Stereotype or grammar? The representation of gender when two-year-old and three-year-old French-speaking toddlers listen to role nouns

LEVY, Arik, et al.

Abstract
Using a preferential looking paradigm, the current study examined the role that grammatical gender plays when preschool French-speaking toddlers process role nouns in the masculine form (e.g., chanteursmasculine ‘singers’). While being auditorily prompted with “Look at the ‘a role noun’!” two- and three-year-olds were presented with two pictures of two characters (‘boy–boy’ versus ‘girl–boy’) with attributes of the given role noun (e.g., singers with microphone and music notes). All role nouns were presented in the masculine plural form, which, despite its use to refer to mixed-gender groups, can be interpreted as referring to men. We expected toddlers to be biased by stereotypes, yet when non-stereotypical role nouns were presented, toddlers were not influenced by grammatical gender, but by their own sex (even more so for three-year-old toddlers). The absence of sensitivity to grammatical cues for either age group is discussed in terms of the developmental awareness of grammatical gender.

Reference

DOI: 10.1017/S030500091500063X
Stereotype or grammar? The representation of gender when two-year-old and three-year-old French-speaking toddlers listen to role nouns
Stereotype or grammar? The representation of gender when two-year-old and three-year-old French-speaking toddlers listen to role nouns*

ARIK LÉVY AND PASCAL GYGAX
University of Fribourg, Switzerland

UTE GABRIEL
Norwegian University of Science and Technology, Trondheim, Norway

AND

PASCAL ZESIGER
University of Geneva, Switzerland

(Received 22 December 2014 – Revised 3 June 2015 – Accepted 1 October 2015)

ABSTRACT

Using a preferential looking paradigm, the current study examined the role that grammatical gender plays when preschool French-speaking toddlers process role nouns in the masculine form (e.g., chanteursmasculine ‘singers’). While being auditorily prompted with “Look at the ‘a role noun’!!”, two- and three-year-olds were presented with two pictures of two characters (‘boy–boy’ versus ‘girl–boy’) with attributes of the given role noun (e.g., singers with microphone and music notes). All role nouns were presented in the masculine plural form, which, despite its use to refer to mixed-gender groups, can be interpreted as referring to men. We expected toddlers to be biased by stereotypes, yet when non-stereotypical role nouns were presented, toddlers were not influenced by grammatical gender, but by their own sex (even more so for three-year-old toddlers). The absence of sensitivity to grammatical cues for either age group is discussed in terms of the developmental awareness of grammatical gender.

* Special thanks to nursery referees and tenants, as well as all toddlers and their parents, for taking part in the study. We also thank Prof Billard for having lent us the head-mounted eye-tracker system. This research was funded by the Swiss National Science Foundation, (Project No 100014_135271). Address for correspondence: Arik Lévy, Département de Psychologie, University de Fribourg, Rue Faucigny 2, 1700 Fribourg, Switzerland. e-mail: arik.levy@unine.ch
In grammatical gender languages such as French or German, (i) all nouns are grouped into gender and number classes, (ii) grammatical agreement is made between a noun and its associated elements (e.g., articles and adjectives), and (iii) the class membership of a noun shows some semantic correlation with sex (Dixon, 1982). As such, the choice of grammatical gender for role nouns can be indicative of the biological sex of the persons the nouns refer to. For example, when referring to a group of football players, speakers of French might use “Regarde les joueurs_{masculine} de football!” or “Regarde les joueuses_{feminine} de football!” ‘football players’. In this latter example, and in most cases, the feminine form is a combination of the masculine form and an added or changed suffix (e.g., un joueur_{masculine} vs. une joueuse_{feminine} ‘a player’), causing morphological and (often) phonological differences. Importantly for the present paper, whilst the use of the feminine form (i.e., joueuses) exclusively indicates that the footballers are female, the use of the masculine form (i.e., joueurs) is not necessarily informative regarding the gender composition of the group, as masculine forms can be used specifically, meaning that they specifically refer to male referents, as well as generically, meaning that they refer to a group composed of individuals of both genders or a group where gender is unknown or irrelevant. In principle, when gender is not explicitly stated with other semantic cues, one should assume that the masculine form is used in a generic way (Baudino, 2001).

However, previous research suggests that speakers of French (e.g., Gygax & Gabriel, 2008; Gygax, Gabriel, Sarrasin, Oakhill, & Garnham, 2008; Gygax, Gabriel, Lévy, Pool, Grivel, & Pedrazzini, 2012) and other grammatical gender languages (e.g., in German: Irmeng, 2007; Stahlberg Braun, Irmeng, & Sczesny, 2007; in Spanish: Carreiras, Garnham, Oakhill, & Cain, 1996; Flaherty, 2001) tend to interpret the grammatical masculine information as indicative of a specific use, which hence leads to interpreting masculine role nouns as referring to male persons.

Furthermore, research has revealed an additional influence of gender stereotypical expectations (e.g., playing football might be more strongly associated with being male than with being female) on adult readers’ gender representations in grammatical gender languages. Esaulova, Reali, and von Stockhausen (2014), for example, found that anaphor resolution, whereby an anaphor (e.g., he) is linked to an antecedent (e.g., the football player), was also influenced by gender stereotypicality in German (although only in late temporal eye-tracking measures). In all, from the research on adults, we know that readers of grammatical gender languages (such as French) can base their mental representation of gender on both grammatical and stereotypical information, yet grammatical
gender seems to be dominant in this process. However, we know little about
the age at which these two sources start to shape an individual’s
language-related representations of gender. This knowledge gap is partially
due to most developmental research focusing either on general language
development without paying much attention to the specific case of the
generic use of masculine forms, or on the general development of gender
stereotypes without paying much attention to varieties in the child’s
linguistic input that result from different grammatical gender systems.
More specifically, whereas English-learning toddlers are exposed to
semantic (and few morphological) cues that initiate gender categories (e.g.,
“Look at the boys!” instead of “Look at the children!”), and associate them
with activities (e.g., “Look at the boys playing football!”; Zosuls, Ruble,
Tamis-LeMonda, Shrout, Bornstein, & Greulich, 2009), French-learning
toddlers are additionally exposed to grammatical cues and to ambiguity
when using masculine forms to refer to people.

To our knowledge, only two studies have actually explored the
development of the processes involved in how stereotypical and
grammatical information may or may not influence the way schoolchildren
construct gender representations of role nouns, namely Chatard,
Guimond, and Martinot (2005) and Liben, Bigler, and Krogh (2002). In
short, these two studies show that, when processing role nouns, BOTH
gender stereotypes and lexical markings influence English-speaking
schoolchildren (aged 5;0–6;0 onwards; Liben et al., 2002), and BOTH
gender stereotypes and grammatical gender influence French-speaking
adolescents (aged 14;0–15;0; Chatard et al., 2005). While these findings
echo what was found in adults (as described above), they are rather
uninformative regarding the onset of an interplay between stereotypical
information and grammar cues of gender in language development in
toddlerhood.

In order to answer such a question, one has to first determine the age onset
of sensitivity to those features. Against this background, the current study
aims to explore the development of grammatical- and stereotype-related
information processing for language-based gender representations in
toddlers.

The onset of sensitivity to gender stereotypes
The acquisition of a gender stereotype requires awareness and the deliberate
application of social categories (female vs. male) as well as knowledge of the
attributes that are typically associated with these categories (e.g., female –
pink, male – blue). Based on a literature review, Martin and Ruble (2010)
concluded that there is empirical evidence “that most children develop the
ability to label gender groups and to use gender labels in their speech
between 18 and 24 months” (p. 356). In other words, children show an awareness of the social ‘female’ and ‘male’ categories. Furthermore, rudimentary stereotypes develop by about two (Kuhn, Nash, & Brucken, 1978), and basic stereotypes by three years of age (Signorella, Bigler, & Liben, 1993). In the latter case, children actually associate attributes with these categories.

Of particular interest to our study are studies that focused on the associations between particular activities and gender, such as Serbin, Poulin-Dubois, and Eichstedt (2002) or Eichstedt, Serbin, Poulin-Dubois, and Sen (2002), who presented two-year-old toddlers with pairs of photographs of (male or female) actors performing a stereotypical activity (masculine, feminine, or gender-neutral) in a preferential-looking paradigm. In these studies, two-year-old toddlers spent more time looking at activities that were inconsistent with gender norms than at those that were consistent. More recently, Hill and Flom (2007) presented toddlers with pairs of videos showing an actor performing a stereotypical activity (e.g., putting on lipstick or shaving), and found that toddlers age 2;6 but not 1;6 looked at activities that were inconsistent for longer than at activities that were consistent with gender norms. Together, these results indicate that two-year-old toddlers have acquired some knowledge of gender-typed activities.

In sum, empirical evidence suggests that two-year-old toddlers are not only aware of gender categories but even link gender categories to certain attributes and activities.

The onset of sensitivity to grammatical gender
Recent research from van Heugten and Shi (2009) indicates that when comprehending nouns (e.g., ‘a balloon’), French-learning toddlers age 2;1 are already sensitive to grammatical gender information carried by determiners. Earlier, and pertaining to our study, Karmiloff-Smith (1979) ran a series of seminal experiments on the acquisition of grammatical gender knowledge (suffixing) in French speakers aged 3;0 to 11;0. In one of her experiments (Exp. 10, p. 156), she studied how children would resolve a conflict between a natural gender cue (i.e., the sex of two human-like Martians on a picture) and a phonological clue from the suffix of the (invented) noun to refer to them (e.g., a female character with a masculine phonological clue). Children up to the age of 10;0 used the phonological clue to determine an appropriate determiner (i.e., unmasculine or uneféminine) irrespective of the gender of the persons depicted. In the non-conflict condition (i.e., an arbitrary and non-gendered suffix was used) children relied on the gender of the character, although this was more pronounced when referring to male persons: to refer to the picture of a
male that had been introduced with a noun with an arbitrary and non-gendered suffix, three-year-old toddlers used the masculine article in 80% of the trials and the feminine article in 20% of the trials. To refer to the picture of a female character, the same age group used masculine and feminine articles equally (50%).

More recently, Royle and Valois (2010) showed that three-year-old French-speaking toddlers, though only capable of producing correct easy adjective inflexions in a puzzle task, could still recognize and understand feminine forms of more complex adjectives. Similar results have been reported in Spanish (e.g., Pérez-Pereira, 1991) and German (e.g., Mills, 1986). These results suggest that three-year-old toddlers have started to link grammatical gender marks to human gender.

Our current study
Regarding the development of grammatical and stereotype-related information processing for language-based gender representations, we can conclude from previous research that toddlers from age 2;0 can be expected to be sensitive to gender stereotypes, and from age 3;0 to process grammar cues. Therefore, the present study was designed to investigate the gender representation of French role nouns in two- and three-year-old toddlers.

We employed a preferential looking paradigm and relied on auditorily presented stereotypically male (e.g., ‘police officers’), female (e.g., ‘dancers’), and non-stereotypical (e.g., ‘musicians’) role nouns in the grammatically masculine plural form (i.e., the form that is used to refer to gender-mixed groups). Stereotypically female and male role nouns will be indicative of an influence of stereotype-related information, whereas non-stereotypical role nouns will be indicative of an influence of grammatical information.

More specifically, we examined how two- and three-year-old toddlers—both girls and boys—react to ‘boy–boy’ and ‘girl–boy’ pictures when prompted by a voice asking them to “Look at ‘the role noun’!” We assumed toddlers of both ages to be driven by stereotypes when processing male and female stereotype role nouns, yet we were particularly interested in toddlers’ behaviour when presented with non-stereotypical role nouns.

A challenge when using a preferential looking paradigm lies in how to interpret gaze behaviours. As discussed by Houston-Price and Nakai (2004; but see also Zesiger, Dupuis Lozeron, Lévy, and Frauenfelder, 2011), especially when examining gaze properties in toddlers, it is difficult to determine whether toddlers prefer (and therefore look longer at) familiar or unfamiliar items. Though novelty effects, reflecting toddlers’
interest in what is new or unexpected, are frequently reported in the literature on preferential gazes (e.g., Serbin et al., 2002). Familiarity effects, reflecting toddlers’ interest in what is familiar to them, are also often reported (e.g., Hirsh-Pasek, Golinkoff, & Hollich, 1999).

Houston-Price and Nakai (2004) suggest that familiarity and novelty effects often cancel each other out within an experiment, especially when testing young children who vary in age, because the strategies of preferential looking that sustain the effects are dynamic. In other words, a given pair of visual stimuli in a preferential-looking paradigm will be gazed at differently depending on age development or the developmental stage of each individual child. Because of these issues, data scrutinizing and transformation are highly recommended. The present study included young children of varying ages; therefore, data were transformed ad hoc to prevent a false-negative effect. We transformed the data to comply with a stereotypical gaze pattern, as suggested by the literature presented above, and closely examined the resulting pattern on the non-stereotypical condition, the latter of which were most indicative of possible early grammatical effects. The details of this particular transformation are presented in the ‘Results’ section.

METHOD
Participants
Forty-three French-speaking toddlers (eighteen age 2;0 and twenty-five 3;0; ±0;1; 20 girls and 23 boys) recruited from nurseries in Fribourg, Lausanne, and Geneva (Switzerland) took part in the experiment. An additional fifteen toddlers were excluded from analyses because of either a technical failure to properly record their gazes (N = 4) or because their parents did not provide information on the child’s role noun comprehension (N = 11).

Materials
The setting. The experiment was conducted either in the nursery (36 toddlers) or in the BabyLab at the University of Fribourg (7 toddlers). Both locations were set up as similarly as possible. In the nursery, toddlers were tested in a separate and quiet room. The setting was composed of a chair for the referent in front of a 15-inch presentation screen and a loudspeaker put on a table. The computer for controlling the timecourse of the item presentations was placed behind the referent, so that the child was not distracted. The BabyLab’s setting was composed of a laboratory room where the experiment took place, and an attendant control room where the experimenter could control the timecourse of experiment. The laboratory room consisted of a three-walled box that included a chair for the parent, a computer screen, and a loudspeaker below a 21·5-inch
presentation screen in front of the chair. The laboratory also had a playing corner with children’s toys to ensure that all toddlers would be at ease in the laboratory. The distance between the child and the presentation screen was adapted in order to reach similar eye-movement amplitude in both the nursery and laboratory.

Eye-movements were filmed with a head-mounted camera (Pomelo Sàrl <www.pomelo-technologies.ch>) kindly lent by Professor A. Billard from Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. The head-mounted system was composed of two 768×640 pixel cameras of a 74° x 100° field of view that allowed successful recording of the stimuli display and eye-gaze with the help of mirrors at a sample rate of 25 images per seconds. Cameras were connected by a cable connector to an A/D converter, which was connected to the computer for controlling the stimuli display. Synchronization of the two recordings (i.e., display field and eye-gaze) was guaranteed by frame numeration. Third-party software from the same group further allowed realignment of the visual field and eye-gaze records and extraction of the gaze position on each frame throughout experiment.

Stimuli building. Participants were presented with thirty-four role nouns (see Table 1), which were each accompanied by a picture and an auditory prompt.

Role noun selection was very challenging as the two-year-old toddler’s lexicon is very small. We first tried to pick up role nouns from Kern and Gayraud’s (2010) Communicative Developmental Inventory (CDI), but the items were not numerous enough. We therefore chose to select role nouns from the adults’ database and to check for toddlers’ comprehension. Role nouns were thus female stereotyped (e.g., ‘dancers’), male stereotyped (e.g., ‘police officers’), or non-stereotyped (e.g., ‘musicians’), twenty-three of which were based on Gabriel, Gygax, Sarrasin, Garnham, and Oakhill (2008, on adult ratings), and eleven were chosen from the CDI (Kern & Gayraud, 2010). The non-stereotypical role nouns taken from Gabriel et al. (2008) were characterized by a low standard deviation (i.e., their neutrality was not a result of some people rating them as male and others as female). All role nouns were additionally normed by a subset of sixteen students (see Table 1) to ensure no disparities with the norms provided by Gabriel et al. (2008). Note that almost all role nouns (91%) in this present study follow the principle mentioned earlier by which the feminine form is a combination of the masculine form and an added or changed suffix.

The pictures were each composed of four characters divided into two pairs, which always consisted of a pair of two boys (referred as ‘boy–boy’) and a pair of one girl and one boy (referred as ‘girl–boy’). Characters were avatars, created with the www.doppelme.com toolbox and modified with Gimp 2.6.11, that held attributes of a given role noun (e.g., white blouse
for ‘doctors’, guitar for ‘musicians’) that were not trivial at first glance but became obvious once the role noun was voiced (i.e., as confirmed by a panel of five judges). This was important, inasmuch as we did not want participants to activate role nouns and/or show preferential gazes before the prompt. Eye colour (blue or brown), hair colour (black or brown), and skin colour (light or tanned) was randomly assigned to each avatar. In essence, all avatars for a given role noun were similar but not identical, yet their main differences were whether they were portraying girls or boys (eyes and mouth and hair style) and the nature of their attributes (see

<table>
<thead>
<tr>
<th>French</th>
<th>English translation</th>
<th>Mean (SD)</th>
<th>Stereotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chasseurs</td>
<td>‘Hunters’</td>
<td>27 (10)</td>
<td>Male</td>
</tr>
<tr>
<td>Footballers</td>
<td>‘Football players’</td>
<td>29 (9)</td>
<td>Male</td>
</tr>
<tr>
<td>Mécaniciens</td>
<td>‘Mechanics’</td>
<td>29 (8)</td>
<td>Male</td>
</tr>
<tr>
<td>Pompoms</td>
<td>‘Firefighters’</td>
<td>29 (14)</td>
<td>Male</td>
</tr>
<tr>
<td>Pêcheurs</td>
<td>‘Anglers’</td>
<td>31 (13)</td>
<td>Male</td>
</tr>
<tr>
<td>Constructeurs</td>
<td>‘Constructors’</td>
<td>35 (12)</td>
<td>Male</td>
</tr>
<tr>
<td>Conducteurs</td>
<td>‘Drivers’</td>
<td>39 (15)</td>
<td>Male</td>
</tr>
<tr>
<td>Pilotes</td>
<td>‘Pilots’</td>
<td>41 (15)</td>
<td>Male</td>
</tr>
<tr>
<td>Magiciens</td>
<td>‘Magicians’</td>
<td>43 (16)</td>
<td>Male</td>
</tr>
<tr>
<td>Policiers</td>
<td>‘Police officers’</td>
<td>44 (16)</td>
<td>Male</td>
</tr>
<tr>
<td>Clowns</td>
<td>‘Clowns’</td>
<td>45 (15)</td>
<td>Male</td>
</tr>
<tr>
<td>Fermiers</td>
<td>‘Farmers’</td>
<td>46 (13)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Jardiniers</td>
<td>‘Gardeners’</td>
<td>48 (11)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Prisonniers</td>
<td>‘Prisoners’</td>
<td>48 (15)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Facteurs</td>
<td>‘Post Officer / Post Carrier’</td>
<td>53 (12)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Balayeurs</td>
<td>‘Sweepers’</td>
<td>54 (18)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Coureurs</td>
<td>‘Runners’</td>
<td>54 (10)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Docteurs</td>
<td>‘Doctors (medical)’</td>
<td>54 (12)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Boulangers</td>
<td>‘Bakers’</td>
<td>58 (10)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Cuisiniers</td>
<td>‘Cooks’</td>
<td>58 (13)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Dessinateurs</td>
<td>‘Cartoonists (or someone who draws)’</td>
<td>50 (4)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Voisins</td>
<td>‘Neighbours’</td>
<td>60 (6)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Skieurs</td>
<td>‘Skiers’</td>
<td>61 (11)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Musiciens</td>
<td>‘Musicians’</td>
<td>63 (9)</td>
<td>Female</td>
</tr>
<tr>
<td>Nageurs</td>
<td>‘Swimmers’</td>
<td>63 (8)</td>
<td>Female</td>
</tr>
<tr>
<td>Maîtres</td>
<td>‘Teachers (primary school)’</td>
<td>65 (16)</td>
<td>Female</td>
</tr>
<tr>
<td>Promeneurs</td>
<td>‘Walkers’</td>
<td>66 (10)</td>
<td>Female</td>
</tr>
<tr>
<td>Patineurs</td>
<td>‘Ice-skaters’</td>
<td>68 (17)</td>
<td>Female</td>
</tr>
<tr>
<td>Serveurs</td>
<td>‘Waiters’</td>
<td>68 (11)</td>
<td>Female</td>
</tr>
<tr>
<td>Vendeurs</td>
<td>‘Salespeople’</td>
<td>68 (12)</td>
<td>Female</td>
</tr>
<tr>
<td>Danseurs</td>
<td>‘Dancers’</td>
<td>69 (9)</td>
<td>Female</td>
</tr>
<tr>
<td>Chanteurs</td>
<td>‘Singers’</td>
<td>71 (13)</td>
<td>Female</td>
</tr>
<tr>
<td>Caisseurs</td>
<td>‘Cashiers’</td>
<td>77 (11)</td>
<td>Female</td>
</tr>
<tr>
<td>Nettoyeurs</td>
<td>‘Cleaners’</td>
<td>83 (15)</td>
<td>Female</td>
</tr>
</tbody>
</table>

TABLE 1. Mean percentages (and standard deviations) of perceived women in the occupations referred to by the role nouns in the pilot study.
Figure 1). The characters’ sex for each role noun was successfully recognized in a sample of eight toddlers in a pilot study.

The prompt “Look at the ‘role noun’!” was voiced simultaneously with the presentations of the characters (see below for the exact timing details). Prompts were recorded (44.1 kHz, 16 bits, stereo) either with a male voice or with a female voice with Sound Studio 3 in a hearing test booth. As length varied from item to item, the shorter ones were filled with silence at the beginning so that all items lasted 1.8 seconds.

Role nouns were presented in two semi-random sequences to control for order effects as well as to ensure that across the sample each role noun would be equally often presented in the first and in the second part of the experiment (i.e., given an expected relatively small number of participants). For each participant, the position (i.e., left or right) of the picture including a female character was random for each role noun. As a result, four lists were formed to control for item order and gender voice of prompts (i.e., the two semi-random sequences in two different voices). Participants were randomly assigned to a list.

Questionnaires to parents. The parents completed a questionnaire on (i) the toddler’s sex and birthdate and (ii) the toddler’s comprehension and production patterns of each of the role nouns used in the experiment, as well as of any of their semantic derivatives (e.g., ‘dancer’, ‘dance’, ‘to dance’). Additional questions about the milestones on motor and language development, and about known language disorders amongst family members, targeted participants with a possible language delay who would have been excluded from analysis.
Procedure

**The experiment timeline.** Each stimulus was composed of a pre-prompt phase that preceded the prompt and that was followed by a post-prompt phase. To catch the child’s attention and to centre its gaze, each item was preceded by three coloured lights flashing in the centre of the screen with a bell ringing. Once the child’s gaze was centred, the experimenter manually initiated the presentation of a stimulus on the monitor. In the pre-prompt phase, each pair of characters of a given combination was displayed one after the other, in the centre of the screen, for two seconds each. Note that for the first eighteen toddlers, we presented both pairs at the same time in the pre-prompt phase. It appeared that, on some occasions, some of the pairs received no attention at all, with seven toddlers showing no gazes on one or more stereotyped groups; hence, we decided to modify the procedure of the pre-prompt phase to ensure that the toddlers would equally process both pairs before getting the prompt, and only maintained eleven toddlers from the first procedure in the analyses.

Subsequent to the pre-prompt, the prompt was presented. Immediately after the prompt, both pairs were displayed simultaneously for 3.8 seconds. In all, a trial lasted 7.8 seconds (5.8 seconds for the toddlers from the initial procedure). Gaze fixations were recorded during the pre-prompt phase and post-prompt phase. See [Figure 1](#) for an illustration of the timeline display of an item.

**General procedure.** Parents of toddlers that met the age requirements (i.e., $2;0 \pm 0.2$ and $3;0 \pm 0.2$) were recruited through nurseries. Toddlers whose parents indicated willingness to participate by signing an agreement form were scheduled for the experiment.

After arrival in the lab, the caretakers (parent or nursery personnel) were introduced to the experiment, while the toddler was given the opportunity to play in order to become familiarized with the setting. Parents were asked to sign a consent form declaring that they had received all necessary explanations on the pros and cons of the experiment, fulfilling local ethical requirements. Most importantly, it was made clear that the toddler’s face was going to be recorded for data collection. During the session each toddler sat on the caretaker’s lap facing the screen where a cartoon was playing to capture the attention of the toddler. Before starting the experiment, a 9-point calibration was performed. During the experiment, caretakers wore a mask as well as headphones playing classical music to ensure that they would not bias toddlers’ gazes.

After the experiment, parents (via the caretakers when the toddlers were tested at the nursery) were given the questionnaire and were requested to return them within two weeks. The whole session lasted approximately 15 minutes, including the experimental session, which lasted five minutes.
Data processing of gaze time in the prompt phase. Recordings from the display field and eye-gaze were reassembled with third-party software developed by Professor A. Billard’s team at EPFL. The location of the visual fixations for each frame of each trial was coded into MIDDLE, RIGHT, LEFT, or ELSEWHERE by examining eye-movement throughout each item based on the eye-gaze record. By recording one frame every 40 ms for 7.8 seconds, one trial contained 195 frames (145 frames for the toddlers from the initial procedure). If a toddler did not look at a pair in the pre-prompt phase at least 25% of the time (and to each picture 50% of the time for the initial set-up for the first eighteen toddlers), the trial was removed from the analysis. In essence, it meant that 35% of the data had to be removed. In all, data were satisfactory (i.e., gazes in all three conditions of interest) for thirty-five toddlers (7 girls and 7 boys of 24 months; 9 girls and 12 boys of 36 months).

Gaze location was decided based on eye-gaze recordings only. The first author of this paper coded eye location of half of the participants, and a collaborator blind to the experiment coded the other half. To estimate inter-rater agreement, both coders independently coded the data of five participants. The inter-rater reliabilities were above .81, which we deemed as highly satisfactory (as suggested by Landis & Koch, 1977).

Data transformation for the critical exploration of grammatical effects. Prior to data transformation, role nouns that – according to parents’ report – were not understood by a participant were discarded from the analyses. The thirty-five toddlers that were analyzed understood on average 18.29 of the 36 role nouns ($SD = 8.49$), equally spread across stereotype conditions (Female: 5.97; Male: 5.49; Neutral: 6.83).

Gaze times spent on each pair (‘boy–boy’ and ‘girl–boy’ pictures) varied strongly across items and participants. We therefore decided to focus on the proportion of gaze time spent on each pair instead. Only the proportion of gazes on the boy–boy picture, computed as the gaze time spent on the ‘boy–boy’ picture relative to the total amount of gazes spent on both pairs (the ‘boy–boy’ and the ‘girl–boy’), were included in the analyses, as the proportions are complementary (i.e., a proportion of 0.3 on the ‘boy–boy’ meant a proportion of 0.7 on the ‘girl–boy’).

In the next step, we used the data from the stereotyped role noun conditions to check for the toddler’s preferential gaze (novelty vs. familiarity). Based on the notion that children are already sensitive to gender stereotypes by the age of two, we expect toddlers who prefer novel stimuli to have more gazes on the ‘boy–boy’ pair (than the ‘girl–boy’ pair) if prompted by a female stereotyped role noun, whereas we expect toddlers who prefer familiar stimuli to have more gazes on the ‘girl–boy’ pair (than the ‘boy–boy’ pair) if prompted by a female stereotyped role noun, and the other way around if prompted by a male stereotyped role noun.
For each toddler the proportion of gazes to the ‘boy–boy’ picture when presented with female stereotypical role nouns (higher values indicating a preference for novel stimuli) was subtracted from the proportion to the ‘boy–boy’ picture when presented with male stereotypical role nouns (higher values indicating a preference for familiar stimuli). Positive differences would then indicate a relatively stronger preference for familiar stimuli, and negative differences a relatively stronger preference for novel stimuli. For example, a given participant, exemplifying a familiarity pattern by scoring 0.8 on the ‘boy–boy’ when presented with male stereotyped role nouns and 0.3 when presented with female stereotyped ones would get a positive resulting value of 0.5 (i.e. 0.8–0.3 = 0.5). In contrast, a participant exemplifying a novelty pattern by scoring 0.2 on the ‘boy–boy’ when presented with male stereotyped role nouns and 0.7 when presented with female stereotyped ones would get a negative resulting value of −0.5 (i.e. 0.2–0.7 = −0.5). In our data, 22 toddlers complied with a familiarity pattern (positive difference) and 13 toddlers complied with a novelty pattern (negative difference). The two groups did not differ in terms of age ($\chi^2 (1, N = 35) < 1$). Slightly more boys showed a familiarity pattern than a novelty pattern, whereas an equal number of girls was assigned to each pattern ($\chi^2 (1, N = 35) = 2.1, p = .15$).

The above calculation allowed us to determine toddlers’ patterns of preferential gazes. The following step was implemented to merge the results from both patterns in order to simplify the analyses. To do this, we simply inverted the gaze proportions for each toddler complying with the novelty gaze pattern to mimic a familiarity gaze pattern for all conditions. For example, the data from the participant complying with a novelty gaze pattern presented above would be scored as 0.8 (i.e., $1 \cdot 0_{\text{full proportion}} - 0.2 = 0.8$) on the ‘boy–boy’ when presented with male stereotyped role nouns and 0.3 (i.e., $1 \cdot 0_{\text{full proportion}} - 0.7 = 0.3$) when presented with female stereotyped ones. Note that the same transformation was applied to scores associated with non-stereotyped role nouns. All analyses were hence conducted on assumed familiarity-bound gazes.

RESULTS
A general 3 (Stereotype: Female vs. Male vs. Neutral) × 2 (Age: 2;0 vs. 3;0) × 2 (Sex of respondent: Girls vs. Boys) full factorial ANOVA with Age and Sex of respondent as between-subject factors and Stereotype as a within-subject factor was initially run on the proportion of gaze time on the ‘boy–boy’ pictures. In line with the data preparation to determine the preferential gaze (the group splitting was based on gaze pattern to male and female stereotyped role nouns), the analysis revealed a main effect of Stereotype ($F(2,62) = 14.97; p < .01$), with toddlers’ proportion of gaze to
the ‘boy–boy’ picture being lower for female stereotypes (0.37, see Figure 2 – solid line) and higher for male stereotypes (0.62). For neutral role nouns, which were not used to determine the preferential gaze, toddlers’ proportion of gaze to the ‘boy–boy’ picture corresponded to a random gaze distribution (0.49). All three pairwise comparisons t-tests on proportion of gaze to the picture composed of ‘boy–boy’ were significant (ps < 0.05), suggesting that there was no effect of the grammatically masculine form, which would have been reflected in higher scores (i.e., preference to ‘boy–boy’ for neutral role nouns).

There was also a Sex of respondent by Age interaction effect (F(1,31) = 5.28; p < 0.05), indicating that three-year-old toddlers tended to show an own-sex preference (age 3;0: F(1,19) = 3.70; p = 0.07; girls: .46; boys: .55), whereas the two-year-old toddlers did not (age 2;0: F(1,12) = 3.04; p = 11; girls: .51; boys: .46). There were no other main or interaction effects.

As mentioned earlier, since our data transformation complied with a stereotypical gaze pattern, we wanted to more closely examine the resulting pattern on the neutrally stereotyped condition, as the latter is the most indicative of possible early grammatical effects. We therefore conducted a 2 (Age: 2;0 vs. 3;0) × 2 (Sex of respondent: Girls vs. Boys) full-factorial ANOVA on the proportion of gaze time on the ‘boy–boy’ pictures in the neutral role noun condition only, with Age and Sex of respondent as between-subject factors. The analysis revealed a main effect of Sex (F(1,35) = 4.48; p < 0.05, η² = .13), suggesting that children displayed a preference for their own sex (girls: .42; boys: .56). To illustrate this effect, the means for all three role noun conditions split by sex of respondent were plotted into Figure 2. There were no other main (i.e., age) or interaction effects.

**DISCUSSION**

The present study examined – in a French language context – two- and three-year-old toddlers’ preferential gazes in response to auditorily presented role nouns in the grammatically masculine plural form (i.e., the grammatical form that can be used to refer to gender-mixed groups). The role nouns were either stereotypically male (e.g., police officers), female (e.g., dancers), or non-stereotypical (e.g., musicians). Based on previous research (e.g., Hill & Flom, 2007), we expected the toddlers to be sensitive to gender stereotypes and we used this sensitivity to qualify the toddlers’ preferential gaze. Focusing on the toddlers’ responses to non-stereotypical (neutral) role nouns, the main result of our study is that toddlers aged 2;0 to 3;0 do not seem, in contrast to adults (e.g., Gygax et al., 2012) and older children (e.g., Chatard et al., 2005), to be biased by the masculine grammatical form, but instead seem to refer to their own sex at least at age
This preference for one’s own sex is well in line with previous research (e.g., Ruble & Martin, 1998) and thus lends validity to our data preparation procedure.

Our main finding, that toddlers aged 2;0 to 3;0 do not seem to integrate grammatical (masculine) information, at first glance contrasts with previous research (e.g., Royle & Valois, 2010). However, we believe that it does not necessarily mean that our toddlers were not yet aware of (the phonological correlates of) grammatical gender. Indeed, in our research, we did not really compare grammatical forms, but instead used grammatically masculine forms only. We might speculate that although toddlers might have developed an awareness of gender categories in terms of lexicon, as well as an awareness of phonological differences (here: masculine vs. feminine suffixes), these processes may not yet be sufficiently semantically grounded (i.e., feminine form = female person and masculine form = male person) to impact on language comprehension. It could nevertheless be reflected in language production (e.g., Karmiloff-Smith, 1979), although we did not test this. Nonetheless, based on the notion that the feminine, but not the masculine, forms carry (semantic) gender information (i.e., feminine forms as gender marked and masculine forms as gender unmarked), an alternative explanation could be that toddlers might well link feminine forms of role nouns to female referents (i.e., a
feminine = female link), but that this association does not interfere with linking female referents to masculine forms (i.e., a masculine = person link). Our data do not allow us to distinguish between these explanations, as we neither have data on our participants’ conceptual and grammatical gender awareness, nor on their understanding of feminine forms of role nouns. Future research should also assess the children’s current developmental state in these respects. Our data are further limited by the use of toddlers’ gaze responses to stereotypic role nouns to determine their preferential gaze without using an additional task that would directly give us this information, provided such a task exists.

We believe that future research should focus on children older than age 3;0, most preferably in a longitudinal design, to shed light on the developmental timecourse of the use of grammatical gender when comprehending discourse, and maybe to avoid the gaze preference issues present in the current study. Parallel to this question is the identification of the sequence by which gender stereotypes are learnt. Some authors suggest three steps, from (i) BEGINNING AWARENESS (i.e., construction / information gathering; two–five years old), (ii) RIGIDITY (i.e., consolidation / schema confirmation; five–six years old), and (iii) flexibility (i.e., integration / schema development; seven–eight years old) (e.g., Trautner, Ruble, Cyphers, Kirsten, Behrendt, & Hartmann, 2005). The toddlers in our experiment (i.e., age 2;0 and 3;0) were clearly in the first step: beginning awareness.

As underlined by Kohlberg (1966), the development of gender stereotypes is further related to the development of gender identity, or more precisely, to the CONSISTENCY of gender identity. Following Slaby and Frey’s (1975) research, there are three developmental stages of gender consistency: (i) GENDER IDENTITY (i.e., knowledge of their own sex and that of others; around age 2;0); (ii) GENDER STABILITY (i.e., understanding that sex is stable over time; age 4;0–5;0); and (iii) GENDER PERSISTENCY (i.e., assuming that sex is invariant in spite of physical variations; age 6;0–7;0). These development sequences are also to be understood as illustrating the way children start to use gender as a social norm (Martin, Ruble, & Szkrybalo, 2002). Focusing on one’s own sex initiates gender reference, and understanding variability in behaviours compels children to consider gender references as social norms (Martin & Ruble, 2004).

Our toddlers seem to demonstrate both gender stereotypes as well as own-sex preference (especially at age 3;0), yet seem to be unaffected by grammatical gender, at least when associated to human referents. Future research should extend the age groups tested to older ones, to better encompass the notion that the acquisition of gender concepts may be interwoven with language learning processes (as hinted by Guioara, Beit-Hallahmi, Fried, & Yoder, 1982). If this is the case, future research
should also further clarify the developmental timeline of the impact of both stereotypical and grammatical sources of information. Younger children seem to be affected only by stereotypes (and by their own sex), whereas older ones (i.e., adolescents aged 14;0–15;0, as in Vervecken, Gygax, Gabriel, Guillod, & Hannover, 2015) seem to be mainly affected by grammatical cues. Note that on top of these possible future directions, it would be interesting to examine whether the form and frequency not only of the masculine form but also of the feminine form may have an impact on the way the masculine form is perceived. Namely, some role nouns in the masculine form may have feminine counterparts that are not often used and also not clearly audible (e.g., employé vs. employée ‘employee’ with the final e generating just a difference in vowel quantity). In these cases, it might be that the masculine form is more likely to be interpreted as a true generic form. In our study, we could not examine this issue, as 91% of our role nouns had clearly audible feminine counterparts.

To conclude, our study showed that toddlers already spontaneously integrate the gender of a whole range of daily activities into their mental representations at age 2;0, and that their representations are not biased by grammatical gender but rather based on their own sex at age 3;0. Methodologically, two gaze patterns seemed to be present in our results, confirming previous claims of caution when investigating gaze patterns at these ages. Thus, although our experimental paradigm seems valid, future research should include additional measures to further validate the procedure, as well as children of a wider age range to better monitor when grammatical information starts to impact on the mental representation of stereotypical role nouns.

REFERENCES


