

Development of Abstract Grammatical Categorization in Infants

Marilyn Cyr and Rushen Shi
Université du Québec à Montréal

This study examined abstract syntactic categorization in infants, using the case of grammatical gender. Ninety-six French-learning 14-, 17-, 20-, and 30-month-olds completed the study. In a preferential looking procedure infants were tested on their generalized knowledge of grammatical gender involving pseudonouns and gender-marking determiners. The pseudonouns were controlled to contain no phonological or acoustical cues to gender. The determiner gender feature was the only information available. During familiarization, some pseudonouns followed a masculine determiner and others a feminine determiner. Test trials presented the same pseudonouns with different determiners in correct (consistent with familiarization gender pairing) versus incorrect gender agreement. Twenty-month-olds showed emerging knowledge of gender categorization and agreement. This knowledge was robust in 30-month-olds. These findings demonstrate that abstract, productive grammatical representations are present early in acquisition.

When learning a language, children face many challenges. For example, they need to segment word forms from continuous speech, map meaning to word forms, and learn the grammatical properties of the words. To acquire syntax, children must learn that sentences are not simply composed of specific words, but rather are composed of abstract grammatical categories that can be combined to generate an infinite number of possible utterances. The acquisition of grammatical categories has been a major focus in child language research. In this study we examine abstract categorization of grammatical classes in infants, specifically grammatical gender in French.

Theoretical models have been proposed to address the question of how infants acquire the system of grammatical categories. According to the semantic bootstrapping model (e.g., Pinker, 1984), infants begin by using semantic concepts (e.g., objects, actions) and syntax–semantics linking rules to associate word forms in the input with innate syntactic categories (such as nouns and verbs), a process that is essential for infants' analysis of the distributional regularities of these categories. Based

on another model, certain phonological and prosodic cues in the input mark grammatical categories and constituents, and infants can use these cues to discover initial grammatical categories and syntactic relations (e.g., Christophe, Guasti, & Nespor, 1997; Morgan, 1986; Shi, Morgan, & Allopenna, 1998). They may notice the phonological or prosodic marking of a class of words, and use it to analyze the co-occurrence distribution of these words with other classes or with specific close-class items (e.g., determiners, auxiliaries). A third type of theory (e.g., Maratsos & Chalkley, 1980; Mintz, 2003) focuses on the direct support of distributional properties among words in the input (e.g., co-occurrence patterns and relative positions of words in sentences) for initial grammatical categorization. Thus, the three types of theories differ with respect to the kinds of cues that can bootstrap infants into the grammatical system, but converge on the use of word distributional patterns during category learning.

There had been a limited number of empirical studies on initial categorization. There is evidence that grammatical categorization begins at an early age. It has been shown that 1- to 3-day-old newborns can use prosodic and phonological information to categorize content words versus function words (Shi, Werker, & Morgan, 1999), the most fundamental, broad grammatical categorization. This is consistent with the observation that, across

This research was supported by NSERC, SSHRC, and CFI grants to the second author, and an NSERC scholarship to the first author. Parts of this study were presented at the BUCLD 34 meeting (Boston, 2009). We thank Amy Arnott, Andréane Melançon, Mireille Babineau, and Elena Koulaguina for their assistance. We also thank the four anonymous reviewers and the editor for their helpful comments.

Correspondence concerning this article should be addressed to Rushen Shi, Département de psychologie, Université du Québec à Montréal, C.P. 8888, Succursale Centre-Ville, Montréal, QC, Canada H3C 3P8. Electronic mail may be sent to shi.rushen@uqam.ca.

© 2012 The Authors
Child Development © 2012 Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2012/xxxx-xxxx
DOI: 10.1111/j.1467-8624.2012.01869.x

languages, the two classes of words are distinct in their spoken forms: In comparison to content words, function words are universally minimized, both acoustically and phonologically (e.g., Cutler, 1993; Monaghan, Christiansen, & Chater, 2007; Shi et al., 1998). It is therefore logical that the categorization of these two global, universal classes is the first to emerge in infants.

Shortly after their first birthday, infants begin to show knowledge about more specific categories, such as nouns or verbs. Höhle, Weissenborn, Kiefer, Schulz, and Schmitz (2004) showed noun categorization in German-learning infants aged 14–16 months. After being familiarized with pseudowords that were preceded by a determiner in German, infants discriminated sentences containing the same pseudowords in new contexts appropriate for nouns from sentences containing the pseudowords in contexts appropriate for verbs. The authors concluded that by 14 months of age infants possess a representation of the co-occurrence pattern between determiners and nouns, and use this cue for categorization.

Mintz (2006) reported verb categorization in 12-month-old English-learning infants. Infants were familiarized with sentences, each containing a pseudoword. Certain pseudowords were used as nouns (i.e., surrounded by frequent frames of function words supporting the noun category, such as *the_in*) and others were used as verbs (e.g., in the frame of *to_it*). In the test phase, the same pseudoword and function word frames appeared, but the pairings were novel, in that some were grammatical and others were ungrammatical. Infants discriminated grammatical trials versus ungrammatical trials for pseudowords that had been familiarized in verb-use sentences but not for those that had been familiarized as nouns.

Evidence of noun categorization was also shown with 14-month-old French-learning infants in Shi and Melançon (2010). In that study, one group of infants was familiarized with pseudowords preceded by determiners (noun condition) and another group with pseudowords preceded by pronouns (verb condition). Both groups of infants were then tested on the same pseudowords co-occurring with other determiners versus with other pronouns, that is, different function words from those presented during familiarization. Infants in the noun condition showed evidence of categorization (by discriminating grammatical vs. ungrammatical trials). The stimuli were carefully controlled so that the pseudowords contained no prosodic (pitch, amplitude, or duration) or phonological differences when used as nouns versus verbs. Thus, noun categorization was solely based on

the distribution of the preceding determiners. That study also showed that infants perceived different determiners as an equivalent class.

Grammatical category learning has also been investigated in artificial languages and unknown natural languages. Gómez and Lakusta (2004) conducted an experiment with an artificial language using an *aX bY* paradigm. They found that 12-month-olds were able to form categories by associating the distributional information of *a* and *b* words with a phonological feature (i.e., monosyllabic or disyllabic) of *X* and *Y* words. Moreover, infants generalized the categorization to novel *X* and *Y* words. Gerken, Wilson, and Lewis (2005) familiarized English-learning 17-month-olds with a Russian grammatical gender paradigm, and found that infants were able to form the categories when two cues were copresent in the input: distributional information provided by gender suffixes and phonological information at the end of the noun stem. Infants required the co-occurrence of the two cues to form the categories. Interestingly, once they learned the categories using the two cues, they no longer needed both. They generalized the categorization to test stimuli that contained only the gender suffixes.

The previous experimental findings indicate that grammatical categorization begins early in life, first with the global bifurcation of content words versus function words at birth. Categorization of nouns and verbs begins some time after 1 year of age. Initial grammatical categorization does not depend on semantic cues, as either the targets or all the stimuli in those studies were unknown words to the infants. The evidence thus far suggests that phonological cues along with distributional information support initial category learning.

It appears that function words may constitute an efficient kind of distributional cue to initial categorization in infants. This is reasonable given that function words, which occupy crucial syntactic positions, are a small set of words that occur extremely frequently in natural languages. Function words and their distributional regularities with adjacent nonfunction words may therefore be particularly salient to infants. In an adult artificial language learning experiment, grammatical categorization performance was optimal when participants were presented with a small set of highly frequent words (i.e., resembling function words) relative to a large set of infrequent open-class words (Valian & Coulson, 1988). Given that specific function words are perceived and encoded early in acquisition, starting from about 6 to 8 months of

age (Hallé, Durand, & Boysson-Bardies, 2008; Höhle & Weissenborn, 2003; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008; Shi, Marquis, & Gauthier, 2006; Shi, Werker, & Cutler, 2006), we suggest that they may serve as the distributional anchor for infants to categorize less frequent open-class content words.

In this study we examined the development of the ability to use function words alone as the distributional anchor for categorizing content words. We addressed this question by testing French-learning infants' categorization of nouns according to grammatical gender. In French, nouns are divided into masculine and feminine categories. The determiner and the noun within a noun phrase agree in gender (e.g., *la*_{feminine} *maison*_{feminine}, "the house"; *le*_{masculine} *ballon*_{masculine}, "the ball"). Crucially, the assignment of gender for nouns is largely arbitrary semantically. For example, there is no particular reason why *lit* ("bed") is masculine and *table* ("table") is feminine. Hence, grammatical gender categorization cannot be acquired through semantics. Furthermore, nouns in French do not contain consistent phonological markings for gender, unlike in languages with clear gender morphophonological markings on the noun (e.g., Spanish, Russian). Although certain markings have been observed for some nouns in corpus and dictionary analyses, the markings were only weakly probabilistic (e.g., Tucker, Lambert, & Rigault, 1977). Thus, infants must learn to categorize genders of nouns based on their distributional relations with other words in the sentence, especially with determiners.

To unambiguously test if infants can categorize the gender of nouns on the sole basis of distributional cues related to determiners, we used novel nouns and strictly controlled their forms to avoid all possible cues to gender. In particular, we did an adult prestudy and an input analysis to ensure that the novel words that we constructed contained no phonological marking to gender. In addition, acoustic control was necessary to avoid potential distinct acoustic cues to gender in stimuli production. This consideration was based on the finding that English-speaking mothers produce noun-verb ambiguous words (e.g., *talk*, *walk*, *kiss*) differently, in that nouns tend to be longer in duration, higher in pitch, and more dynamic in pitch changes than verbs (Conwell & Morgan, 2012). Shi and Moisan (2008) also found prosodic differences in French-speaking mothers' productions of noun-verb ambiguous words. In previous infant perceptual studies of grammatical categorization (reviewed

above), the same pseudowords in different categories might have been produced with subtle acoustic differences. We therefore used the cross-splicing technique in this study to completely remove any possible speech cue to gender. Specifically, a same pseudoword token produced in a gender-neutral condition was spliced and conjoined with masculine and feminine determiners, respectively, thus preserving only the distributional cues of the gender feature in the determiners.

Grammatical genders of nouns are subcategories of the larger noun category. It is plausible for gender categorization to emerge after the higher order noun-verb categorization has been learned. Distributionally, nouns of different genders seem to exhibit a higher degree of overlap in sentences than nouns versus verbs. For example, nouns and verbs typically appear in distinct syntactic contexts, whereas nouns of masculine and feminine genders both occur in noun phrases, can precede the same auxiliaries and prepositions, and can follow the same gender-neutral determiners (plurals), numerals, and adjectives. Thus, gender categorization may be somewhat delayed in comparison to the categorization of the larger noun class. Indeed, gender knowledge has been observed in production (e.g., Bassano, 1998; Bassano, Maillochon, & Eme, 1998; Corrêa & Name, 2003) and comprehension (Johnson, 2005; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009) studies in children after age 2, later than for nouns and verbs. Nevertheless, French determiners might offer a certain degree of distributional support for acquiring noun genders because determiners are structurally required in front of all common nouns, making French more restrictive than other languages such as German and Dutch. For example, sentences like *Hunde rennen schnell* ("Dogs run fast") and *Nahrung ist wichtig* ("Food is important") contain no determiner and are grammatical in German, whereas in French, the equivalent expressions require a determiner before the noun (i.e., *Les chiens courent vite* and *La nourriture est importante* rather than **Chiens courent vite* and **Nourriture est importante*). Thus, infants have many chances to hear gender-marking determiners in front of nouns in French (in comparison to German and Dutch), which may provide considerable support for gender categorization at a young age.

In short, we hypothesized that during the 2nd year of life infants can categorize grammatical gender classes of nouns based strictly on the distributional support of determiners. We tested infants' abstract ability to categorize gender, using

pseudonouns and a purely syntactic task (i.e., with no semantic component). The use of pseudonouns was important for revealing whether the categorization was productive, that is, whether the knowledge involved abstract equivalent classes that were generalizable to novel instances. Our study thus differed from previous gender studies that used familiar nouns (Johnson, 2005; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009), and we reasoned that infants should have already acquired abstract grammatical gender categorization at an earlier stage. In a preferential looking procedure, we first familiarized infants with pseudonouns following gender-marked French determiners. Infants were then tested with the pseudonouns co-occurring with other gender-marked determiners in French that were not used in familiarization. The determiner and pseudonoun pairings were grammatical (i.e., consistent with the familiarized gender pairings) in half of the test trials and ungrammatical (i.e., inconsistent with the familiarized gender pairing) in the other trials. Looking times were measured during grammatical and ungrammatical test trials.

Experiment 1

Method

Participants. Participants were sixteen 14-month-old ($M_{\text{age}} = 1;2;21$; range = 1;2;6–1;3;2; 9 males, 7 females), sixteen 17-month-old ($M_{\text{age}} = 1;5;20$; range = 1;5;8–1;5;30; 6 males, 10 females), sixteen 20-month-old ($M_{\text{age}} = 1;8;24$; range = 1;8;1–1;9;11; 11 males, 5 females), and sixteen 30-month-old ($M_{\text{age}} = 2;6;21$; range = 2;6;3–2;7;14; 6 males, 10 females) monolingual French-learning infants. An additional eighteen infants were tested but were not included in the analyses because of fussiness (eight), parental interference (three), experimental error (three), equipment failure (one), and ceiling looking (three, i.e., never turned fixation away from the display monitor throughout the entire experiment).

Stimuli. The stimuli consisted of pseudonouns and French determiners. There were several reasons for the use of pseudonouns in this study. First, using pseudonouns allowed us to examine infants' capacity to assign novel words to abstract gender categories. Second, it allowed us to eliminate potential phonological or acoustical biases for gender. Finally, using pseudonouns allowed us to control for infants' prior exposure to the stimuli, as the stimuli in our experiment were equally unfamiliar to all infants.

Four pseudonouns were used in the stimuli set: *mouveille*, *ravole*, *cagère*, and *gombale*. They were chosen from a prestudy with French-speaking adults. We designed the prestudy to determine which pseudonouns were equally probable in both masculine and feminine genders. The goal was therefore to neutralize any phonological cues to gender in the pseudonouns for our experiments with infants. In this prestudy, 10 participants were orally presented with 26 pseudowords that we constructed. Following the presentation of each pseudoword, the participant had to orally create a sentence using the pseudoword as a noun. Next, he or she was asked to decide whether the pseudoword was noun-like in French, and whether it was a typical-sounding noun. As determiners generally occur in noun phrases in French, the participants spontaneously produced a determiner before the pseudoword in every sentence. Therefore, we were able to implicitly induce gender production from the adults for each pseudoword. The rationale is that if a pseudoword is phonologically biased toward one of the two genders, it should be produced more often with a determiner of that gender, whereas if a pseudoword is not phonologically gender biased, it should be equally likely to be produced with a determiner of either gender. The four pseudoword listed above were produced with comparable numbers of masculine versus feminine determiners across participants (*mouveille*: 40% masculine, 60% feminine; *ravole*: 50% masculine, 50% feminine; *cagère*: 50% masculine, 50% feminine; *gombale*: 60% masculine, 40% feminine). They were also judged as typical-sounding nouns in French.

In addition, a corpus analysis of infant-directed speech in French was conducted to ensure that the specific endings of the four pseudonouns contained no phonological marking to gender in children's input. We counted the frequencies of endings in all masculine and feminine nouns in sample transcripts from several corpora of the CHILDES database (MacWhinney, 2000): Anais, Marie, Nathan, and Tim (Demuth & Tremblay, 2008); Léonard, Madeleine, and Theophile (Morgenstern & Parisse, 2007), Pauline (Bassano & Maillachon, 1994), and Grégoire (Champaud corpus), and from the Cécycy and Shi (2005) corpus. The ages of the infants ranged from 0;08;09 to 1;09;18, and only adult speech was analyzed. We found that of a total of 562 noun types (1,919 tokens), few carried the particular endings that we used for our pseudonouns. Specifically, the /ɛʁ/ending appeared in 14 feminine noun types (33 tokens) and 11 masculine noun types (26 tokens), the /ɔl/ending in one feminine

noun type (13 tokens) and four masculine noun types (24 tokens), the /al/ ending in one feminine noun type (4 tokens) and no masculine noun types, and the /ɛj/ ending in one feminine noun type (4 tokens) and one masculine noun type (1 token). Overall, nouns carrying those endings were very infrequent (< 1% of all noun types and tokens). In fact, we observed over 30 other noun endings, some of which were more frequent than the endings that we used. It is therefore unlikely that the four pseudoword endings carry phonological information that could bias gender categorization.

We then constructed noun phrases, each containing a French determiner: *un* ("a," masculine), *une* ("a," feminine), *le* ("the," masculine), and *la* ("the," feminine). The following phrases were used as the familiarization stimuli: *un mouveille, une mouveille, un ravole, une ravole, un gombale, une gombale, un cagère, une cagère*. The stimuli for the test phase were as follows: *le mouveille, la mouveille, le ravole, la ravole, le gombale, la gombale, le cagère, la cagère*.

The auditory materials were recorded in an acoustical chamber by a native French female speaker in the infant-directed speech style. The stimuli were recorded using a digital audio recorder Sound Device 702T (44 kHz sampling frequency, 24 bits bit rate) and then transferred digital-to-digital to a computer. The four pseudonouns were each recorded multiple times in isolation and in noun phrases with each of the four determiners (*un, une, le, and la*). The pseudonouns were produced with three different intonations (flat, rising, and falling). In total, three tokens of every noun phrase (one for each intonation) were selected. To eliminate any possible acoustic or prosodic cues to gender in the pseudonouns, the final stimuli were created within Praat software (version 4.5.26; Boersma & Weenink, 2007) using a cross-splicing technique: Each pseudonoun produced in isolation with one of the three intonations was conjoined with determiners of both genders. The latter were spliced from noun phrases containing the corresponding pseudonouns in the same intonation as the citation forms. Specifically, the isolated pseudonoun and the corresponding noun phrase sound files were both cut at matching points on the waveforms at the beginning of the pseudonoun. For the pseudonouns *gombale* and *cagère*, the cutting point was just before the release of the initial consonant. For the pseudonouns *mouveille* and *ravole*, the cutting point was in the middle of the initial consonant. This was done to prevent the spliced phrases from sounding artificial.

The visual stimulus file consisted of a cartoon puppet standing in the center of the screen. We synchronized the mouth movements of the puppet to match the auditory stimuli so that the puppet "spoke" the noun phrases. While speaking, the puppet also moved her hands, head, and body. The visual stimuli were created using Adobe® Flash® CS3 Professional software (Adobe Systems Incorporated, San Jose, CA) and exported in a QuickTime format with a 30 fps frame rate and a resolution of 1360 × 768 pixels. The purpose of this animation was to make the task interesting for the infants.

Design. The experiment consisted of a familiarization phase and a test phase. During familiarization, infants were presented with noun phrases in which the four pseudonouns described above were paired with two of the four determiners (i.e., *un*—masculine, *une*—feminine). Two of the four pseudonouns were preceded by the masculine determiner *un* (e.g., *un mouveille, un ravole*) and two were preceded by the feminine determiner *une* (e.g., *une gombale, une cagère*). The grammatical gender pairings were counterbalanced across infants, forming two different familiarization conditions (see the familiarization stimuli for Group A vs. Group B in Table 1). Each familiarization condition consisted of multiple trials, and tokens of the four noun phrases were presented quasi-randomly within and across trials, ensuring variability. In particular, adjacent noun phrases were always different, except in the few cases where only one sequential repetition occurred. The same variability applied for the intonations of the stimuli. Each familiarization trial had a maximum length of 15.1 s, with an interphrase interval of 700 ms. The familiarization materials were presented until the infant accumulated 90 s of looking time. As the presentation of the stimuli was contingent upon infants' looking to the audiovisual presentation, infants only received attentive exposure. Following familiarization, the test phase began automatically.

In the test phase, the same four pseudonouns were again presented in noun phrases, this time paired with the other two determiners (*le*—masculine, *la*—feminine) that were not used during familiarization. There were two trial types: grammatical and ungrammatical trials. In the grammatical trials, two pseudonouns from familiarization (one that had been paired with the masculine determiner *un* and one that had been paired with the feminine determiner *une*) were now paired with the nonfamiliarized determiners that were gender consistent with the familiarization (e.g., *le mouveille, la gombale*). In the ungrammatical trials, the remaining

Table 1

Familiarization Conditions (Groups A and B) and Test Conditions (Subgroups 1–4) for Experiment 1

| Familiarization—Group A <i>un mouveille, une gombale, un ravole, une cagère</i> | | | | |
|--|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Familiarization—Group B <i>une mouveille, un gombale, une ravole, un cagère</i> | | | | |
| Test | | | | |
| | Subgroup 1 | Subgroup 2 | Subgroup 3 | Subgroup 4 |
| Trial | <i>le mouveille</i> | <i>le cagère</i> | <i>le ravole</i> | <i>le gombale</i> |
| | <i>la gombale</i> | <i>la ravole</i> | <i>la cagère</i> | <i>la mouveille</i> |
| | (Grammatical-A, ungrammatical-B) | (Ungrammatical-A, grammatical-B) | (Grammatical-A, ungrammatical-B) | (Ungrammatical-A, grammatical-B) |
| Trial | <i>le cagère</i> | <i>le mouveille</i> | <i>le gombale</i> | <i>le ravole</i> |
| | <i>la ravole</i> | <i>la gombale</i> | <i>la mouveille</i> | <i>la cagère</i> |
| | (Ungrammatical-A, grammatical-B) | (Grammatical-A, ungrammatical-B) | (Ungrammatical-A, grammatical-B) | (Grammatical-A, ungrammatical-B) |

Note. Table 1 contains different familiarization stimuli presented to two groups of infants (Groups A and B). The same test stimuli were presented to both familiarization groups, but the grammaticality of test materials is reversed (i.e., grammatical trials for Group A are ungrammatical for Group B) because of the different determiner–noun gender pairings during the familiarization. In addition to this grammaticality counterbalancing, the subgroups of the test conditions represent the counterbalancing of the first test trial (grammatical first vs. ungrammatical first) within each familiarization group, and the counterbalancing of the specific pseudonouns as grammatical versus ungrammatical stimuli, that is, Subgroups 1–2 versus 3–4.

two pseudonouns were paired with the nonfamiliarized determiners that were inconsistent in gender with the familiarization (e.g., *le cagère*, *la ravole*). Tokens of the noun phrases of each trial type were presented quasi-randomly to ensure variability. In particular, any given noun phrase never occurred more than twice consecutively. The same variability applied for the intonations of the stimuli. During each test trial the interphrase interval was 700 ms. In total, the test phase consisted of 10 trials, with a maximum length of 15.2 s for each trial. Grammatical and ungrammatical trials (five trials each) were presented in alternation. The type of the first trial (grammatical or ungrammatical) was counterbalanced across participants. The specific pseudonouns used as grammatical and ungrammatical stimuli were also counterbalanced across participants (as shown in Table 1).

Procedure. The participants were tested individually in a preferential looking paradigm. Upon arrival, the caregiver(s) were informed of the procedure and were asked not to interfere during the experiment (e.g., talking to the child or pointing at the screen). Then, the caregiver and the infant were led to a double-walled IAC acoustical chamber (interior dimension 9 ft 8 in. × 9 ft 8 in.). Here, the infant sat on the caregiver's lap approximately 1.5 m in front of a 42-in. LG TV monitor. The caregiver listened to masking music through noise

cancellation headphones. The experimenter in the adjacent room, blind to the audiovisual stimuli, observed the infant's eye movements through a closed-circuit TV and pressed a computer key whenever the infant looked at the display monitor. Experimental software (Cohen, Atkinson, & Chaput, 2000) presented the stimuli in the test chamber and recorded all looks. The experiment was fully infant controlled: Each trial was initiated when the infant looked toward the TV monitor and stopped when the infant looked away from the TV for 2 s or more, or when the maximum trial length was reached. Between trials, an attention getter appeared on the display monitor to attract the infant's attention. The attention getter consisted of an image of a bird zooming in and out, synchronized with a whistle sound.

Results and Discussion

The four age groups of infants reached the familiarization criterion in eight trials on average (14 months: 9.75 trials; 17 months: 8.44 trials; 20 months: 7.88 trials; 30 months: 8.25 trials). For each infant, average looking (i.e., listening) times per trial were calculated for grammatical test trials and for ungrammatical test trials, respectively. Figure 1 shows the looking times for the two test trial types. We conducted a 2 × 4 × 2 mixed analysis of

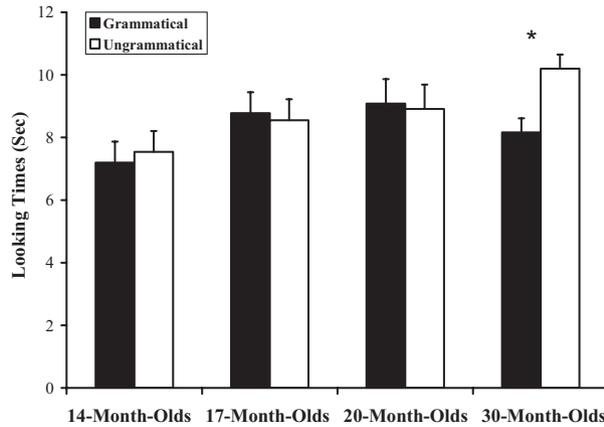


Figure 1. The 14-, 17-, 20-, and 30-month-old infants' average looking (listening) times (means and standard errors) for the two test trial types (grammatical vs. ungrammatical) in Experiment 1. Note. The task in these experiments was more complex than that in Experiment 2. Thirty-month-olds discriminated the two trial types. * $p < .05$.

variance (ANOVA), with Grammaticality of the test trials as the within-subject factor (grammatical vs. ungrammatical), and Age (14, 17, 20, 30 months) and Sex (male vs. female) as the between-subject factors. No main effect of Grammaticality was observed, $F(1, 56) = 1.056$, $p = .308$, $\eta^2 = .064$. There was no main effect of Sex $F(1, 56) = 0.427$, $p = .516$, $\eta^2 < .001$; no interaction between Grammaticality and Sex, $F(1, 56) = 0.633$, $p = .430$, $\eta^2 = .039$; between Age and Sex, $F(3, 56) = 0.409$, $p = .747$, $\eta^2 = .002$; or between Grammaticality, Age and Sex, $F(3, 56) = 1.449$, $p = .238$, $\eta^2 = .265$. Therefore, the factor Sex was not included in subsequent analyses. There was no main effect of Age, $F(3, 56) = 1.465$, $p = .234$, $\eta^2 = .007$. Crucially, there was a significant Grammaticality \times Age interaction, $F(3, 56) = 3.448$, $p = .023$, $\eta^2 = .631$. Follow-up simple effects analyses were conducted for each age. The results revealed that the 30-month-old infants discriminated between grammatical and ungrammatical trials. They listened longer to ungrammatical trials ($M = 10.20$ s, $SE = 0.58$ s) than to grammatical ones ($M = 8.16$ s, $SE = 0.57$ s), $F(1, 15) = 20.443$, $p < .001$, $\eta^2 = .577$. No effect of Grammaticality was found for the 14-month-olds, $F(1, 15) = 0.259$, $p = .618$, $\eta^2 = .017$; the 17-month-olds, $F(1, 15) = 0.112$, $p = .742$, $\eta^2 = .007$; or the 20-month-olds, $F(1, 15) = 0.050$, $p = .826$, $\eta^2 = .003$. These three younger age groups did not respond differently toward grammatical versus ungrammatical trials.

These results suggest that 30-month-old infants can categorize novel nouns into gender classes on the sole basis of their distribution with determiners.

These infants perceived the gender feature in the determiners and encoded the feature onto the novel nouns in the corresponding categories. The robust discrimination of grammatical versus ungrammatical trials shown by these infants contrasts with the results of the younger infants, for whom no discrimination was observed.

The possibility remains, however, that gender categorization emerges earlier in development, but that our learning task did not allow younger children to show this ability. The learning task that our infants faced may have been too complex. Our design required infants to learn to associate two pseudonouns with a masculine determiner and two other pseudonouns with a feminine determiner. Then, they were tested with the same four pseudonouns preceded by different determiners. The pseudonouns in the grammatical trials were not the same as those in ungrammatical trials. To succeed in the present task, infants have to categorize all four nouns in different gender classes based on the determiners during the familiarization phase, generalize the gender pairings to other determiners, and perceive the grammaticality of the test stimuli. Hence, the learning task required significant attention and memory resources, and thus may have been too challenging for younger infants to show their categorization ability. Interestingly, 30-month-olds showed a striking novelty preference in this difficult learning task. As novelty preference may indicate a more advanced ability (as shown in Hunter & Ames, 1988), a simpler task might reveal the knowledge of gender categorization in younger infants. We therefore simplified our design in Experiment 2 by using fewer pseudonouns, requiring the learning of a smaller number of determiner–noun pairings, and by increasing the familiarization time. We hypothesized that younger infants would show knowledge of noun gender categorization based on gender-marking determiners.

Experiment 2

Method

Participants. Participants were sixteen 14-month-old ($M_{\text{age}} = 1;2;13$; range = 1;2;1–1;2;30; 8 males, 8 females) and sixteen 20-month-old ($M_{\text{age}} = 1;8;29$; range = 1;8;16–1;9;19; 9 males, 7 females) monolingual French-learning infants. An additional eleven infants were tested but were not included in the analyses because of fussiness (six), experimental error (one), and ceiling looking (four).

Stimuli, design, and procedure. The stimuli consisted of two of the pseudonouns used in Experiment 1, *cagère* and *ravole*, paired with the determiners *un*, *une*, *le*, and *la*. The same tokens of these noun phrases from the final stimuli set (i.e., after cross-splicing) of the previous experiment were used: three tokens of *un cagère*, *une cagère*, *un ravole*, *une ravole* as the familiarization stimuli, and three tokens of *le ravole*, *la ravole*, *le cagère*, *la cagère* as the test stimuli.

The design consisted of a familiarization phase and a test phase. Familiarization trials presented the two pseudonouns, *cagère* and *ravole*, each preceded by one of the determiners, *un*—masculine or *une*—feminine (e.g., *un ravole*, *une cagère*). The pairings of the grammatical gender of the determiner with the pseudonouns were counterbalanced across infants, forming two different familiarization conditions, as shown in Table 2.

Each familiarization condition consisted of multiple trials in which the two noun phrases were presented in a quasi-random fashion across trials, with the restriction that the same noun phrase did not occur more than twice consecutively. The noun

phrases with each intonation were presented a comparable number of times across trials. The maximum trial length was 15.2 s, with an interphrase interval of 700 ms. The familiarization trials were presented until the infant accumulated 120 s of looking time, after which the experiment advanced to the test phase automatically.

In the test phase, the same two pseudonouns were again presented, this time preceded by the nonfamiliarized determiners (*le*—masculine, *la*—feminine). The test phase consisted of grammatical and ungrammatical trials. In the grammatical trials, the two pseudonouns were each paired with a determiner consistent in gender with the familiarization (e.g., *le ravole*, *la cagère*). In the ungrammatical trials, the same two pseudonouns were paired with determiners that were inconsistent in gender with the familiarization (e.g., *le cagère*, *la ravole*). Unlike Experiment 1, in which the nouns for grammatical trials were different from those for ungrammatical trials, both test trial types in Experiment 2 consisted of the same two nouns.

As in the previous experiment, tokens of the two noun phrases in a test trial were presented quasi-randomly to ensure variability. In particular, any given noun phrase never occurred more than twice consecutively. The same variability applied for the intonations of the stimuli. During each test trial the interphrase interval was 700 ms. In total, the test phase consisted of 10 trials, with a maximum length of 15.2 s for each trial. Grammatical and ungrammatical trials (5 trials each) were presented in alternation. The type of the first trial (grammatical vs. ungrammatical) was counterbalanced across participants. The grammaticality of the noun phrases was also counterbalanced across participants (see Table 2).

The procedure was identical to that of Experiment 1.

Results and Discussion

On average, infants reached the familiarization criterion within 9–10 trials (14 months: 10.19 trials; 20 months: 9.25 trials). For each infant, average looking time per trial was calculated for grammatical versus ungrammatical trials. Figure 2 shows the looking times for the two test trial types in Experiment 2. A $2 \times 2 \times 2$ mixed ANOVA was conducted, with Grammaticality of the test trials as the within-subject factor (grammatical vs. ungrammatical), and Age (14 vs. 20 months) and Sex (male vs. female) as the between-subject factors. No main effect of Grammaticality was observed, $F(1, 28) = 0.387$,

Table 2
Familiarization Conditions (Groups A and B) and Test Conditions (Subgroups 1 and 2) Across Different Groups of Infants for Experiment 2

| Familiarization—Group A <i>un ravole, une cagère</i> | | |
|---|-------------------------------------|-------------------------------------|
| Familiarization—Group B <i>une ravole, un cagère</i> | | |
| Test | | |
| | Subgroup 1 | Subgroup 2 |
| Trial | <i>le ravole</i> | <i>le cagère</i> |
| | <i>la cagère</i> | <i>la ravole</i> |
| | (Grammatical-A, ungrammatical-B) | (Ungrammatical-A, grammatical-B) |
| Trial | <i>le cagère</i> | <i>le ravole</i> |
| | <i>la ravole</i> | <i>la cagère</i> |
| | (Ungrammatical-A, grammatical-B) | (Grammatical-A, ungrammatical-B) |

Note. Table 2 contains different familiarization stimuli presented to two groups of infants (Groups A and B). The same test stimuli were presented to both familiarization groups, but the grammaticality of test materials is reversed (i.e., the grammatical trials for Group A are ungrammatical for Group B) because of the different determiner–noun gender pairings during the familiarization. In addition to this grammaticality counterbalancing, the subgroups of the test conditions represent the counterbalancing of the first test trial (grammatical first vs. ungrammatical first) within each familiarization group.

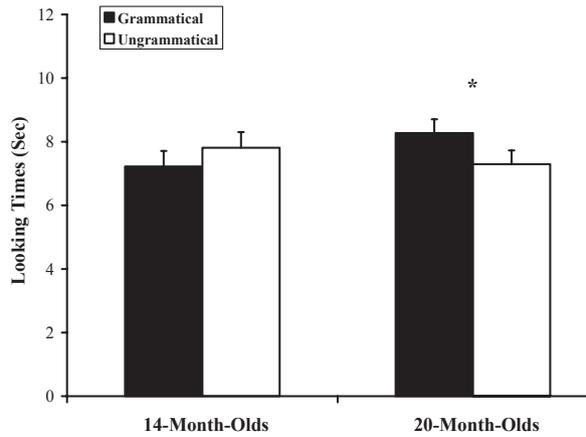


Figure 2. The 14- and 20-month-old infants' average looking (listening) times (means and standard errors) for the two test trial types (grammatical vs. ungrammatical) in the simpler task of Experiment 2.

Note. Twenty-month-olds discriminated the two trial types.
* $p < .05$.

$p = .539$, $\eta^2 = .011$. There was no main effect of Sex $F(1, 28) = 0.009$, $p = .926$, $\eta^2 < .001$; no interaction between Grammaticality and Sex, $F(1, 28) < 0.001$; $p = .988$, $\eta^2 < .001$; between Age and sex, $F(1, 28) = 0.030$, $p = .864$, $\eta^2 \leq .001$; or between Grammaticality, Age and Sex, $F(1, 28) = 0.432$, $p = .516$, $\eta^2 = .013$. Therefore, the factor Sex was not included in subsequent analyses. There was no main effect of Age, $F(1, 28) = 0.047$, $p = .830$, $\eta^2 < .001$. Importantly, there was a significant Grammaticality \times Age interaction, $F(1, 28) = 5.628$, $p = .025$, $\eta^2 = .163$. Follow-up simple effects analyses on each of the two ages revealed that the 20-month-old infants listened longer to grammatical trials ($M = 8.28$ s, $SE = 0.93$ s) than to ungrammatical ones ($M = 7.30$ s, $SE = 0.84$ s), $F(1, 15) = 5.229$, $p = .037$, $\eta^2 = .259$. No effect of Grammaticality was found for the 14-month-olds, $F(1, 15) = 1.434$, $p = .250$, $\eta^2 = .087$, indicating that infants of this group did not respond differently to grammatical versus ungrammatical trials.

The 20-month-olds tested with this simplified learning task showed knowledge of gender in determiners and the ability to use this knowledge to categorize the gender of novel nouns. The reduced number of novel nouns to be learned and the increased familiarization time allowed us to reveal younger infants' ability. Unlike these 20-month-olds, the 14-month-olds in the present experiment failed to show evidence of categorization despite the simplified learning task and longer familiarization. The results suggest that the ability to categorize gender of novel nouns based solely on

their distributional relations with gender-marked determiners emerges around 20 months of age or slightly before.

General Discussion

The goal of the present study was to examine the development of grammatical gender knowledge in infants. Previous studies have shown that categorization of nouns emerges shortly after 1 year of age (Höhle et al., 2004; Shi & Melançon, 2010). We show in this study that by 20 months of age, infants can divide nouns into grammatical gender categories, using determiners. Furthermore, we demonstrate that infants can perform the categorization abstractly and productively. Infants showed knowledge of an abstract grammatical category that included multiple masculine determiners (e.g., *un*, *le*) and another that included multiple feminine determiners (e.g., *une*, *la*), and they treated the determiners within each class as equivalent, as occurring before the same nouns. They used the gender features of determiners to categorize the gender of novel nouns.

In our experiments, the stimuli were carefully manipulated to remove all potential phonological and acoustic or prosodic cues. We used pseudonouns that carry no phonological cues to gender drawn from our prestudy with adults. In addition, we removed all possible acoustic and prosodic cues by using cross-spliced gender-neutral productions of the pseudonouns. Therefore, children must rely exclusively on the distributional and featural information associated with the determiners to succeed in our gender-categorization task. This ability thus closely reflects knowledge that is syntax-like. We found that 14-month-olds (Experiments 1 and 2) and 17-month-olds (Experiment 1) did not show evidence of gender categorization, whereas 20- and 30-month-olds did. The 20-month-olds showed success when the learning task was relatively simple (Experiment 2). Infants aged 30 months showed robust gender knowledge even in the complex learning task (Experiment 1).

It is noteworthy that the 30-month-olds in Experiment 1 showed a novelty preference (i.e., preference for the ungrammatical over the grammatical trials), whereas the 20-month-olds from Experiment 2 showed a familiarity preference (i.e., preference for the grammatical over the ungrammatical trials). In the literature, novelty preference has often been associated with a longer exposure time to experimental stimuli and with older infants, and conversely, familiarity preference has been associated

with younger infants and relatively complex stimuli (e.g., Hunter & Ames, 1988; Thiessen & Saffran, 2003). One interpretation is that novelty preference indicates better encoding or better knowledge of the material presented than does familiarity preference (Hunter & Ames, 1988). In line with this interpretation, the familiarity preference shown by our 20-month-olds in Experiment 2 suggests an emerging ability to use the distribution of determiners for noun gender categorization, whereas the novelty preference shown by the 30-month-olds in Experiment 1 indicates a robust knowledge. In fact, the familiarity preference in the 20-month-olds is particularly revealing as the stimuli were simpler and the familiarization exposure was longer in that experiment (Experiment 2) than in Experiment 1. When the stimuli and familiarization duration were both challenging (Experiment 1), no preference was observed in the 20-month-olds. Therefore, the combined results of Experiments 1 and 2 provide clear evidence that knowledge of distribution-based gender categorization emerges around 20 months of age, or slightly earlier, and becomes strong by 30 months of age.

Our results and those of the previous online noun comprehension studies show a coherent developmental progression of grammatical gender. As mentioned earlier, for several languages, there is evidence that children use gender information in the comprehension of the meaning of familiar nouns, starting from 2 years of age (Johnson, 2005; Van Heugten & Shi, 2009). However, in those studies, the possibility remained that the children simply learned to associate specific determiners with specific familiar nouns, without having generalized the abstract knowledge of noun gender classes. For example, children may have learned that *le* and *un* (masculine determiners) go with *ballon* (a masculine familiar noun) because they heard the co-occurrences of those specific determiners with that specific noun often in their natural environment, without knowing that *le* and *un* belong to the same category (masculine determiners) and apply to all masculine nouns. The positive results with pseudonouns obtained in the present study suggest that children possess an abstract representation of grammatical gender features for determiners and nouns at an early age. Indeed, the 30-month-olds of Experiment 1 and the 20-month-olds of Experiment 2 were able to generalize the appropriate agreement between determiners and nouns to novel nouns.

The categorization ability shown in our study is consistent with the theoretical view that grammati-

cal categories and relations are represented during early language acquisition (e.g., Valian, 2009). Our results are inconsistent with the position that there are no syntactic representations for infants under 2 years and that children's early utterances are formulaic exemplars with no abstract grammatical structure (e.g., Pine & Lieven, 1997). In our experiments the phrases contained novel nouns that could not have been previously stored. Infants encoded the gender feature in determiners and generalized it to the novel nouns, demonstrating abstract grammatical representations.

We note here that although our study shows evidence of abstract gender knowledge, the full representations of grammatical gender systems is more complex than what we showed here. In French, gender agreement applies to multiple categories within noun phrases, including determiners, nouns, adjectives, and quantifiers. Moreover, gender agreement may apply across phrases or sentences, for example, between the subject and the predicate that contains an adjective (*la_{feminine} maison_{feminine} est petite_{feminine}*—"the house is small"), and between the noun and its modifying relative clauses (e.g., ... *la_{feminine} réponse_{feminine} qui a été faite_{feminine}*—"the response that was made"). Whether the knowledge of determiner-noun gender agreement that we demonstrated here extends to (or even triggers) other gender agreement elements across phrasal structures is an empirical question that has implications to linguistic theories. Further experiments are necessary to better understand the precise degree of complexity of the syntactic knowledge of gender in toddlers.

Given that the pseudonouns in our experiments contained no phonological or acoustical cues to grammatical gender, we were able to conclude that categorization was strictly distribution based and did not require the support of any phonological or acoustical cue. Our study did not test whether such speech cues may assist gender categorization. In the literature, phonological and prosodic cues have in fact been found to correlate with grammatical categories such as nouns and verbs (e.g., Cassidy & Kelly, 1991; Kelly, 1992; Sereno & Jongman, 1990; Shi & Moisan, 2008). In a corpus study Monaghan et al. (2007) found that parental speech contains both distributional and phonological cues that can support the distinction between nouns and verbs. In addition, mothers tend to produce noun-verb homophones with different acoustic patterns, and infants are sensitive to the cues and can discriminate the forms intended for one category versus the other (Conwell & Morgan, 2012). Shi and Moisan (2008) also showed that infant-directed speech in French contains prosodic cues that distinguish noun

and verb productions. With respect to grammatical gender, it is unknown if there are acoustical cues to genders in the input. It has been suggested, however, that gender attribution in French can be partially predicted based on certain phonological properties (Lyster, 2006; Tucker et al., 1977). An analysis of nouns of a French dictionary revealed that many nouns carry an ending that associates in a systematic way with gender (Tucker et al., 1977), such as the *-ette* ending (e.g., *poussette*) that co-occurs with feminine gender and the *-in* ending (e.g., *bain*) that co-occurs with masculine gender. We should note that such associations are far from perfect, as violations are common (e.g., *main*—feminine noun), and many nouns carry a neutral ending that do not indicate either gender (e.g., *-ale*, *-ère*; Lyster, 2006). The inconsistencies are also clear in our frequency analysis of infant-directed speech in French (as described in Experiment 1 above). Hence, purely distribution-based gender categorization is necessary for French. Nevertheless, the possibility remains that the partially regular phonological cues to gender in French nouns and the link of these cues with the gender feature of determiners might be perceived and used for categorization by infants.

It is pertinent to mention that phonological markings to grammatical gender are variable and specific to individual languages. In this respect, gender is unlike content words versus function words. The content–function–word bifurcation is fundamental in natural languages, and it is supported reliably by a large cohort of phonological and acoustical cues in a language-universal fashion (e.g., Shi et al., 1998). Consistent with this characteristic, the ability to categorize the two broad classes is present at birth (Shi et al., 1999). Phonological or acoustical cues to gender and to other lexical categories such as nouns and verbs, however, are language-specific, varying both in terms of how regularly markings appear and what types of markings a language employs. Thus, discovering the phonological or acoustical cues to these categories may require a certain amount of learning experience by the infant. Further experiments can reveal whether and how the phonological or acoustic cues may play a role in the categorization of grammatical gender at the early and later stages of learning.

In conclusion, through careful control of all possible cues in the stimuli, we tested infants' ability to categorize nouns into gender classes on the sole basis of determiners. We showed that by 20 months of age infants have abstract knowledge of determiner gen-

der classes. They are able to encode the grammatical gender feature of determiners onto novel nouns in noun phrases, a process that reflects representations that are syntactic. By 30 months of age, the knowledge of gender feature and agreement is robust. This study is the first to demonstrate the development of this abstract grammatical knowledge.

References

- Bassano, D. (1998). Sémantique et syntaxe dans l'acquisition des classes de mots: l'exemple des noms et des verbes en français. *Langue française*, 118, 26–48.
- Bassano, D., & Maillolochon, I. (1994). Early grammatical and prosodic marking of utterance modality in French: A longitudinal case study. *Journal of Child Language*, 21, 649–675. doi:10.1017/S0305000900009491
- Bassano, D., Maillolochon, I., & Eme, E. (1998). Developmental changes and variability in early lexicon: A study of French children's naturalistic productions. *Journal of Child Language*, 25, 493–531. doi:10.1017/S0305000998003547
- Boersma, P., & Weenink, D. (2007). Praat: Doing phonetics by computer (Version 4.6.16) [Computer software]. Retrieved August 22, 2007, from <http://www.praat.org/>
- Cassidy, K., & Kelly, M. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, 30, 348–369. doi:10.1016/0749-596X(91)90041-H
- Cécylre, M.-J., & Shi, R. (2005). Speech rate in maternal speech to French-learning infants. Presented at the 15th meeting of the Canadian Society for Brain, Behaviour and Cognitive Science. Montréal, Canada.
- Christophe, A., Guasti, T., & Nespor, M. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12, 585–612. doi:10.1080/016909697386637
- Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2000). Habit 2000: A new program for testing infant perception and cognition. (Version 2.2.5c) [Computer software]. Austin: University of Texas.
- Conwell, E., & Morgan, J. L. (2012). Is it a noun or is it a verb? Resolving the ambicategoricity problem. *Language Learning and Development*, 8, 87–112.
- Corrêa, L. M. S., & Name, M. C. (2003). The processing of determiners—noun agreement and the identification of the gender of nouns in the early acquisition of Portuguese. *Journal of Portuguese Linguistics*, 2, 19–44.
- Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research*, 22, 109–131. doi:10.1007/BF01067827
- Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, 35, 99–127. doi:10.1017/S0305000907008276

- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249–268. doi:10.1017/S0305000904006786
- Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, 7, 567–580. doi:10.1111/j.1467-7687.2004.00381.x
- Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles?. *Language and Speech*, 51, 23–44. doi:10.1177/00238309080510010301
- Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, 6, 122–127. doi:10.1111/1467-7687.00261
- Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5, 341–353. doi:10.1207/s15327078in0503_5
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 5, pp. 69–95). Norwood, NJ: Ablex.
- Johnson, E. K. (2005). Grammatical gender and early word recognition in Dutch. In A. Brugos, M. R. Clark-Cotton & S. Ha (Eds.), *BUCLD 29: Proceedings of the 29th annual Boston University conference on language development* (Vol. 1, pp. 320–330). Somerville, MA: Cascadilla.
- Kelly, M. H. (1992). Using sound to solve syntactic problems. The role of phonology in grammatical category assignments. *Psychological Review*, 99, 349–364. doi:10.1037/0033-295X.99.2.349
- Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18, 193–198. doi:10.1111/j.1467-9280.2007.01871.x
- Lyster, R. (2006). Predictability in French gender attribution: A corpus analysis. *French Language Studies*, 16, 69–92. doi:10.1017/S0959269506002304
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
- Maratsos, M., & Chalkley, M. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language* (Vol. 2, pp. 127–214). New York: Gardner.
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, 90, 91–117. doi:10.1016/S0010-0277(03)00140-9
- Mintz, T. 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. doi:10.1093/acprof:oso/9780195170009.003.0002
- Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55, 259–305. doi:10.1016/j.cogpsych.2006.12.001
- Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press/Bradford Books.
- Morgenstern, A., & Parisse, C. (2007). *Codage et interprétation du langage spontané d'enfants de 1 à 3 ans*. *Corpus*, 6, 55–78. Retrieved May 5, 2011, from <http://corpus.revues.org>
- Pine, J., & Lieven, E. (1997). Slot and frame patterns in the development of the determiner category. *Applied Psycholinguistics*, 18, 123–138. doi:10.1017/S0142716400009930
- Pinker, S. (1984) *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Sereno, J. A., & Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research*, 19, 387–404. doi:10.1007/BF01068886
- Shafer, V., Shucard, D., Shucard, J., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research*, 41, 874–886.
- Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, 119, EL61–EL67. doi:10.1121/1.2198947
- Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, 11, 407–413. doi:10.1111/j.1467-7687.2008.00685.x
- Shi, R., Marquis, A., & Gauthier, B. (2006). Segmentation and representation of function words in preverbal French-learning infants. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30: Proceedings of the 30th annual Boston University conference on language development* (Vol. 2, pp. 549–560). Boston, MA: Cascadilla.
- Shi, R., & Melançon, A. (2010). Syntactic Categorization in French-Learning Infants. *Infancy*, 15, 517–533. doi:10.1111/j.1532-7078.2009.00022.x
- Shi, R., & Moisan, A. (2008). Prosodic cues to noun and verb categories in infant-directed speech. In H. Chan, H. Jacob, & E. Kapia (Eds.), *BUCLD 32: Proceedings of the 32th annual Boston University conference on language development* (Vol. 2, pp. 450–461). Somerville, MA: Cascadilla.
- Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, 25, 169–201. doi:10.1017/S0305000997003395
- Shi, R., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, 10, 187–198. doi:10.1111/j.1532-7078.2009.00022.x
- Shi, R., Werker, J., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72, B11–B21. doi:10.1016/S0010-0277(99)00047-5
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Statistical and stress cues in infant word segmentation.

- Developmental Psychology*, 39, 706–716. doi:10.1037/0012-1649.39.4.706
- Tucker, G. R., Lambert, W. E., & Rigault, A. A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. Hague, the Netherlands: Mouton.
- Valian, V. (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.
- Valian, V., & Coulson, S. (1988). Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language*, 27, 71–86. doi:10.1016/0749-596X(88)90049-6
- Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, 12, 419–425. doi:10.1111/j.1467-7687.2008.00788.x