Developmental Science

Developmental Science 12:3 (2009), pp 419–425 DOI: 10.1111/j.1467-7687.2008.00788.x

PAPER

French-learning toddlers use gender information on determiners during word recognition

Marieke van Heugten[1,2] and Rushen Shi[1]

1. Département de psychologie, Université du Québec à Montréal, Canada
2. Department of Psychology, University of Toronto, Canada

Abstract

In gender-marking languages, the gender of the noun determines the form of the preceding article. In this study, we examined whether French-learning toddlers use gender-marking information on determiners to recognize words. In a split-screen preferential looking experiment, 25-month-olds were presented with picture pairs that referred to nouns with either the same or different genders. The target word in the auditory instruction was preceded either by the correct or incorrect gender-marked definite article. Toddlers' looking times to target shortly after article onset demonstrated that target words were processed most efficiently in different-gender grammatical trials. While target processing in same-gender grammatical trials recovered in the subsequent time window, ungrammatical articles continued to affect processing efficiency until much later in the trial. These results indicate that by 25 months of age, French-learning toddlers use gender information on determiners when comprehending subsequent nouns.

Introduction

Infants face many challenges at the early stage of language acquisition. Not only do they need to segment the speech stream into smaller units, they also have to organize those units into syntactic categories such as noun phrases (NPs) and verb phrases (VPs). Major phrasal units tend to contain co-occurrences of functional and lexical categories (e.g. *the dog*). It has been proposed that infants' sensitivity to function words (e.g. determiners, auxiliaries) plays an important role in language acquisition (e.g. Christophe, Guasti, Nespor, Dupoux & Van Ooyen, 1997; Morgan, Shi & Allopenna, 1996; Shi, 2005). More specifically, function words are frequent in the speech infants hear and tend to occur at the edges of utterances (e.g. Shi, Morgan & Allopenna, 1998), positions that are favorable for word segmentation (e.g. Seidl & Johnson, 2006). Extracting function word forms from utterances may therefore be relatively easy. Once the sound patterns of function words are stored, this information may be used to break the signal down further, to find boundaries of previous or following words, and to distinguish between different phrasal units. Function words may further assist children in determining specific syntactic and semantic properties of upcoming words (Bernal, Lidz, Millotte & Christophe, 2007; Gelman & Taylor, 1984; Fisher, Klingler & Song, 2006; Höhle, Weissenborn, Kiefer, Schulz & Schmitz, 2004; Katz, Baker & MacNamara, 1974) and in analyzing larger morphosyntactic structures (Höhle, Schmitz, Santelmann & Weissenborn, 2006).

While function words are usually not among the first words in children's productions (e.g. Brown, 1973), children do perceive them at an early age. In fact, function word acquisition is an incremental process. At birth, infants possess the ability to categorically discriminate between function words and content words (Shi, Werker & Morgan, 1999), most likely due to the acoustical differences between these two types of words (Shi *et al.*, 1998). During the second half of the first year of life, the phonetic detail with which function words are stored starts to increase (Shady, 1996; Shafer, Shucard, Shucard & Gerken, 1998; Shi, Marquis & Gauthier, 2006; Shi, Werker & Cutler, 2006). This improved sensitivity to the acoustic properties of individual function words is acquired as a function of frequency, and frequent function words allow infants to segment adjacent potential word forms at an earlier age than do less frequent function words (Shi, Cutler, Werker & Cruickshank, 2006; Shi & Lepage, 2008). Around the same age, infants begin to attend to the basic article-noun word order (Shady, Gerken & Jusczyk, 1995), and at 16 months of age infants discriminate passages containing grammatical versus ungrammatical function words (Shady, 1996). Thus, infants learn some aspects of function word placement around or shortly after their first birthday.

It is not until 18 months of age, however, that children start using function words to comprehend sentences. As a starting point, Shipley, Smith and Gleitman (1969) and Petretic and Tweney (1977) showed that children of approximately 2 years of age responded more appropriately to grammatical commands containing real function words

Address for correspondence: Marieke van Heugten, Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road N., Mississauga, Ontario, Canada, L5L 1C6; e-mail: marieke.vanheugten@utoronto.ca

(e.g. *Throw me the ball*) compared to commands containing nonsense function words (*Throw ronta ball*). These results were replicated in a pointing task (Gerken & McIntosh, 1993). Pointing was less accurate when the article in the instruction was replaced by an ungrammatical auxiliary or a nonsense article (e.g. *Find was\gub bird for me*). Recently, similar results were obtained with younger children (Kedar, Casasola & Lust, 2006; Zangl & Fernald, 2007). Eighteen-month-old English-learning toddlers looked at target pictures more reliably when the target word was preceded by a grammatical as opposed to a nonsense or ungrammatically positioned function word (e.g. *Can you see the\el\and ball?*). Taken together, these studies have shown that English-learning children, who generally fail to produce determiners, nevertheless attend to them, and by 18 months of age use function words to understand sentences.

Despite the fact that function words differ greatly between languages, little research exists on infants' processing of function words in languages other than English. In gender-marking languages, the form of the determiner depends on the gender of the following noun. Thus, while all determiners occupy the same syntactic position, they are subdivided into different gender classes. In French and Spanish, for example, masculine articles precede masculine nouns (e.g. *le ballon*, 'the ball', *un ballon*, 'a ball') whereas feminine articles precede feminine nouns (e.g. *la table*, 'the table'; *une table*, 'a table'). This principle is similar in Dutch, where some nouns are preceded by *de* and other nouns by *het*. Adult studies have shown that this gender information on articles is integrated during word comprehension and that it is used to constrain lexical candidates (e.g. Dahan, Swingley, Tanenhaus & Magnuson, 2000). The question of whether gender information activates congruent gender words and whether this gender influence is pre- or post-lexical is still unresolved (e.g. Colé & Segui, 1994; Grosjean, Dommergues, Cornu, Guillelmon & Besson, 1994; Dahan *et al*., 2000; see also Friederici & Jacobson, 1999, for an overview).

How do infants in gender-marking languages acquire their article system? Interestingly, production and corpus studies suggest language-specific patterns (e.g. Rozendaal & Baker, 2008; Van der Velde, 2003). This raises the question of how infants in these languages start using multiple article forms in comprehension. In one recent study in Dutch, Van Heugten and Johnson (submitted) compared the processing efficiency of target words in sentences containing a real article (*de*, *het*) versus a nonsense article (*se*). Note that unlike the larger acoustic/prosodic changes employed in English studies (Zangl & Fernald, 2007), the nonsense article *se* is acoustically reduced and similar to real function words. Dutch-learning 19- to 24-month-olds were shown to better comprehend sentences with real articles as opposed to sentences with nonsense articles, indicating that Dutch learners begin using definite articles in comprehension without apparent delay compared to English learners. No evidence was obtained, however, that these children had learned about gender (also see Johnson & Diks, 2005). Older Spanish-learners, in contrast, do show evidence of using gender information on articles in online word comprehension (Lew-Williams & Fernald, 2007). Children in this study were presented with two pictures while hearing a noun preceded by a gender-marked article (e.g. *Encuentra* la_{FEM} $pelota_{\mathrm{FEM}}$, 'Find the ball'). The two pictures on the screen referred to nouns with either the same gender or with different genders. Children were faster in recognizing the target when the two pictures were different in gender, a performance comparable to the Spanish adults.

The children in Lew-Williams and Fernald (2007) were between 34 and 42 months of age, an age most likely beyond the early stages of gender learning. An interesting question thus concerns how and when this ability emerges in younger infants, that is, whether infants at the early stage of vocabulary learning encode gender information and how they use this information during online comprehension of nouns. There is some evidence that Dutch learners start using gender information on definite articles in an adult-like fashion around 28 months of age (Johnson, 2005). However, more cross-linguistic data are needed to better understand the acquisition of gender.

In this study, we inquired whether French-learning toddlers around 2 years of age encode and use gender information on determiners. This age group was chosen because it is conceivable that children in a gender-marking language need slightly more time to acquire the full range of aspects of gender-marked articles than English-learners do (cf. Van Heugten & Johnson, submitted). In addition, this age group enables us to examine the use of gender-marked articles at the earlier stages of word learning than do previous gender studies. In a split-screen preferential looking experiment, toddlers were presented with two pictures on a TV screen. An auditory sentence instructed them to look at one of the pictures. If gender-marked articles indeed assist word recognition, French learners should recognize target words more efficiently (i.e. display a higher proportion of looking time to target) when the two pictures on the screen are different as opposed to the same in gender (i.e. when the gender-marked article preceding the target noun disambiguates the target and distractor picture). In addition, recognition should be impeded (i.e. a lower proportion of looking time to target) when words are preceded by the incorrect as opposed to the correct gender-marked article. That is, although it may be demanding to recognize targets when the gender information is not informative (i.e. when the target and distractor nouns are of the same gender), it should be even harder to overcome ungrammatical information.

Method

Participants

Twenty-four monolingual Quebec-French-learning 25-month-olds (mean: 760 days, range: 735–788 days) completed

this experiment. An additional four toddlers were tested, but excluded from the analyses because of fussiness.

Materials

Four familiar [ba]-initial target words were selected, two carrying masculine (*ballon* 'ball', *bateau* 'boat') and two carrying feminine gender (*banane* 'banana', *balançoire* 'swing'). The onset consonants of the second syllable of the targets were all coronals. The choice of '*ba*+coronal' targets eliminates possible coarticulation effects on the preceding determiner, which may affect children's responses. Besides the target nouns, four filler nouns were selected (*crayon*$_{\text{MASC}}$ 'pencil', *livre*$_{\text{MASC}}$ 'book', *table*$_{\text{FEM}}$ 'table', *fleur*$_{\text{FEM}}$ 'flower'). The definite articles were *le* (masculine) and *la* (feminine).

The auditory materials were recorded in an IAC booth by a female Quebec-French speaker in a child-directed register (44.1 kHz sampling frequency, 16 bits). To avoid possible artifacts due to pronouncing ungrammatical utterances, we recorded only grammatical utterances. Both grammatical and ungrammatical utterances were created via cross-splicing. First, the target word was selected from an utterance with the correct article (e.g. *ballon* from *Regarde, le*$_{\text{MASC}}$ *ballon*$_{\text{MASC}}$ or *banane*$_{\text{FEM}}$ from *Regarde, la*$_{\text{FEM}}$ *banane*$_{\text{FEM}}$). The noun was spliced from the utterance just before the release of the [b]. Then, two different utterance carriers were selected for each target noun, one containing *le* and the other one containing *la* (e.g. *Regarde, le* from *Regarde, le*$_{\text{MASC}}$ *ballon*$_{\text{MASC}}$ and *Regarde, la* from *Regarde, la*$_{\text{FEM}}$ *banane*$_{\text{FEM}}$). The two carriers were combined with the target noun that was spliced from another utterance (as described above), such that both grammatical (e.g. *Regarde, le*$_{\text{MASC}}$ *ballon*$_{\text{MASC}}$) and ungrammatical utterances (e.g. *Regarde, la*$_{\text{FEM}}$ *ballon*$_{\text{MASC}}$) were created by cross-splicing. The durations of grammatical (mean: 1.720 s) and ungrammatical utterances (mean: 1.738 s) were closely matched, as were the durations of the articles (on average .322 s within the grammatical and .317 s within the ungrammatical trials). One token of the target noun was used across conditions.

The filler utterances, which were not cross-spliced, had a structure identical to the test utterances. The four target words were also recorded in isolation without a preceding article, to be used in the practice trials.

The visual stimuli were colorful images that clearly represented the target words. All test words referred to inanimate objects familiar to children. The picture sizes were matched, as was the brightness of the pictures.

The auditory and visual stimuli were combined to form the trials. The four practice trials were each 5 seconds long, containing one center picture that was named without an article. We used these practice trials to familiarize infants with the pictures of the targets before the test phase. In each test trial, two pictures were simultaneously presented for 7 seconds, one of which always represented the noun in the speech stream. The distractor picture served as the target in other trials. The pictures were gradually decreased in size for 3 seconds and subsequently enlarged for 4 seconds. After approximately 3.3 seconds, the auditory stimulus (in the form *Regarde*, [article] [noun]!) started, and at exactly 4 seconds after the onset of the trial, the article began. Between trials, a 2-second colorful zooming star together with a cricket sound attracted the toddler's attention.

Design

By varying the gender combinations of the two pictures and the correctness of the definite article in the auditory instruction, three conditions were created. In the different-gender grammatical condition (informative), one picture represented a masculine and one a feminine noun. Toddlers listened to a grammatical utterance, with either the masculine or the feminine word as the target. In the same-gender grammatical condition (uninformative), the two pictures were matched in gender, and the auditory sentence named one of them grammatically. Ungrammatical trials consisted of two different-gender objects combined with an ungrammatical utterance. Table 1 displays the conditions and examples in more detail.

There were 16 trials plus four practice trials. We created four orders, each consisting of four instances of each condition and four fillers. The trials were quasi-randomized in blocks of three trials such that all three conditions occurred within any block. Each block was followed by a filler trial. Only four of the 20 trials in the experiment contained an incorrect article.

Within an order, a picture always appeared on the same side of the screen. This position was counterbalanced between orders. Targets occurred an equal number of times on each side. Participants were randomly assigned to one of the four orders.

Procedure

Toddlers were tested individually in a split-screen preferential looking paradigm (e.g. Fernald, Pinto, Swingley, Weinberg & McRoberts, 1998). In an IAC booth, the child sat on the parent's lap, approximately 2 meters in front of a 42-inch TV. Outside the booth, the researcher started the experiment when the child looked towards the screen. The sessions were videotaped. Parents wore headphones delivering masking music. The experiment lasted approximately 3 minutes.

Coding eye movements

The videotaped sessions were imported with 30 frames per second on a computer, and were coded off-line frame by frame with SuperCoder (http://hincapie.psych.purdue.edu/Splitscreen/home.html). For all frames between the onset and the offset of a trial, the eye gazes were judged to be left, right or away. The coder was blind to the stimuli. Seventeen percent of the sessions (four) were also coded

Table 1 *Examples of the three conditions*

Condition	Video	Audio
Different-gender grammatical trials (informative)	*ballon – banane* $ball_M$ – $banana_F$ *ballon – banane* $ball_M$ – $banana_F$ *balançoire – bateau* $swing_F$ – $boat_M$ *balançoire – bateau* $swing_F$ – $boat_M$	*Regarde, le ballon!* 'Look, the_M $ball_M$!' *Regarde, la banane!* 'Look, the_F $banana_F$!' *Regarde, la balançoire!* 'Look, the_F $swing_F$!' *Regarde, le bateau!* 'Look, the_M $boat_M$!'
Same-gender grammatical trials (uninformative)	*ballon – bateau* $ball_M$ – $boat_M$ *ballon – bateau* $ball_M$ – $boat_M$ *banane – balançoire* $banana_F$ – $swing_F$ *banane – balançoire* $banana_F$ – $swing_F$	*Regarde, le ballon!* 'Look, the_M $ball_M$!' *Regarde, le bateau!* 'Look, the_M $boat_M$!' *Regarde, la banane!* 'Look, the_F $banana_F$!' *Regarde, la balançoire!* 'Look, the_F $swing_F$!'
Ungrammatical trials	*ballon – banane* $ball_M$ – $banana_F$ *ballon – banane* $ball_M$ – $banana_F$ *balançoire – bateau* $swing_F$ – $boat_M$ *balançoire – bateau* $swing_F$ – $boat_M$	*Regarde, la ballon!* 'Look, the_F $ball_M$!' *Regarde, le banane!* 'Look, the_M $banana_F$!' *Regarde, le balançoire!* 'Look, the_M $swing_F$!' *Regarde, la bateau!* 'Look, the_F $boat_M$!'

by a second coder. The agreement between the two coders was high (mean correlation = .99; *SD* = .009).

Results

Following previous studies (e.g. Swingley, 2003; Van Heugten & Johnson, submitted), we used the proportion of looking time (LT) to target as the dependent variable. Starting 500 ms after article onset, the mean proportion of LT to target was calculated for a 1.5-second period. Eye gazes occurring before 500 ms after the article onset were likely initiated before the target could be (partially) processed and hence were not analyzed. We used a slightly later time window than previous studies (e.g. Fernald *et al.*, 1998; Lew-Williams & Fernald, 2007; Zangl & Fernald, 2007) because all target words started with the same initial syllable approximately 200 ms long. LT to target is reported as a proportion of the total LT to both pictures. Looks away from the screen are thus not included. To better understand infants' incremental processing, we subsequently analyzed three successive time windows of 500 ms each, starting from 500 ms after article onset. The NPs were approximately 1000 ms long. Because children need time to program a saccade (e.g. Fernald *et al.*, 1998), the complete first time window and part of the second time window represented looking times before the offset of the noun. The third time window occurred after the NP.

First, we verified that the children recognized the targets. The proportion of LT to target was calculated for a 1.5-second window before (1000–2500 ms from the beginning of the trial) and after (500–2000 ms from article onset) article onset. On average, the proportion of looking time to target was .44 (SEM = .020) before and .59 (SEM = .019) after target onset. A one-tailed paired-sample *t*-test showed that this difference was statistically significant ($t(23) = -5.561$; $p < .001$). In addition, a one-way ANOVA showed that the three conditions did not differ before article onset ($F(2, 22) < 1$).

We then examined whether the proportion of LT to target after article onset differed across conditions. A two-way ANOVA with Instruction Type (different-gender grammatical, same-gender grammatical, ungrammatical) and Target Gender (masculine, feminine) as factors revealed only a main effect of Instruction Type ($F(1, 23) = 3.720$; $p = .041$) and no interaction ($F(2, 22) = 1.524$; $p = .241$). Thus, given the lack of this interaction, there is no evidence that the pattern of LT is modulated by the gender of the target word. The mean proportion of LT to target was .66 (SEM = .036) in the different-gender grammatical condition, .61 (SEM = .024) in the same-gender grammatical condition, and .54 (SEM = .024) in the ungrammatical condition. As expected, planned comparisons showed a significant difference between different-gender grammatical and ungrammatical trials ($t(23) = 2.716$; $p = .012$) and between same-gender grammatical and ungrammatical trials ($t(23) = -2.287$; $p = .032$). Unexpectedly, the difference between same- and different-gender grammatical trials was not significant ($t(23) = 1.060$; $p = .300$, all analyses are two-tailed).

Using a relatively large time window may have caused possible effects to be averaged out. Indeed, the looking patterns for the smaller time windows showed interesting effects (see Figure 1). For the initial time window, mean proportion of LT to target was .61 (SEM = .038) for different-gender grammatical trials, .46 (SEM = .041) for same-gender grammatical trials, and .45 (SEM = .044)

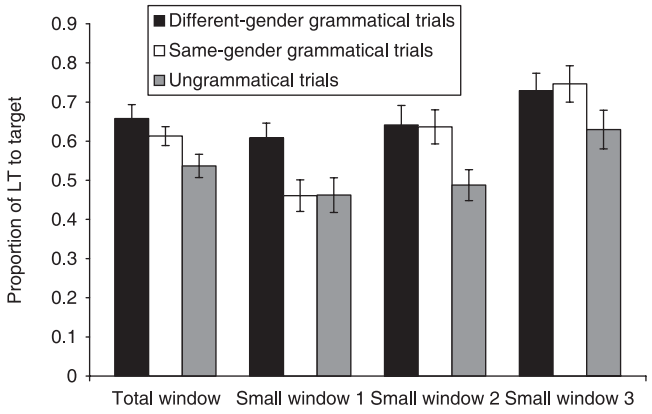

Figure 1 *Mean proportion of looking time to target for each condition (error bars indicate standard errors) for the larger total time window and each of the small time windows.*

for ungrammatical trials. A one-way ANOVA with three levels (different-gender grammatical, same-gender grammatical, ungrammatical) revealed a significant difference ($F(2, 22) = 3.717$; $p = .041$). Both same-gender grammatical and ungrammatical trials differed from different-gender grammatical trials ($t(23) = 2.615$; $p = .015$; $t(23) = 2.156$; $p = .042$, respectively) with a higher value in the different-gender grammatical condition than in the other two conditions. No differences were observed between same-gender grammatical and ungrammatical conditions ($t(23) = .028$; $p = .978$). This indicates that the target word is processed more efficiently in different-gender grammatical trials than in same-gender grammatical and ungrammatical trials.

In the second time window, the mean proportion of LT to target was .64 (SEM = .050) for different-gender grammatical trials, .64 (SEM = .044) for same-gender grammatical trials, and .49 (SEM = .040) for ungrammatical trials. Again, a one-way ANOVA indicated that the three conditions differed ($F(2, 22) = 5.211$; $p = .014$). The proportion of LT to target was significantly higher in different- and same-gender grammatical trials than in ungrammatical trials ($t(23) = 2.631$; $p = .015$; $t(23) = -2.765$; $p = .011$, respectively), while no differences were found between the two types of grammatical trials ($t(23) = .072$; $p = .943$). Thus, the lower recognition rate in same-gender grammatical trials as compared to different-gender grammatical trials found in the previous time window has now disappeared.

The mean proportion of LT in the third time window was .73 (SEM = .044) for different-gender grammatical trials, .75 (SEM = .046) for same-gender grammatical trials, and .63 (SEM = .050) for ungrammatical trials. Although there was no main effect for Instruction Type ($F(2, 22) = 2.044$; $p = .153$), planned comparisons showed similar tendencies of looking patterns as in the second time window (different-gender grammatical versus ungrammatical, $t(23) = 1.804$; $p = .084$; same-gender grammatical versus ungrammatical, $t(23) = 1.933$; $p = .066$; same- versus different-gender grammatical, $t(23) = -.350$; $p = .729$).

Discussion

This study shows that French-learning toddlers use gender cues on determiners in language comprehension. Nouns were recognized more efficiently when correct and informative articles (i.e. different-gender grammatical trials) were used rather than when incorrect (i.e. ungrammatical trials) or uninformative (i.e. same-gender grammatical trials) articles were used. Moreover, ungrammatical articles impeded comprehension for a longer time than uninformative, grammatical articles. Thus, although nouns following uninformative gender information are processed less efficiently than nouns following informative gender information, it is even harder to overcome ungrammatical articles. Together, these results demonstrate that by 25 months of age toddlers have already learned at least some article–noun dependencies and that they can use the form of determiners to constrain possible word candidates, similar to French adults (Dahan *et al.*, 2000).

One may wonder what underlies the acquisition of determiners. French learners use gender-marking information on definite articles to recognize words at 25 months of age, only a few months after English learners have been shown to use gender-absent articles (e.g. Gerken & McIntosh, 1993; Kedar *et al.*, 2006; Zangl & Fernald, 2007). In contrast, it is not until 28 months of age that Dutch learners have acquired adult-like gender-marking competence (Johnson & Diks, 2005; Van Heugten & Johnson, submitted). Compared to French learners, Dutch learners thus seem to need more time to fully develop the determiner system, a pattern that is also reflected in children's production data (Rozendaal & Baker, 2008; Van der Velde, 2003). One possible explanation could be that definite articles are acoustically less reduced in French than in Dutch. This difference might make definite articles prosodically more salient in French, hence leading to a faster rate of acquisition. Alternatively (or additionally), distributional properties may guide the initial learning process. French differs from Dutch in that the co-occurrence of articles and nouns is more consistent in French. Like English, Dutch allows nouns to occur without articles (e.g. *Water flows* or *Moeders klagen niet* 'Mothers don't complain'), whereas this possibility is highly restricted in French (e.g. *La garderie est intéressante* 'Daycare is interesting' or *Les mères ne se plaignent pas* 'Mothers don't complain', but not **Garderie est intéressante* or **Mères ne se plaignent pas*). In fact, a recent corpus study comparing Dutch and French learners' determiner input has shown that bare nouns occur more frequently in Dutch than in French (Rozendaal & Baker, 2008). Thus, the more obligatory distribution of French determiners may contribute to infants' relatively early perception and use of determiners in noun comprehension.

Distributional properties of determiners might also explain why previous comprehension studies in English and Dutch found mixed results with regard to toddlers' understanding of noun targets when the preceding function word was omitted. In some studies, 18–24-month-olds were not affected by the absence of an article (Gerken & McIntosh, 1993; Zangl & Fernald, 2007; Van Heugten & Johnson, 2006), whereas in other studies, article omission led to impaired comprehension (Petretic & Tweney, 1977; Shipley *et al*., 1969), indicating that 18–24-month-olds might not yet have robust knowledge regarding when an article is required within an NP. In both English and Dutch, determiners are not obligatory before mass nouns or plurals. Thus, infants might consider determiners as optional until they have learned the specific conditions in which they are required (e.g. Valian, 1986). In French, however, the omission of determiners is highly restricted. One might predict that French learners are sensitive to the presence of an article at an earlier age than English and Dutch children. Future studies should test whether the absence of a determiner affects French toddlers' comprehension differently.

What is the nature of toddlers' use of gender cues on determiners? Several possible mechanisms may underlie the effect. Given that the gender of most words in French is arbitrary, it is unlikely that gender-marked determiners activate or select certain semantic features, yielding faster access to words possessing these features. Do our results then indicate that 25-month-olds have acquired grammatical knowledge about gender-marked function words? According to the grammatical view, children have made the generalization that certain nouns are preceded by masculine functors, while other nouns are preceded by feminine functors. For example, *ballon* is a masculine noun and should be preceded by *le* or by another masculine functor, such as *un*, *mon*, or *son*. Alternatively, according to the co-occurrence view, the specific co-occurrence between a particular article form and a noun determines the efficiency of recognition (e.g. the probability of encountering *ballon* given *le* is higher than *banane* given *le*). This effect is merely local and no generalizations are made. Once two specific items (e.g. *le* and *ballon*) co-occur frequently enough, the first item starts facilitating the second one (or inhibiting others). This experiment does not address this question, and our results are compatible with both the grammatical and the co-occurrence views. We are conducting further experiments to test if toddlers can generalize a whole gender class of pre-noun function words from a subset of functors during word learning.

In conclusion, this study has shown that by 25 months of age, French learners use gender information on definite articles to recognize subsequent nouns. In comparison with learners of gender-free determiner systems, French learners need to acquire the additional gender-marking information on articles. It is therefore remarkable that only a few months after English learners have been shown to use gender-absent definite articles in sentence comprehension, French learners have acquired sufficient knowledge about gender-marked articles to use them in an adult-like fashion.

Acknowledgements

We would like to thank all the families that participated, Louis-Pierre Sarrazin for his work in a pilot experiment, Alexandra Marquis for speech stimulus recording, Andréane Melançon and Annick Moisan for their help with recruitment, Elizabeth Johnson for helpful discussions and comments on earlier versions of this paper, and the three anonymous reviewers for their feedback. Parts of this research were presented at the 16th International Conference on Infant Studies (Vancouver, Canada, March 2008). This research was funded by an HSP Talent Programme grant to the first author and by NSERC, SSHRC, and CFI grants to the second author.

References

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development*, **3**, 325–341.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Christophe, A., Guasti, T., Nespor, M., Dupoux, E., & Van Ooyen, B. (1997). Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, **12**, 585–612.

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

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

Fernald, A., Pinto, J.P., Swingley, D., Weinberg, A., & McRoberts, G.W. (1998). Rapid gains in speed of verbal processing by infants in the second year. *Psychological Science*, **9**, 228–231.

Fisher, C., Klingler, S.L., & Song, H.-J. (2006). What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition*, **101**, B19–B29.

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

Gelman, S.A., & Taylor, M. (1984). How two-year-old children interpret proper and common names for unfamiliar objects. *Child Development*, **55**, 1535–1540.

Gerken, L.A., & McIntosh, B. (1993). The interplay of function morphemes and prosody in early language. *Developmental Psychology*, **29**, 448–457.

Grosjean, F., Dommergues, J.-Y., Cornu, E., Guillelmon, D., & Besson, C. (1994). The gender-marking effect in spoken word recognition. *Perception and Psychophysics*, **56**, 590–598.

Höhle, B., Schmitz, M., Santelmann, L.M., & Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development*, **2**, 277–300.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: the role of determiners in segmentation and categorization of lexical elements. *Infancy*, **5**, 341–353.

Johnson, E.K. (2005). Grammatical gender and early word recognition in Dutch. In A. Brugos, M.R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development* (Vol. 1, pp. 320–330). Sommerville MA: Cascadilla Press.

Johnson, E.K., & Diks, M. (2005). On-line processing of grammatical gender in Dutch-learning toddlers. Paper presented at the Xth International Congress for the Study of Child Language, Berlin, Germany.

Katz, B., Baker, G., & MacNamara, J. (1974). What's in a name? On the child's acquisition of proper and common nouns. *Child Development*, **45**, 269–273.

Kedar, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development*, **77**, 325–338.

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.

Morgan, J.L., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories: toward a broader conceptualization of bootstrapping. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 263–283). Mahwah, NJ: Lawrence Erlbaum Associates.

Petretic, P.A., & Tweney, R.D. (1977). Does comprehension precede production? The development of children's responses to telegraphic sentences of varying grammatical adequacy. *Journal of Child Language*, **4**, 201–209.

Rozendaal, M.I., & Baker, A.E. (2008). A cross-linguistic investigation of the acquisition of the pragmatics of indefinite and definite reference in two-year-olds. *Journal of Child Language*, **35**, 773–807.

Shady, M.E. (1996). Infants' sensitivity to function morphemes. Unpublished PhD dissertation, State University of New York at Buffalo, Buffalo, NY.

Shady, M.E., Gerken, L.A., & Jusczyk, P.W. (1995). Some evidence of sensitivity to prosody and word order in ten-month-olds. In D. MacLaughlin & S. McEwan (Eds.), *Proceedings of the 19th Boston University Conference on Language Development* (Vol. 2, pp. 553–562). Sommerville, MA: Cascadilla Press.

Shafer, V., Shucard, D., Shucard, J., & Gerken, L.A. (1998). 'The' and the brain: an electrophysiological study of infants' sensitivity of English function morphemes. *Journal of Speech, Language and Hearing Research*, **41**, 874–886.

Seidl, A., & Johnson, E.K. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science*, **9**, 565–573.

Shi, R. (2005). Early syntactic categories in infants. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 481–495). Amsterdam: Elsevier.

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.

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, **11**, 407–413.

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.), *Proceedings of the 30th annual Boston University Conference on Language Development* (Vol. 2, pp. 549–560). Sommerville, MA: Cascadilla Press.

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.

Shi, R., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, **10**, 187–198.

Shi, R., Werker, J., & Morgan, J. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, **72**, B11–B21.

Shipley, E., Smith, C., & Gleitman, L. (1969). A study in the acquisition of language. *Language*, **45**, 322–342.

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, **46**, 265–294.

Valian, V.V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, **22**, 562–579.

Van der Velde, M. (2003). Déterminants et pronoms en néerlandais et en français: syntaxe et acquisition. Unpublished PhD dissertation, Université de Paris 8, Paris, France.

Van Heugten, M., & Johnson, E.K. (2006). Eye movements reveal Dutch-learning 19-month olds' sensitivity to definite articles. Poster presented at the 12th Annual Conference on Architectures and Mechanisms for Language Processing, Nijmegen, the Netherlands.

Van Heugten, M., & Johnson, E.K. (submitted). Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. Manuscript submitted for publication.

Zangl, R., & Fernald, A. (2007). Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development*, **3**, 199–231.

Received: 17 July 2007
Accepted: 9 April 2008