current Spanish language input and output are expected to be predictive factors (Dussias et al., 2013; Gillon-Dowens et al., 2010; Keating, 2009). Guillelmon and Grosjean (2001) noted that early bilinguals showed gender sensitivity while late bilinguals did not. Thus, we expect age of first exposure to English and cumulative exposure to English to be predictors of gender sensitivity as well.

Vocabulary size has been strongly associated with speed of comprehension in looking tasks and gender processing (Fernald, Perfors, & Marchman, 2006; Lew-Williams & Fernald, 2007; Marchman & Fernald, 2008). Additionally, in a previous study with bilingual children, number of different words (a measure of vocabulary from a narrative sample) was a predictor of gender sensitivity (different gender trials vs. same gender trials) (Baron & Griffin, in prep). Others have found that receptive vocabulary scores were not determining factors in gender processing (Brouwer, Sprenger, & Unsworth, 2017). These differences, however, may be due to differences in the language pairs studied. As such, participants' vocabulary score was hypothesized to be significantly predictive of gender sensitivity in the current study.

While comprehension precedes production in typical language development, the two processes follow different timetables with respect to onset, mastery, and rate and trajectory of development. Different measures of comprehension and production can yield varied results in regards to language skill, as comprehension and production are dissociated psycholinguistic properties that draw on different skills and neurological bases (Bates, 1993). Bornstein and Hendricks (2012) examined over 100,000 2 to 9-year-old children from 16 developing countries to determine the relation between comprehension and production. In Bornstein and Hendricks' sample, comprehension slightly exceeded production, and the two were positively correlated. Given the close link

between comprehension and production, article-noun naming accuracy was hypothesized to be a significant predictor as it is the only gender production task.

Lastly, grammaticality judgment accuracy was hypothesized to be significantly predictive of gender sensitivity as it captures the participants' metalinguistic awareness and offline comprehension of gender processing, while the eye-tracking task as the outcome variable, looks at gender from an online comprehension perspective.

Chapter 2: Methods

PARTICIPANTS

Thirty-four bilingual Spanish-English neurotypical children and 39 heritage Spanish-speaking adults, ages 18-35 were recruited from the central Texas community. Seventeen 5-6 year-olds and 17 8-9 year-olds qualified to participate. All participants were exposed to Spanish since birth. They had normal or corrected-to-normal vision and normal hearing as determined by parent and participant report prior to the start of the study. The exclusionary factors were: focal brain injury, severe social-emotional problems, genetic syndromes, mental retardation, autism spectrum disorder, hearing loss, and speech or language disorders. Four child participants and 3 adult participants were additionally excluded due to difficulty tracking their eye movements. Mother's education level was used as a proxy for socioeconomic status (SES) for children. The adult's education level was used for SES. Using the Hollingshead (1975) index, education was scored from 0 to 7 (0 = no formal education, 1 = less than a 7th grade education, 2 = junior high school, 3 = partial high school, 4 = high school graduate, 5 = partial college, 6 = university degree, and 7 = graduate or professional degree).

Child participants were paid for their participation and adult participants were either paid or received university class credit. Young child and older child participant analyses are based on data from 15 participants each and 36 adult participants. The participant characteristics are included in Table 1. As current daily input and output in Spanish were correlated at .94, they were combined into one composite variable.

Table 1. *Participant characteristics in means and standard deviations*

Characteristic	Young	Older	Adults
Sex	8F, 7M	6F, 9M	27F, 9M
Age (years;months)	6;1 (0;6)	9;3 (0;6)	20;9 (2;2)
SES (Mother's Education)	4.07 (1.94)	3.73 (2.22)	4.30 (1.90)
Age of first exposure to English (years)	1.67 (1.72)	2.60 (1.80)	3.97 (2.20)
Spanish Input (percent)	65.94% (9.36)	65.67% (8.20)	22.05% (13.96)
Spanish Output (percent)	55.62% (10.09)	58.35% (12.61)	17.70% (14.67)

MATERIALS

Standardized Measures

Inventory to Assess Language Knowledge (ITALK) and Bilingual Input Output Survey (BIOS)

The ITALK and BIOS are questionnaire subtests within the Bilingual Spanish English Assessment (BESA; Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2018). In the ITALK, adult participants and parents answered questions regarding the participant's language abilities (i.e. vocabulary use, speech production, comprehension, grammar, and sentence production). Each area of language ability was rated in both languages on a 5-point scale, where 1 = minimal proficiency in the area and 5 = high proficiency in the area. If the parent was unable to answer a question regarding one of the areas, it was marked as 'unknown' and was not included in the average. The 5 scores were averaged in each language. In the BIOS, adult participants and parents detailed the history of exposure to both languages at home and school/work environments since birth to calculate the child's age of first exposure to English, languages spoken at each year of

the participant's life, and reported the language input and output at home and at school/work on an hourly basis during the week and on weekends. The interview was administered in person to adult participants and parents.

Expressive One Word Picture Vocabulary Test (EOWPVT)

The Expressive One-Word Picture Vocabulary Test – Fourth Edition (EOWPVT-4; Martin & Brownell, 2011) and the Expressive One-Word Picture Vocabulary Test-4: Spanish Bilingual Edition (EOWPVT-4 SBE; Martin, 2013) are norm-referenced tests of single-word expressive vocabulary that were used to provide a gross measure of cumulative vocabulary knowledge in each language and an index of proficiency. Each test was administered one time to all participants. For the current study, the EOWPVT-4 and EOWPVT-4 SBE were administered as English-only and Spanish-only versions, respectively. If the participant responded in the non-target language, the examiner prompted the participant to respond in the target language. Per the test manuals, basal was achieved when 8 correct consecutive items were obtained and ceiling was achieved when 6 consecutive incorrect items were obtained. Subsequently, raw scores were computed for each language.

Eye-tracking task

Design

Thirty-six Spanish nouns (18 masculine and 18 feminine) were included as experimental stimuli. Twelve filler items were included for a total of 48 stimuli. Two practice items were presented at the beginning of the task to allow participants to become accustomed to the nature of the experiment. The target location was counterbalanced such that targets appeared in each quadrant on the screen the same number of times.

Three groups of experimental stimuli were prepared; one group with informative (different-gender) articles, one group with uninformative (same-gender) articles, and one group with incorrect (ungrammatical) articles. A target appeared only once in each presentation list, preceded by the correct (or incorrect) gender marking surrounded by distractors with the same or different gendered articles. There were 3 presentation lists such that each target occurred in each condition across the lists. For example, *el conejo*, occurred in the same gender condition in List 1, in the different gender condition in List 2, and the ungrammatical condition in List 3. Each subject completed one presentation list.

Participants heard a sentence 'enséñame + el/la + target noun'. A male speaker of Mexican Spanish recorded each sentence. For the ungrammatical sentences, the article was spliced from a grammatical sentence with the same initial consonant of a different target noun and was inserted into the ungrammatical sentence so that the sentence sounded natural. The articles *el* and *la* were unstressed within the sentence. 'Enséñame (show me)' was used as the instruction because in Spanish, if one says 'mira a (look at) + el', the a + el is combined to form 'al' while a + la stays unchanged. Therefore, 'enséñame', which has also been used in previous gender processing studies, was selected for the instructions (Lew-Williams & Fernald, 2007, 2010).

For example, the participants heard 'enséñame la cama' [show me the.FEM bed] where *la cama* was the target. For the same gender condition, *la casa* [the.FEM house] was the phonological competitor and *la pelota* [the.FEM ball] and *la jirafa* [the.FEM giraffe] were the distractors (Figure 2). For the different gender condition, *el carro* [the.MASC car] was the phonological competitor and *el guante* [the.MASC glove] and *el tenedor* [the.MASC fork] were the distractors (Figure 3). For the ungrammatical condition, participants heard 'enséñame el cama' [show me the.MASC bed] were *la cama*

was still the target, *el carro* was the phonological competitor and *el guante* and *el tenedor* were the distractors (Figure 3). Although the pictures were the same for the different gender and ungrammatical conditions, the auditory stimulus was different. The different gender condition had a target that was grammatical (la cama), while the ungrammatical condition had a target that was ungrammatical (el cama). Please see Appendix A for an example stimuli presentation list.

Figure 2. Example of same gender trial in the eye-tracking task

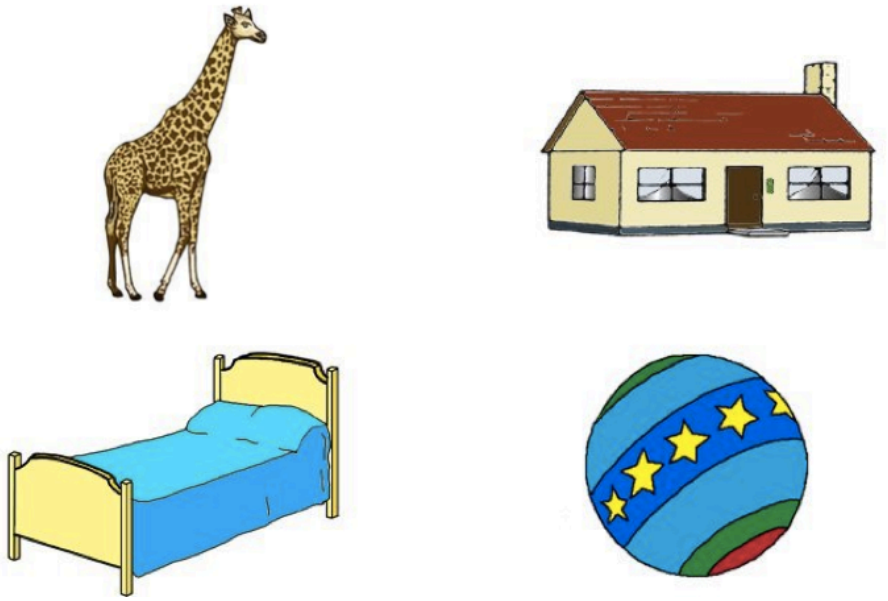
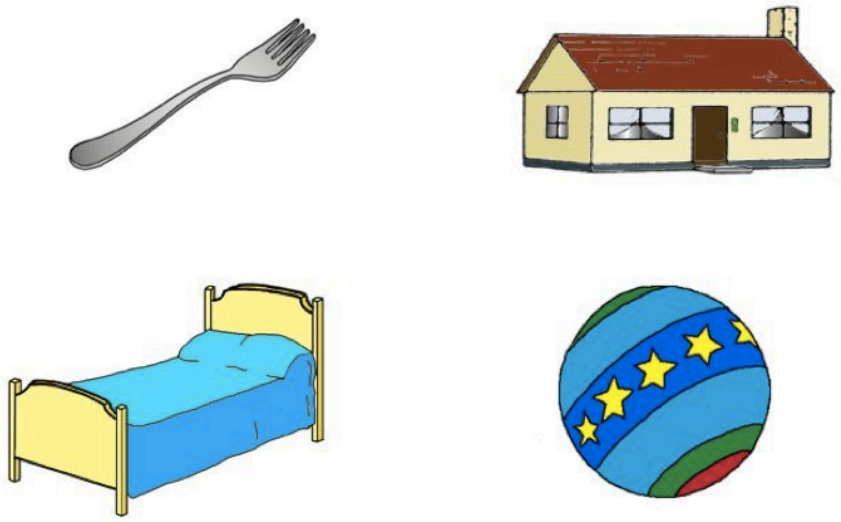


Figure 3. Example of different gender/ungrammatical gender trial in the eye-tracking task



Stimulus characteristics

Selection of noun target and phonological competitors was restricted to words with the same initial consonant-vowel. The phonological competitor in each item had a similar syllable length to the target noun (+/- 1 syllable), which led to comparable salience across the 2 objects.

The other two distractors in each stimulus also began with consonants (except for /l/). The two distractors in each stimulus used for the different gender and ungrammatical gender conditions, stayed together for a different target stimulus in the same gender condition across stimuli lists. Additionally, items from different categories (animals, foods, furniture, transportation, musical instruments, clothing, etc.) were presented together with the exceptions of cuchillo/cuchara (knife/spoon), pavo/pato (turkey/duck), tiburón/tigre (shark/tiger), zanahoria/salsa (carrot/salsa), and canguro/caballo (kangaroo/horse), as they were phonological competitors and were the most closely matched selections based on noun frequency and same initial consonant-vowel.

Additionally, animals are high frequency words that most 5-year-old children are familiar with, and thus some of the distractors were animals even if the target or phonological competitor was an animal.

Based on the Corpus del Español, the phonological competitors had an average lexical frequency that was higher (M = 20418) than the frequency of the target nouns (M = 11326). The phonological competitor lexical frequency between item sets was not significantly different (List 1 vs. 2 $t(35) = 1.55$, $p = .13$; List 2 vs. 3 $t(35) = -1.72$, $p = .09$; List 1 vs. 3 $t(35) = .44$, $p = .66$). The distractors' lexical frequency was similar across the same (List 1 vs. 2 $t(11) = 1.67$, $p = .12$; List 2 vs. 3 $t(11) = -0.65$, $p = .53$; List 1 vs. 3 $t(11) = 0.82$, $p = .43$), different (List 1 vs. 2 $t(11) = -0.82$, $p = .43$; List 2 vs. 3 $t(11) = 1.19$, $p = .26$; List 1 vs. 3 $t(11) = 0.57$, $p = .58$), and ungrammatical (List 1 vs. 2 $t(11) = -0.57$, $p = .58$; List 2 vs. 3 $t(11) = -0.81$, $p = .43$; List 1 vs. 3 $t(11) = -1.19$, $p = .26$) gender conditions between item sets.

One hundred and eleven colored pictures were selected from Snodgrass and Vanderwart (1980). A subset of the items had imageability, concreteness, and familiarity ratings (subjective ratings) using EsPal (Duchon, Perez, Sebastián-Gallés, Martí, & Carreiras, 2013; retrieved from www.bcbl.eu/databases/espal/) (81% of the feminine items, 70% of the masculine items). All 3 categories use a scale from 1 to 7, with 7 indicating fully imageable, concrete, or familiar. The feminine items had average ratings of 6.07, 6.16, and 5.88 on familiarity, imageability, and concreteness respectively. The masculine items had average ratings of 5.89, 6.16, and 5.99 on familiarity, imageability, and concreteness respectively. Ratings across feminine and masculine items were not significantly different. Pictures were normed for naming agreement in Spanish, and pictures above 80% were used as targets, while pictures below 80% were used in filler items. The pictures were edited to fit within 462 x 334 pixels.

Procedure

The experiment was built in Eyelink Experiment Builder software and eye movements were recorded on an Eyelink 1000 head-mounted tracking device designed by SR Research. Viewing was binocular but eye movements were recorded from one eye only. Participants were tested individually. Stimuli were presented on a 27-inch monitor, with participants seated approximately 67cm from the monitor with chins on a chin rest. Sampling rate was 500 Hz. At the beginning of the task, participants were instructed to use a mouse to click on the object on the screen that was referred to in the sentence. To begin each trial, participants looked at a fixation point in the center of the computer screen before the 4 images were displayed. Participants could not see the images if they did not fixate on the point in the center of the screen. Once they fixated on the point, 4 pictures appeared on the screen and 500 ms later the sentence was auditorily presented. Participants needed to click on a picture to end the trial. If necessary, recalibration of the eye-tracker was performed between trials. In each trial, fixations were recorded from the onset of the images on the screen until the participant clicked on an image. Latencies were recorded for mouse-click responses.

Article-noun pair naming task

Participants completed a familiarization task, which included all 234 of the pictures used in the eye-tracking experiment. Participants were instructed to name each picture in Spanish with its corresponding definite article (el [the.MASC] or la [the.FEM]). This task was included to see if participants were able to assign the correct name and gender to each picture. Participants were provided with the correct response if they did not label the item in Spanish or were prompted if they provided a non-target name. Participants were not asked to repeat the target name after the model was given,

but many did so spontaneously. Each response was recorded and participants' production accuracy on gender-marked articles was calculated where a score of 1 was given to correctly named pictures and a score of 0 was given to incorrectly named pictures or if the participant was unable to name the picture.

Grammaticality judgment task

Thirty-six targets and 12 fillers from the eye-tracking task were presented in sentences. The target noun was embedded in a simple 5-10 word sentence. Two practice sentences were presented at the beginning of the task to allow participants to become accustomed to the nature of the experiment (e.g. Monté el camello en el desierto [I rode the camel in the desert]. Half of the sentences were presented as grammatical sentences and half as ungrammatical sentences. On a keyboard, participants were asked to press one of two buttons indicating if the sentence they heard was grammatical or ungrammatical. The grammaticality across the eye-tracking task and grammaticality judgment task were the same in order to directly compare participants' ability to identify grammaticality of the targets offline within sentences. If a target noun was presented in the eye-tracking task as grammatical, participants heard the noun in a simple grammatical sentence within the grammaticality judgment task. For example, if the target in the eye-tracking task was la cama [the.FEM bed.FEM], then the sentence participants heard during the grammaticality judgment task was "Todos los días, duermo en la cama" [Every day, I sleep in the.FEM bed.FEM]. If however, participants heard el cama [the.MASC bed.FEM], which is ungrammatical, then the sentence participants heard during the grammaticality judgment task was "Todos los días, duermo en el cama" [Every day, I sleep in the.MASC bed.FEM]. Filler ungrammatical sentences were presented from the target nouns of the filler items within the eye-tracking task. There were 3 grammaticality

judgment lists to mirror that of the eye-tracking task as the grammaticality of the sentence presented depended on the grammaticality of the target in the eye-tracking task. There were 12 sentences from each of the conditions (same gender, different gender, ungrammatical) and 12 filler ungrammatical sentences. Participants were asked to listen to the sentences carefully and focus on the grammaticality of the sentence. The participant's button press was recorded for accuracy and reaction time. Please see Appendix B for an example list of grammaticality judgment sentences.

STATISTICAL ANALYSES

Behavioral Measures

One-way analyses of variance (ANOVAs) were conducted for all variables to examine the effect of participant age (3 age groups) on each behavioral tasks (EOWPVT raw scores, article-noun naming accuracy, grammaticality judgment accuracy) and experimental tasks (time to first fixation, and reaction times of the grammaticality judgment and eye-tracking tasks). ANOVAs were followed by pairwise comparisons to identify significant differences between groups. As there were multiple comparisons for each behavioral task, the Bonferroni correction of .025 was employed. Due to computer error, one young child is missing a score for grammaticality judgment accuracy and one young child is missing an article-noun naming accuracy score. Therefore, there is slight variation in the degrees of freedom for the ANOVAs and pairwise comparisons.

Research Question 1

To address Research Question 1, initially an omnibus 3 (age) × 3 (trial type) analysis of variance (ANOVA) was conducted to identify any main effects of age or trial type or an age by trial type interaction on the variable of eye-tracking first fixation. An interaction of age and trial type was predicted, as young children are expected to be faster

at looking at the target in the different gender condition than the same gender condition, older children are not expected to show a difference in looking time to the target after article onset between the conditions, and adults are expected to show a difference in looking time to the target. Subsequently, two 2 (age) × 2 (Informativeness) mixed ANOVAs were conducted to identify: any main effect of the between-subjects factor of age: (1) young vs. older children and (2) older children vs. adults), any main effect of within-subjects factor of Informativeness: (1) different gender trials vs. (2) same gender trials on the dependent variable of eye-tracking first fixation to the target noun after article onset, and any (Age x Informativeness) interaction. To address Research Question 1b, two 2 (age) x 2 (Grammaticality) mixed ANOVAs were also conducted to identify: any main effect of the between-subjects factor of age: (1) young vs. older children and (2) older children vs. adults), any main effect of the within-subjects factor of Grammaticality: (1) same gender trials vs. (2) ungrammatical trials on the dependent variable of eye-tracking first fixation to the target noun after article onset, and any (Age x Grammaticality) interaction.

Research Question 2

To address Research Question 2, predictors related to language experience and language knowledge were entered into 4 regression models. In a preliminary inspection of the data, correlations were examined to ensure that the multicollinearity assumption was not violated and to see the strength of association between variables. Outcomes for the regressions were based on difference scores, which were calculated by subtracting mean of first fixation on informative (different-gender) articles from uninformative (same-gender) articles and mean of first fixation on uninformative (same-gender articles) from incorrect (ungrammatical) articles. Experiential (age of first exposure to English,

cumulative exposure to English, and percent current daily input and output in Spanish) and language (Spanish vocabulary raw score on the EOWPVT and percent grammaticality judgment accuracy) measures were included. Three types of regression models, forward, backward, and stepwise, were employed to address the question of which combination of language experience and language knowledge measures influenced gender sensitivity as indicated by the fewest number of independent variables that explained the most variance.

Chapter 3: Results

BEHAVIORAL MEASURES

Table 2 shows performance data for all behavioral tasks. In order to understand the behavioral task scores across groups, means and standard deviations were calculated for EOWPVT in Spanish and English, percent article noun naming accuracy, and percent grammaticality judgment accuracy. Young bilingual children in this study had similar vocabulary knowledge in English and Spanish $t(14) = 1.54$, $p = .15$ and older bilingual children and adults had a higher English vocabulary score than Spanish $t(14) = 4.42$, $p < .001$, $t(35) = 7.33$, $p < .001$ respectively. For the EOWPVT Spanish and English, article-noun naming accuracy, and grammaticality judgment accuracy, scores and percentages were numerically greater for each increasing age group. Young children and adults had higher grammaticality judgment accuracies than article-noun naming accuracies $t(12) = 3.15$, $p = .008$ and $t(35) = 3.91$, $p < .001$ respectively. The grammaticality judgment accuracy and article-noun naming accuracy did not differ for older children $t(14) = .05$, $p = .96$.

Table 2. *Mean values (and standard deviations) for language measures*

Measure	Young Children	Older Children	Adults
EOWPVT Spanish (raw score)	45.93 (14.86)	66.53 (16.59)	107.92 (20.41)
EOWPVT English (raw score)	58.8 (24.37)	95.2 (21.00)	139.58 (14.63)
Article-noun naming accuracy (percent)	47.66 (19.27)	79.63 (16.70)	90.06 (.08)
GJ accuracy (percent)	60.32 (14.31)	79.81 (15.87)	94.14 (6.89)

Note. EOWPVT = Expressive One Word Picture Vocabulary Test, GJ = grammaticality judgment.

Analyses of variance (ANOVAs) were conducted to compare all language measures (EOWPVT raw scores, article-noun naming accuracy, grammaticality judgment accuracy) across the 3 age groups and are included, with pairwise comparisons, in Table 3. ANOVAs for EOWPVT English and Spanish, article-noun naming accuracy and grammaticality judgment accuracy were all significant. Older children had higher scores and accuracies than young children on all behavioral measures. Adults had higher scores on the EOWPVT Spanish and English and grammaticality judgment accuracy than older children. Adults and older children were not significantly different from one another on article-noun naming accuracy.

Table 3. *ANOVAs and pairwise comparisons of behavioral measures*

Behavioral Measure	ANOVA		Pairwise comparisons			
		p-value	Young/Old	p-value	Old/Adult	p-value
EOWPVT Spanish	F(2,63) = 68.62	<.001	t(14) = -4.13	0.001	t(32.1) = -7.57	<.001
EOWPVT English	F(2,63) = 106.10	<.001	t(14) = -4.20	<0.001	t(19.9) = -7.47	<.001
Article-noun naming accuracy	F(2,62) = 50.48	<.001	t(13) = -5.33	<0.001	t(17.0) = -2.30	0.034
Grammaticality Judgment accuracy	F(2,62) = 46.21	<.001	t(13) = -4.74	<0.001	t(16.2) = -3.32	0.004

Note. Young = 5-6 year-old children, Old = 8-9 year-old children, EOWPVT = Expressive One Word Picture Vocabulary Test

RESEARCH QUESTION 1

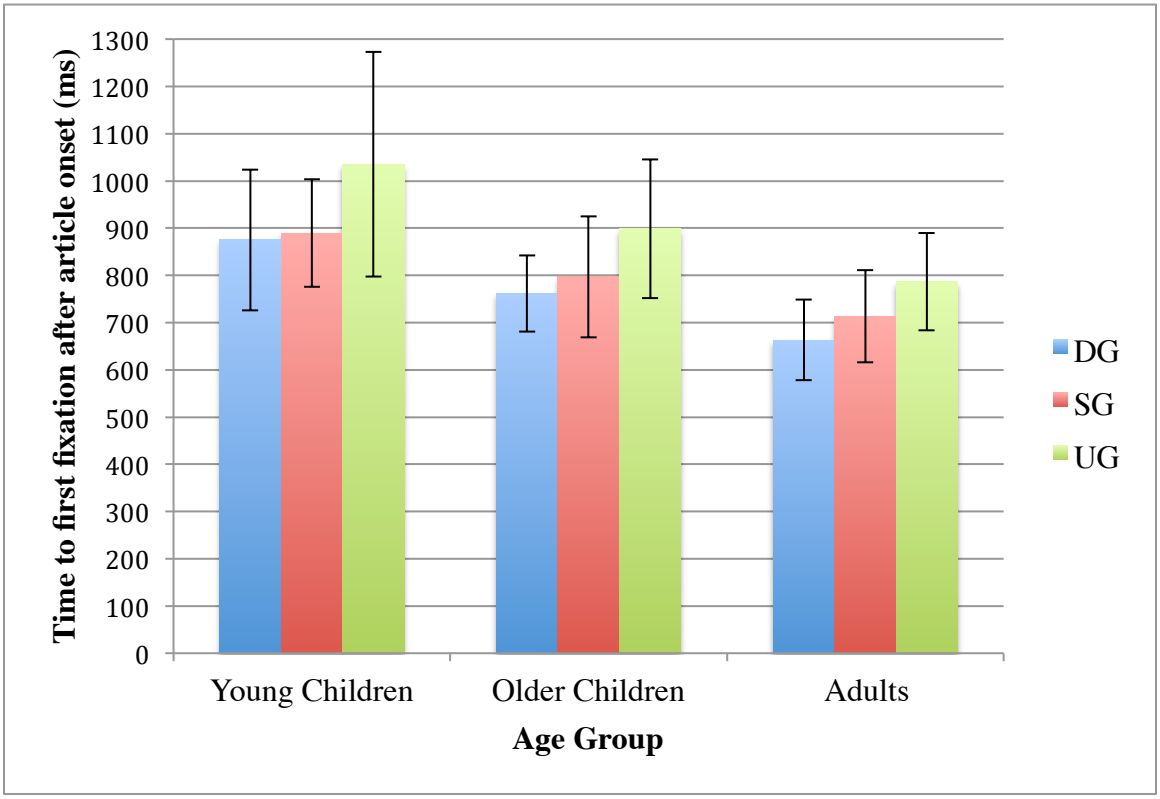
First fixation to the target was calculated to examine whether bilinguals responded faster to informative than uninformative than ungrammatical trials, the first research question. First fixation was defined as a look after 200 ms from article onset as it requires approximately 200 ms to plan and execute a saccade (Fukushima, Hatta, & Fukushima, 2000; Martin, Shao, & Boff, 1993). Mean time of first fixation, and reaction times for the grammaticality judgment and eye tracking tasks were calculated for each subject for each trial type and shown in Table 4. Figure 4 depicts first fixations to the target after article onset by age group.

Table 4. *Mean values (and standard deviations) for experimental measures*

Measure	Young Children	Older Children	Adults
Eye-tracking first fixation to target noun			
DG	875 (149)	759 (81)	663 (85)
SG	890 (114)	792 (128)	715 (98)
UG	1035 (238)	908 (147)	786 (103)
Reaction times			
Eye-tracking task			
DG	4624 (768)	3626 (767)	2890 (341)
SG	4077 (556)	3636 (508)	2932 (301)
UG	4833 (1004)	4131 (1089)	3222 (294)
Grammaticality judgment task			
DG & SG	5425 (3367)	4097 (1066)	2877 (629)
UG	5557 (3254)	4048 (1004)	2981 (736)

Note. DG = different gender, SG = same gender, UG = ungrammatical gender

Figure 4. Mean time to first fixation to target noun after article onset by age group (error bars represent the standard deviation)



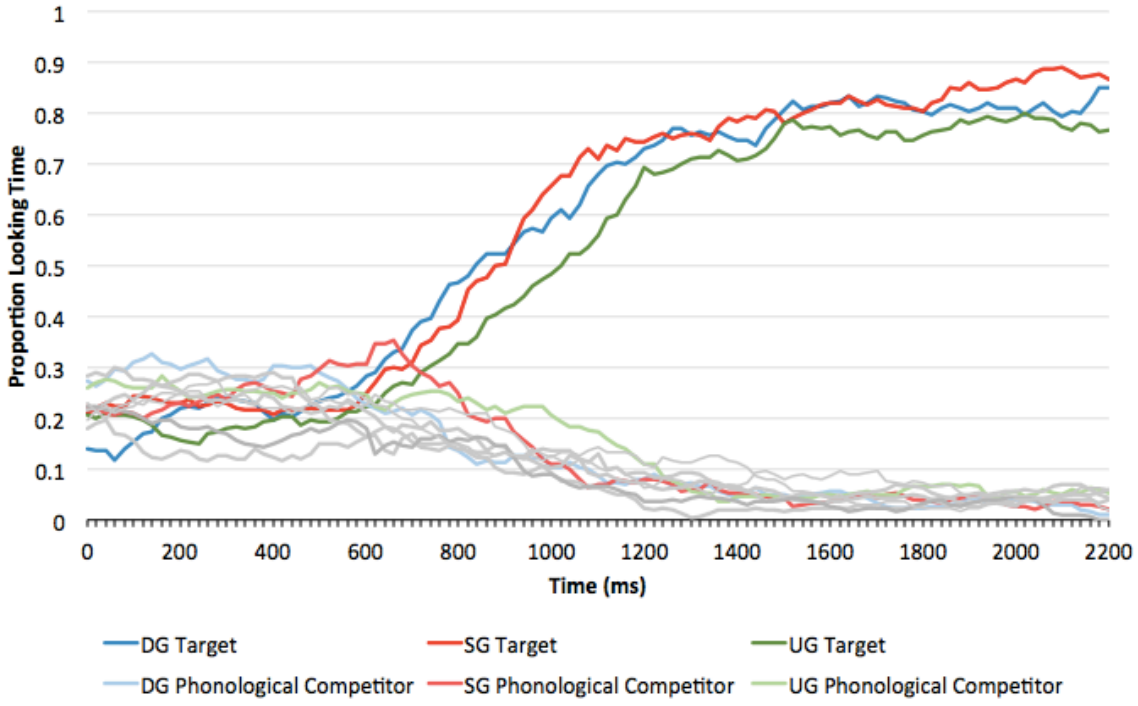
Note. DG = different gender, SG = same gender, UG = ungrammatical gender

VISUAL REPRESENTATION

Although time to first fixation was calculated, it is also important to visualize eye-tracking fixations as they show the temporal aspect of the data for all three experimental conditions, their distractors, and their phonological competitors, for each age group. To be able to directly visualize the data to be used for statistical analysis is beneficial. Visualization is useful for detecting unexpected patterns, finding alternative explanations and generating new hypotheses (Fox & Hendler, 2011). Eye-tracking trials were excluded if there was no target fixation (1.20%, 1.08% of total trials for children and adults respectively), if it took more than 2000 ms for first fixation to target (1.9%,

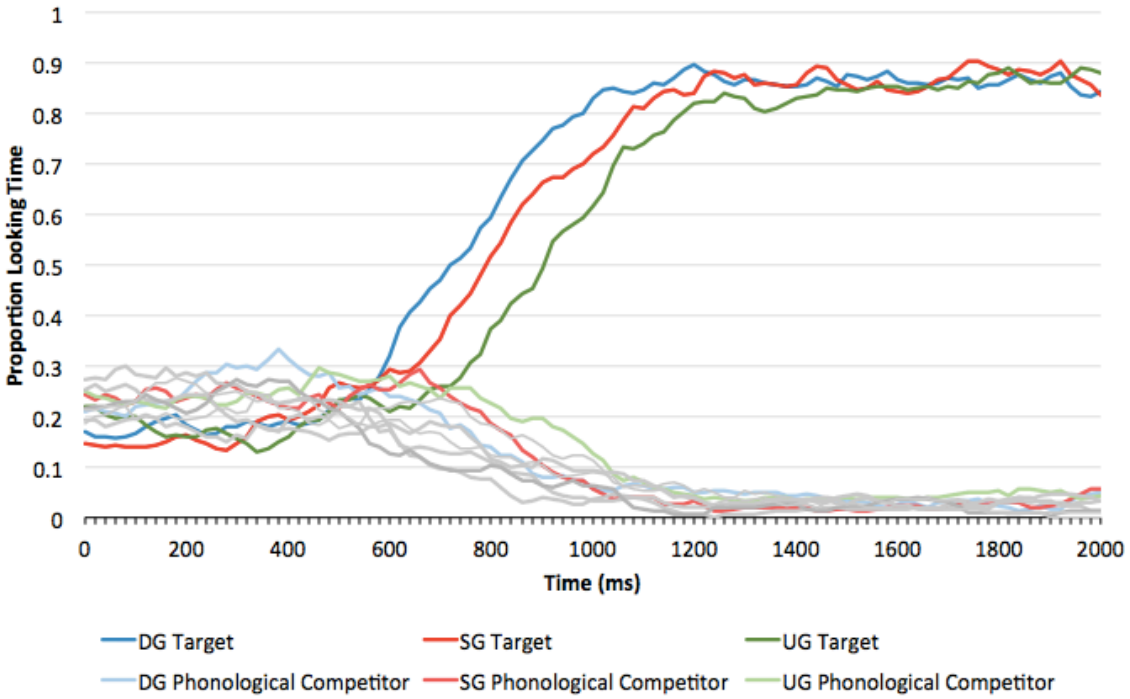
0.31% of total trials for children and adults respectively), if the mouse click response time was more than 10,000 ms (3.06%, 0% of total trials for children and adults respectively), if target fixation occurred before 200 ms post article onset and was maintained (11.39%, 0.46% of total trials for children and adults respectively), or if the participant clicked on an object other than the target (2.59%, 0% of total trials for children and adults respectively). The percent of total trials was calculated individually although there were some trials that had no target fixation and had a longer response time and/or clicked on an object other than the target were also excluded. Time courses of orienting to the target on informative (different-gender trials), uninformative (same-gender trials) and ungrammatical articles were plotted for young children, older children and adults (Figures 4, 5, & 6) using bins in 20 ms intervals.

Figure 5. Time course of young children looking to target



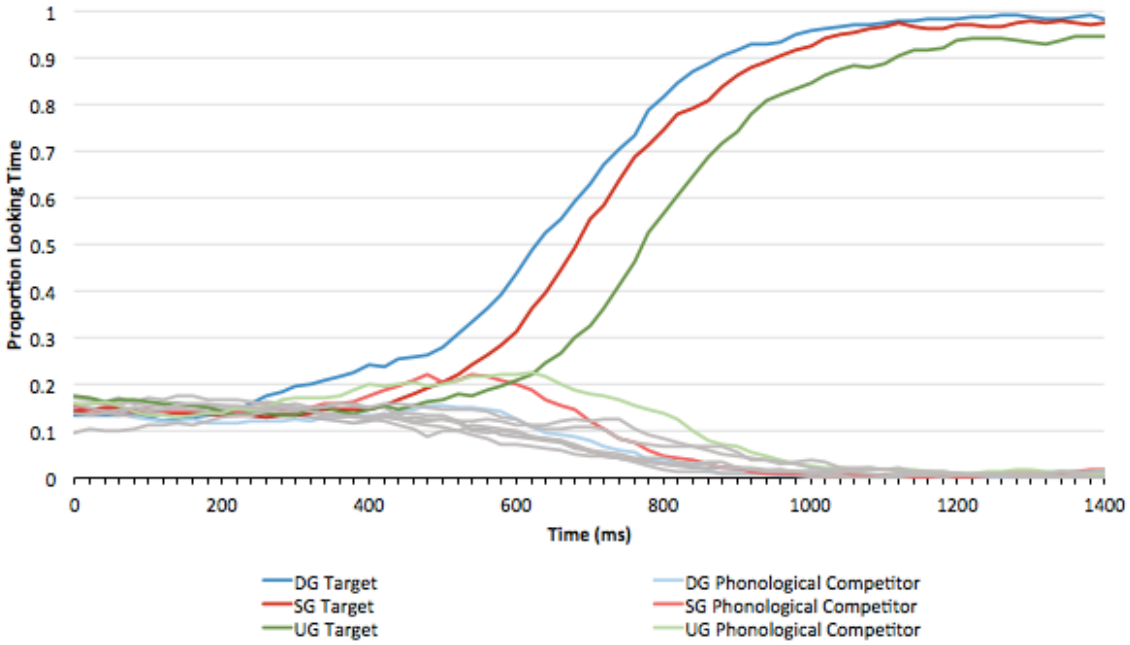
Note. DG = different gender, SG = same gender, UG = ungrammatical gender

Figure 6. Time course of older children looking to target



Note. DG = different gender, SG = same gender, UG = ungrammatical gender

Figure 7. Time course of adults looking to target



Note. DG = different gender, SG = same gender, UG = ungrammatical gender

Time in milliseconds is on the x-axis and begins at article onset. Proportion to looking time is shown on the y-axis showing proportionally how much time each group looks at each image. As participants heard the beginning of the article, they all looked at the 4 images equally. As participants continued to hear the rest of the article and then the noun, they looked at the distractor images (in gray) proportionally less and the target image proportionally more.

In the different gender condition, children looked at the phonological competitor (light blue line) slightly more than the distractors, but then looked at the target persistently and proportionally more of the time (dark blue line). In the same gender condition, participants looked at the phonological competitor (light red line) proportionally more than distractors and for a longer period of time. For a time,

participants looked at the phonological distractor the same amount or proportionally more time than the target (dark red line). Then participants looked proportionally more to the target than all other images.

For the ungrammatical condition, as the auditory stimulus continued to unfold, participants tended to spend proportionally more time looking at the phonological competitor (light green line) than they had for the phonological competitors in the other two conditions. They spent more time looking at the phonological competitor than the target before proportionally more looks occurred towards the target (dark green line).

An omnibus 3 (age group) × 3 (trial type) analysis of variance (ANOVA) was conducted where age group and trial type were the independent variables and eye-tracking first fixation was the dependent variable shown in Table 5. There were main effects for age group $F(2,63) = 26.57$, $p < .0001$, and trial type $F(1,126) = 32.72$, $p < .0001$. There was no interaction of age and trial type $F(4,126) = 0.78$, $p = .54$.

Table 5. *3X3 ANOVA of eye-tracking first fixation by age group and trial type*

Source	df	SS	MS	F	*p-value*
Group	2	1475012.18	737506.09	26.57	<.0001
Residuals	63	1748590.96	27755.41		
Trial Type	2	599449.78	299724.88	32.72	<.0001
Group*Trial type	4	28568.04	7142.01	0.78	.54
Residuals	126	1154105.48	9159.57		

As no interaction of age and trial group was seen in the omnibus test above as predicted, follow-up ANOVAs were conducted to disambiguate how age and Informativeness (different gender trials vs. same gender trials) and age and

Grammaticality (same gender trials vs. ungrammatical trials) impact time to first fixation after article onset. Two 2 (age group) × 2 (trial type) mixed ANOVAs were conducted where age group was the between subjects variable and Informativeness (different gender trials vs. same gender trials) was the within subjects variable and eye-tracking first fixation to the target noun after article onset was the dependent variable. Table 6 shows the results for a 2 × 2 ANOVA between the young and older children and Table 7 shows the results between the older children and adults.

Older children fixated on targets significantly earlier than younger children, as reflected in the main effect of age group $F(1,28) = 10.95$, $p = .003$. Although children fixated on targets earlier in the different gender condition than the same gender condition, the difference did not approach significance, $F(1,28) = 0.62$, $p = .44$; nor did the interaction of Age and Informativeness $F(1,28) = 0.10$, $p = .75$.

Table 6. *2X2 Mixed ANOVA of eye-tracking first fixation by young vs. older children and Informativeness*

Source	df	SS	MS	F	*p-value*
Group	1	159407.7	159407.7	10.95	0.003
Residuals	28	407632.06	14556.28		
Informativeness	1	9307.6	9307.6	0.62	0.44
Group*Informativeness	1	1539.05	1539.05	0.10	0.75
Residuals	28	420614.16	15021.93		

Adults fixated on targets significantly earlier than older children, as reflected in the main effect of age group $F(1,49) = 14.13$, $p < .001$. Adults and older children fixated on targets significantly earlier in the different gender condition than the same gender

condition, as seen in the main effect of Informativeness $F(1,49) = 5.90$, $p = .02$. The interaction of Age and Informativeness was not significant $F(1,28) = 0.10$, $p = .75$.

Table 7. *2X2 Mixed ANOVA of eye-tracking first fixation by older children vs. adults and Informativeness*

Source	df	SS	MS	F	*p-value*
Group	1	174601.83	174601.83	14.13	<.001
Residuals	49	605621.97	12359.63		
Informativeness	1	38822.03	38822.03	5.90	0.02
Group*Informativeness	1	1280.9	1280.9	0.20	0.66
Residuals	49	322607.13	6583.82		

As seen in the two ANOVAs above, when young children and older children were compared, there was no effect of Informativeness. However, when older children were compared with adults, there was an effect of Informativeness. This suggests that the effect of Informativeness in the older children is weak. Thus, 2 follow-up t-tests for Informativeness were conducted comparing the different gender condition and same gender condition within the older children and within the adults. Older children were 35 ms faster on the different gender condition than the same gender condition but the difference was not significant $t(14) = 0.75$, $p = .47$. Adults fixated on the targets a significant 50 ms earlier on the different gender condition than the same gender condition $t(35) = 4.22$, $p = .0002$.

Two 2 (age group) × 2 (trial type) mixed ANOVAs were also conducted where age group was the between subjects variable, Grammaticality (same gender trials vs. ungrammatical trials) was the within subjects variable and eye-tracking first fixation to the target noun after article onset was the dependent variable. Table 8 shows the results

for a 2 × 2 ANOVA between the young and older children and Table 9 shows the results comparing the older children and adults.

Older children fixated on targets significantly earlier than young children, as reflected in the main effect of age group $F(1,28) = 4.84$, $p = .04$. Both groups of children fixated on targets significantly earlier in the same gender condition than the ungrammatical condition, as seen in the main effect of Grammaticality $F(1,28) = 16.38$, $p < .001$. The interaction of Age and Grammaticality was not significant $F(1,28) = 0.50$, $p = .75$.

Table 8. *2X2 Mixed ANOVA of eye-tracking first fixation by young vs. older children and Grammaticality*

Source	df	SS	MS	F	*p-value*
Group	1	197316.24	197316.24	4.84	0.04
Residuals	28	1140624.81	40736.6		
Grammaticality	1	230104.45	230104.45	16.38	< .001
Group*Grammaticality	1	7085.38	7085.38	0.50	0.48
Residuals	28	393319.34	14047.12		

Adults fixated on targets significantly earlier than older children, as reflected in the main effect of age group $F(1,49) = 9.89$, $p = .003$. Adults and older children fixated on targets significantly earlier in the same gender condition than the ungrammatical condition, as seen in the main effect of Grammaticality $F(1,49) = 29.76$, $p < .001$. The interaction of Age and Grammaticality was not significant $F(1,49) = 0.84$, $p = .36$.

Table 9. *2X2 Mixed ANOVA of eye-tracking first fixation by older children vs. adults and Grammaticality*

Source	df	SS	MS	F	*p-value*
Group	1	202434.79	202434.79	9.89	0.003
Residuals	49	1003230.9	20474.1		
Grammaticality	1	161669.36	161669.36	29.76	< .001
Group*Grammaticality	1	4605.44	4605.44	0.84	0.36
Residuals	49	266169.49	5432.03		

The analyses indicate that all age groups looked to the target earlier when the gendered article was grammatical than ungrammatical. In contrast, the young children and older children did not show an ability to make use of informative gender contexts whereas adults did.

RESEARCH QUESTION 2

To address Research Question 2, we conducted correlation analyses for all potential variables of interest with all three age groups, shown in Table 10. The correlation table shows us the degree of association between variables.

Table 10. *Correlation Table*

Measure	Cum exp	SP inout	EOWPVT SP	Article Acc	GJ Acc	Diff DG & SG	Diff SG & UG
First exp	0.11	-0.26[1]	0.44[3]	0.46[3]	0.48[4]	0.02	-0.12
Cum exp		-0.81[4]	0.71[4]	0.56[4]	0.60[4]	0.12	-0.21
SP inout			-0.60[4]	-0.43[3]	-0.52[4]	-0.05	0.21
EOWPVT SP				0.81[4]	0.79[4]	0.12	-0.16
Article Acc					0.85[4]	0.19	-0.24
GJ Acc						0.05	-0.15
Diff DG & SG							-0.33[2]

Note. 4 = $p < .0001$, 3 = $p < .001$, 2 = $p < .01$, 1 = $p < .05$. Exp = exposure, SP inout = Spanish input/output percentage, EOWPVT SP = Expressive One Word Picture Vocabulary Test in Spanish, Acc = accuracy, GJ = grammaticality judgment, DG = different gender, SG = same gender, UG = ungrammatical gender, Diff = difference score

I hypothesized that first exposure to English, cumulative exposure to English, current Spanish language use, EOWPVT Spanish, and percent grammaticality judgment accuracy would be most predictive of difference scores of first fixation in the eye-tracking task. Difference scores were calculated by subtracting mean of first fixation on informative (different-gender) articles from uninformative (same-gender) articles and mean of first fixation on uninformative (same-gender articles) from incorrect (ungrammatical) articles.

The results of the correlation analysis were used to evaluate and identify which variables would be entered into the regressions. Regression analyses were conducted to indicate which experiential and language variables best predicted Informativeness and Grammaticality. Variables that were considered included: experiential (age of first exposure to English, cumulative exposure to English, percent current daily input and output in Spanish) and language (Spanish vocabulary raw score on the EOWPVT, article-noun naming accuracy, percent grammaticality judgment accuracy). To make sure that the multicollinearity assumption was not violated, article-noun naming accuracy was not included as a predictor since this metric and grammaticality judgment accuracy were correlated at .85 and both measure aspects of gendered article accuracy for the same nouns. Thus, grammaticality judgment accuracy was included in the regression models and article-noun naming accuracy was not. EOWPVT measures vocabulary while grammaticality judgment accuracy and article-noun naming accuracy measure grammar. Thus, EOWPVT in Spanish remained a variable in the regression models. Ultimately, 5 variables were entered into the regression models – experiential variables (age of first exposure to English, cumulative exposure to English, Spanish input/output), a vocabulary measure (EOWPVT Spanish score), and a measure of article knowledge (grammaticality judgment accuracy).

Forward, backward, and stepwise regression analyses were conducted for variable selection in a best fit model for each of the dependent variables of interest for children and adults separately. The first set of regression analyses focused on the child data. To check the multicollinearity assumption for these data, the variance inflation factor (vif) was reviewed. Generally, if the vif is larger than 10, then the multicollinearity assumption is violated. The vif for English first exposure is 3.44, for English cumulative exposure it is 2.37, for Spanish input and output it is 2.67, for EOWPVT Spanish it is 3.95 and for grammaticality judgment accuracy it is 2.45. Thus, including these 5 variables in the subsequent regressions does not violate the multicollinearity assumption. To check the independence assumption, the Durbin Watson test was conducted. The Durbin Watson statistic tests for autocorrelation in the residuals from a regression analysis and is a number between 0 and 4. A value around 2 means there is no autocorrelation in the sample. The Durbin Watson statistic is 1.75 for the child data.

In all three regression models (forward, backward, stepwise), no combination of variables explained the difference between informative (different-gender) and uninformative (same-gender) trials. The model, including all of the predictor variables, is presented in Table 11.

Table 11. *Regression 1 full model for children in the difference score between first fixation of informative and uninformative trials*

	Estimate	Standard error	t-value	*p-value*	Adj. R^2	F	*p-value*
Intercept	-12.47	394.37	-0.03	0.978	-0.167	0.21	0.96
English first exposure	23.71	36.92	0.64	0.53			
English cumulative exposure	15.60	25.77	0.61	0.55			
Spanish input/output	-0.40	6.20	-0.06	0.95			
EOWPVT Spanish	1.39	3.75	0.37	0.72			
GJ accuracy	-2.20	3.10	-0.71	0.49			

In the forward and stepwise regression models, there was no combination of variables for the difference between uninformative (same-gender) and ungrammatical trials. The backwards regression model included two variables, first exposure to English and current Spanish language use (Spanish input/output), $F(2,27) = 4.21$, $p = .03$, and explained 18.11% of the variance. As age of first exposure to English decreases and current Spanish language use increases, the larger the time difference between uninformative and ungrammatical trials on first fixation to the target after article onset. The full model containing all predictor variables is presented in Table 12 while the backwards regression model is shown in Table 13.

Table 12. *Regression 2 full model for children in the difference score between first fixation of uninformative and ungrammatical trials*

	Estimate	Standard error	t-value	*p-value*	Adj. R^2	F	*p-value*
Intercept	-146.20	338.90	-0.43	0.67	0.03	1.18	0.35
English first exposure	-63.61	31.73	-2.01	0.06			
English cumulative exposure	-13.83	22.15	-0.62	0.54			
Spanish input/output	6.55	5.33	1.23	0.23			
EOWPVT Spanish	0.17	3.22	0.05	0.96			
GJ accuracy	0.98	2.66	0.37	0.72			

Table 13. *Regression 2 backwards regression model for children in the difference score between first fixation of uninformative and ungrammatical trials*

	Estimate	Standard error	t-value	*p-value*	Adj. R^2	F	*p-value*
Intercept	-214.48	198.27	-1.08	0.29	0.18	4.21	0.03
English first exposure	-50.7	18.12	-2.8	0.01			
Spanish input/output	7.27	3.48	2.09	0.05			

For the following 2 regressions analyses, only adults were included. To check the multicollinearity assumption, the variance inflation factor (vif) was once again reviewed.

The vif for English first exposure is 2.56, for English cumulative exposure it is 2.49, for Spanish input and output it is 1.07, for EOWPVT Spanish it is 1.62 and for grammaticality judgment accuracy it is 1.65. Thus, including these 5 variables in the subsequent regressions does not violate the multicollinearity assumption. The independence assumption was addressed using the Durbin Watson test. A value around 2 means there is no autocorrelation in the sample. The Durbin Watson statistic is 2.20 for the adult data. In all three regression models (forward, backward, stepwise), no combination of variables explained the difference between informative (different-gender) and uninformative (same-gender) trials. The full model with all predictor variables is shown in Table 14.

Table 14. *Regression 1 full model for adults in the difference score between first fixation of informative and uninformative trials*

	Estimate	Standard error	t-value	*p-value*	Adj. R^2	F	*p-value*
Intercept	175.21	234.22	0.75	0.46	-0.09	0.48	0.79
English first exposure	-8.22	9.52	-0.86	0.40			
English cumulative exposure	-1.16	7.00	-0.17	0.87			
Spanish input/output	0.65	0.99	0.66	0.51			
EOWPVT Spanish	0.40	0.87	0.46	0.65			
GJ accuracy	-1.37	2.45	-0.56	0.58			

In the forward and backward regression models, there was no combination of variables for the difference between uninformative (same-gender) and ungrammatical trials. The stepwise regression model included only one variable, first exposure to English $F(1,34) = 4.54$, $p = .04$, and explained 9.19% of the variance. As age of first exposure to English increases, the time difference between uninformative and ungrammatical trials on first fixation to the target increases. Table 15 presents the full model with all predictor variables and Table 16 shows the stepwise regression model.

Table 15. *Regression 2 full model for adults in the difference score between first fixation of uninformative and ungrammatical trials*

	Estimate	Standard error	t-value	*p-value*	Adj. R^2	F	*p-value*
Intercept	-212.88	247.69	-0.86	0.40	0.02	1.14	0.36
English first exposure	15.68	10.07	1.56	0.13			
English cumulative exposure	2.96	7.40	0.40	0.69			
Spanish input/output	-0.12	1.04	-0.12	0.91			
EOWPVT Spanish	0.02	0.92	0.03	0.98			
GJ accuracy	1.91	2.56	0.74	0.47			

Table 16. *Regression 2 stepwise regression model for adults in the difference score between first fixation of uninformative and ungrammatical trials*

	Estimate	Standard error	t-value	*p-value*	Adj. R^2	F	*p-value*
Intercept	20.81	27.69	0.75	0.46	0.09	4.54	0.04
English first exposure	13.04	6.12	2.13	0.04			

Chapter 4: Discussion

The aim of the current study was to examine bilingual children and adults' use of grammatical gender during online sentence comprehension. Additionally, we investigated which experiential and language variables predicted an individual's gender sensitivity by calculating a difference score for Informativeness (different gender trials vs. same gender trials) and another difference score for Grammaticality (same gender trials vs. ungrammatical trials. We tested 5 to 6-year-old and 8 to 9-year-old bilingual children and adults aged 18-35 on an article-noun naming task, an eye-tracking task (online comprehension), expressive vocabulary tests, and a grammaticality judgment task. Spanish sentences were heard in combination with visual displays showing 4 pictures (one of which was a phonological competitor to the target noun) representing nouns of informative, uninformative, or ungrammatical conditions.

BEHAVIORAL MEASURES

For the participants in this study, as age increased across groups (5-6, 8-9 and 18-35), raw vocabulary scores in English (Mean (SD) = 59(24), 95(21), 140(15) by group respectively) and Spanish (46(15), 67(17, 108(20)) increased significantly, article-noun naming accuracy (production) (48(19), 80(17), 90(0.1)) increased numerically and grammaticality judgment accuracy increased (comprehension) (60(14), 80(16), 94(7)) significantly. The 5 to 6-year-old bilingual children in this study were about 60% accurate in a grammaticality judgment task. This is similar to a previous study by Unsworth and Hulk (2010), who found that 6-year-old children were 70% accurate on a grammaticality judgment task and some were still at chance. Although the procedures between the tasks in these studies were different, it is worth noting that another gendered language like Dutch also showed percentages that were not at ceiling for this age group.

A grammaticality judgment task requires children to make meta-linguistic judgments about gender incongruency. Older children may rely on explicit knowledge when responding and thus, grammaticality judgments may over- or underestimate a child's knowledge of gender. Children learn articles early on as part of early word learning and they appear to pay more attention to phonological cues present in nouns when assigning gender (Pérez-Pereira, 1991). Spanish dominant bilingual children in the US typically start using articles early, but acquire articles in their productive language between 5;0 and 6;10, which is later than most monolinguals (e.g., Gutiérrez-Clellen, Restrepo, & Simon-Cereijido, 2006; Morgan, Restrepo, & Auza, 2013; Pérez-Pereira, 1989). As article-noun naming accuracy for the young children are around chance, these accuracy levels suggest that only some of the children have acquired gender-marked articles in their productive language. They showed significantly higher scores on the grammaticality judgment task, thus implying that they have a slightly greater metalinguistic understanding of gender relative to productive use of gendered articles. Task demands also differ and the possibility of being correct by chance for the grammaticality judgment task is much higher (answering correct vs. incorrect) than for the article-noun naming task as children needed to come up with the noun as well as the correct gendered article. There is also a possibility that children listening to the sentences in the grammaticality judgment task were not specifically and explicitly focused on the accuracy of the gendered article, but rather that as a whole, something in the sentence did not sound right to them. As children engage in higher level discourse and reading, they may begin to appreciate the informativeness of articles.

RESEARCH QUESTION 1

Research question 1 focused on (a) whether participants oriented faster to the target when the article was informative vs. uninformative (different gender trials vs. same gender trials) and (b) whether participants oriented faster to the target when the article's gender matched vs. did not match that of the target (same gender trials vs. ungrammatical trials) First fixation to the target noun after article onset was faster in older children than younger children and adults were faster than older children. There was no effect of Informativeness in the child groups, but there was an effect in the adults. There was an effect of Grammaticality (matched vs. mismatched) in young and older children as well as older children and adults. The Spanish-English bilingual children studied here did not show earlier looks to target referents when gender-marked articles were informative relative to those that were uninformative (Informativeness). This lack of gender sensitivity is dissimilar to native Spanish speaking toddlers and adults (Dussias et al., 2013; Kroff et al., 2016; Lew-Williams & Fernald, 2007; Morales et al., 2016). This result however, converges with our previous findings of a lack of gender sensitivity in bilingual children (Baron & Griffin, in preparation). Notably, the previous study included bilingual children who learned Spanish since birth and learned English between 0-5 years as well as children who grew up in an English-speaking home and began learning Spanish in kindergarten. In this study however, all children were exposed to Spanish since birth. Thus, these bilingual children do not use informativeness of gendered information to narrow down their search space and attend to the referent faster. The young children's inability to show gender sensitivity within the eye-tracking task could also be explained by their low performance on the article-noun naming task and the grammaticality judgment task. This points to the fact that these young children may not explicitly know the gender of enough of the nouns to make strategic use of this

knowledge in spoken word recognition. Perhaps gender sensitivity develops in steps where the broader gender agreement is first understood and utilized at a metalinguistic level (implicitly) and then, with more language exposure, the gender cue strengthens and children are able to assign gender accurately (explicitly) which in turn, leads to using gender informatively.

The adults in this study were also exposed to Spanish since birth and had a slightly higher age of first exposure to English relative to the children in this study. The adult Spanish-English bilinguals in this study did show earlier looks to referents when gender-marked articles were informative in contrast to the child participants. This is similar to native Spanish speakers (Dussias et al., 2013; Kroff, et al., 2016; Lew-Williams & Fernald, 2007; Morales et al., 2016) as well as monolingual adults of other languages (Dahan et al., 2000; Brouwer, et al., 2017). As the adults in this study are early sequential bilinguals, their gender sensitivity is also in line with Guillelmon and Grosjean's (2001) study regarding early bilinguals' ability to process gendered information.

In respect to Grammaticality, all groups were faster at orienting to the same gender trials (matched) than the ungrammatical trials (mismatched). All groups showed sensitivity to the grammaticality of the article-noun pair and in turn, demonstrated inhibitory effects suggesting that their processing is significantly slowed. This has been supported by previous research (Colé & Segui, 1994; Faussart et al., 199; Jacobsen, 1999; Jakubowicz & Faussart, 1998; Lew-Williams & Fernald, 2007; Wicha et al., 2005). Additionally, the findings in this study also support Gillelmon and Grosjean (2001) study as early bilinguals also oriented to targets faster when the article and noun matched in gender than when the article and noun were mismatched. van Heugten and Shi (2009) also observed that the ungrammatical nature of the articles slowed processing efficiency. In a reading comprehension study, Foote (2011) also noted that early bilinguals had

slower reading times in ungrammatical noun-adjective agreement sentences than in grammatically correct sentences.

The young children did not use the informativeness of gender to orient to the target noun, but they did use the grammaticality of the article-noun pair to attend to the referent noun faster. This may speak to explicit versus implicit knowledge. Implicitly, from a young age, children are able to identify when a sentence is ungrammatical. They may not understand what structure within a sentence was specifically ungrammatical, but they have an implicit awareness of such an error. Using puppets in a forced choice paradigm, children as young as 3 can choose between 2 sentences (one grammatical and one ungrammatical) to show their implicit understanding of grammaticality (e.g. Grinstead et al., 2013; Unsworth, 2013). Additionally, children learn to parse words through word frequencies and probabilities. As children hear ungrammatical article-noun pairs with a very low frequency compared to grammatical pairings, they would be expected to be slower at word recognition due to this low frequency of exposure (Garlock, Walley, & Metsala, 2001). The young children's low performance on the article-noun naming task in this study may indicate a lack of explicit knowledge of gender classification of the nouns. This ties back to lexical access (Bölte & Connine, 2004; Spinelli, et al., 2006) and syntactic processing (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Foucart & Frenck-Mestre, 2011; Gunter et al., 2000) as children understand gender agreement at the syntactic level, but are continuing to learn gender assignment at the lexical level and are therefore unable to use gender informatively at this time. Accordingly, for articles to be informative, lexical knowledge is required.

RESEARCH QUESTION 2

Research question 2 focused on identifying which experiential and language predictors explained gender sensitivity through difference scores reflecting effects of Informativeness (different gender trials vs. same gender trials) and Grammaticality (same gender trials vs. ungrammatical trials). Five variables were entered into the regressions: 3 experiential variables (age of first exposure to English, cumulative exposure to English, Spanish input/output), a vocabulary measure (EOWPVT Spanish score), and a measure of article knowledge (grammaticality judgment accuracy). There were no significant models for either children or adults in predicting Informativeness difference scores (different gender trials vs. same gender trials). Since the number of children in each group is small (N=15), we may not have enough power to find subtle effects in Informativeness.

Age of first exposure to English and current Spanish language use were significant predictors of Grammaticality difference scores (same gender trials vs. ungrammatical trials). Therefore, age of first exposure and current Spanish language use are important measures that help to explain the Grammaticality – matched versus mismatched inhibitory effect. The significance of these predictors is supported by other research that has found that gender sensitivity is influenced by proficiency with the gendered language (Dussias et al., 2013; Gillon-Dowens et al., 2010; Keating, 2009) and documenting gender sensitivity in early bilinguals but not in late bilinguals (Guillelmon & Grosjean, 2001).

PHONOLOGICAL COMPETITORS

Based on the visual representations shown in Figures 5, 6, and 7, it appears that participants initially looked at all four images on the screen for the same amount of time and did not show a bias to look at a particular image more than another. As the auditory

stimulus unfolded, participants narrowed down their competitor set to the phonological competitor and target noun. Children did not take advantage of the Informativeness of the gendered article and waited to hear the target noun auditorily before looking for it on the screen. Adults however, did use the Informativeness of the gendered article to begin looking at the target noun in the visual display. Participants looked at the phonological competitor proportionally more and for a longer period of time in the same gender condition than the different gender condition. Previous studies have also found that phonological competitors delay fixations to the target but that the delay is mediated by grammatical gender (Allopenna et al., 1998; Dahan et al., 2000; Tanenhaus, et al., 1996). One would expect that as the gender is the same for the phonological competitor and target in the same gender condition, that participants would spend proportionally more time looking at the phonological competitor in this condition.

All participants were slower on the ungrammatical trials and spent more time looking at the phonological competitor (more so in this condition than in the different and same gender conditions) as it was the most viable option with the same article and initial consonant and vowel as the auditory stimulus. As the rest of the target noun unfolded auditorily, the target noun needed to re-enter the competitor set. Thus, participants were significantly slower in settling on the target noun as the correct response as it may have already been discarded from the competitor set earlier. Quantitative analyses should be undertaken in future analyses in order to more explicitly discern how the phonological competitor enters or exits the competitor set and when participants look proportionally more to the phonological competitor in comparison to the distractors and the target.

THEORETICAL IMPLICATIONS

Within the Competition Model, younger children were expected to show gender sensitivity (both in Informativeness and Grammaticality) as they are less proficient in English and therefore gender sensitivity as a cue should be strong and not weakened by English. In regards to the grammaticality effect, the results of this study supports the Full Access Full Transfer Model. All groups showed a grammaticality effect where they were slower on the ungrammatical trials than the same gender trials. Thus, all participants were negatively affected by an article-noun mismatch where the sentence was ungrammatical. They were faster at orienting to the target when the sentence was grammatical. As article-noun congruence is a grammatical feature that exists in all of the participants' L1, the Full Access Full Transfer Model is supported.

As Spanish-speaking children in the U.S. context learn English, they will gain more experience with English and have a higher proficiency, but may still have high resource demands, which may weaken the gender sensitivity cue. Due to their increased proficiency and ability to use both languages in their daily environment, adults show gender sensitivity. As neither group of children showed an Informativeness effect, it is possible that the cue's strength is merging L1 and L2 (MacWhinney, 1987) where they rely on the informativeness of the article but not necessarily the gender of the article. In children, the two forms (gendered article in Spanish and no gender-marking in English) may be competing. As the gendered article is more structurally complex, the additional gender component may be slowing children down (MacWhinney et al., 1984). Perhaps they are no longer using the gender for informativeness to narrow down possible future referent nouns as they may find it unnecessary. However, they must take note of gender within a sentence as they still have a need to hear and produce sentences that are grammatically correct. Thus, they may be using gender in a broader sense of conveying

information in ways that follow the conventions of the language (which contributes to grammaticality) rather than informativeness. Adults, however, are skilled in both languages, as they have a high degree of cumulative exposure to both. They are able to quickly and accurately process the incoming gender information and are therefore able to demonstrate both an Informativeness effect as well as a Grammaticality effect.

Research on the real-time processing of gender in production and comprehension has shown that the degree of difficulty with gender is moderated by various factors, including L1 (Dussias et al., 2013; Foucart & Frenck-Mestre, 2011), proficiency (Gillon-Dowens et al., 2010; Hopp, 2013), lexical knowledge of gender classes (Lemhöfer et al., 2014), and lexical access (Hopp, 2013). Previous investigators have noted that bilingual children and adults are slower and less accurate in naming pictures than monolinguals (e.g. Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ivanova & Costa, 2008; Yan & Nicoladis, 2009). As bilinguals have divided input and output across their two languages, bilingual speakers use each language less often than monolingual speakers and have less accumulated practice overall. Thus, the frequency values of the corresponding lexical representations are lower for the bilinguals than the monolinguals. Accordingly, the differential frequency with which words are used may affect the respective availability of lexical items and thus bilinguals may exhibit weaker lexical links and lexical representation than their monolingual peers. Bybee's model of network associations (1995, 1999) can be useful to understand how input affects gender assignment. Bybee explains that as children hear and begin to store specific words in their memory, networks (or lexical schemas) of phonological and semantic similarities are developed. High frequency use of specific lexical items in the speech input results in independent representations of those items. However, as exposure frequency decreases, access is more dependent upon activation of components of the schemas. Children hear

nouns preceded by an article and in a developing network of associations, one would expect that articles and nouns would emerge as subparts with links among articles, among nouns, and between the two. The extent to which a schema will generalize to other items is based on the type frequency of the pattern and whether it relates to the construction of gender agreed article-noun pairs by children during the developmental period.

METHODOLOGICAL LIMITATIONS/WEAKNESSES

The design of the present study differs from that used with toddlers in several ways that might have made it more challenging for children to use gender-marked determiners to anticipate referents. Our displays contained four objects rather than two. So, rather than rule out one distractor based on gender, participants needed to rule out three. This adds an additional cognitive load while participants are narrowing down the competitor set; however, it is a task that is more representative of what occurs naturally when one hears spontaneous language. Thus, it is possible that children in this study would have shown gender sensitivity if there was a two picture rather than a four-picture display as in previous studies children showed gender sensitivity but were significantly slower than the adults. Therefore, one must take methodological differences into account when comparing monolingual toddlers in previous studies and bilingual children in this study.

There are many dialects in Spanish. For example, the ball in Mexican Spanish is referred to as ‘la pelota’ while in Puerto Rican Spanish it is ‘la bola’. Therefore, if ‘the ball’ had been used as a target (which it was not), then the phonological competitor would no longer function as a competitor. Another example of dialectal differences also leads to different gender assignment. For example, some speakers of Mexican Spanish use ‘el calcetín’ for ‘the sock’, while other speakers of Mexican Spanish say ‘la calceta’.

Participant recruitment was not restricted to a specific dialect; therefore, there were participants from multiple backgrounds (i.e. Mexican, Cuban, Peruvian, etc.). Although naming agreement was above 80% for target nouns for the norming population of 15 adults and 1 child, these variations in object naming may have added additional variability to the data.

In regards to the tasks that were included, perhaps a receptive vocabulary task would have been warranted. In this study, the Expressive One Word Picture Vocabulary Test was included, however the Receptive One Word Picture Vocabulary Test could also have been administered. Findings across studies diverge with regard to the correlation between gender processing and vocabulary. As the eye-tracking task is receptive in nature, perhaps there would be a stronger correlation with scores from a receptive vocabulary test than an expressive vocabulary test. Different measures of comprehension and production can yield different results as they represent dissociated psycholinguistic properties that draw on different skills and neurological bases (Bates, 1993).

FUTURE STUDIES

It would be beneficial to replicate this study with a smaller number of items with multiple exposures. This would make the study more similar to previous toddler studies, which have shown an Informativeness effect (faster on different gender trials than same gender trials) whereas the children in this study do not. Perhaps with multiple exposures of a smaller set of items, children in this study would have also shown an Informativeness effect.

There were two groups of children in this study, 5 to 6-year-olds and 8 to 9-year-olds and adults 18-35. It would also be beneficial to add 11 to 17-year-olds to see a more complete developmental picture between the children and adults in the current study.

This would help identify when children or adolescence begin to use gender predictively like the adults in this study. Additionally, obtaining a monolingual cohort would show differences between monolingual and bilingual processing on this task.

Appendix A

Condition	Target	Phonological Competitor I	Phonological Competitor II	Distractor 1	Distractor 2
training	libro	avión	avión	anillo	elefante
training	lápiz	estrella	estrella	hamburguesa	luna
SG	camello	cangrejo	camisa	guante	tenedor
DG	bota	bosque	boca	pez	camión
DGFILLER	fuego	lechuga	lechuga	estufa	ardilla
DG	cuchillo	cuchara	cuadrado	pulsera	mochila
UG	planta	plato	playa	globo	dedo
SG	patineta	paleta	payaso	chancla	mosca
UG	canguro	calabaza	caballo	bolsa	llanta
SGFILLER	oso	círculo	círculo	mono	lobo
SG	tesoro	teléfono	televisión	soldado	pingüino
SG	montana	mochila	monstruo	bruja	cereza
DGFILLER	iglesia	escritorio	escritorio	pastel	huevo
UG	pavo	palma	pato	trompeta	silla
SG	conejo	corazón	corona	tambor	parque
DG	campana	castillo	canasta	chaleco	bebe
SG	cama	casa	carro	pelota	jirafa
SGFILLER	oveja	oruga	oruga	escalera	lámpara
SG	sartén	zapato	sandía	búho	queso
UG	ballena	baño	bandera	sombrero	papel
UG	regalo	regla	reloj	flauta	naranja
DG	piña	piano	pizza	ratón	volcán
DG	pantalón	palomita	pañal	cartera	muñeca
DGFILLER	león	alfombra	alfombra	araña	sal
SG	pala	papa	palo	mesa	fresa
UG	tiburón	tijera	tijera	foca	chaqueta
DG	galleta	garaje	gallina	tren	pie
DGFILLER	hoja	elote	elote	helado	espagueti
SGFILLER	ojo	limón	limón	pelo	cerdo
UG	corbata	cocodrilo	computadora	bigote	brazo
UG	maleta	marcador	mantequilla	gato	cohete
DGFILLER	hormiga	sobre	sobre	águila	cacahuate
SG	cepillo de dientes	semáforo	cebolla	dinosaurio	columpio
SG	mariposa	manzana	martillo	tarjeta	boca
SGFILLER	triangulo	murciélago	murciélago	hueso	árbol

DG	bate	vaca	vaso	nube	roca
SG	bufanda	burbuja	burrito	guitarra	motocicleta
SGFILLER	rosa	piscina	piscina	roca	rana
UG	zanahoria	sándwich	salsa	burro	nido
DG	tocino	toalla	tomate	bicicleta	nariz
DG	pesa	perro	pera	violín	suéter
UG	botón	botella	bombero	pluma	llave
UG	fabrica	faro	falda	robot	caracol
SGFILLER	olla	mujer	mujer	almohada	granja
SG	venado	vestido	ventana	refrigerador	rinoceronte
DG	pan	pasta de dientes	pájaro	flecha	lluvia
UG	puente	puerta	pulpo	cebra	concha
DG	cámara	casco	caja	pato	dragón
DGFILLER	espejo	escoba	escoba	calceta	libreta
DG	zorro	sopa	sol	vela	flor

Appendix B

Target	Grammaticality Judgment sentence
el camello	Monte el camello en el desierto.
la patineta	El usó la patineta afuera.
el tesoro	El pirata enterró el tesoro.
el conejo	El conejo brinca todos los días.
la montana	La montana es muy grande.
la cama	Todos los días, duermo en la cama.
la pala	Ella escarbó un hoyo con la pala.
el sartén	Cocine los vegetales en el sartén.
la mariposa	La mariposa vuela en el jardín.
la bufanda	Ella lleva la bufanda en el invierno.
el cepillo de dientes	Ella se lava los dientes con el cepillo de dientes.
el venado	El venado vive en el bosque.
la bota	El busca la bota en su cuarto.
el cuchillo	Se usa el cuchillo para cortar fruta.
la campana	La campana suena cada día
la piña	Mi fruta favorita es la piña.
el pantalón	El se cayó y rompió el pantalón.
la galleta	La galleta tiene pedacitos de chocolate.
el bate	Usé el bate durante un partido de beisbol.
el tocino	Comí el tocino que cocino mi mama
el pan	Ella come el pan con leche.
la pesa	El sacó la pesa para sus ejercicios.
la cámara	Me encanta sacar fotos con la cámara.
el zorro	El zorro les da miedo a las gallinas
la planta	La planta crece en el sol
el canguro	El canguro vive en Australia
el pavo	Se come el pavo para día de acción de gracias.
la ballena	La ballena vive en el océano.
el regalo	El regalo que trajé es para ti.
el tiburón	El tiburón come los peces.
la corbata	Para la boda, el se puso la corbata.
la maleta	Puse toda mi ropa en la maleta.
la zanahoria	Corté la zanahoria para la ensalada.
el botón	El botón de mi chaqueta se cayo.
la fabrica	Se hace chocolate en la fabrica.
el puente	Las personas cruzan el puente.
la fuego	Apagué la fuego después de cocinar.
la león	La león vive en la pradera.
la espejo	La espejo rojo esta en mi cuarto.

el iglesia	Voy al iglesia los domingos.
el hormiga	El hormiga es un insecto.
el hoja	El hoja se cayó del árbol.
la ojo	Ella tiene una pestana en la ojo.
la oso	La oso duerme todo el invierno.
la triangulo	La triangulo tiene tres lados.
el oveja	El oveja tiene mucha lana.
el rosa	El eligió el rosa mas bonita en la tienda.
el olla	La mama prepara la sopa en el olla.

References

Alarcón, I. (2009). The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction time data. *Hispania*, 814-828.

Baron, A., & Griffin, Z. M. (in preparation). Predictive use of gender-marked articles in Spanish-English bilingual children.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137-153.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perceptual Psychophysiology, 58,* 992–1004.

Bedore, L. M., & Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics*, *26*(2), 195–225.

Bölte, J., & Connine, C. M. (2004). Grammatical gender in spoken word recognition in German. *Perception & Psychophysics*, *66*(6), 1018-1032.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417-436.

Bosch, L., & Serra, M. (1997). Grammatical morphology deficits of Spanish-speaking children with specific language impairment. *Amsterdam Series in Child Language Development*, *6*(69), 33-45.

Brouwer, S., Sprenger, S., & Unsworth, S. (2017). Processing grammatical gender in Dutch: Evidence from eye movements. *Journal of Experimental Child Psychology*, *159*, 50-65.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes 10*(5), 425–55.

Bybee, J. (1999). Use impacts morphological representation. *Behavioral and Brain Sciences 22*(6), 1016–17.

Cacciari, C., Carreiras, M., & Barbolini Cionini, C. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory & Language, 37*, 517-532.

Carroll, S. (1989). Second-language acquisition and the computational paradigm. *Language Learning, 39,* 535-594.

Castilla-Earls, A. P., Restrepo, M. A., Pérez-Leroux, A. T., Gray, S., Holmes, P., Gail, D., & Chen, Z. (2016). Interactions between bilingual effects and language impairment: Exploring grammatical markers in Spanish-speaking bilingual children. *Applied Psycholinguistics*, *37*(5), 1147–1173.

Colé, P., & Segui, J. (1994). Grammatical incongruency and vocabulary types. *Memory & Cognition, 22,* 387 – 394.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Dahan, D., Swingley, D., Tanenhaus, M. K., Magnuson, J.S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42,* 465-480.

Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, *45*(2), 200-224.

Domínguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, *28*(5), 485-498.

Dussias, P. E. (2010). Uses of eye-tracking data in second language sentence processing research. *Annual Review of Applied Linguistics*, *30*, 149-166.

Faussart, C., Jakubowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista di Linguistica, 11*, 75 – 101.

Fernald, A., & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, *9*(3).

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42,* 98–116.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*(4), 1003-1015.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32,* 187-220.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, *14*(3), 379-399.

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(1), 226-248.

Fox, P., & Hendler, J. (2011). Changing the equation on scientific data visualization. *Science*, *331*(6018), 705-708.

Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, *28*(5), 467-484.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and language*, *45*(3), 468-492.

Grinstead, J., Baron, A., Vega-Mendoza, M., De la Mora, J., Cantú-Sánchez, M., & Flores, B. (2013). Tense marking and spontaneous speech measures in Spanish specific language impairment: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 56*(1), 352-363.

Guillelmon, D. & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition, 29*(3), 503-511.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*(4), 556-568.

Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. *Language Acquisition and Language Disorders*, *32*, 175-206.

Hoff, E., & Core, C. (2013). Input and language development in bilingually developing children. *Seminars in speech and language*, *34*(4), 215-226.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1-27.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33-56.

Jacobsen, T. (1999). Effects of grammatical gender on picture and word naming: Evidence from German. *Journal of Psycholinguistic Research, 28,* 499 – 514.

Jackson-Maldonado, D., & Maldonado, R. (2017). Grammaticality differences between Spanish-speaking children with specific language impairment and their typically developing peers. *International Journal of Language & Communication Disorders*, *52*(6), 750–765.

Jakubowicz, C., & Faussart, C. (1998). Gender agreement in the processing of spoken French. *Journal of Psycholinguistic Research*, *27*(6), 597-617.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, *57*(1), 1-33.

Kroff, J. R. V., Dussias, P. E., Gerfen, C., Perrotti, L., & Bajo, M. T. (2016). Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism*, *7*(2), 163-198.

Lew-Williams, C. & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18,* 193- 198.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, *63*(4), 447-464.

Leonard, L. B. (1998). *Children with Specific Language Impairment*. Cambridge, MA: MIT Press.

MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics*, *8*(4), 315-327.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 127-150.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11,* F9–F16.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition, 25*(1), 71-102.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, *10*(1), 29-63.

Martin, N. A. (2013). *Expressive one-word picture vocabulary test: Spanish-bilingual edition* (4th ed.). Novato, CA: Academic Therapy Publications.

Martin, N. A., & Brownell, R. (2011). *Expressive one-word picture vocabulary test* (4th ed.). Novato, CA: Academic Therapy Publications.

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 621.

Merino, B. (1983). Language development in normal and language handicapped Spanish-speaking children. *Hispanic Journal of Behavioral Sciences, 5(4),* 379-400.

Morales, L., Paolieri, D., Dussias, P. E., Kroff, J. R. V., Gerfen, C., & Bajo, M. T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition*, *19*(2), 294-310.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, *58*(3), 503-553.

Morgan, G. P., Restrepo, M. A., & Auza, A. (2013). Comparison of Spanish morphology in monolingual and Spanish–English bilingual children with and without language impairment. *Bilingualism: Language and Cognition*, *16*(3), 578-596.

New, B, Ferrand, L, Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45-52.

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*(1), 41-58.

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). *BESA: Bilingual English-Spanish Assessment*. Brooks Publishing: Baltimore, MD.

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2‐year‐olds' bilingual proficiency. *Child Development*, *82*(6), 1834-1849.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: a comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22(5),* 1188-1200.

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research*, *24*(3), 397-430.

Spinelli, E., Meunier, F., & Seigneuric, A. (2006). Does gender information influence early phases of spoken word recognition. *The Mental Lexicon*, *1*, 277-297.

Sabourin, L., & Haverkort, M. (2003). Neural substrates of representation and processing of a second language. *Language Acquisition and Language Disorders*, *30*, 175-196.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40-72.

Segalowitz, N. (2003). Automaticity and second language acquisition. In C. Doughty & M. Long (Eds.), *The Handbook of Second Language Acquisition*. (pp. 382-408) Oxford: Blackwell Publishers.

Tanenhaus, M. K., & Trueswell, J. C. (2006). Eye movement and spoken language comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 835 – 862). Amsterdam: Elsevier.

Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics, 1,* 115-132.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition*, *27*(2), 173-204.

Trenkic, D. (2009). Accounting for patterns of article omissions and substitutions in second language production. *Second language acquisition of articles: Empirical findings and theoretical implications*, *49*, 115-143.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition, 16(1),* 86-110.

Unsworth, S., & Hulk, A. (2010). L1 acquisition of neuter gender in Dutch: production and judgement. In J. Costa, A. Castro, M. Lobo, & F. Pratas (Eds.). *Language acquisition and development: proceedings of GALA 2009* (pp. 483-492). Newcastle: Cambridge Scholars.

U.S. Census Bureau (2015). QuickFacts United States. Retrieved from https://www.census.gov/quickfacts/table/POP815215/00.

van Heugten, M., & Johnson, E. K. (2011). Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. *Journal of Child Language*, *38*(01), 87-100.

van Heugten, M., & Shi, R. (2009). French‐learning toddlers use gender information on determiners during word recognition. *Developmental Science*, *12*(3), 419-425.

Weber, A., & Paris, G. (2004, January). The origin of the linguistic gender effect in spoken-word recognition: Evidence from non-native listening. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 26, No. 26).

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272-1288.

Wicha, N. Y. Y., Orozco-Figueroa, A., Reyes, I., Hernández, A., Gavaldón de Barreto, L., & Bates, E. (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.

White, L. (2003). On the nature of interlanguage representation: Universal grammar in the second language. In C. J. Doughty & M. H. Long (Eds.). *The handbook of second language acquisition* (pp. 19-42). New York: Blackwell.

White, L., Valenzuela, E., Kozlowska–Macgregor, M., & Leung, Y. K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied psycholinguistics*, *25*(1), 105-133.