Annual Review of Applied Linguistics, **37** (2017), pp. 128–147.
© Cambridge University Press, 2017
doi: 10.1017/S0267190517000101

Specific Referential Contexts Shape Efficiency in Second Language Processing: Three Eye-Tracking Experiments With 6- and 10-Year-Old Children in Spanish Immersion Schools

Casey Lew-Williams
Department of Psychology, Princeton University,
caseylw@princeton.edu

ABSTRACT

Efficiency in real-time language processing generally poses a greater challenge to adults learning a second language (L2) than to children learning a first language (L1). A notoriously difficult aspect of language for L2 learners to master is grammatical gender, and previous research has shown that L2 learners do not exploit cues to grammatical gender in ways that resemble L1 speakers. But it is not clear whether this problem is restricted to grammatical gender or whether it reflects a broader difficulty with processing local relations between words. Moreover, we do not know if immersive L2 environments, relative to typical L2 classrooms, confer advantages in learning regularities between words. In three eye-tracking experiments, 6- and 10-year-old children who were enrolled in Spanish immersion elementary schools listened to sentences with articles that conveyed information about the grammatical gender (Experiment 1), biological gender (Experiment 2), and number of referents in the visual field (Experiment 3). L1 children used articles to guide their attention to target referents in all three experiments. L2 children did not take advantage of articles as cues to grammatical gender, but succeeded in doing so for biological gender and number. Interpretations of these findings focus on how learning experiences interact with the nature of specific referential contexts to shape learners' efficiency in language processing.

INTRODUCTION

To investigate the challenges faced by second language (L2) learners relative to first language (L1) learners, researchers have often turned to grammatical gender, "the most puzzling of the grammatical categories" (Corbett, 1991, p. 1). Grammatical gender is a system of assigning nouns to different categories and requiring nearby words (such as articles, adjectives, and verbs) to be marked for agreement. In some languages, grammatical gender categories are organized by semantic criteria, but in other languages, they are semantically arbitrary. Spanish is an example of a language with weak semantic correlates to its binary genders, feminine and masculine. Definite articles like $la_{[fem]}$ and $el_{[masc]}$ are usually paired with nouns that have no transparent semantic relations, and the difficulty of learning linguistic rules associated with such arbitrary gender categories has been a focus of previous research.

By categorizing nouns into often nonintuitive subclasses, grammatical gender systems introduce vast complexities to learners at any age. Despite this complexity,

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

young children are capable of exploiting gender-marked articles in language comprehension, by using words like *la* and *el* in Spanish to anticipate the identity of referents. In a study using the looking-while-listening (LWL) procedure (Fernald, Zangl, Portillo, & Marchman, 2008) with Spanish-learning 2- and 3-year-olds, Lew-Williams and Fernald (2007) presented children with pairs of pictures showing objects with names of either the same grammatical gender (e.g., *pelota, galleta*, "ball$_{[f]}$, cookie$_{[f]}$") or different grammatical gender (*pelota, zapato*, "ball$_{[f]}$, shoe$_{[m]}$"). Eye movements were monitored as children viewed the pictures and heard a sentence referring to one picture (*Encuentra la pelota*, "Find the ball"). Children took advantage of gender-marked articles by shifting their eyes to target pictures more rapidly on different-gender than on same-gender trials, revealing an impressive capacity for making use of a complex morphosyntactic system in real time. Similar findings have been reported in 2-year-old children learning French (van Heugten & Shi, 2009), Dutch (van Heugten & Johnson, 2011), and Czech (Smolík, 2014).

While native speakers of gender-marking languages show relative ease in learning noun assignment, processing gender-marked words, and marking other words for gender agreement, adult L2 learners typically do not. L2 learners are often capable of explicitly categorizing nouns into masculine and feminine categories, but they rarely achieve native-like fluency in processing cues to grammatical gender in real time. In studies using behavioral tasks that require metalinguistic judgments, such as auditory naming (Guillelmon & Grosjean, 2001) and lexical decision (Scherag, Demuth, Rösler, Neville, & Röder, 2004), adult L2 learners showed an inability to process article-noun and adjective-noun phrases in ways that resemble native speakers. When tested in the same eye-tracking task used previously with Spanish-learning toddlers, Lew-Williams and Fernald (2010) found that moderately proficient L2 Spanish learners were not able to take advantage of informative articles to identify familiar nouns. A subsequent study tested participants who had learned Spanish after age 11 but who spoke it regularly as adults in their daily family and professional lives, and even these L2 participants failed to exploit gender-marked articles to process familiar nouns more rapidly (Grüter, Lew-Williams, & Fernald, 2012). These findings suggest that the ability to use grammatical gender marking to anticipate subsequent words in a sentence is a hallmark of fluency, differentiating native from nonnative speakers. In several investigations, L2 learners have been shown to achieve some degree of fluency in processing grammatical gender, but this may only be the case when L1 and L2 have similarly structured grammatical gender systems (Foucart & Frenck-Mestre, 2011; Tokowicz & MacWhinney, 2005). Across these studies—as in the enterprise of L2 research more generally—it is difficult to disentangle the effects of language learning environments, age of exposure to L2, and cross-linguistic transfer of processing strategies from L1 to L2.

Overall, previous research indicates that the complicated regularities associated with grammatical gender systems pose a challenge to L2 learners under the demands of real-time language processing, but not to L1 learners. However, the scope of this L1-L2 disparity is not well understood. The goals of

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

the current research were to address two important dimensions of this L1-L2 disparity.

First, the experiments aimed to understand the specificity of L2 learners' difficulties processing cues to grammatical gender. While difficulties could reflect a narrow problem with grammatical gender per se, they could instead be symptomatic of a pervasive problem with real-time processing of local morphosyntactic relations, such as agreement between articles and nouns more generally. To address these alternatives and refine current knowledge about subsecond L2 processing, the experiments manipulated the informative value of Spanish articles in different referential contexts. In many grammatical gender systems, including that of Spanish, some nouns are *not* arbitrarily assigned. Feminine and masculine articles are sometimes aligned with semantically transparent categories, such as biological gender. In Spanish, *la* and *el* (and their indefinite equivalents) are used in combination with nouns referring to female and male humans, such as *niña* ("girl") and *niño* ("boy"). Further semantic markers are provided by number marking: the difference between one and more than one can be conveyed using the contrastive articles $la_{[sing.]}$ versus $las_{[plu.]}$ and $el_{[sing.]}$ versus $los_{[plu.]}$ (and their indefinite equivalents), which convey whether a singular or plural noun will follow, such as *zapato* ("shoe") and *zapatos* ("shoes"). Thus, Spanish articles can serve as cues to a range of information about nouns.

Second, conclusions about L2 learning have largely been made based on a particular category of participants: 18- to 22-year-olds who learn their L2 in adolescence through formal classroom instruction. They differ in several key dimensions from L1 speakers: They learn their L2 later in life, they learn it against the backdrop of L1, and they have less cumulative exposure to regularities between words in the L2. In contrast, infants are immersed in language through interactions with caregivers throughout the first years of life. To bridge these striking contrasts, the experiments reported here tested L1 and L2 children enrolled in Spanish immersion programs and Spanish-English dual immersion programs in the United States. Immersion is an increasingly popular model for elementary school education, and it is designed to combine elements of L1 and L2 learning environments by using immersive learning to expose children to an L2 in the classroom. School systems in many linguistically diverse regions of the world have established immersion programs to promote bilingualism, biliteracy, and multicultural proficiency among both language-majority students and language-minority students. Generally, immersion education engenders gains in L2 learning, yet students sometimes show limited proficiency in certain domains of language (Tedick & Wesley, 2015).

Different contexts for language learning—learning from parents, high school instructors, and immersion teachers—yield substantial differences in the nature and amount of language experience, giving learners different kinds of exposure to words and sentences that may be critical in laying the foundation for fluency in understanding. As L2 learners progress through their Spanish immersion schooling, do they achieve the same fluency in real-time processing as L1 speakers, or do they more closely resemble L2 adults?

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

OVERVIEW OF EXPERIMENTS 1, 2, AND 3

Participants in the three experiments were children enrolled in Spanish-English dual immersion programs in elementary schools in the United States. There were four participant groups in each experiment: 6-year-old native Spanish-speaking children, 10-year-old native Spanish-speaking children, 6-year-old nonnative Spanish-speaking children, and 10-year-old nonnative Spanish-speaking children.

Experiment 1 used an experiment design from previous research on the processing of grammatical gender (Lew-Williams & Fernald, 2007, 2010) to assess whether L1 and L2 children can exploit Spanish articles (*la* and *el*) as cues to grammatical gender in real-time processing. On *same-gender trials*, participants viewed side-by-side pictures of objects and animals with names belonging to the same grammatical gender categories (e.g., two objects with feminine names, such as *pelota*, "ball$_{[f]}$" and *galleta*, "cookie$_{[f]}$"), while on *different-gender trials*, participants viewed pictures of objects and animals belonging to different grammatical gender categories (e.g., *pelota*, "ball$_{[f]}$" and *zapato*, "shoe$_{[m]}$"). Meanwhile, participants heard a simple sentence referring to one of the objects (e.g., *Encuentra la pelota*, "Find the$_{[f]}$ ball$_{[f]}$"). Experiment 2 investigated the processing of Spanish articles that conveyed information about the biological gender of referents in the visual field. On *same-gender trials*, participants viewed two pictures of people belonging to the same biological sex category (e.g., two females), while on *different-gender trials*, participants viewed pictures of people belonging to different sexes (e.g., one female, one male). Participants simultaneously heard a simple sentence referring to one of the people (e.g., *Encuentra la niña*, "Find the$_{[f]}$ girl$_{[f]}$"). Experiment 3 used a parallel design, but examined listeners' use of Spanish number-marked articles (*el*$_{[sing]}$ vs. *los*$_{[plu]}$) to anticipate whether a speaker would refer to pictures showing one or five objects. On *same-number trials,* participants viewed side-by-side pictures showing the same number of objects (e.g., five cars, five shoes), and on *different-number trials*, they viewed pictures showing a different number of objects (e.g., five cars, one shoe), while hearing a sentence such as *Encuentra los carros*, "Find the$_{[plu]}$ cars$_{[plu]}$." Only on different-gender trials (Experiments 1 and 2) and different-number trials (Experiment 3) was the article useful in signaling which picture would be referred to.

Predictions for L1 children were consistent across the three experiments: They should more rapidly shift gaze to target referents when articles were informative. Predictions for L2 children varied by experiment. For Experiment 1, L2 6-year-olds should show equivalent speed of processing on different- and same-gender trials, as they would still be relative newcomers to the subtleties of the Spanish language. There were two possible outcomes for L2 10-year-olds. If their experience with Spanish was immersive in ways that resemble immersive language environments in infancy, then they should exploit articles as cues to grammatical gender. However, if their learning of cues to grammatical gender was hindered by the lack of gender marking in their first language (English), or if their immersion was not on par with that of typical language environments in infancy, they should identify referents

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

TABLE 1. *Participant Information*

	L1 Children		L2 Children	
	6-year-olds	10-year-olds	6-year-olds	10-year-olds
Number of Participants (*N*)*	14	14	18	26
Number of Females	8	7	7	14
Age (years)				
Mean	6.2	10.2	6.2	10.1
Range	5.3–7.4	9.4–11.3	5.4–7.3	9.1–11.7
Mean % of Home Speech in Spanish	64	70	5	4

* Note that we tested more participants in the older group of L2 children in order to match the sample sizes used in Lew-Williams and Fernald (2007, 2010).

with equivalent speed on different- and same-gender trials. In contrast to previous studies on L2 adults' processing of cues to grammatical gender, L2 children should succeed in exploiting articles that serve as cues to biological gender and number information.

EXPERIMENT 1: PROCESSING ARTICLES AS CUES TO GRAMMATICAL GENDER

Method

Participants. Participants in Experiments 1, 2, and 3 were 72 typically developing school-age children (Table 1). All participants were U.S. elementary school students enrolled in Spanish-English dual immersion programs (in which language input was 50% Spanish) or Spanish immersion programs (in which input was initially 100% Spanish, with an incremental shift to 50% Spanish by the final year). Children were either Spanish-English bilinguals who began learning Spanish from birth (L1 children) or native English speakers who were not exposed to Spanish before kindergarten (L2 children). Two ages were included in each group: children enrolled in kindergarten or first grade (hereafter referred to as 6-year-old children) and children enrolled in fourth or fifth grade (hereafter referred to as 10-year-old children). Four additional L1 children were tested but excluded from analyses due to showing unexpectedly low proficiency in Spanish (>1 *SD* below the mean on a standardized assessment of Spanish vocabulary knowledge). As a thank-you for participation, families received gift cards to Target and a local bookstore.

To help ensure that children's home language environments matched their L1 or L2, parents completed a questionnaire about typical language input in Spanish and English on each day of the week from a wide variety of people. Moreover, to ensure that both L1 and L2 participants were reasonably proficient in Spanish, children were tested on the Receptive One-Word Picture Vocabulary Test (Brownell, 2000), a standardized assessment of Spanish receptive vocabulary ($M = 100$, $SD = 15$; see Table 2). Children were also tested on a custom gender production task. For each

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

TABLE 2. *Scores on Offline Tests*

	L1 Children		L2 Children	
	6-year-olds	10-year-olds	6-year-olds	10-year-olds
Receptive One-Word Picture Vocabulary Test				
Standard Score	101.8	96.3	95.8	101.2
SD	10	11	21	20
Range	83–123	73–115	55–117	55–137
Percentile	53.9	42.0	50.9	51.2
Gender Production				
% Correct	90.8	92.0	66.3	76.7
SD	14.3	13.8	12.9	14.3
Range	54–100	59–100	44–90	54–100
Leiter-R Forward Memory				
Mean	12.7	10.9	13.7	12.3
SD	3.8	3.6	2.8	2.8
Range	6–17	1*–15	9–17	8–17

* One participant did not understand the working memory test, but scored within the normal range on other tests.

of 52 pictures of objects, animals, and people, the Spanish-speaking experimenter elicited an article-noun sequence from the child by asking for a completion of the sentence. Finally, children were tested the Leiter-R Forward Memory subtest (Roid & Miller, 1997), a standardized assessment of working memory span ($M = 10$, $SD = 3$). Reaction times (RTs) in the eye-tracking experiments were not related to performance on offline tasks, thus data are not reported.

Stimuli. A list of speech stimuli from all experiments is provided in Appendix 1. The stimuli in Experiment 1 were identical to those used by Lew-Williams and Fernald (2007). In a soundproof booth, a female native Spanish speaker recorded sentences consisting of a simple frame (*Encuentra*, "Find" or *¿Dónde está*, "Where is"), followed by one of eight article-noun pairs, half feminine and half masculine (*la pelota*, "ball," *la galleta*, "cookie," *la vaca*, "cow," *la rana*, "frog," *el zapato*, "shoe," *el carro*, "car," *el pájaro*, "bird," or *el caballo*, "horse"). Articles were unstressed, as is typical in Spanish conversation. Each sentence was followed by a short attention-getting phrase (e.g., *¿Te gusta?*, "Do you like it?"). Three tokens of each sentence were recorded using intonation characteristic of child-directed speech. Minor editing of the waveform of each sentence was conducted to control for the duration of the sentence frame ($M = 914$ ms, range = 900–931 ms), gender-marked article ($M = 280$, range = 268–299), and noun ($M = 720$, range = 670–770). Visual stimuli were colorful digital pictures of the named objects on gray backgrounds. Each picture served as target on four trials and as distracter on four trials, with side of target picture presentation counterbalanced. Two tokens of each object/animal were used in testing to enhance the variety of visual stimuli.

Procedure. Prior to testing, the Spanish-speaking experimenter showed L1 and L2 children a picture book introducing the target object names, because some

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

pictures could be named in different ways (e.g., the feminine noun *rana*, "frog," vs. the masculine noun *sapo*, "toad"). Objects were labeled using indefinite articles (*una* or *un*), in order to avoid priming the particular article-noun sequences that would be used in the subsequent test session.

Following familiarization, participants were tested using the LWL procedure, a child-friendly technique that yields online measurements of the time course of spoken language processing (Fernald et al., 2008). Participants were simply asked to look at pictures as they listened to Spanish sentences. Participants were seated 150 cm from a screen on which they viewed two pictures side by side on each trial. Pictures were about 36 × 50 cm, with 60 cm between pictures. On each trial, one of the pictures was named using a prerecorded sentence. The target noun was always preceded by a definite article. Participants were presented with 16 *same-gender trials*, where pictures depicted objects with names of the same grammatical gender, and 16 *different-gender trials*, where pictures depicted objects with names of different grammatical gender, interspersed in four counterbalanced, quasi-random orders. Test trials were interspersed with 24 filler trials.

On each trial, pictures were visible for 2 s prior to the speech signal, for the duration of the 3-s speech signal, and for 1 s following the speech signal. The participant's face was recorded from a centrally located camera. Using custom software, eye movements were coded offline, frame by frame, with 33-ms resolution. A digital time-code was time-locked to the acoustic onset of the article on each trial, and coders blind to trial type indicated at each frame whether the participant was looking left, right, between the pictures, or away from both. To assess the reliability of eye-movement coding, two independent observers coded sessions from 15% of participants. The calculation of interobserver agreement was based only on the 33-ms frames surrounding shifts. The two observers agreed within a single frame on 98.1% of the frames across Experiments 1–3.

Measurement of Language Processing Efficiency

Children were given no specific instructions regarding looking behavior; thus they could not know in advance which picture would be named. There were two primary ways of responding, contingent on which picture they happened to be fixating initially. If they were already looking at the correct picture (target-initial trials), they should maintain fixation; but if they were looking at the distracter picture (distracter-initial trials), they should shift their eyes to the named picture as the article-noun sequence unfolded. Distracter-to-target shifts were expected to begin more rapidly on different-gender trials, where the article provided a prenominal cue to the identity of the target referent. Distracter-initial trials were used to calculate reaction time (RT), the latency to initiate a shift toward the target picture. RT was calculated from article onset—the first moment in the unfolding sentence when acoustic information relevant to referent identification became available to participants. Shifts initiated within the first 300 ms were not included in the analyses, because they were likely to represent random shifting initiated prior to the possible influence of the article (Haith, Wentworth, & Canfield, 1993). RTs

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

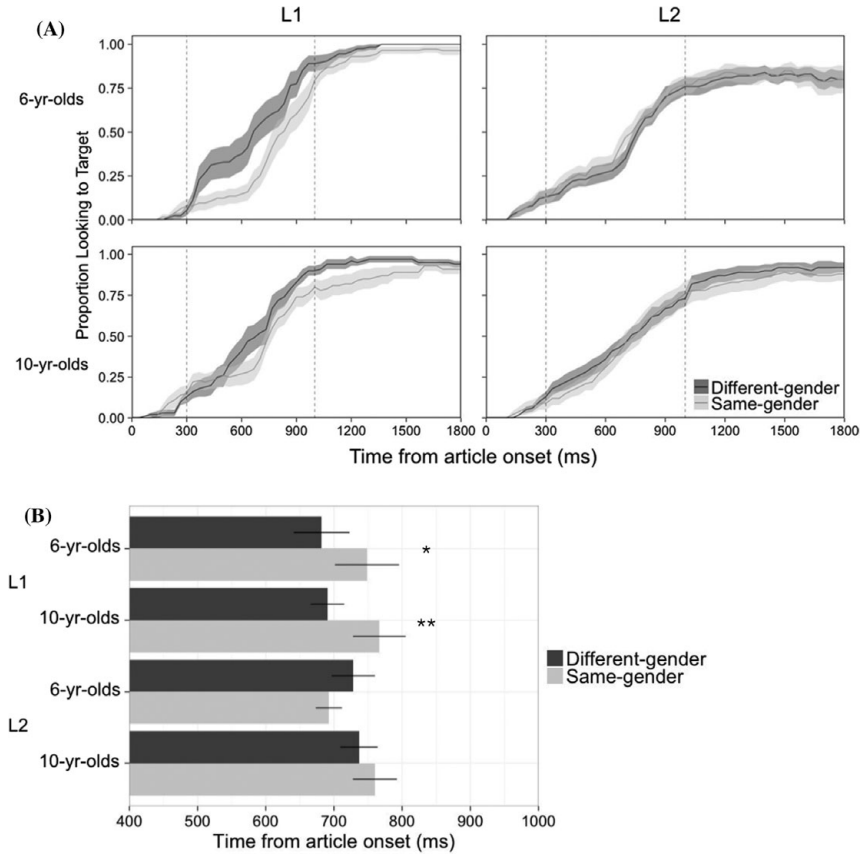


FIGURE 1. (A) Time course of looking to the target on different-gender and same-gender trials in Experiment 1. Data are presented separately for the four groups of participants: L1 6-year-olds, L1 10-year-olds, L2 6-year-olds, and L2 10-year-olds. Plots only show trials in which participants were looking at the distracter at article onset, that is, the trials used to calculate reaction times (RTs). Vertical dashed lines indicate acoustic offsets of the article and noun. (B) Mean RTs in shifting from the distracter to the target picture. Error bars indicate +/−1 *SEM*.
$^{*}p < .10$ $^{**}p < .05$

were included in analyses if they occurred between 300 and 1,300 ms from article onset, as in previous studies of article-noun processing.

Results and Discussion

The primary question in Experiment 1 was whether English-speaking children learning Spanish as L2 in elementary school immersion programs can exploit informative grammatical gender-marked articles in ways that resemble L1 Spanish speakers. The plots in Figure 1a depict the time course of speech processing on

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

same-gender and different-gender trials for L1 and L2 children. Only those trials where participants were looking at the distracter picture at article onset are included in these plots; thus the curve begins at zero and rises over time as participants shift their eyes to the target picture. Time from article onset is shown on the *x*-axes. The *y*-axes show the proportion of trials on which participants were looking at the correct picture at each 33-ms frame after shifting from distracter to target, averaged across participants. Figure 1b shows summary data of RTs on each trial type for L1 6-year-olds (different-gender: $M = 682$ ms; same-gender: $M = 748$), L1 10-year-olds (different-gender: $M = 691$; same-gender: $M = 767$), L2 6-year-olds (different-gender: $M = 728$; same-gender: $M = 693$), and L2 10-year-olds (different-gender: $M = 737$; same-gender: $M = 760$).

To compare performance of L1 and L2 children, mean RTs were analyzed in a $2 \times 2 \times 2$ mixed ANOVA, with language group (L1, L2) and age group (6-year-olds, 10-year-olds) as between-subjects factors, and trial type (different-gender, same-gender) as a within-subjects factor.[1] This analysis revealed no main effect of trial type, $F(1,66) = .6, p = .43, \eta_p^2 = .01$, showing that participants responded with the same speed on the two trial types; no main effect of language group, $F(1,66) = .2$, $p = .65$, $\eta_p^2 = .003$, indicating that L1 children showed the same overall processing speed as L2 children; no main effect of age group, $F(1,66) = .3$, $p = .59$, $\eta_p^2 = .004$, suggesting that the younger and older participants showed the same speed of processing; and a significant interaction between trial type and language group, $F(1,66) = 8.0$, $p = .006$, $\eta_p^2 = .11$, indicating that the trial type effect was different across L1 and L2 participants. There were no other significant interactions (all Fs < 1.2, ps $> .25$). Planned one-tailed paired-samples *t*-test comparisons explored whether participants in each group were faster on different-gender trials than on same-gender trials (Figure 1b). There was a marginally significant trial type difference among L1 6-year-olds, $t(12) = 1.3$, $p = .095$, $d = .42$, and a significant trial type difference among L1 10-year-olds, $t(12) = 2.2$, $p = .035$, $d = .68$, but no significant differences for L2 6-year-olds, $t(17) = .85$, $p = .24$, $d = .33$, or L2 10-year-olds, $t(25) = .06$, $p = .96$, $d = .04$.

The looking behaviors of L1 children—across grade groups—were consistent with earlier results showing that native Spanish-speaking toddlers (Lew-Williams & Fernald, 2007) and adults (Grüter et al., 2012; Lew-Williams & Fernald, 2010) take advantage of the gender-marked articles *la* and *el* to more rapidly identify referents of familiar nouns. Interestingly, findings were also consistent between the L2 children tested here and the L2 adults tested in our previous research. Native English-speaking children enrolled in Spanish immersion programs, like L2 adults, did *not* identify target referents more rapidly when articles were informative about the identity of candidate referents. Response times were similar on same- and different-gender trials even for the children who had been immersed in Spanish from kindergarten through fourth or fifth grade (i.e., for approximately 5 years). Together, results using this experimental paradigm with many groups of L1 and L2 participants suggest that native speakers—but not L2 learners—can use grammatical gender-marked articles and nouns to their advantage in real-time language interpretation. Alternatively, articles may simply be too subtle and short

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

in duration for L2 adults to process in real time. In Experiment 2, L1 and L2 children listened to the same Spanish articles, *la* and *el,* but referents in the visual context contrasted in biological gender, and thus articles marked more semantically transparent categories than grammatical gender.

EXPERIMENT 2: PROCESSING ARTICLES AS CUES TO BIOLOGICAL GENDER

Method

Participants. L1 and L2 children were the same participants from Experiment 1.

Stimuli. Stimuli in Experiment 2 were designed to facilitate comparison between L1 and L2 children's processing of linguistic cues to biological gender. A female native Spanish speaker recorded sentences consisting of a simple frame (*Encuentra*, "Find") followed by one of eight nouns referring to people, half female and half male, and each preceded by a gender-marked article (*la niña*, "the girl$_{[f]}$," *el niño*, "the boy$_{[m]}$," *la señora*, "the woman$_{[f]}$," *el señor*, "the man$_{[m]}$," *la princesa*, "the princess$_{[f]}$," *el príncipe*, "the prince$_{[m]}$," *la doctora*, "the doctor$_{[f]}$," *el doctor*, "the doctor$_{[m]}$"). Each sentence was recorded using intonation characteristic of child-directed speech and was followed by a short attention-getting phrase (e.g., *¿La/Lo ves?*, "Do you see her/him?"). Minor editing of the waveform of each sentence was conducted to control for the duration of the sentence frame (M = 909 ms, range = 900–918 ms), gender-marked article (M = 299, range = 297–299), and noun (M = 750, range = 718–786). Visual stimuli were colorful images of female and male humans that included each individual's head and shoulders. Two different pictures were used for each target noun. Pictures contained prototypical and stereotyped features of each social category: Males had short hair; females had longer hair; princes and princesses wore royal clothing and crowns; doctors wore white lab coats and stethoscopes.

Procedure. On each trial, participants heard a sentence while viewing two pictures side by side, as in Experiment 1. Twelve trials were *same-gender trials*, where each picture depicted a person of the same gender, and 12 trials were *different-gender trials*, where pictures depicted one female and one male. Each noun served as target on three trials and as distractor on three trials, with side of target picture presentation counterbalanced. Test sentences were interspersed among 46 filler trials (from a word learning experiment not reported here) in two counterbalanced orders.

Results and Discussion

In contrast to Experiment 1, L2 10-year-olds in Experiment 2 succeeded in using the articles *la* and *el* to identify female and male humans. Figure 2a depicts the time course of language interpretation by each group. Figure 2b shows summary data of RTs on each trial type for L1 6-year-olds (different-gender: M = 581 ms,

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

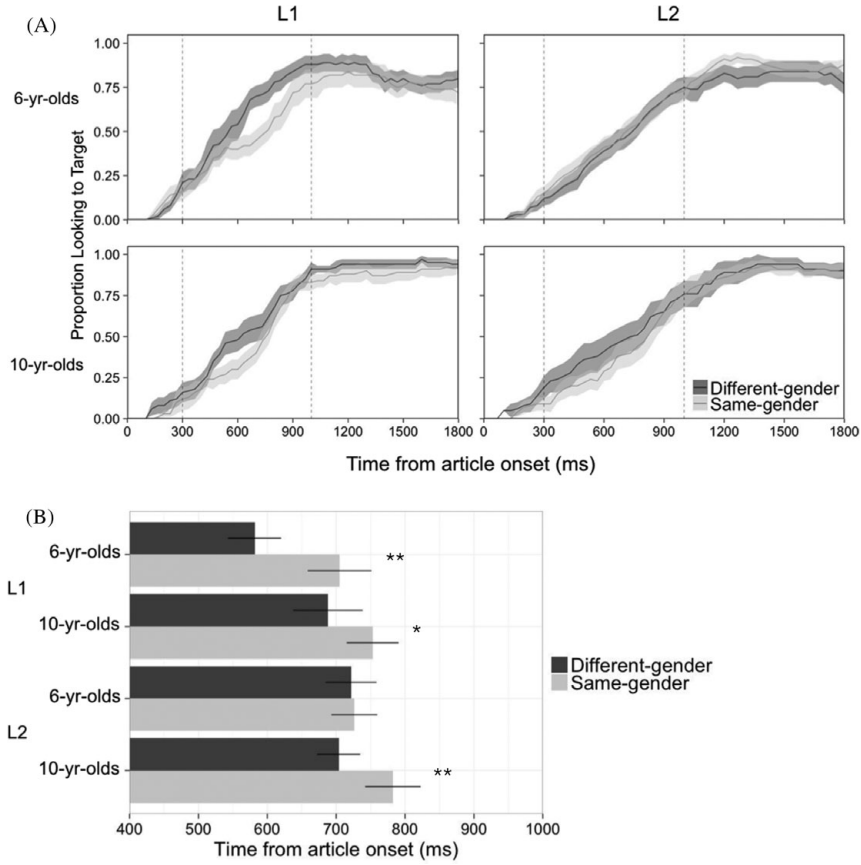


FIGURE 2. (A) Time course of looking to the target on different-gender and same-gender trials in Experiment 2, by group. Plots only show trials in which participants were looking at the distracter at article onset. Vertical dashed lines indicate acoustic offsets of the article and noun. (B) Mean RTs in shifting from the distracter to the target picture. Error bars indicate +/−1 *SEM*.
$^{*}p < .10$ $^{**}p < .05$

SD = 180; same-gender: *M* = 705, *SD* = 99), L1 10-year-olds (different-gender: *M* = 688; same-gender: *M* = 753), L2 6-year-olds (different-gender: *M* = 722; same-gender: *M* = 726), and L2 10-year-olds (different-gender: *M* = 704; same-gender: *M* = 782).

To compare performance by L1 and L2 children in Experiment 2, mean RTs were analyzed in a 2 × 2 × 2 mixed ANOVA, with language group (L1, L2) and age group (6-year-olds, 10-year-olds) as between-subjects factors, and trial type (different-gender, same-gender) as a within-subjects factor. This analysis revealed a significant main effect of trial type, $F(1,66) = 7.2$, $p = .009$, $\eta_p^2 = .10$, showing that participants overall were faster on different-gender trials than on same-gender trials; no main effect of language group, $F(1,66) = 2.2$, $p = .14$, $\eta_p^2 = .01$,

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

indicating that L1 and L2 children initiated eye movements to the target noun with approximately equal speed; no main effect of age group, $F(1,66) = 2.2$, $p = .14$, $\eta_p^2 = .01$, suggesting that younger and older children responded with the same overall speed; and no significant interactions (all Fs < 2.5, ps $> .13$). Planned t-tests revealed a significant trial type difference for L1 6-year-olds, $t(13) = 2.2$, $p = .024$, $d = .84$, a marginally significant difference for L1 10-year-olds, $t(13) = 1.3$, $p = .10$, $d = .39$, no significant difference for L2 6-year-olds, $t(16) = .11$, $p = .91$, $d = .03$, and a significant difference for L2 10-year-olds, $t(24) = 1.8$, $p = .04$, $d = .72$. Experiments 1 and 2 captured a disparity in the processing skills of L2 learners: Here, they made use of articles that distinguished easily recognizable females and males, a finding that contrasts with previous evidence of failure to use articles that indexed membership in largely arbitrary noun classes. In our next experiment, articles were again informative about upcoming nouns, but in terms of number information as opposed to biological gender.

EXPERIMENT 3: PROCESSING ARTICLES AS CUES TO NUMBER

Method

Participants. L1 and L2 children were the same participants from Experiments 1 and 2.

Stimuli. In the LWL procedure, children heard sentences consisting of a sentence frame (*Encuentra*, "Find," for inanimate referents; *Mira*, "Look at," for animals) followed by one of 12 article-noun pairs, half singular and half plural (*el carro*, *los carros*, "the$_{[sing]}$/the$_{[plu]}$ car/s," *el zapato*, *los zapatos*, "the shoe/s," *el caballo*, *los caballos*, "the horse/s," *el pájaro*, *los pájaros*, "the bird/s," *el gato, los gatos*, "the cat/s," *el perro, los perros*, "the dog/s"). On every trial, the pairs of objects consisted either of two animals or two inanimate objects. Only masculine nouns were used in this experiment, because the masculine definite articles *el* and *los* diverge phonologically at article onset, unlike the feminine definite articles *la* and *las*, which diverge acoustically partway through the article. The waveform of each sentence was edited to control for duration of the sentence frame ($M = 802$ ms, range = 734–930 ms), definite article ($M = 293$, range = 280–299), and noun ($M = 766$, range = 690–831). Visual stimuli were colorful pictures of the named objects. Half of the pictures showed one object (e.g., a car), and half of the pictures showed five objects (e.g., five cars). In five-object pictures, identical objects were arranged in an X-shaped array against a gray background, with each object appearing approximately one-fifth the size of the object in one-object pictures. Two different pictures were used for each target noun.

Procedure. On 12 *same-number trials*, two pictures depicted the same number of objects (e.g., one car vs. one shoe, or five cars vs. five shoes); on 12 *different-number trials*, each picture depicted a different number of objects (e.g.,

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

five cars vs. one shoe). Each picture served as target and as distractor on two trials (once as a distractor on same-number trials and once as a distractor on different-number trials), with side of target picture presentation counterbalanced. Test sentences were interspersed with 40 filler trials in two counterbalanced, quasi-random orders.

Results and Discussion

The main finding in Experiment 3 was that all four participant groups succeeded in using number-marked articles to identify pictures showing one versus more than one referent. Figure 3a shows the time course of looking on same- and different-number trials. Figure 3b shows summary data of RTs on each trial type for L1 6-year-olds (different-gender: $M = 647$ ms; same-gender: $M = 731$), L1 10-year-olds (different-gender: $M = 680$; same-gender: $M = 777$), L2 6-year-olds (different-gender: $M = 664$; same-gender: $M = 739$), and L2 10-year-olds (different-gender: $M = 670$; same-gender: $M = 794$).

Mean RTs were analyzed in a $2 \times 2 \times 2$ mixed ANOVA, with language group (L1, L2) and age group (K+1, 4+5) as between-subjects factors, and trial type (different-number, same-number) as a within-subjects factor. This analysis revealed a significant main effect of trial type, $F(1,65) = 11.7$, $p = .001$, $\eta_p^2 = .15$, showing that participants were faster overall on different-number trials than on same-number trials; no main effect of language group, $F(1,65) = .35$, $p = .56$, $\eta_p^2 = .005$, indicating that L1 and L2 children initiated eye movements to target referents with similar overall speed; no main effect of age group, $F(1,65) = .35$, $p = .56$, $\eta_p^2 = .005$, suggesting that younger and older children showed similar overall speed of processing; and no significant interactions (all Fs < 1, ps $> .8$). Planned t-tests revealed a marginally significant trial type differences in the two groups of 6-year-old children [L1: $t(12) = 1.47$, $p = .083$, $d = .56$; L2: $t(17) = 1.57$, $p = .067$, $d = .49$], and significant trial type differences in the two groups of 10-year-old children [L1: $t(13) = 1.88$, $p = .041$, $d = .72$; L2: $t(24) = 2.15$, $p = .021$, $d = .79$]. Overall, both L1 and L2 children used informative number-marked articles to more rapidly establish reference. Performance was generally similar to that in Experiment 2, except that even the younger group of L2 children—those with only 1–2 years of exposure to Spanish—already showed native-like efficiency in processing number-marked articles.

GENERAL DISCUSSION

Three experiments revealed that native and nonnative Spanish-speaking children's efficiency in processing article-noun phrases varies in different referential contexts. In previous research, native Spanish speakers—both toddlers and adults—used informative grammatical gender-marked articles like *la* and *el* to more rapidly identify objects in the visual scene, but L2 adults failed to do so with any consistency (Grüter et al., 2012; Lew-Williams & Fernald, 2007, 2010). While these findings might suggest that L2 learners are unable to process local

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

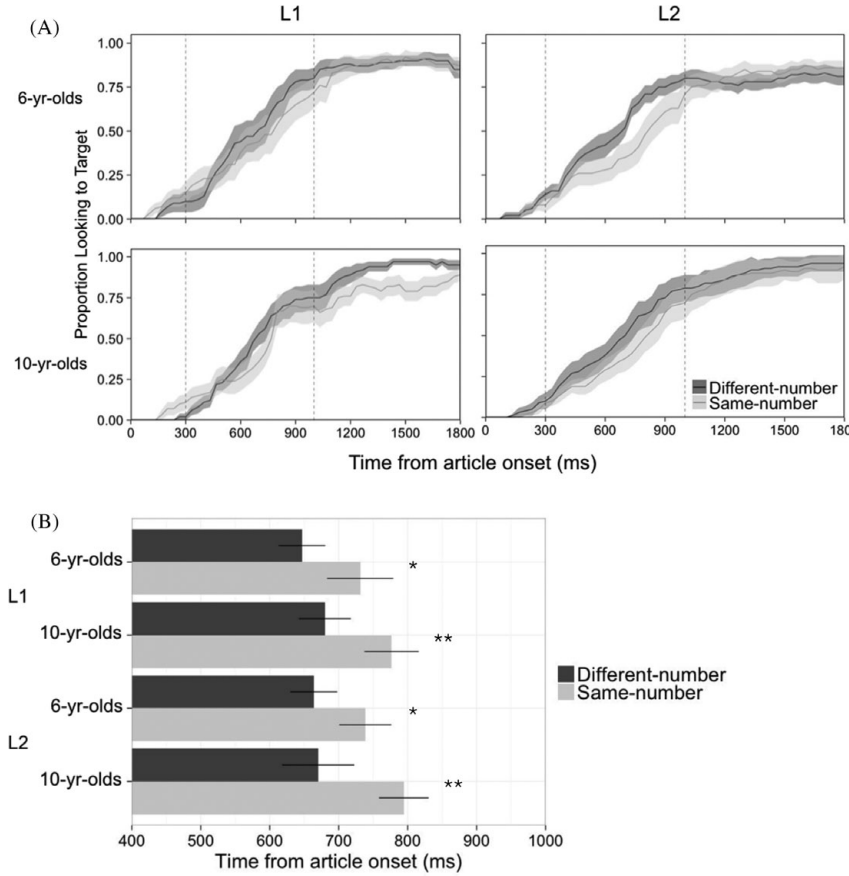


FIGURE 3. (A) Time course of looking to the target on different-number and same-number trials in Experiment 3, by group. Plots only show trials in which participants were looking at the distracter at article onset. Vertical dashed lines indicate acoustic offsets of the article and noun. (B) Mean RTs in shifting from the distracter to the target picture. Error bars indicate +/−1 *SEM*.
$^{*}p < .10$ $^{**}p < .05$

morphosyntactic relations in ways that resemble L1 speakers, the experiments reported here examined a broader range of the richness of Spanish grammatical morphology in order to examine the nature of L2 learners' difficulty.

Using nearly identical experiment designs about the processing of articles as cues to grammatical gender, biological gender, and number, this research indicates that language proficiency cannot be understood without studying how specific language structures interact with specific referential contexts. Participants were 6- and 10-year-old children enrolled in Spanish immersion programs in U.S. elementary schools. Across all three experiments, L1 children took advantage of articles to more rapidly identify referents in the visual field. However, data for L2 children

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

varied across experiments. In Experiment 1, 6- and 10-year-old L2 children did not gain leverage from articles as cues to grammatical gender. Experiment 2 revealed that 10-year-old (but not 6-year-old) L2 children could use articles as cues to look at females versus males in the visual field. In Experiment 3, both 6- and 10-year-old L2 children used number-marked articles to more rapidly identify pictures of one versus five objects. Researchers often explain L1-L2 differences in terms of L2 learners' deficient grammatical knowledge (e.g., Clahsen & Felser, 2006; Hawkins & Chan, 1997) or problems accessing and retrieving linguistic information in situations that demand rapid processing (e.g., Jiang, 2004). But across our experiments on grammatical gender, biological gender, and number, sentence structures were nearly identical, yet L2 children achieved native-like processing in only a subset of cases, that is, when the visual scene contained rich information about biological gender and number. This suggests that L2 learners' previously documented failure to exploit articles as cues to grammatical gender was entirely not due to deficient underlying grammar or to the demands of real-time processing.

The findings are more compatible with models of sentence processing that emphasize dynamic interaction between different types of cues in real time. For example, in the competition model (Bates & MacWhinney, 1981; MacWhinney, 2005), interpretation of a phrase or sentence depends not only on grammatical competence, but also on how cues converge at particular moments. Learning of forms in this model depends on many exposures to patterns of words in different contexts, and interpretations of incoming sentences are determined by the salience of each cue in particular contexts. It is this context-dependent process through which L2 learners deploy their use of particular cues that determines processing, as in Experiments 2 and 3. Next, I focus on three classes of explanations for L1-L2 differences in processing cues to grammatical gender, biological gender, and number: semantic transparency of noun/object categories, transfer of processing skills from L1 to L2, and differences in language input.

While biological gender is a salient feature of the social world, grammatical gender is not. The pictures of stereotypically female or male humans from Experiment 2 are easily recognizable by both L1 and L2 children. Even 3- to 4-month-old infants are capable of perceptual discrimination between stereotypical females and males (Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002). In contrast, inanimate objects do not fit intuitively into feminine and masculine categories in Spanish (Corbett, 1991). Thus, in Experiment 2, L2 learners had a semantically based means of processing article-noun sequences. Indeed, linguistic markers of biological gender have been argued to serve as the foundation for learning grammatical gender classes (Imai, Schalk, Saalbach, & Okada, 2014; Pinker, 1982; Vigliocco, Vinson, Paganelli, & Dworzynski, 2005). Thus, it is most useful to posit a dynamic interplay between grammatical structures, processing speed, and specific referential contexts: For transparent categories like biological gender, the real-time nature of speech is not an obstacle for native-like processing of Spanish morphosyntax. This interactive perspective, emphasizing the simultaneous processing of different types of information in the linguistic signal, has received considerable attention in L1 psycholinguistic literature (e.g., MacDonald & Seidenberg, 2006).

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

Similarly, the salience of number in the nonlinguistic domain may facilitate learning of linguistic markers of number, particularly if a language marks number in relatively transparent ways, as in English and Spanish. Some researchers have proposed that basic "number sense" is a core concept, deployed automatically at the beginning of life (Feigenson, Dehaene, & Spelke, 2004; Xu & Spelke, 2000), and that grammars of languages are targeted to accommodate representations of small versus large numbers of objects. Presumably, according to this argument, minimal experience with linguistic number marking is needed to detect the difference between one and more than one, making the task in Experiment 3 of recognizing one versus five objects relatively straightforward for L2 learners.

A second factor that may account for L2 children's variable success in exploiting articles is the transfer of processing skills from L1 to L2. A consistent and robust finding in language research is that properties of the L1 influence the ability to learn properties of the L2 (e.g., Bates & MacWhinney, 1981; Sabourin, Stowe, & de Haan, 2006). Transfer effects in reading abilities have been documented previously among children in immersion schools (Feinauer, Hall-Kenyon, & Davison, 2013). In Experiment 2, L2 children were native speakers of English, a language that marks biological gender on pronouns but does not overtly mark grammatical gender. Although English and Spanish mark biological gender in distinct ways (e.g., Spanish includes gender marking on articles and adjectives, which English does not), L2 children may have transferred processing strategies from their L1 to their L2. Years of experience using English pronouns to interpret sentences about females and males (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Song & Fisher, 2005) may have heightened attention to the distinct suite of markers of biological gender in Spanish. Similarly, experience with the predictive value of number cues in English (between determiners and nouns, subjects and verbs, and pronouns and antecedents) may have supported L2 children's learning of number marking and concord morphology in Spanish. In both L1 and L2, the use of English plural markers in the presence of multiple exemplars of an object may occur with such regularity that learners can achieve efficient number processing skills from early stages of language learning (Robertson, Shi, & Melançon, 2012). In contrast, the lack of markers of grammatical gender in English may have derailed L2 children's abilities to achieve native-like processing of grammatical gender in Spanish.

Finally, the quality and quantity of exposure to language varies dramatically in home versus classroom environments. Clark (2003) estimated that infants may hear their L1 for 3,640 hours per year (70 hours per week), but that adult L2 learners may hear their L2 for just 180 hours per year (6 hours per week for a 30-week academic year). This results in a striking 20:1 ratio in hours of language exposure between infant and adult learners. In elementary schools in the United States, children are in school for 1,192 hours per year, on average (National Center for Education Statistics, 2008). Thus, children in immersion programs hear their L2 for a range of 596 hours (for children in 50:50 dual immersion programs) to 1,073 hours per year (for children in 90:10 immersion programs). Immersion provides notably more L2 exposure than traditional secondary school classrooms,

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

which confers many benefits for L2 proficiency (Dahl & Vulchanova, 2014; Fortune & Tedick, 2015; Kaushanskaya, Gross, & Buac, 2014; Padilla, Fan, Xu, & Silva, 2013; Westerveld, 2014), but it nonetheless pales in comparison to typical infant learning environments. There is a 3:1 or up to 6:1 ratio in hours of exposure between infants and L2 immersion students, which may help explain why L2 children do not always achieve native-like proficiency (Fortune & Tedick, 2015; Kovelman, Baker, & Petitto, 2008; Potowski, 2007; Tedick & Young, 2014). The density of language exposure is critical to understanding these ratios, and while data for infants and toddlers are available (Weisleder & Fernald, 2013; Hirsh-Pasek et al., 2015), comparable analyses of language use in L2 immersion classrooms are not available. However, it is reasonable to conclude that infants gain intensive linguistic practice with their caregivers, and L2 children have relatively fewer opportunities to hear common article-noun sequences. Clark (2003) suggested that "it may well be that the general bar to learning another language well is less a matter of age-of-acquisition than a matter of willingness-to-invest-enough-time" (p. 47). Differences in input quantity, in addition to many qualitative factors, offer a compelling explanation for L1-L2 differences in processing cues to Spanish grammatical gender. As revealed by Experiments 2 and 3, inequality in input can be at least somewhat overridden when semantically grounded information in the visual context supports L2 children's processing of article-noun sequences.

In summary, the three experiments used Spanish articles to explore the nature of L1 and L2 learners' efficiencies (and previously documented inefficiencies) in real-time language processing. Like L2 adults in previous research, L2 children in Spanish immersion schools failed to exploit cues to Spanish grammatical gender when establishing reference, but they showed efficient processing of articles that conveyed biological gender and number information. Each experiment used sentences with similar structure: a simple sentence frame, a definite article, and a familiar noun. L1 children showed a continuous uptake of speech and visual information in each circumstance, using incoming information to guide attention to referents that were most likely to be talked about. But L2 children's success was (a) contingent on whether referents belonged to semantically transparent categories and (b) likely supported by their history of processing English and Spanish—particularly markers of biological gender and number in each language. The experiments broaden our knowledge of what is and is not problematic in processing a second language, revealing that interpretation of the incoming linguistic signal interacts not only with the proficiency of the listener, but also with the categorical transparency versus opacity of referents in the visual field.

Acknowledgments

This research was supported by grants from the National Institutes of Health to Casey Lew-Williams (R03HD079779) and Anne Fernald (R01DC008838). Thank you to Anne Fernald, Virginia A. Marchman, Adriana Weisleder, Nereyda Hurtado, Theres Grüter, Arielle Borovsky, Susan C. Bobb, Renate Zangl, Grace Van

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

Berkel, Ricardo A. H. Bion, Ana Luz Portillo, Lucia Rodriguez Mata, Christine E. Potter, Lily Kornbluth, Narges Afshordi, and Eva Fourakis. Special thanks to the participating families from many Spanish immersion elementary schools, as well as Eve V. Clark, Ellen M. Markman, Herbert H. Clark, and Kenji Hakuta.

NOTE

1. To explore whether participants responded with different speed for feminine versus masculine article-noun sequences, or on the first versus second half of test trials, two additional within-subjects factors (noun gender and block) were entered into the mixed ANOVA analyses for each experiment. But there were no main effects of gender or block and no significant interactions between these factors and trial type, thus statistics are not reported. Moreover, no significant differences were observed between L2 children enrolled in dual immersion versus Spanish immersion programs.

REFERENCES

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76*, 13–26.

Bates, E., & MacWhinney, B. (1981). Second language acquisition from a functionalist perspective: Pragmatic, semantic, and perceptual strategies. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences: No. 379. Native language and foreign language acquisition* (pp. 190–214). New York, NY: New York Academy of Sciences.

Brownell, R. (2000). *Receptive one-word picture vocabulary test: Spanish-bilingual edition*. Novato, CA: Academic Therapy.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Clark, E. V. (2003). Critical periods, sensitive periods, and lateralization. *University of Pennsylvania Working Papers in Linguistics, 9*(2), 39–48.

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

Dahl, A., & Vulchanova, M. D. (2014). Naturalistic acquisition in an early language classroom. *Frontiers in Psychology, 5*, 329.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.

Feinauer, E., Hall-Kenyon, K. M., & Davison, K. C. (2013). Cross-language transfer of early literacy skills: An examination of young learners in a two-way bilingual immersion elementary school. *Reading Psychology, 34*, 436–460.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam, The Netherlands: John Benjamins.

Fortune, T. W., & Tedick, D. J. (2015). Oral proficiency assessment of English-proficient K–8 Spanish immersion students. *Modern Language Journal, 99*, 637–655.

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*, 379–399.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*, 191–215.

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

Haith, M. M., Wentworth, N., & Canfield, R. L. (1993). The formation of expectations in early infancy. *Advances in Infancy Research, 8*, 251–297.

Hawkins, R., & Chan, C. Y.-H. (1997). The partial availability of universal grammar in second language acquisition: The "failed functional features hypothesis." *Second Language Research, 13*, 187–226.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., ... Suma, K. (2015). The contribution of early communication to low-income children's language success. *Psychological Science, 26*, 1071–1083.

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

Imai, M., Schalk, L., Saalbach, H., & Okada, H. (2014). All giraffes have female-specific properties: Influence of grammatical gender on deductive reasoning about sex-specific properties in German speakers. *Cognitive Science*, *38*, 514–536.
Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.
Kaushanskaya, M., Gross, M., & Buac, M. (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Developmental Science, 17*, 564–583.
Kovelman, I., Baker, S., & Petitto, L. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition, 11*, 203–223.
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*, 447–464.
MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler, & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd ed., pp. 581–611). London, UK: Elsevier.
MacWhinney, B. (2005). Extending the competition model. *International Journal of Bilingualism, 9*, 69–84.
National Center for Education Statistics. (2008). *Schools and Staffing Survey (SASS)* [Public school data file]. Retrieved from U.S. Department of Education website: https://nces.ed.gov/surveys/sass
Padilla, A. M., Fan, L., Xu, X., & Silva, D. (2013). A Mandarin/English two-way immersion program: Language proficiency and academic achievement. *Foreign Language Annals, 46*, 661–679.
Pinker, S. (1982). A theory of the acquisition of lexical interpretive grammars. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 655–726). Cambridge, MA: MIT Press.
Potowski, K. (2007). *Language and identity in a dual immersion school*. Clevedon, UK: Multilingual Matters.
Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31*, 1109–1121.
Robertson, E. K., Shi, R., & Melançon, A. (2012). Toddlers use the number feature in determiners during online noun comprehension. *Child Development, 83*, 2007–2018.
Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting.
Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research, 22*, 1–29.
Scherag, A., Demuth, L., Rösler, F., Neville, H. J., & Röder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition, 93*, B97–B108.
Smolík, F. (2014). The use of gender information in lexical processing in Czech 23-month-olds: An eye-tracking study. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the 38th Annual Boston University Conference on Language Development* (pp. 451–462). Somerville, MA: Cascadilla Press.
Song, H., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, *52*, 29–57.
Tedick, D. J., & Wesley, P. M. (2015). A review of research on content-based foreign/second language education in US K–12 contexts. *Language, Culture and Curriculum, 28*, 25–40.
Tedick, D. J., & Young, A. I. (2014). Fifth grade two-way immersion students' responses to form-focused instruction. *Applied Linguistics*, *37*, 784–807.
Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar. *Studies in Second Language Acquisition, 27*, 173–204.
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*, 87–100.
van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science, 12*, 419–425.
Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implication for language learning and language use. *Journal of Experimental Psychology: General, 134*, 501–520.
Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 20*, 1–10.

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press

Westerveld, M. F. (2014). Emergent literacy performance across two languages: Assessing four-year-old bilingual children. *International Journal of Bilingual Education and Bilingualism, 17*, 526–543.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

APPENDIX I: SPEECH STIMULI

Experiment 1: Grammatical Gender

Encuentra la pelota. ¿La ves?	(Find the ball. Do you see it?)
Encuentra la galleta. ¿Te gusta?	(Find the cookie. Do you like it?)
Encuentra el zapato. ¿Lo ves?	(Find the shoe. Do you see it?)
Encuentra el carro. ¿Te gusta?	(Find the car. Do you like it?)
¿Dónde está la vaca? ¿La ves?	(Where is the cow? Do you see it?)
¿Dónde está la rana? ¿Te gusta?	(Where is the frog? Do you like it?)
¿Dónde está el pájaro? ¿Lo ves?	(Where is the bird? Do you see it?)
¿Dónde está el caballo? ¿Te gusta?	(Where is the horse? Do you like it?)

Experiment 2: Biological Gender

Encuentra la niña. ¿La ves?	(Find the girl. Do you see her?)
Encuentra el niño. ¿Lo ves?	(Find the boy. Do you see him?)
Encuentra la señora. ¿La ves?	(Find the woman. Do you see her?)
Encuentra el señor. ¿Lo ves?	(Find the man. Do you see him?)
Encuentra la princesa. ¿La ves?	(Find the princess. Do you see her?)
Encuentra el príncipe. ¿Lo ves?	(Find the prince. Do you see him?)
Encuentra la doctora. ¿La ves?	(Find the doctor$_{[f]}$. Do you see her?)
Encuentra el doctor. ¿Lo ves?	(Find the doctor$_{[m]}$. Do you see him?)

Experiment 3: Number

Encuentra el carro. ¿Te gusta?	(Find the car. Do you like it?)
Encuentra los carros. ¿Te gustan?	(Find the cars. Do you like them?)
Encuentra el zapato. ¿Te gusta?	(Find the shoe. Do you like it?)
Encuentra los zapatos. ¿Te gustan?	(Find the shoes. Do you like them?)
Mira el caballo. ¡Que bonito!	(Look at the horse. How pretty!)
Mira los caballos. ¡Que bonitos!	(Look at the horses. How pretty!)
Mira el pájaro. ¡Que bonito!	(Look at the bird. How pretty!)
Mira los pájaros. ¡Que bonitos!	(Look at the birds. How pretty!)
Mira el gato. ¡Muy bien!	(Look at the cat. Very good!)
Mira los gatos. ¡Muy bien!	(Look at the cats. Very good!)
Mira el perro. ¡Bien hecho!	(Look at the dog. Well done!)
Mira los perros. ¡Bien hecho!	(Look at the dogs. Well done!)

https://doi.org/10.1017/S0267190517000101 Published online by Cambridge University Press