The Gender Puzzle: Toddlers’ Use of Articles to Access Noun Information

Natalia Arias-Trejo*, Alberto Falcón, Elda A. Alva-Canto

Universidad Nacional Autónoma de México

Grammatical gender embedded in determiners, nouns and adjectives allows indirect and more rapid processing of the referents implied in sentences. However in a language such as Spanish, this useful information cannot be reliably retrieved from a single source of information. Instead, noun gender may be extracted either from phono-morphological, semantic or syntactic cues, such as determiner-noun frames. This experimental work sought to explore toddlers’ ability to use feminine and masculine determiners to infer a referent whose name was marked for grammatical gender, ending in ‘o’ or ‘a’, as well as a referent whose name was unmarked. Using the intermodal preferential looking paradigm, 24-, 30- and 36-month-old children were presented with pairs of images, target and distracter, of familiar objects while, at the same time, they heard a feminine or masculine definite (Experiment 1) or indefinite (Experiment 2) article that could only refer to one of the images displayed. Half of the trials presented objects whose target names ended in a final-vowel indicative of grammatical gender and the other half had another ending. The results demonstrated toddlers’ ability to use the determiners to infer a target. Differences between toddlers’ use of definite and indefinite determiners were found. The ability to use indefinite articles preceded the ability to use definite articles. In general, the results showed that anticipation of the noun from the determiner is mostly a function of phono-morphological cues embedded in marked nouns. Nevertheless, 36-month-old children were also able to associate the articles to targets whose names were unmarked for grammatical gender.

* Acknowledgments: The research described in this article was supported by a research grant awarded to Natalia Arias-Trejo by CONACyT-CVU39175 and by a CONACYT-101752 and a PAPIIT-IN306010-3 research grant to Elda A. Alva Canto. We thank the parents and toddlers who participated in these experiments as well as two anonymous reviewers for their comments on an earlier draft of the manuscript. We are also indebted to our research assistants Tania Jasso López and Irma Vázquez López for their assistance in data collection. Corresponding author: Natalia Arias-Trejo, Facultad de Psicología, División de Investigación y Estudios de Posgrado, Universidad Nacional Autónoma de México, Cubículo 6, Edificio C, Sótono; Ciudad Universitaria, Coyoacán, Ciudad México, CP 04510. E-mail: nariast@unam.mx
In Spanish, nouns are either masculine or feminine. Accessing gender information allows more efficient and rapid linguistic processing as speakers can rely on gender information to track or infer a referent. For example, if someone has both apples and bananas and asks, ‘Do you want one?’ if the ‘one’ ends in ‘a’ as in ‘una’, a Spanish speaker can be sure that the person is offering an apple (manzana); however, if ‘one’ is marked as masculine (typically ending in ‘o’, that is ‘uno’), a banana (plátano) has been offered. Are toddlers learning Spanish using grammatical gender cues to access noun information? Can they use gender information embedded in definite or indefinite articles to infer a referent? We present two experiments that aim to answer these questions. We exposed 24-, 30- and 36-month-olds to gender information in the form of definite and indefinite articles to explore whether they could use their gender status to anticipate a referent.

Grammatical gender is assigned in different ways depending on the linguistic system. There are some formal rules which are determined by phonological and morphological cues. Morphological rules require access to word structure, typically to inflectional class information. For example, in Russian, gender may depend on inflectional class (Smelser & Baltes, 2001). In languages such as Spanish and French, morphological classes are predictable largely from phonological information (Harley, 1998; Karmiloff-Smith, 1979; Tucker, Lambert, & Rigault, 1977). Previous studies have revealed that children take advantage of the phonomorphological patterns in Hebrew, German and Russian to learn noun gender (Levy, 1983; Mills, 1985; Tucker et al., 1977). Evidence shows that infants are particularly sensitive to word-initial and word-final phonological information (Fais, Kajikawa, Amano, & Werker, 2009; Safran, Aslin, & Newport, 1996). In specifying grammatical gender, noun-vowel ending is especially relevant (Corbett, 1991). A clear example of assignment depending on phonological information in Spanish is that nouns ending in ‘a’ tend to be feminine and nouns ending in ‘o’ tend to be masculine. In an elicitation task, Pérez-Pereira (1991) found that 4- to 11-year-old Spanish children use phonological cues to ascribe gender. Participants were asked to assign gender to novel words, when phonological cues (ending ‘a’ or ‘o’) were in conflict with semantic cues (masculine or female referent), children ascribe gender by means of the phonological cues. Thus, they ascribed feminine gender to a novel word ending in ‘a’ regardless of whether it was presented within a masculine or feminine semantic context. Previous studies have also found that children are more accurate at assigning gender when it is linguistically determined through formal cues than when attention to the sex of the referent is required (Bohme & Levelt, 1979; Hernández-Pina,
Spanish Article-Noun Agreement

Nevertheless phonological regularities for gender assignment in Spanish are far from being absolutely reliable, since there are many exceptions, for instance the noun ‘mano’ (hand), although being feminine, ends in ‘o’. Moreover, there are many nouns whose endings are different from ‘a’ or ‘o’. Although there are some derivational suffixes which contain a morphological feature that marks gender in Spanish, such as the endings "-ción" / "-sión", that mark feminine gender, there are many exceptions and many other irregular endings: for example, whilst the noun ‘lápiz’ (pencil) is masculine, the noun ‘lombriz’ (earthworm) is feminine.

Therefore, whilst young learners may use the most predominant rule - ending -a/-o- to assign gender to familiar nouns, this rule is not sufficient to assign gender to all of the familiar nouns constituting their early vocabularies. According to the MacArthur Communicative Development Inventory employed to identify the vocabulary of infants learning Spanish, between 18 and 30 months of age, (Jackson-Maldonado, Thal, Marchman, Bates, & Gutierrez-Clellen, 1993; Thal, Jackson-Maldonado, & Acosta, 2000), 75% of the nouns follow the formal rule ‘o/a’. Thus, for every four nouns infants learn, the gender of approximately one of them cannot be discovered by accessing its ending. As a consequence, the gender of unmarked nouns needs to be learned by other means.

Alternatively, infants’ sensitivity to the statistical patterns in the input should also allow them to exploit information about the probabilistic co-occurrence between the gender of articles and nouns, or articles and adjectives. Thus, an alternative possibility is that infants rely on determiner-noun pairs (independent of the phonological ending of the noun) to learn gender. Comprehension tasks suggest that article-noun information can influence language processing from as early as 18 months of age (Fernald & Hurtado, 2006; Kedar, Casasola, & Lust, 2006; Zangl & Fernald, 2007). In terms of gender-marked determiners, infants learning German rely on the preceding determiner to assign gender to a noun (MacWhinney, 1978). Szagun, Stumper, Sondag and Franik (2007) found that German infants made fewer mistakes when producing article-noun frames when the regularities applied than when they did not. Importantly, infants learn gender assignment faster for the nouns produced more frequently by adults in their environment (Szagun et al; 2007). A study by Seigneuric and her colleagues (2007) showed that from a mean age of 3.5 years, French children classified pseudowords by using the masculine article ‘un’ or the feminine article ‘une’, according to their ending.
In Spanish, all nouns preceded by the definite article ‘la’ or by the indefinite article ‘una’ are feminine, whilst all nouns preceded by the definite article ‘el’ or by the indefinite article ‘un’ are masculine. Thus, it is possible that access to both the article plus the noun ending (‘a/o’) act together to accelerate and reiterate the gender of a noun. Nonetheless, not all of the articles are gender-marked: the masculine cases ‘el’ and ‘un’ have no phono-morphological gender mark as the feminine cases ‘la’ and ‘una’ do. For unmarked nouns, infants might learn to pay attention to articles as the most reliable cues for gender attribution. For example, for unmarked nouns such as ‘pastel’ (cake) infants can access information regarding its gender through a sentence frame such as ‘el pastel’ (the cake) by extracting its gender from the article. This may also apply to the article-noun agreement exceptions for nouns beginning with ‘a’ in which cases the feminine determiner is replaced by the masculine determiner (e.g., ‘el águila’) to avoid pronouncing double ‘a’. As previously suggested by Boloh and Ibernon (2010), the system can be flexible and work in two ways: when the noun ending is not informative of gender, infants pay attention to the determiner but rely on phonology when this cue is regular for the ending of the noun. Besides, in Spanish, the article can also be introduced without a noun (e.g., ‘el rojo’ – the red one, in masculine) when referring to one specific item amongst others. These replacements suggest that determiners may be particularly salient to Spanish speakers when learning gender.

In general, research has focused on young children’s ability to attribute gender in production tasks or on children’s use of noun phrase (NP) gender agreement (Karmiloff-Smith, 1979; Mulford, 1985; Pérez Pereira, 1991; Tucker et al., 1977). Research has identified children’s ability to produce grammatically correct Spanish NPs from about the age of three (e.g., Hernández-Pina, 1984; Karmiloff-Smith, 1979; Levy, 1983; Mills, 1986; Pérez Pereira, 1991), just as French and Italian learners do around the same age (Pizutto & Caselli, 1992). Although early identifications of Spanish determiner-noun production (Hernández-Pina, 1984; Idiazabal, 1995; Smith et al., 2003), it is possible that infants produce the appropriate determiner for the nouns they are most familiar with, as they have memorised familiar article-noun frames. In a longitudinal study with four infants, Mariscal (2009) found that initially (age 1;10-2;02) article production was linked to particular nouns for each child, and its use did not generalise immediately to other tokens. Various concerns can be raised regarding production data. One is the controversy around coding incomplete utterances of words. It may be more accurate to interpret an incomplete
noun than an incomplete determiner; thus, the error rate of interpretations increases when computing article-noun production.

Whilst production data can show us when a child is using gender correctly or when she has made mistakes, it is possible that some children do not produce the determiner but nevertheless are sensitive to its gender status. Thus, we examined young children’s use of gender markers during sentence comprehension. Under these circumstances, this research aims on the one hand to shed light regarding the early contribution of gender knowledge to facilitate linguistic processing; on the other hand, this research aims to evaluate whether young children are able to assign gender to both marked and unmarked nouns.

Comprehension tasks have found divergent patterns of response across different languages regarding infants’ sensitivity to gender marked article-noun frames. Van Heugten and Johnson (2011) found in a Preferential Looking task that 19- to 24-month-old Dutch learners’ sentence processing was impaired when a nonsense article as opposed to a real article preceded target words. However, gender cues embedded in the articles, which are integrated by two gender-marked forms, did not help Dutch learners to recognise target nouns more efficiently, indicating that gender has yet to be acquired. In contrast, in a very similar study with French children, it was found that by 25 months of age, French learners identified a target more efficiently when presented with two objects of different gender (Van Heugten & Shi, 2009) than when presented with two objects of the same gender. These differences in the ages in which sensitivity to article-noun agreement is present can be partially attributed to the specificities of how article-noun agreement works. In Dutch, there are two determiners for three genders; whereas in French there are two determiners for two genders. Thus, the Dutch gender system may be less transparent to learn than the French system. A similar pattern has been found by Lew-Williams and Fernald (2007) with 34- to 42-month-old Spanish learners. Gender agreement seems to be more salient in languages in which words are reiteratively gender-marked, such as in French and Spanish, than in languages, such as Dutch, in which adjectives, for example, are not always marked for gender. It is worth noting that previous research on early grammatical-gender sensitivity has mainly employed nouns which ending is indicative of gender, testing their ability to rely on the most common source of information of gender: noun ending. In contrast, the experiments presented here explore infants’ use of articles to infer a target referent that is marked for gender as well as a referent which ending is not marked for gender.
In summary, acquiring grammatical gender in a language such as Spanish is very likely to imply a not straightforward learning. Phonological cues are the most reliable source of information, but still children have to deal with an important number of irregular cases (unmarked ending nouns) and thus rely on article-noun frames. Are children biased to use one kind of cue? Are they flexible enough to take advantage of these different cues? Are there developmental changes in learning grammatical gender?

We performed two experiments by means of an intermodal preferential looking (IPL) task with the aim of testing whether 24-, 30- and 36-month-olds take advantage of the article gender to more efficiently process nouns. Furthermore, a seminal question for this study is whether infants comprehend the article-noun gender agreement for both marked and unmarked nouns. Previous research has focused on marked nouns only. The answer to this question would point to the sources of information preferentially used by toddlers to access noun gender, as well as to developmental changes that could determine this possible preference. We chose to test 24-, 30- and 36-month-olds because previous studies have reported that it is around this period that children start using gender information in comprehension tasks (Cyr & Shi, 2010; Johnson, 2005; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009).

We use the terms ‘regular’ or ‘marked’ interchangeably when referring to nouns ending in ‘a’ for feminine and ‘o’ for masculine; and ‘irregular’ or ‘unmarked’ when referring to nouns whose endings do not follow this rule. We acknowledge that these terms are not exactly the same, as an ‘irregular ending’ can be marked, for example the ending ‘-ión’ generally refers to feminine nouns. However, there are unmarked and irregular cases in which an ending, for example ‘e’, can correspond to feminine or masculine nouns (for example, ‘el elefante’ for the elephant, ‘la fuente’ for the fountain). The variety in the terminology previously employed reflects the different explanations and analyses that have been proposed in the field.

**EXPERIMENT 1**

**METHOD**

**Participants.** A total of thirty-six children (18 females) were tested in each age group. The data of nine participants were not analysed due to: experimental error (n = 3), children’s refusal to participate (n = 3), parental interaction (n = 2) and lack of knowledge with at least 50% of the presented
words \((n = 1)\). The data analysis includes thirty-three 24-month-olds, thirty-four 30-month-olds and thirty-two 36-month-olds. The mean ages were: 24 months 1 day (range = 23 months 14 days to 24 months 24 days), 30 months 1 day (range = 29 months 20 days to 30 months 28 days) and 36 months 16 days (range = 35 months 26 days to 37 months 8 days). All toddlers were monolingual and came from homes where Mexican Spanish was the only spoken language; they were born full term and had no known hearing or visual problems. Participants came from Mexico City and surroundings; they were recruited from local clinics and playgroups; as well as from printed adverts. Children came from varied socioeconomic backgrounds. However, all parents had completed at least nine years of education \((M = 14.62\) for mothers and 13.98 for fathers).

**Stimuli.** Twenty-two nouns familiar to 24-month-olds with frequencies above 60% as indicated by previous studies (Contreras Wilcox, Arias-Trejo, & Alva Canto, 2006; Jackson-Maldonado et al., 2003) were selected. The list of words grouped by their grammatical gender and characteristics of being marked or unmarked is the following:


Note that we worked with twelve different marked nouns and ten different unmarked nouns. This was due to the available nouns familiar to children from the age of 24 months; as previously stated, one of every four familiar nouns to two-year-olds is unmarked. The design takes this difference into account by presenting across children the same number of times the ten unmarked nouns and by controlling the number of times that each unmarked noun was paired with other marked or unmarked noun.

**Auditory stimuli:** the twenty-two nouns, two articles ‘*el/la*’ (the) and two carrier phrases ‘*Mira/Ve*’ (Look/See) were digitally recorded in the same session by a female voice in child-directed speech. The stimuli were edited to remove background noise and to match for peak-to-peak amplitude.

**Visual stimuli:** twenty-two coloured computer images of real items were selected from digital commercial sources and public libraries of images. All images were the same size (800 x 600 pixels) and had ten
percent of grey in the background to diminish brightness on the screen. The target-distracter nouns were all highly imageable (Bird, Franklin, & Howard, 2001; Cortese & Fugett, 2004). Table 1 lists an example of target-distracter pairings used in Experiments 1 and 2.

**Table 1. Example sequence in Experiments 1 and 2. The condition of the pairings is indicated in the regularity of nouns column. The article varies according to Experiment 1 (definite) and 2 (indefinite).**

<table>
<thead>
<tr>
<th>Trial</th>
<th>Target picture</th>
<th>Distracter picture</th>
<th>Regularity of nouns</th>
<th>Definite article/Indefinite article</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>manzana (f)</td>
<td>plátano (m)</td>
<td>regular-regular</td>
<td>LA/UNA (f)</td>
</tr>
<tr>
<td>2</td>
<td>teléfono (m)</td>
<td>paleta (f)</td>
<td>regular-regular</td>
<td>EL/UN (m)</td>
</tr>
<tr>
<td>3</td>
<td>muñeca (f)</td>
<td>caballo (m)</td>
<td>regular-regular</td>
<td>LA/UNA (f)</td>
</tr>
<tr>
<td>4</td>
<td>carro (m)</td>
<td>moto (f)</td>
<td>regular-irregular</td>
<td>EL/LA (m)</td>
</tr>
<tr>
<td>5</td>
<td>pelota (f)</td>
<td>avión (m)</td>
<td>regular-irregular</td>
<td>LA/UNA (f)</td>
</tr>
<tr>
<td>6</td>
<td>zapato (m)</td>
<td>televisión (f)</td>
<td>regular-irregular</td>
<td>EL/UN (m)</td>
</tr>
<tr>
<td>7</td>
<td>pastel (m)</td>
<td>galleta (f)</td>
<td>irregular-regular</td>
<td>EL/UN (m)</td>
</tr>
<tr>
<td>8</td>
<td>flor (f)</td>
<td>globo (m)</td>
<td>irregular-regular</td>
<td>LA/UNA (f)</td>
</tr>
<tr>
<td>9</td>
<td>árbol (m)</td>
<td>vaca (f)</td>
<td>irregular-regular</td>
<td>EL/UN (m)</td>
</tr>
<tr>
<td>10</td>
<td>moto (f)</td>
<td>pastel (m)</td>
<td>irregular-irregular</td>
<td>LA/UNA (f)</td>
</tr>
<tr>
<td>11</td>
<td>pie (m)</td>
<td>mano (f)</td>
<td>irregular-irregular</td>
<td>EL/UN (m)</td>
</tr>
<tr>
<td>12</td>
<td>llave (f)</td>
<td>calcetín (m)</td>
<td>irregular-irregular</td>
<td>LA/UNA (f)</td>
</tr>
</tbody>
</table>

**Design.** The experiment consisted of 24 trials divided into two blocks. In each block, participants saw twelve different pairs of target-distracter images. In all trials, toddlers saw one object whose gender was masculine and one object whose gender was feminine. The visual stimuli were displayed throughout the 6000 ms duration of each trial. The trials were divided into three phases: a baseline phase, an article phase and a post-naming phase. Each trial phase lasted 2000 ms. In the baseline phase of the trials, toddlers heard one of the carrier phrases ‘Mira’ or ‘Ve’ 200 ms after the onset of the visual stimuli. Toddlers’ looking at each image during this phase served as a measure of baseline image preference. This baseline phase concluded at 2000 ms ending with the definite article ‘la or el’. In the article phase, from 2000 ms to 4000 ms, the effect of the article was measured. If
young Spanish-speaking learners begin by using definite or indefinite articles to recognise upcoming nouns, they should choose to look at a target that matches the article gender. In the post-naming phase, from 4000 ms to 6000 ms, toddlers heard the target label (e.g., ‘manzana’). This phase served to confirm that toddlers were familiar with the names of the objects. Inter-trial intervals lasted around 1000 ms.

The artificial silence in between the determiner and the noun implied that we exposed toddlers to an uncommon prosody. There is evidence suggesting that children are better at recognizing words in sentence frames than in isolation (Fernald & Hurtado, 2006). However, this manipulation permitted us to examine the effect of article gender-information on referent identification. Removal of acoustic and prosodic cues to gender would make our results more robust as children would not be able to rely on any of these. Studies looking into the use of gender cues tend to manipulate the silence between words in order to measure the specific contribution of a gender marker rather than of prosodic cues (see Cyr & Shi, 2010).

On average, for each toddler each image appeared in two different trials (once in each block): once as a target and secondly as a distracter. However, for each toddler, there were a few images corresponding to unmarked nouns that appeared in three trials. We acknowledge that having an equal number of different nouns for all types of trials would have been ideal. However, due to vocabulary restrictions, this was not possible. We aimed to expose toddlers to as many different types of target-distracter pairs given the limitations of their vocabulary. This resulted in four different sets of pairings, each set containing 24 trials. Approximately, twenty-five per cent of the toddlers in each age group were allocated to each set. The side on which any particular picture was presented (right-left) and the corresponding target and distracter side were balanced within each set. Across participants, each picture in the pairs were used the same number of times as targets and as distracters and each target name was presented the same number of times within each type of trial and each carrier phrase. There were the same numbers of feminine and masculine targets for each child.

The two blocks were formed according to the criteria of increasing the degree of complexity of the assignment of grammatical gender: every three trials the complexity was increased up to the completion of a 12-trial block. The first three trials of a block displayed two images whose names followed the regular system and therefore the target and distracter names ended in ‘a’ or ‘o’. Trials 4 to 6 paired a picture whose name was regular and a picture whose name was irregular, the target always being the regular
one. Trials 7 to 9 presented the same kind of pair (regular and irregular) but in this case the target name was unmarked. Finally, the last three trials of a block introduced two pictures whose names were irregular and the target name was also irregular. The second half of the experiments followed the same logic of the four sub-blocks except that images were paired differently and the named targets were also different.

This design allowed the presentation of an equal number of trials within all the possible combinations that the Spanish gender system allows. Presentation of an increasing degree of difficulty within the task would allow toddlers to show their knowledge of the regular pattern ‘o’-’a’ before being introduced to irregular cases. However, the second block would allow us to test whether regardless of being exposed to irregular nouns; toddlers would still retain their knowledge with regular nouns. We manipulated the status of the target (regular or irregular) with the aim of measuring the impact of retrieving an irregular case when paired with another irregular case or when paired with a regular noun.

**Procedure.** Toddlers were tested using the Intermodal Preferential Looking (IPL) task. The toddler sat centrally on the caregiver’s lap in front of two monitors, placed 30cm apart, measuring 17 inches each. A loudspeaker presenting the auditory message was mounted centrally above the monitors. The images were shown at the toddler’s eye-level, at a distance of approximately 80 cm. Three hidden miniature video cameras, one mounted centrally and two side by side immediately above each picture were connected to a video-mixer that permitted recording of a split screen three angle-image of the toddler’s visual fixation on each picture. Parents were instructed to remain quiet and close their eyes, in order to avoid influencing the toddler’s eye fixations during the presentation of the pictures. The experimenter remained out of the toddler’s sight during the task.

**RESULTS**

*Assessment of visual events: Analyses of Proportion of Target Looking*

The experimenter, blind to which particular images and auditory stimuli were being presented, assessed the digital videos off-line on a frame-by-frame (every 33 ms) system to determine the direction and duration of each fixation (left, right or other). A second skilled coder evaluated the data from 10% of the participants. Agreement between scorers, assessed by computing Pearson’s correlation coefficients was $r =$
Spanish Article-Noun Agreement

.98, \( p < .001 \). Intra-scorer mean reliability was \( r = .99, p < .001 \). The same reliability scores were obtained for Experiment 2. The proportion of total looking time (PTL) was calculated for the 6000 ms of picture presentation. PTL is the proportion of target looking time divided by total looking time at the target and at the distracter \([t/(t+d)]\). We present the statistical analyses for the two phases of interest: the baseline phase and the article phase.

Individual assessment of toddlers’ knowledge of the nouns presented indicated that from the age of 24 months, toddlers were familiar with at least 80% of the nouns. We therefore analysed all the trials that each toddler was presented with. The overall pattern of results did not change when trials containing words that were potentially unknown to the toddlers were included in the analysis. After exclusion of missing trials due to temporary lack of attention there were 700 trials (88%) out of the 792 original trials presented to the 24-month-olds. For the 30-month-olds, there were 731 trials (90%) out of the 816 trials presented. For the 36-month-olds, there were 716 trials (93%) out of the original 768. Each toddler’s looking-times were aggregated by condition.

There were no significant looking-time differences between females and males. Therefore, looking times were collapsed across the two genders in the subsequent analyses. This was the same for Experiment 2. The data were analysed in a 2 x 2 x 2 x 2 x 3 analysis of variance (ANOVA) with Naming (pre-naming vs. post-naming), Block (1 vs. 2), Target gender (Masculine vs. Feminine), and Regularity of Distracter (Regular vs. Irregular) as within-subject factors; and Age (24-month olds vs. 30-month olds vs. 36-month-olds) as a between-subjects factor. The analysis revealed a main effect of Naming \([F(1, 96) = 14.86, \ p = .001, \ \eta^2 = .13]\), and a significant interaction between Naming \times\ Block \times\ Regularity of Distracter \times\ Age \([F(1, 96) = 4.03, \ p = .021, \ \eta^2 = .077]\). The Naming effect indicated an increase in target looking from pre-naming \((M = .4893, \ SD = .16)\) to post-naming \((M = .5202, \ SD = .27)\). However, the significant interaction suggested that toddlers’ target looking was influenced by other factors such as the age of the children, the status of the distracter and the presenting block.

The four-way interaction Naming \times\ Block \times\ Regularity of Distracter \times\ Age was further investigated with mean comparisons: no significant effects were shown in 24-month-olds’ looking in any condition (all \( ps > .05 \)). For the other two age groups, it was found that the increase in target looking from pre- to post-article presentation was significant for the 30-month-olds in Block 2 with Regular Distractors \([t(125) = 2.94, \ p = .004, \ d = .53]\), all other comparisons were not significant (all \( ps > .10 \)). For the 36-month-
olds, significant increases in target looking were found in Block 1 with Regular Distracters ($t(123) = 2.50, p = .014, d = .45$) and in Block 2 with Irregular Distracters ($t(122) = 2.016, p = .046, d = .36$). Figure 1 illustrates this interaction by showing the PTL increase: subtraction of the proportional mean value after the onset of the article from the proportional mean value before the onset of the article. A positive value indicates a looking increase to the target upon disambiguation, while a negative value indicates an increase in looking to the distracter.

![Figure 1](image)

**Figure 1.** Mean ($±1SE$) PTL increase from Pre- to Post-naming in Experiment 1 by Distracter Status, Block and Age. Significant increase at *p < .05 and **p < .01.*

Although this research focused on target looking enhanced by the onset of the article, we confirmed that for all three age groups, target looking was significant from the onset of the target name, all $ps < .01$. 
DISCUSSION

To conclude, these results demonstrated that 30- and 36-month-olds can anticipate a target referent when previously exposed to definite articles for both gender-marked and unmarked target names. Thus, toddlers identify a regular target (e.g., ‘manzana’) as well as an irregular target (‘flor’). However, whether the target is paired with a regular (‘plátano’) or an irregular distracter (‘pastel’) influences toddlers’ ability to map an article onto the correct referent. Specifically, this ability follows a developmental trend as 30-month-olds can only identify a target when it is paired with a regular distracter. In contrast, 36-month-olds identify a target when paired with either regular or irregular distracters. This variable response could be due to the design of the experiment, which sought to account for a more realistic scenario, closer to that which young children, learning Spanish, are actually confronted with when learning concordance of gender between articles and nouns. Would young children perform better with indefinite articles even within the same testing scenario? In Spanish, the articles un/a are frequently used to introduce exemplars, as in ‘Mira, una mariposa’ (Look, a butterfly) when surrounded by members of other categories. Also, the fact that the same words, un-una, can be used as indefinite articles as well as numerals might make the indefinite article quite salient.

A number of studies have explored toddlers’ knowledge of indefinite articles. Mills (1986) and Szagun (2004) report more frequent use of indefinite articles in German. However, (Mills, 1986) reports higher error rates for indefinite than for definite articles. Szagun et al., (2007) found an earlier increment for indefinite articles (at 2;04) than for definite articles (at 2;07) in German learners. In consequence, in Experiment 2 we aimed to test whether being exposed to indefinite articles would facilitate finding a target at the same ages for which we tested the use of definite articles.

EXPERIMENT 2

This experiment is almost identical to Experiment 1. The only difference is that definite articles were replaced by indefinite articles.

METHOD

Participants. We tested the same age groups as for Experiment 1. A total of thirty-six children in each age group were tested. There were 20, 19 and 18 females in the 24-, 30- and 36-month-olds groups respectively. The data from four children was not analysed due to their lack of familiarity
with at least 50% of the presented words as reported by their parents \((n = 1)\), their failure to pay attention to more than 50% of the trials \((n = 1)\) and experimental error \((n = 2)\). The data analysis includes the participation of thirty-four 24-month-olds, thirty-five 30-month-olds and thirty-five 36-month-olds. The mean ages were: 24 months (range = 23 months 15 days to 24 months 16 days), 30 months 8 days (range = 29 months 10 days to 30 months 24 days) and 36 months 2 days (range = 36 months 1 day to 37 months 10 days). Participants had the same characteristics as for Experiment 1. All parents completed at least 9 years of education \((M = 14.86\) for mothers and 14.16 for parents).

**Stimuli.** All visual stimuli were the same as those used in Experiment 1. All auditory stimuli were the same as for Experiment 1, except that the definite articles ‘el’/‘la’ were replaced by the indefinite articles ‘un/una’.

**Design and Procedure.** The design and procedure were the same as in Experiment 1.

**RESULTS**

The data were analysed in the same manner as for Experiment 1. Toddlers were familiar with at least 80% of the nouns, thus we analysed all the trials that each participant was presented with. After exclusion of missing trials due to temporary lack of attention there were 728 trials (89%) out of the 816 original trials presented to the 24-month-olds. For the 30-month-olds, there were 758 trials (90%) out of the 840 presented trials. For the 36-month-olds, there were 729 trials (87%) out of the 840 original trials.

As for Experiment 1, the data were analysed in a 2 x 2 x 2 x 2 x 3 analysis of variance (ANOVA) with the factors Naming (pre-naming vs. post-naming), Block (1 vs. 2), Target gender (Masculine vs. Feminine), Regularity of Distracter (Regular vs. Irregular) as within-subjects factors; and Age (24-month olds vs. 30-month olds vs. 36-month-olds) as a between-subjects factor. The PTL analysis revealed a main effect of Naming \((F(1, 102) = 10.63, \ p = .002, \ \eta^2 = .095)\) and a significant interaction between Naming x Block x Target Gender x Regularity of Distracter \((F(1, 102) = 4.127, \ p = .045, \ \eta^2 = .039)\).

The Naming effect indicated an increase in target looking from pre- \((M = .4857, \ SD = .20)\) to post-naming \((M = .5147, \ SD = .27)\). The four-way interaction, illustrated in Figure 2, showed that toddlers differed in their target preferences, and these differences depended on the Gender of the
Target, the Regularity of the Distracter and the Block. To evaluate this interaction we performed mean comparisons with all of these factors. We found that the comparisons did not reach a level of statistical significance, except for the following. In Block 1, toddlers correctly identified the target upon hearing the indefinite article when it was paired with a regular distracter and when the target was masculine ($t(199) = 3.05, p = .003, d = .43$). In contrast, in Block 2, toddlers identified the target when it was feminine and again, when it was paired with a regular distracter ($t(192) = 2.14, p = .034, d = .31$). This pattern of results indicates that toddlers can use the gender of the indefinite article to identify a target whose name follows the ending rule ‘o/a’ or which does not. However, this ability depends upon the regularity of the distracter. Regular distracters enhance identification of regular and irregular targets whereas irregular distracters disrupt target identification.

Figure 2. Mean (±1SE) PTL increase from Pre- to Post-naming in Experiment 2 by Distracter Status, Block and Gender of Target. Significant increase at *p < .05 and **p < .01.
Thus, in Block 1 toddlers employed the indefinite masculine article to identify either a regular masculine target, such as ‘plátano’, or an irregular masculine target, such as ‘pastel’, when paired with a regular feminine distracter, such as ‘manzana’, but not when paired with an irregular feminine distracter, such as ‘flor’. In Block 2, toddlers used the indefinite feminine article to identify both feminine regular targets, such as ‘manzana’, and feminine irregular targets, such as ‘flor’, when paired with a regular masculine noun, such as ‘plátano’, but not when paired with an irregular masculine noun, such as ‘pastel’. Finally, as for Experiment 1, we confirmed that for all three age groups, target looking was significant from the onset of the target name, all ps < .01.

**DISCUSSION**

These results, as with previous research (Cyr & Shi, 2010; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009), suggest that access to determiners in the form of definite and indefinite articles allows young children to be more efficient at processing speech. Gender information accessible from articles seems to activate congruent candidate words. We have demonstrated that toddlers learning Spanish are capable of using definite and indefinite article gender information to disambiguate a referent. Moreover, this is the first work showing that this ability is functional for Spanish marked nouns, ending in ‘o’ for masculine and ‘a’ for feminine, as well as for unmarked nouns, i.e. with an ending that is not indicative of grammatical gender according to the ‘a/o’ distinction. It is unlikely that our results were driven by natural gender cues as only two of the items presented, doll and teddy bear, contained natural gender cues. Taken together, the results from Experiments 1 and 2 suggest that the use of indefinite articles precedes the use of definite articles to identify a target: at 24 months of age, toddlers were able to use indefinite articles; whereas it was not until 30 months that toddlers used both definite and indefinite articles. The ability to use article information to anticipate a referent was modulated by various factors such as the regularity of the distracter, the gender of the target and the progression of the task.

The pattern of results suggests that by the age of 36 months, children learning Spanish may be using a combination of morpho-phonological cues (noun endings) and syntactic cues (article-noun frames) to access grammatical gender information. The influence of morpho-phonological cues is demonstrated by the relevant role that the status of the distracter noun played when the target name was either marked or unmarked: toddlers were more accurate at finding a target paired with a marked than with an
unmarked distracter. The role of syntactic cues is suggested by toddlers’ ability to associate definite articles with both marked and unmarked target nouns. However, the results suggest that toddlers mainly rely on the regularity of the distracter to correctly associate a determiner to an unmarked referent, suggesting that the mechanism of accessing morphophonological information to anticipate a referent is dominant.

An apparent disparity seems to exist between definite and indefinite articles: at twenty-four months of age, toddlers seem to have problems retrieving useful information from definite articles but not from indefinite articles. This finding parallels previous research indicating some advantage for indefinite articles: Mills (1986) and Szagun (2004) report more frequent use of indefinite articles in German toddlers. This may reflect a correct use of the articles in the input: definite articles are intended to be used with referents that are well enough identified from previous experience whereas indefinite articles are commonly used to refer to any given object.

Apart from comparing the early use of definite and indefinite articles to infer a referent, this research has investigated toddlers’ sensitivity to the agreement between articles and marked and unmarked nouns. To the best of our knowledge, this is the first study that systematically investigates toddlers’ ability to associate articles to unmarked nouns. Our results indicate that in some cases, toddlers are capable of correctly use article information to infer an object which name is marked for gender or which name is unmarked. A consistent finding across Experiments 1 and 2 was that toddlers found it easier to use the gender of the article to infer a target when this was paired with a marked distracter (ending in ‘o’ or ‘a’) than when paired with an unmarked distracter. The only case in which children do not seem to rely on the distracter status is when facing definite articles at 36 months of age. Imagine for a second that toddlers did not know the gender of the irregular target, the fact that they knew the gender of the regular distracter allowed them to disregard the distracter as a good candidate for the article. For example, if the distracter noun ‘manzana’ was paired with the target ‘pastel’, the fact that toddlers knew that ‘manzana’ is a feminine noun would allow them to decide that when hearing the article ‘el’ the correct match was the other picture (‘pastel’). This could be comparable to a mutual exclusivity strategy, in which a target is selected via the rejection of a familiar target, though in this case by means of gender information. This type of response indicates that toddlers were activating at least the gender of one of the objects depicted on screen within the space of 2000ms. This finding parallels previous research reporting 18-month-olds’ rapid mental activation of labels to select a target (Mani & Plunkett, 2010).
We have gone a step further by demonstrating toddlers’ rapid activation of noun-gender to disambiguate two items.

We do not want to argue that toddlers are not able to identify the gender of irregular nouns. In fact, 36-month-olds were able to use definite articles when the distracter names were marked or unmarked implying that they may be using syntactic cues to disambiguate between two unmarked nouns. This ability could be the consequence of having more opportunities to practice a variety of article-noun frames but also of being more exposed to article-noun frames for words that are, at that age, very frequently experienced. Nonetheless, the finding that rapid access to gender identification of familiar nouns is highly dependent on whether the distracter is marked suggests strong reliance on morpho-phonological cues.

It is worth noting that when accessing gender information from indefinite articles (‘un/a’), the gender of the target also impacts toddlers’ ability to disambiguate a target. In Experiment 2, toddlers identified both regular and irregular masculine targets, introduced by ‘un’ paired with a regular distracter in Block 1. Later on, in Block 2, they identify both regular and irregular feminine targets introduced by ‘una’ when paired with regular distracters. This pattern of results suggests that the masculine article ‘un’ is initially easier to match than the feminine article ‘una’. Our results converge with previous production data finding some variability in toddlers’ ability to master both masculine and feminine article-noun agreement (Hernández-Pina, 1984; Perez Pereira, 1991). Results from different studies have found differences in the direction of the error children make when using determiners.

For example, in 1991, Pérez-Pereira reported that 4-11 year-old children tended to ascribe masculine gender more often than feminine gender, particularly in instances where novel words had ambiguous endings. In this respect, according to the theory of markedness (Greenberg, 1966), the masculine is the unmarked term, and, thus easier to acquire. This tendency has also been observed in other languages such as Hebrew (Levy, 1983) and French (Karmiloff-Smith, 1979). In contrast, López-Ornat (1997) found in a single-case longitudinal study fewer errors with feminine than masculine determiner-noun pairs between 2;1 and 2;2. The authors attributed their findings to the phonological simplicity and clarity of the feminine form. Similar results were found with children learning Italian (Pizutto & Caselli, 1992). The gender differences observed in Experiment 2 cannot be explained by a greater knowledge of masculine nouns in comparison to feminine nouns as, before the age of three, Spanish learners know approximately the same number of feminine and masculine nouns.
Research finding a masculine default strategy would suggest that toddlers are more familiar with the masculine article as it works as the default article for novel nouns and for generic nouns that can take feminine gender but that generically are talked of as masculine. However, we must clarify that we selected nouns which have only one gender to avoid this influence. It is inconclusive why toddlers in our experiment were able to use first ‘un’ and, as the task progressed, they were able to use ‘una’.

Taken together, the results from Experiments 1 and 2 indicate that 24-month-olds were able to use the indefinite articles, ‘un’ or ‘una’ but unable to use the definite article ‘el’ or ‘la’ to infer a target. In contrast, 30- and 36-month-olds were able to use both types of articles to correctly infer a target. In a longitudinal study of sixteen Spanish-speaking children (Farell-Rodríguez, Hernández, Suárez-Brito, & Alva, 2009), it was found changes in the use of the articles at four different age points. At 18 months, few productions of the definite articles ‘el’ and ‘la’ were encountered, 3 and 8 times respectively. From 24 months, toddlers introduced indefinite articles, ‘un’ and ‘una’, to their speech, and showed an onset of indefinite articles at 30 months. At 36 months, children were using with very similar frequencies both definite and indefinite articles. The initial predominant production of definite articles was driven by toddlers’ use of article-noun frames for the most familiar nouns: Farell et al. (2009) report that the 18- and 24-month-olds participating in their study were using the articles ‘el’ and ‘la’ repeatedly for a reduced number of familiar nouns. It was not until the age of 30 months that children extended their use of articles to different nouns, coinciding with the onset of indefinite article production. The increase in toddlers’ use of the indefinite articles may be the consequence of the initiation of the counting system in which parents very often introduced the objects, within a frame such as ‘mira un’ or ‘mira una’ versus ‘mira dos’. It is also possible that a more specific distinction between the use of definite and indefinite articles is present in the input from caregivers as children develop.

This research coincides with previous comprehension studies in Spanish and French reporting that seeing two objects of different grammatical gender produces faster target looking than seeing two objects of the same grammatical gender between the ages of 2 to 4 (Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009). Data from Dutch did not reveal this, which may be due to language differences. Nevertheless, Van Heughten and Johnson (2011) found that, although Dutch children do not seem to be using the gender-marking information, between 19 and 24 months of age, they are sensitive to correct or incorrect article-noun pairs. In our design, it was necessary for children to know the gender of the target
and the distracter names to be able to accurately locate the unnamed target. Thus, our study shows an active and efficient way of using the gender information contained in the Spanish articles to infer an unlabeled referent. More recently, Cyr and Shi (2010) found that in French children aged 30 months, but not younger children were sensitive to article-novel noun congruencies, showing evidence that they have abstract knowledge of gender classes for articles and use them to perform formal categorisation of novel nouns.

The current work diverges from previous studies in so far as we have tested both regular and irregular nouns. Previous research in Spanish has focused on infants’ knowledge of the gender of regular nouns. Our design may be more challenging to toddlers as we have introduced as well irregular nouns. Moreover, we have exposed toddlers to 22 different familiar nouns; previous studies have explored toddlers’ sensitivity to the gender of more limited well-known familiar items. We believe that our design allowed us to test in a more robust manner the way in which infants are using gender cues, embedded in determiners, for the purpose of referent identification. To conclude, two to three-year-old children are sensitive to article-noun frames. Initially they seem to rely on morpho-phonological regularities to associate an article to an unlabeled referent but by the age of three, they are able to take advantage of syntactic information when those regularities are not available, which shows toddlers’ ability to exploit the available information, in this case for the purpose of acquiring an important part of the Spanish’s phono-morphological system. The ability to use determiners clearly is allowing toddlers from a very early age to track referents, disambiguate them and infer them. Future research should explore whether toddlers can also learn novel words based on the Spanish article-noun system.

RESUMEN

El Rompecabezas del Género: los Niños Pequeños Usan los Artículo para Acceder a Información del Sustantivo. El género gramatical contenido en los artículos, sustantivos y adjetivos facilita un procesamiento más eficiente de los referentes implicados en los enunciados. Sin embargo, en una lengua como el español, esta información no puede ser extraída de manera confiable de una única fuente. En lugar de ello, el género del sustantivo debe ser determinado a partir claves fono-morfológicas, semánticas, o sintácticas como las combinaciones de artículos y sustantivos. En este trabajo experimental se propuso explorar la habilidad que tienen los infantes para usar los artículos femeninos o masculinos para inferir un referente cuyo nombre está marcado en su género gramatical, terminación
‘a’ u ‘o’, en contraste con un referente cuyo nombre no posee esta marcación. Se empleó el paradigma intermodal de atención preferencial. Se les presentó a niños de 24, 30 y 36 meses de edad pares de imágenes de objetos familiares mientras escuchaban un artículo femenino o masculino, definido (Experimento 1) o indefinido (Experimento 2). La mitad de las imágenes tenían nombres con una terminación indicativa de su género gramatical (con marcación), mientras que en la otra mitad la terminación de sus nombres no era indicativa (sin marcación). Los resultados mostraron que los infantes emplean los artículos para inferir un referente. Asimismo, se observaron diferencias en el uso de artículos definidos versus indefinidos: la habilidad para usar los artículos indefinidos emerge antes que para usar los definidos. En general, se observó que la identificación del referente a partir del artículo está determinada principalmente por las claves fonomorfológicas de los sustantivos. Sin embargo, los niños de 36 meses de edad fueron capaces de asociar los artículos a referentes cuyos nombres no poseían la terminación a/o de género gramatical.

REFERENCES


(Manuscript received: 21 November 2011; accepted: 24 January 2012)