

*BUCLD 35 Proceedings*  
*To be published in 2011 by Cascadilla Press*  
*Rights forms signed by all authors*

## **Online Comprehension of Newly Acquired Nouns and Abstract Knowledge of Grammatical Gender**

**Andréane Melançon and Rushen Shi**

### **1. Introduction**

The system of grammatical gender is exhibited in some human languages. In these languages, grammatical categories such as nouns, adjectives and determiners may be each subdivided into smaller categories such as the masculine and the feminine classes. Certain languages (e.g., German, Russian) have more than two gender classes. When words co-occur to form larger utterances, they need to agree in gender according to appropriate syntactic principles. For instance, in the French noun phrase *la*<sub>FEM</sub> *bonne*<sub>FEM</sub> *banane*<sub>FEM</sub> ('the good banana') the determiner, adjective and noun all agree in gender. Violations of such agreement result in ungrammatical utterances (e.g., *\*le*<sub>MASC</sub> *bonne*<sub>FEM</sub> *banane*<sub>FEM</sub> or *\*la*<sub>FEM</sub> *bon*<sub>MASC</sub> *banane*<sub>FEM</sub>). In French, most determiners and many adjectives have both feminine and masculine forms, for example, *la*<sub>FEM</sub>/*le*<sub>MASC</sub> ('the'), *ma*<sub>FEM</sub>/*mon*<sub>MASC</sub> ('my'), *une*<sub>FEM</sub>/*un*<sub>MASC</sub> ('a'), *cette*<sub>FEM</sub>/*ce*<sub>MASC</sub> ('this'), *petite*<sub>FEM</sub>/*petit*<sub>MASC</sub> ('little'), *chanceuse*<sub>FEM</sub>/*chanceux*<sub>MASC</sub> ('lucky'), although certain words in these two classes are not marked for gender (e.g., *les*-plural 'the', *magnifique*-'magnificent'). Spanish has a similar distinction for determiners (i.e., *la*<sub>FEM</sub>/*el*<sub>MASC</sub>-'the', *una*<sub>FEM</sub>/*uno*<sub>MASC</sub>-'a'), and adjectives (e.g., *buena*<sub>FEM</sub>/*bueno*<sub>MASC</sub>-'good', *cara*<sub>FEM</sub>/*caro*<sub>MASC</sub>-'expensive'). Gender assignment for most common nouns (especially inanimate nouns) behaves differently than for determiners and adjectives. Nouns are either of feminine or masculine form, not both (e.g., in French: *maison*<sub>FEM</sub>-'house', *soulier*<sub>MASC</sub>-'shoe'; in Spanish: *manzana*<sub>FEM</sub>-'apple', *sombrero*<sub>MASC</sub>-'hat'). The present study concerns children's knowledge of grammatical gender in French.

Gender categories are abstract syntactic constructs. The genders of nouns referring to objects are not semantically related. There is no particular reason why the word *canapé* in French ('couch') is masculine rather than feminine. Children therefore cannot rely on semantic features to learn to assign words to gender categories. On the other hand, morphological marking can be a cue to gender categories in some languages. Italian masculine forms end in *-o* or *-i*, whereas feminine ones end in *-a* or *-e* (Bates, Devescovi, Hernandez, &

---

\* Andréane Melançon and Rushen Shi, Université du Québec à Montréal, [andreanemel@yahoo.ca](mailto:andreanemel@yahoo.ca). This research was supported by SSHRC, NSERC and CFI grants to the second author. We thank Mireille Babineau, Jun Gao and Marilyn Cyr for research assistance. We also thank all parents and infants for participating in our study.

Pizzamiglio, 1996). In Spanish nouns endings in *-o* are more often associated with the masculine category, and those ending in *-a* are often feminine forms (Mariscal, 2009). French nouns do not have such kind of regular morphological endings. Although some nouns show gender-consistent phonological cues (e.g., the feminine *-ette* ending: *fourchette*<sub>FEM</sub> ‘fork’, *assiette*<sub>FEM</sub> ‘plate’; the masculine *-eau* ending: *bateau*<sub>MASC</sub> ‘boat’, *rideau*<sub>MASC</sub> ‘curtain’), this is by no means systematic. In fact, a large number of nouns do not have any consistent phonological cue for gender. Therefore, French learners must learn to rely primarily on the distributional regularities between nouns and other categories such as determiners in order to acquire gender agreement.

The idea that determiners may play a primary role for acquiring gender agreement is in accord with the theoretical assumption that they are important in the structure of noun phrases (or strictly speaking, determiner phrases - DPs). There is evidence that determiners are a small set of frequent items that are perceptible and encoded as early as the preverbal stage (e.g., Hallé, Durand, & de Boysson-Bardies, 2008; Höhle & Weissenborn, 2003; Shi, Werker & Cutler, 2006; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008). Determiners, a closed class, thus may serve to cue the gender category of nouns, a vast open class. Indeed, the use of determiners for the categorization of nouns as a global class has been documented in infants shortly after one year of age (Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Shi & Melançon, 2010).

The knowledge of grammatical gender in young children has long been the interest of many language acquisition researchers (e.g., Corrêa & Name, 2003; Johnson 2005; Lew-Williams & Fernald, 2007; van Heugten & Shi, 2009). Several recent studies have used online comprehension tasks to examine children’s processing of the determiner in noun phrases and their representations of gender agreement. In particular, Lew-Williams and Fernald presented Spanish-learning children aged between 34 and 42 months with pictures depicting pairs of familiar objects. In some trials, both objects were of the same grammatical gender (*pelota*<sub>FEM</sub> ‘ball’, *galetta*<sub>FEM</sub> ‘cookie’). In other trials, objects were of different grammatical genders (*pelota*<sub>FEM</sub> ‘ball’, *zapato*<sub>MASC</sub> ‘shoe’). For each pair, children heard a sentence naming one of the two objects (*Encuentra la pelota, la ves?* – ‘Find the ball, do you see it?’). When the two objects were of different genders, the determiner was informative about the gender of the target noun. However, when the objects were of the same gender, the preceding determiner in the instruction sentence was compatible in gender with the nouns for both objects (i.e., not disambiguating), so children needed to wait until the noun unfolded in order to select the target. Indeed, infants were faster in looking at the target when a pair of objects was denoted by nouns of different genders than when objects belonged to nouns of the same gender.

Johnson (2005) explored the perception of gender agreement in Dutch-learning children, a language that has specific determiners for common gender singular nouns (CG) and other determiners for neutral gender singular nouns (NG). Johnson presented 28-month-olds with pairs of objects in a similar way as Lew-Williams and Fernald (2007). For some trials, objects were of the same

gender (*bal*<sub>CG</sub>-‘ball’, *boom*<sub>CG</sub>-‘tree’) while in other trials objects were of different genders (*bal*<sub>CG</sub>-‘ball’, *boek*<sub>NG</sub>-‘book’). Simultaneous to the presentation of objects in each trial, the child was instructed to look at one of them (*Kijk eens naar de*<sub>CG</sub> *bal*<sub>CG</sub>-‘look at the ball’). In addition, there were trials presenting objects of different gender, but the auditory stimuli contained mismatch of gender agreement between the determiner and noun (e.g. \* *Kijk eens naar het*<sub>NG</sub> *bal*<sub>CG</sub>-‘look at the ball’). Trials presenting objects of different genders were analyzed in two ways, and evidence of gender agreement perception for one of the determiners was found: for trials in which looking happened to be on the distractor at the beginning of the determiner, reaction time to target was faster when the determiner and noun were correct in gender agreement than when they were not in agreement, indicating that the gender of the determiner was processed during the recognition of the target noun. For those in which looking happened to be on the target at onset of the determiner, shifting away from target to distractor was faster when gender agreement was mismatched, suggesting that the determiner (which agreed in gender with the distractor) was misleading the children to orient to the distractor.

In addition to the above studies in Spanish and Dutch, van Heugten and Shi (2009) showed gender agreement processing in even younger children. As in Johnson (2005), they presented 25-months-old Quebec-French-learning infants with pairs of familiar objects of different genders (e.g., *bateau*<sub>MASC</sub>-‘boat’, *banane*<sub>FEM</sub>-‘banana’) and the same gender (e.g., *ballon*<sub>MASC</sub>-‘ball’, *bateau*<sub>MASC</sub>-‘boat’). In trials which introduced objects of different genders, auditory instruction contained correct gender agreement in some trials (*Oh regarde! La*<sub>FEM</sub> *banane*<sub>FEM</sub>-‘Oh look! The banana’), but incorrect gender agreement in other trials (\**Oh regarde! Le*<sub>MASC</sub> *banane*<sub>FEM</sub>-‘Oh look! The banana’). Auditory stimuli for trials presenting same-gender objects were always correct in gender agreement. Infants in that study processed the targets more efficiently in different gender trials in which gender agreement was correct. When gender agreement was incorrect, performance was much worse. Moreover, when both objects were of the same gender, the determiner was uninformative about the gender of the target, and the recognition of the target was delayed relative to different-gender correct trials but better than incorrect trials.

These previous online comprehension studies suggest that children learning gender-marked languages are sensitive to the gender system from 25 months of age, and that they can use this knowledge for the online processing of the noun. The findings are compatible with the view that abstract grammatical knowledge is present early in acquisition (e.g., Valian, 2009). However, these results can also be interpreted in terms of memorized exemplars that represent no abstract syntactic categories (e.g., Pine & Lieven, 1997). According to this view, there is no abstract syntactic representation during early acquisition. Young children’s utterances during the early stage (e.g., age two) are formulaic or memorized chunks based on what they have heard in the input, and the utterances reflect conceptual meanings without any abstract syntactic structure. As the materials used in the previous online gender processing studies were all familiar words to

young children, results could be because the grammatical pairings between determiners and those familiar nouns are highly frequent in the input, whereas the ungrammatical pairings are never or rarely heard. It is most likely that children have heard and stored in memory the specific exemplars such as *la banane*, but not the ungrammatical *\*le banane*. Thus, it is unclear whether the processing advantage for correct gender agreement over the incorrect ones shown in previous comprehension studies reflects abstract knowledge of gender agreement.

To tease apart these two interpretations, it is necessary to examine gender agreement knowledge by using novel nouns. This makes it possible to rule out the likelihood that target nouns in correct agreement with gender-informative determiners are better recognized simply because specific determiner-noun sequences have often occurred in the input and are memorized by infants. Therefore, in the present study we examined the abstract syntactic knowledge of gender agreement in French-learning infants, using pseudo-nouns.

Our specific goal in this study is to test the following questions: 1) do infants immediately encode the grammatical gender of the determiner to be an integral part of the representation of a novel noun during word-object mapping? If so, 2) is the gender that is encoded onto the noun an abstract feature (e.g., manifested by the recognition of its generalized agreement with other determiners that share the same gender)? Neither of the two questions can be fully addressed with familiar nouns. We examined these questions using a word-learning task involving determiners and pseudo-nouns. In particular, we first presented infants with novel objects each paired with a pseudo-noun in a context of a gender-marked determiner during a training phase, and then tested how they recognized the nouns in new determiner contexts in which the determiner gender feature was manipulated (present & agreeing, present & non-agreeing, absent).

## **2. Methods**

### **2.1. Participants and auditory stimuli**

Twenty-two 30-month-old Quebec-French-learning infants completed the experiment (mean age: 945 days, range: 910-965 days). Auditory stimuli included five determiners: two indefinite singular: *un* – masculine, *une* – feminine and three definite: *le* – masculine singular, *la* – feminine singular, and *les* – neutral plural (no gender specification; can co-occur with nouns and adjective of both genders). All these determiners are highly frequent in French. In addition, we chose pseudo-nouns that respect the phonological properties of French (i.e., *cagère*, *ravole*). The use of pseudo-nouns ensures that all infants were equally unfamiliar with the words, thus enabling us to better assess their ability to abstractly encode grammatical gender. These words were previously constructed by Cyr and Shi (submitted) in a gender-assignment study with Quebec-French speaking adults. They asked their participants to judge if the pseudo-nouns sounded typically French. They were also asked to create sentences with the pseudo-words as nouns. Their results showed that *cagère* and

*ravole* were among the forms that were comparably used as feminine and masculine. That is, the two forms were phonologically ambiguous in terms of grammatical gender. This control was necessary for testing if infants could rely entirely on the determiner to extract the gender feature and assign it to the novel noun. The pseudo-nouns and the determiners formed determiner phrases (DPs) that served as our stimuli: *un*<sub>MASC</sub> *ravole*, *un*<sub>MASC</sub> *cagère*, *une*<sub>FEM</sub> *ravole*, *une*<sub>FEM</sub> *cagère*, *le*<sub>MASC</sub> *ravole*, *le*<sub>MASC</sub> *cagère*, *la*<sub>FEM</sub> *ravole*, *la*<sub>FEM</sub> *cagère*, *les*<sub>NEUTRAL</sub> *ravoles*, *les*<sub>NEUTRAL</sub> *cagères*.

In addition to the pseudo-nouns, we also selected 12 real nouns, all familiar to two-year-old infants. They were used to replicate the previously published online comprehension studies. Moreover, infants' responses to these words served as a baseline for comparing with their responses to the pseudo-nouns. Six of the familiar words were of the feminine gender: *poule*–'hen', *fleur*–'flower', *pomme*–'apple', *maison*–'house', *banane*–'banana', *grenouille*–'frog'. Six others were of the masculine gender: *singe*–'monkey', *ballon*–'ball', *lapin*–'rabbit', *soulier*–'shoe', *chapeau*–'hat', *mouton*–'sheep'. Six of these familiar words were each paired with an indefinite determiner (e.g., *un lapin*, *une fleur*), and the other six were each paired with a definite determiner (e.g., *le mouton*, *la maison*). Table 1 shows all DPs containing the familiar words.

**Table 1. Familiar nouns presented during the experiment (gender agreement is illustrated by the preceding determiner)**

Familiarization	Test
<i>un</i> <sub>MASC</sub> <i>lapin</i> <sub>MASC</sub> – 'a rabbit'	<i>le</i> <sub>MASC</sub> <i>soulier</i> <sub>MASC</sub> – 'the shoe'
<i>un</i> <sub>MASC</sub> <i>singe</i> <sub>MASC</sub> – 'a monkey'	<i>le</i> <sub>MASC</sub> <i>mouton</i> <sub>MASC</sub> – 'the sheep'
<i>un</i> <sub>MASC</sub> <i>ballon</i> <sub>MASC</sub> – 'a ball'	<i>le</i> <sub>MASC</sub> <i>chapeau</i> <sub>MASC</sub> – 'the hat'
<i>une</i> <sub>FEM</sub> <i>fleur</i> <sub>FEM</sub> – 'a flower'	<i>la</i> <sub>FEM</sub> <i>banane</i> <sub>FEM</sub> – 'the banana'
<i>une</i> <sub>FEM</sub> <i>poule</i> <sub>FEM</sub> – 'a hen'	<i>la</i> <sub>FEM</sub> <i>maison</i> <sub>FEM</sub> – 'the house'
<i>une</i> <sub>FEM</sub> <i>pomme</i> <sub>FEM</sub> – 'an apple'	<i>la</i> <sub>FEM</sub> <i>grenouille</i> <sub>FEM</sub> – 'the frog'

A native Quebec-French female speaker produced multiple tokens of the stimuli in a sound-proof booth. They were digitally recorded using a Sound Device 702T at 48 kHz sampling frequency and 24 bits bit rate. Familiar words were recorded in DPs (Det + Noun) in varying intonations. For the pseudo-nouns, the speaker recorded the phrases containing the neutral-gender determiner *les* (*les*<sub>NEUTRAL</sub> *cagères*, *les*<sub>NEUTRAL</sub> *ravoles*). The phrases containing the pseudo-nouns and the singular determiners were the recordings from the infant experiments of Cyr and Shi (submitted), e.g., *un*<sub>MASC</sub> *ravole*, *un*<sub>MASC</sub> *cagère*, etc. Finally, two different carrier phrases were recorded to introduce the DP in each familiarization trial (*Oh*–'Oh', *tu l'aimes?*–'do you like it?'), and another one to introduce the DP in each test trial (*Oh regarde!* – 'Oh look!'). The same speaker who recorded the stimuli of Cyr and Shi (submitted) recorded all the remaining stimuli for the present study.

It should be mentioned that all pseudo-nouns DPs containing gender-marked determiners were already prepared via a cross-spliced technique (from Cyr & Shi, submitted). Each pseudo-noun (i.e., *cagère* or *ravole*) produced in isolation was cross-spliced and conjoined with a determiner and the beginning part of the onset consonant of the same pseudo-noun that were produced in different DPs, such as *la cagère*, *le cagère*. That is, the same pseudo-noun token followed a feminine determiner and a masculine determiner, respectively. This careful control ensured that all possible acoustical cues to gender in the pseudo-nouns from DPs were removed so that we could examine if infants in our experiment could encode the gender feature for the novel nouns exclusively from the determiner.

The final stimuli set included two tokens of each pseudo-noun DP and one token of each familiar word DP for the familiarization phase, and one token for each pseudo-noun DP and familiar word DP for the test phase. One token for each carrier phrase was used throughout the experiment.

## **2.2. Visual stimuli**

To represent the two pseudo-nouns (*ravole*, *cagère*), we used pictures of objects unknown to children under age three (i.e., a red wrench and a silver shaker). Familiar nouns, such as *maison* or *lapin*, were depicted by colorful pictures representing the words. We also created an animation of a rising rainbow together with a bird song, which were used as the attention getter between all test trials. A moving image showing colorful balls along with a cheerful interjection (i.e., ‘Wow!’) was used at the beginning of the experiment to acquaint the infant with our procedure.

## **2.3. Procedure**

All children were individually tested in an acoustical chamber in a split screen paradigm. The child sat on the knees of his or her parent, facing a 42-inch television screen. Loudspeakers were positioned adjacent to each side of the screen, and a Panasonic video camera recorded the infant. To minimize possible distractions, a black curtain covered the walls, leaving only visible the television screen. Throughout the experiment, the parent listened to masking music from headphones. During the experiment the hidden Panasonic camera sent simultaneous video signals of the infant to a monitor in the neighboring room, where the experimenter observed the looking behavior of the child. The researcher could not hear the auditory stimuli, nor see the visual stimuli, ensuring no bias. As soon as the child was settled and looked at the screen, the experimenter started the experiment. Every test trial was infant-controlled (i.e., initiated only when the infant looked at the screen). All testing sessions were later coded offline.

## **2.4. Design**

The experiment contained a familiarization phase and a test phase. The familiarization phase presented pseudo-noun trials and familiar-noun trials. Each trial presented a single image on either the left or right side of the display monitor, along with the auditory stimuli that named the object. The pause between all familiarization trials was 300 milliseconds. The pseudo-noun trials were each introduced with a carrier phrase (i.e., *Oh, un/une N*; or *Un/une N, tu l'aimes?*), and the trials had a mean duration of 2700 milliseconds. The two pseudo-nouns, *cagère* and *ravole*, were each preceded by an indefinite determiner, one with *un* and the other with *une*. Six familiar nouns were presented, three feminine and three masculine (see Table 1). They were also preceded by an indefinite determiner corresponding to the intrinsic gender of each noun (e.g., *un lapin, une fleur*), each depicted by an image. These known-noun trials served to introduce the word-learning task to the infants, informing them that the image of each novel object was named by the accompanying auditory stimuli. In total, the two pseudo-nouns were each presented seven times and the six familiar nouns one time, all in a random order. The familiarization phase lasted approximately one minute.

Immediately after the familiarization phase, the test phase started. Each trial simultaneously presented pictures of two different objects, on the display monitor, one object on the left and another on the right. One object corresponded to a feminine noun, the other a masculine noun. The auditory stimuli for each test trial named one of the two displayed objects. The speech included a definite determiner *le, la, or les*, which had not been presented during the familiarization phase. The target noun in each trial was produced in the carrier phrase *Oh regarde, le/la/les N*. In trials containing *les*, one side of the screen displayed two copies of one object while the other side displayed two copies of another object.

All test trials were constructed in the same way: First, the two objects appeared on each side of the screen in silence. Exactly 2000 milliseconds afterwards, an auditory instruction started, naming one of the objects with a definite determiner (e.g., *Oh regarde, le + N*). The pictures stayed until the end of the trial. The carrier *Oh regarde* was the same token across all trials, and the onset of the determiner occurred exactly at 3500 milliseconds. The average duration was 298 milliseconds for the determiners, 842 milliseconds for the pseudo nouns, and 975 milliseconds for the known nouns.

We used six familiar nouns in the test phase that were different from those in the familiarization phase, for the purpose of preventing the infant from getting bored (see Table 1). The same pseudo-nouns of the familiarization phase were used in the test phase.

By varying the pairing between determiners and pseudo-nouns in Det + N phrases, three different test trial types were created: Novel-Noun Grammatical, Novel-Noun Ungrammatical, and Novel-Noun Neutral. These trial types correspond to the three ways in which the determiner gender feature was manipulated: present & agreeing, present & non-agreeing, absent. In the

Grammatical trial type, the pseudo-noun was preceded by a determiner of the same gender as in the familiarization phase. For example, if an infant learned in the familiarization phase that *ravole* was a feminine noun (i.e., following a feminine determiner, *une*<sub>FEM</sub> *ravole*<sub>FEM</sub> – ‘a ravole’), test trials presenting *ravole* with another feminine determiner (i.e., *la*<sub>FEM</sub> *ravole*<sub>FEM</sub> – ‘the ravole’) would be grammatical. Thus, the determiner gender feature was present and agreed with the noun gender. The Ungrammatical trial type presented phrases with Det + N pairings incompatible with the familiarization gender assignment. That is, if *ravole* was trained as a feminine noun, the test trials presenting it with a masculine determiner (*\*le*<sub>MASC</sub> *ravole*<sub>FEM</sub> – ‘the ravole’) would be non-agreeing and ungrammatical. Finally, in the Neutral trial type, the auditory instruction included a neutral determiner *les*, which did not contain any gender feature (e.g., *les*<sub>NEUTRAL</sub> *ravoles*<sub>FEM</sub> – ‘the ravoles’).

The familiar-word phrases were presented in two test trial types: Known-Noun Grammatical and Known-Noun Neutral. In the Grammatical trial type, each masculine noun was preceded by a definite masculine determiner *le*, and each feminine noun by a definite feminine determiner *la*. In the Neutral trial type, two identical images of one object appeared on the left side of the screen while two images of another object on the right side. One object was expressed by a masculine noun and the other by a feminine noun. The auditory instruction included the neutral determiner *les* that did not contain any gender feature (e.g., *les*<sub>NEUTRAL</sub> *maisons*<sub>FEM</sub> – ‘the houses’).

Each pseudo-noun was used for each of the three test trial types, forming a total of six test trials (two per trial type). For known-noun test trials, four nouns were selected randomly out of the six known nouns and presented to different groups of infants, also in a random assignment. Two of the known nouns were used for the Grammatical trial type, and another two used for the Neutral trial type. The order stimuli trial presentation was quasi-randomized to ensure variability of determiner use, noun use, and the side for the target across test trials. Word-object mapping and gender-word assignment for the pseudo-nouns were fully counterbalanced. Table 1 shows examples of materials used for one group of infants.

**Table 2. Example stimuli for one group of infants**

<b>Familiarization</b>	<b>Audio Stimuli</b>	<b>Visual Stimuli</b>
	<i>un</i> <sub>MASC</sub> <i>lapin</i> <sub>MASC</sub>	<i>lapin</i> – ‘rabbit’
	<i>une</i> <sub>FEM</sub> <i>maison</i> <sub>FEM</sub>	<i>maison</i> – ‘house’
	<i>un</i> <sub>MASC</sub> <i>ravole</i> <sub>MASC</sub>	<i>ravole</i> – red wrench
	<i>une</i> <sub>FEM</sub> <i>cagère</i> <sub>FEM</sub>	<i>cagère</i> – silver shaker
<b>Test Trials</b>	<b>Audio Stimuli</b>	<b>Visual Stimuli</b>
Novel-Noun Grammatical	<i>le</i> <sub>MASC</sub> <i>ravole</i> <sub>MASC</sub>	<i>ravole</i> – <i>cagère</i>
Novel-Noun Ungrammatical	<i>*la</i> <sub>FEM</sub> <i>ravole</i> <sub>MASC</sub>	<i>cagère</i> – <i>ravole</i>
Novel- Noun Neutral	<i>les</i> <sub>NEUTRAL</sub> <i>ravoles</i> <sub>MASC</sub>	<i>ravoles</i> – <i>cagères</i>
Known-Noun Grammatical	<i>la</i> <sub>FEM</sub> <i>maison</i> <sub>FEM</sub>	<i>maison</i> – <i>soulier</i>
Known-Noun Neutral	<i>les</i> <sub>NEUTRAL</sub> <i>chapeaux</i> <sub>MASC</sub>	<i>bananes</i> - <i>chapeaux</i>

## 2.5. Coding Eye Movements

All infants were off-line coded frame by frame at a rate of 30 frames/sec by a blind researcher. Using the SuperCoder Software (Hollich, 2005), we coded each frame as either left look, right look or away (i.e., looking elsewhere).

## 3. Predictions

The Known-Noun test trials were designed to examine the processing of grammatical gender during infants' online recognition of familiar nouns, as has been done in previous comprehension studies (Johnson, 2005; Lew-Williams & Fernald, 2007; van Heugten & Shi, 2009). In addition, these trials could serve as a baseline to be compared with the Novel-Noun test trials. We predicted that recognition of the target in the Known-Noun Grammatical trials should be better than in the Known-Noun Neutral trials, where the neutral determiner *les* was uninformative of the gender of the target noun. Positive results with known nouns, however, cannot definitively show if infants have abstract gender knowledge and encode the feature from the determiner onto the noun when learning a new word. We tested this question with novel nouns.

During the familiarization phase, the training input included novel objects each paired with a novel noun, and the gender feature of the noun was only expressed in the preceding determiner. If infants at 30 months of age have no abstract gender knowledge, or if they have such knowledge but simply fail to show it in online processing, we should then expect to see no difference in looking responses to the three test trial types (Grammatical, Ungrammatical, Neutral). In that case, infants may either show equally good recognition of the novel words in all trial types, or show equally poor recognition across trial types (if they fail to learn the word-object mappings).

On the other hand, if infants have abstract knowledge of the gender feature in determiners and automatically encode the feature onto novel nouns as an integral part of the noun representation during our word learning task, they should then show better recognition of the target in the Grammatical test trials than in the other two trial types, in which the feature is either missing or incorrect in agreement. This prediction applies in our study only if infants can map word forms to objects. But if infants only encode the gender feature onto nouns without achieving word-object mapping, we should observe confusion in the recognition of the target noun across all three types of test trials. This outcome (failure in word-object mapping) is unlikely since the existing literature suggests that two-year-olds are generally good at mapping words to objects.

## 4. Results

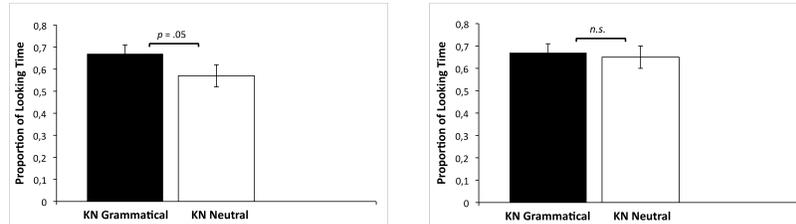
We first examined the Known-Noun trials. The analysis window was one second starting at 300 milliseconds from the onset of the determiner, following Lew-Williams & Fernald, 2007. The Proportion of looking to target (PLT) was

calculated as the total looking time to the target divided by the sum of the looking time to the target and that to the distractor. Looks away from the screen were not considered. A two-tailed paired t-test showed that infants looked longer at the target picture during Grammatical trials ( $M = .67$ ,  $SE = .03$ ) than Neutral trials ( $M = .57$ ,  $SE = .04$ ,  $t(21) = 2.029$ ,  $p = .055$ ). We also compared each PLT with the chance level 0.50 to assess whether infants recognized the target. The PLT for the Grammatical trials was significantly above chance ( $t(21) = 4.596$ ,  $p < .001$ ), but this was not the case for the Neutral trials ( $t(21) = 1.670$ ,  $p = .110$ ). The results suggest that familiar noun recognition was better in the Grammatical trials than in the Neutral trials, and that there was no recognition of the target in the Neutral trials within this time window.

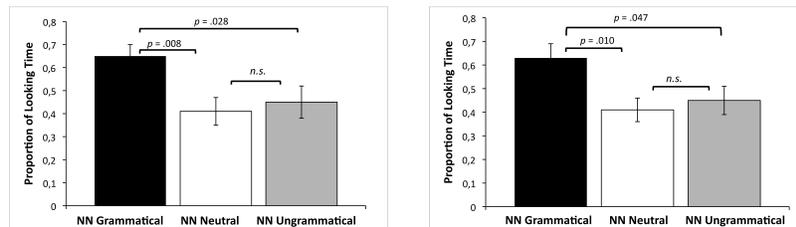
These results were comparable to those of the earliest time window in van Heugten and Shi (2009). But van Heugten and Shi also showed that familiar noun recognition in gender-neutral trials reached the level of the grammatical trials in a slightly later time window. We therefore analyzed the PLTs in a slightly later one-second window, starting at 800 milliseconds after the determiner onset, corresponding to the later part of the noun and a post-noun period. A two-tailed paired t-test revealed that PLT in Grammatical trials ( $M = .67$ ,  $SE = .05$ ) no longer differed from that in Neutral trials ( $M = .64$ ,  $SE = .05$ ),  $t(21) = .443$ ,  $p = .662$ . Both looking times were significantly different from chance level (Grammatical:  $t(21) = 3.487$ ,  $p = .002$ ; Neutral:  $t(21) = 2.846$ ,  $p = .01$ , respectively). Thus, when the determiner contained no gender information, familiar words were recognized, but more slowly than when the determiner was gender-informative.

Subsequently, we analyzed the Novel-Noun trials. A repeated-measure ANOVA with Trial Type as the factor (Grammatical, Ungrammatical, Neutral) revealed a significant effect ( $F(2,42) = 3.839$ ,  $p = .029$ ). Subsequent pairwise comparisons showed that infants looked significantly longer in Grammatical trials ( $M = .65$ ,  $SE = .06$ ) compared to Ungrammatical ( $M = .45$ ,  $SE = .07$ ),  $p = .028$ , and to Neutral trials ( $M = .41$ ,  $SE = .07$ ),  $p = .008$ . Ungrammatical and Neutral trials were not different ( $p = .730$ ). Further comparisons to the chance level 0.50 revealed that PLT for Grammatical trials was significantly different from chance ( $t(21) = 2.669$ ,  $p = .014$ ), whereas PLTs for Ungrammatical and Neutral trials were not ( $t(21) = -.710$ ,  $p = .486$ ,  $t(21) = -1.171$ ,  $p = .255$ , respectively).

For the later time window, the same ANOVA analysis again showed an effect of Trial Type ( $F(2,42) = 4.003$ ,  $p = .026$ ). Pairwise comparisons revealed that, as in the first time window, infants maintained a longer PLT for Grammatical trials ( $M = .63$ ,  $SE = .06$ ) than for the Ungrammatical ( $M = .46$ ,  $SE = .06$ ,  $p = .047$ ) and Neutral ( $M = .41$ ,  $SE = .06$ ,  $p = .010$ ) trials. Again, PLTs for Ungrammatical and Neutral trials were not different ( $p = .625$ ). Comparisons with the chance level revealed that PLT was again significant above chance in Grammatical trials ( $t(21) = 2.180$ ,  $p = .041$ ), but not in Ungrammatical ( $t(21) = -.803$ ,  $p = .431$ ) and Neutral trials ( $t(21) = -1.402$ ,  $p = .176$ ).



**Fig. 1. Proportion of Looking Time (means and standard errors) to the Known Noun trials in the first time window (left graph) and the second time window (right graph).**



**Fig. 2. Proportion of Looking Time (means and standard errors) to the Novel Noun trials in the first time window (left graph) and the second time window (right graph).**

## 5. Discussion

The present study tested the processing of grammatical gender in 30-month-old French-learning infants. We replicated the findings of van Heugten and Shi (2009), confirming that infants indeed use gender-marked determiners in online comprehension of familiar nouns. Crucially, the results of Novel-Noun trials revealed that infants encoded the abstract gender feature in determiners and mapped it onto the novel nouns in DPs during the training period. When gender agreement was respected in the test trials, infants showed successful recognition of the newly acquired nouns. Recognition of the novel nouns failed when the preceding determiner contained no gender feature or carried a gender feature that did not agree with the gender of the noun that infants had learned. Our results showed that infants' gender feature representation was abstract because 1) the nouns in our experiment were novel, i.e., those phrases could not have been memorized unanalyzed exemplars, and 2) the determiners in the test phase were different from those of the familiarization phase, thus requiring the infants

to perceive the abstract gender feature shared by determiners of the same gender and to compare it with the gender feature of the novel noun.

The training phase in this study presented infants with several different kinds of information: the semantic properties of a novel object, the phonological form of a novel noun, and the grammatical gender feature carried by the determiner within the DP. In the mature grammar of French, all these aspects are encoded in the lexical representation of nouns. The results of the Novel-Noun trials suggest that infants encode all three kinds of information when encountering a new noun. Importantly, abstract grammatical information is immediately represented during word learning. In the test phase of our experiment infants automatically activated the gender feature of the determiner and that of the noun, as well as the semantic representation of the noun. The presence and the value of the gender feature of determiners yielded a major impact on the online recognition of nouns in DPs. These findings are consistent with the theoretical notion that abstract syntactic structures are represented during the early stage of language acquisition (e.g., Valian, 2009).

Our study demonstrated the use of abstract grammatical gender category knowledge for word learning in 30-month-olds. This knowledge must have been acquired at an earlier age. Indeed, a number of studies have shown that categorization learning occurs shortly after one year of age. For example, English infants as young as 12 months of age have been shown to categorize verbs (Mintz, 2006), and German 14- to 16-month-olds achieve noun categorization (Höhle et al., 2004). Categorization of determiners has been observed in French-learning infants aged 14 months (Shi & Melançon, 2010). With regards to grammatical gender, Gerken, Wilson and Lewis (2005) showed that English-learning 17-month-old infants categorized gender categories after a brief exposure to stimuli representing a Russian gender paradigm. In a recent study on grammatical gender in French, determiner phrases containing pseudo-nouns (including those in the present study) were used to test Quebec-French-learning infants, and infants by 20 months of age categorized the gender of nouns on the sole basis of their distribution with gender-marked determiners (Cyr & Shi, submitted). All those studies used preferential looking tasks testing the assignment of word forms to grammatical categories, and presented no semantic information (neither known words, nor images for concepts), thus suggesting that categorization occurs at an early age based on the properties of the word forms and/or their distributions. The knowledge demonstrated in the present study follows logically from the ability shown in those previous studies, especially that in Cyr and Shi. Infants do not need to accumulate a sizable meaningful vocabulary before acquiring their grammatical categories. Rather, abstract grammatical representations such as the gender feature are acquired early in infancy without meaning. As demonstrated in the present study, the knowledge is readily activated and used in subsequent language acquisition tasks, including the learning of novel words and the syntactic relations among words in phrases.

## References

- Bates, Elizabeth, Devescovi, Antonella, Hernandez, Arturo, & Pizzamiglio, Luigi (1996). Gender priming in Italian. *Perception & Psychophysics*, *58*, 992-1004.
- Corrêa, Leticia M. Sicuro, & Name, Maria Cristina L. (2003). The processing of Determiner – Noun agreement and the identification of the gender of Nouns in the early acquisition of Portuguese. *Journal of Portuguese Linguistics*, *1*, 19-43.
- Cyr, Marilyn, & Shi, Rushen (submitted). Development of Distribution-Based Grammatical Categorization in Infants.
- Gerken, Louann, Wilson, Rachel, & Lewis, William (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.
- Hallé, Pierre A., Durand, Catherine, & de Boysson-Bardies, Bénédicte (2008). Do 11-month-old French Infants Process Articles? *Language and Speech*, *51*, 23-44.
- Höhle, Barbara, & Weissenborn, Jurgeb (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*, 154-159.
- Höhle, Barbara, Weissenborn, Jurgen, Kiefer, Dorothea, Schulz, Antje, & Schmitz, Michaela (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, *5*, 341-353.
- Hollich, George (2005). *Supercoder: A program for coding preferential looking* (Version 1.5). [Computer Software]. West Lafayette: Purdue University.
- Johnson, Elizabeth K. (2005). Grammatical gender and early word recognition in Dutch. In *Proceedings of the 29th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Lew-Williams, Casey, & Fernald, Anne (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*, 193-198.
- Mariscal, Sonia (2009). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language*, *36*, 143-171.
- Mintz, Toben H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action Meets Word: How Children Learn Verbs*, (pp. 31-63). New York: Oxford University Press.
- Pine, Julian M., & Lieven, Elena V.M. (1997). Slot and frame patterns in the development of the determiner category. *Applied Psycholinguistics*, *18*, 123-138.
- Shi, Rushen, Cutler, Anne, Werker, Janet F., & Cruickshank, Marisa (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, *119*, EL61-EL67.
- Shi, Rushen, & Lepage, Mélanie (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, *11*, 407-413.
- Shi, Rushen, & Melançon, Andréane (2010). Syntactic Categorization in French-Learning Infants. *Infancy*, *15*, 517-533.
- Shi, Rushen, Werker, Janet F., & Cutler, Anne (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*, 187-198.
- Valian, Virginia (2009). Abstract linguistic representations and innateness: The development of determiners. In W.D. Lewis, S. Karimi, H. Harley, & S. Farrar. (Eds.), *Time and Again: Theoretical Perspectives on Formal Linguistics in Honor of D. Terence Langendoen* (pp: 189-206). Amsterdam: John Benjamins.
- van Heugten, Marieke, & Shi, Rushen (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, *12*, 419-425.