COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2013) 1–23

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12074

All Giraffes Have Female-Specific Properties: Influence of Grammatical Gender on Deductive Reasoning About Sex-Specific Properties in German Speakers

Mutsumi Imai,[a] Lennart Schalk,[b] Henrik Saalbach,[c] Hiroyuki Okada[d]

[a]*Faculty of Environment and Information Studies, Keio University at Shonan-Fujisawa*
[b]*Institute for Behavioral Sciences, ETH Zurich*
[c]*Department of Education, Saarland University*
[d]*Department of Engineering, Tamagawa University*

Received 24 May 2012; received in revised form 20 November 2012; accepted 14 February 2013

Abstract

Grammatical gender is independent of biological sex for the majority of animal names (e.g., any giraffe, be it male or female, is grammatically treated as feminine). However, there is apparent semantic motivation for grammatical gender classes, especially in mapping human terms to gender. This research investigated whether this motivation affects deductive inference in native German speakers. We compared German with Japanese speakers (a language without grammatical gender) when making inferences about sex-specific biological properties. We found that German speakers tended to erroneously draw inferences when the sex in the premise and grammatical gender of the target animal agreed. An over-generalization of the grammar–semantics mapping was found even when the sex of the target was explicitly indicated. However, these effects occurred only when gender-marking articles accompanied the nouns. These results suggest that German speakers project sex-specific biological properties onto gender-marking articles but not onto conceptual representations of animals per se.

Keywords: Language and thought; Deductive reasoning; Property inference; Grammatical gender; Linguistic relativity

1. Introduction

Many languages of the world have a system of grammatical gender, in which nouns are assigned to one of a limited number of gender classes (Corbett & Fraser, 2000). Unlike languages that mark gender semantically only (e.g., English), languages with

Correspondence should be sent to Mutsumi Imai, Keio University at Shonan-Fujisawa, Fujisawa, Kanagawa 252-8520, Japan. E-mail: imai@sfc.keio.ac.jp

grammatical gender assign gender to all nouns regardless of whether the referents have a biological sex. The link between gender assignment and conceptual properties of non-human referents has widely been said to be arbitrary (e.g., Aikhenvald, 2000; Fox, 1990), as grammatical gender is not relevant to biological sex for a majority of words. For example, different languages assign different gender categories to the same object (e.g., "table" is masculine in German [*der Tisch*], but feminine in Spanish [*la mesa*]). The separation between the grammatical category and the meaning is most evident in basic-level animal names. For example, in German, the word *Giraffe* is grammatically feminine and *Elefant* is masculine, but of course it is not the case that all giraffes are female or that all elephants are male. Nonetheless, the feminine article *die* ($_{[\text{FEM}]}$) must be applied when one refers to a grammatically feminine noun, and the feminine pronoun *sie* must be used for anaphoric reference, whether the referent is biologically female or male (e.g., *die männliche* [male] *Giraffe*).

Here, an interesting question is whether speakers link grammatical gender to biological sex even though they are in general orthogonal, and if yes, to what extent speakers are able to separate the two when necessary. From the perspective of a speaker of a language without grammatical gender, it appears confusing that one has to use the feminine article and the female pronoun even when the giraffe one is talking about is actually male. Of course, speakers of a language with a grammatical gender system must know that grammatical gender does not directly reflect biological sex. However, speakers may still consciously or unconsciously link grammatical gender and biological sex based on a small number of salient exemplars such as human terms. In German, for example, salient female terms such as *woman*, *aunt*, and *mother* are grammatically feminine, whereas salient male terms such as *man*, *uncle*, and *father* are grammatically masculine (*Natural Sex Principle*, cf. Zubin & Köpcke, 1986). Thus, speakers may falsely generalize this rather *exceptional* mapping between gender class and biological sex to words for animated entities in general. This assumption is consistent with Vigliocco and colleagues' (Vigliocco, Vinson, Paganelli & Dworzynski, 2005) *sex and gender hypothesis,* which proposes that a conceptual influence of grammatical gender originates in speakers' first noticing the correspondence between grammatical classes and corresponding conceptual classes. In other words, acknowledging the link between biological sex and the grammatical gender class in the case of some salient human-specific terms leads speakers to develop a general (probably unconscious) anticipation that even non-human animals from the same grammatical gender class are more similar to one another than animals from different grammatical gender classes.

The question above of course is deeply related to linguistic relativity, which asks how linguistic categories are related to some form (or domain) of thought. Most of the previous research asked whether grammatical gender influences speakers' concepts of entities in terms of typically feminine/masculine attributes or whether any sense of similarity among otherwise conceptually unrelated objects is heightened just because they belong to the same gender category. Konishi (1993), for example, looked at how Spanish and German speakers construe the femininity or masculinity of non-animal objects. Participants rated various nouns on semantic differential potency scales which are

related to sex-role stereotypes (e.g., weak vs. strong; tender vs. vigorous). German speakers rated the noun meaning "moon" (which is masculine in German, i.e., *der*$_{[MASC]}$ *Mond*, and feminine in Spanish, i.e., *la*$_{[FEM]}$ *luna*) higher in masculinity than the word for "sun" (which is feminine in German, i.e., *die*$_{[FEM]}$ *Sonne*, and masculine in Spanish, *el*$_{[MASC]}$ *sol*), whereas Spanish speakers showed the reverse pattern. Sera and colleagues (Sera, Berge, & del Castillo Pintado, 1994; Sera et al., 2002) asked Spanish and French speakers to assign either a female or a male voice to artifact objects and reported that the judgments tended to agree with the grammatical gender of the objects (see also Boroditsky, Schmidt, & Phillips, 2003; Flaherty, 2001; Ramos & Roberson, 2010). Influence of grammatical gender has been identified in tasks that do not involve explicit judgements as well (e.g., Kousta, Vinson, & Vigliocco, 2008; Vigliocco & Franck, 2001; Vigliocco, Vinson, Indefrey, Levelt, & Hellwig, 2004; Vigliocco et al., 2005). For example, Vigliocco et al. (2005) elicited semantic substitution errors during rapid naming of depicted objects and animals and found that native Italian speakers' substitution errors were influenced by (Italian) grammatical gender class, whereas English native speakers were not.

In the present research, we approach the issue concerning the relation between gender grammar and thought from a different angle. Instead of asking whether grammatical gender influences non-linguistic representations of objects, we ask whether German speakers are able to stay unaffected by grammatical gender when they draw inferences about biological sex-specific properties of animals. This question is important and goes beyond what has been revealed by previous studies using the voice attribution paradigm (Ramos & Roberson, 2010; Sera et al., 2002), the adjective rating paradigm (Boroditsky & Schmidt, 2000; Konishi, 1993), the similarity or categorization paradigm (Vigliocco et al., 2005; Phillips & Boroditsky, 2003; see also Saalbach & Imai, 2007), or the semantic substitution error paradigm (Kousta et al., 2008; Vigliocco et al., 2004, 2005), all of which involve implicit or explicit judgments about the femininity or masculinity of objects. Asking whether judgment of the femininity/masculinity of inanimate objects varies as a function of grammatical gender assignment in the speaker's language is interesting, but this judgment may not be especially relevant to everyday cognitive activities, nor be connected to deep inferences about the object. In contrast, inference about biological properties is extremely important for understanding animals. Although our task does not intend to assess purely non-linguistic representation or cognitive processes, our approach to the issues concerning the relation between language and thought may be claimed to be more meaningful to everyday cognition: We often hear about the dilemma experienced by German speakers in choosing the pronoun to refer to a pet animal when the animal's basic-level name and its biological gender are in conflict, yet, to our knowledge, this natural question has not been addressed in the literature.

Whether the grammatical gender categorization system affects speakers' cognitive processes in any fashion is not so clear when gender–sex mapping is not straightforward in the speakers' language (i.e., masculine—male/feminine—female), as in the case of German (see Boroditsky & Schmidt, 2000; Konishi, 1993; Phillips & Boroditsky, 2003 for results showing some influence of grammatical gender on German speakers; but see

Sera et al., 2002 and Vigliocco et al., 2004, 2005). For example, Vigliocco et al. (2004, 2005) found that gender class influenced semantic substitution errors in German speakers when they name depicted objects and animals using a noun phrase including gender-marked determiners. However, no such influence was found when using bare nouns. Furthermore, Vigliocco et al. (2005), using an odd-one-out categorization task, found an effect of grammatical gender on Italian speakers' construal of similarity among animals, but not on German speakers.' They suggested that the relation between grammatical gender and speakers' concepts is weaker for languages with more than two gender classes, such as German which additionally has a neutral gender.

However, inference about biological sex-specific properties is more directly linked to grammatical gender categories. It is thus conceivable that German speakers are aware of the motivated link between grammatical gender and biological sex, even though the mapping between the gender categories and biological sex is irregular, as human males are clearly mapped to the masculine gender and human females are mapped to the feminine gender (with very few exceptions such as *das Mädchen* ["girl"]). Nonetheless, when thinking about animals at the level of generic species (dog, mouse, giraffe, etc.), speakers should separate grammatical gender and biological sex.

Saalbach, Imai and Schalk (2012) found that German-speaking children have difficulties in separating grammatical gender and biological sex. They taught Japanese- and German-speaking children a novel sex-specific property of animals (e.g., "All mama/papa animals have *X* inside," where *X* is an unknown biological property). The children were shown a picture of an animal whose sex could not be determined by the appearance. They were then asked whether the animal in the picture would have the property. While Japanese children, whose language has no gender grammar system, showed no particular response bias, German children's response was clearly affected by the grammatical gender class of the target animal. For example, when the taught property was female-specific and the grammatical gender of the target animal was feminine, German children tended to accept the deduction, while they reject the deduction for the target animal whose grammatical gender was masculine.

Thus, German-speaking children link grammatical gender to biological sex even though the grammar–concept mapping is irregular, and this awareness affects their reasoning about sex-specific properties. Given this result, it is important to examine whether adults are also influenced by the grammatical gender when they draw deductive inferences about sex-specific properties.

Here, we use the term "deductive reasoning" in an informal sense, referring to a process of reasoning about a specific case from a general fact or statement about a class of things to create a conclusion. In formal logic, a typical form of deductive reasoning is syllogism, in which a major premise and minor premise are both given (e.g., all animals have a brain, dogs are animals) and the conclusion is deduced (dogs have brains). However, in everyday contexts, deductive reasoning involves both instantiation—judgment of whether the fact you know is true of a certain class of things also applies to your case—and deductive inference. In fact, to draw sound deductive reasoning, it is critical that you correctly judge if your case belongs to the class. If you misidentify the class membership,

then your conclusion will be unsound. After this inference, you then try to draw a deductive conclusion.

For example, you may have heard that male pure-breed cats have a high risk of urethral calculus, which may cause urine tube blockage and then uremia. You might then think that you should buy cat food which claims to help prevent this problem. Here, you have made an inference that your pet cat may have this risk because you identified your cat as belonging to the class of "male pure-breed cats." Hence, the statement true of this class will be applied to your cat. In this study, we ask whether grammatical gender of animals' basic-level names affects the process of deductive reasoning like this.

It is well established that linguistic framing affects reasoning or decision making. For example, in a classic study, Tversky and Kahneman (1981) asked people the same problem framed in different ways (e.g., 200 of 600 people will be saved vs. 400 of 600 people will be killed), people shifted their preference for the choice. If grammatical gender affects deductive reasoning about sex-specific biological properties, then this would indicate that much simpler, obligatory language functions such as grammatical categories can sometimes bias reasoning and decision making.

Of course, adults speaking a language with gender classes must consciously understand that the grammatical gender of basic-level animal names is independent of animals' biological sex. However, as the link between grammatical gender to biological sex (masculine to male and feminine to female) was established during childhood, as demonstrated by Saalbach et al. (2012), adults may still be affected when they reason about sex-related properties of things such that grammatically feminine animals *in general* have a female-specific biological property. The issue here is not whether German speakers are aware that grammatical gender should not be used as a cue—of course, they must be—but whether they can successfully suppress the link between grammatical gender and biological sex.

2. Overview of the present research

We tested German speakers to investigate the potential influence of grammatical gender on deductive reasoning about animal-specific properties. Japanese participants served as a control sample, as Japanese does not have a grammatical gender-marking system. In Experiment 1, we set up five conditions (three experimental and two control conditions) within participants so that we could specify in what contexts we might find the influence of grammatical gender. In all conditions, we asked the participants to indicate whether a target had the property specified in the premise by pressing designated keys for "Yes" and "No." Here, we presented the target nouns in the *gender-marking determiner + noun* construction. Importantly, we explicitly instructed the participants that they should give a "No" response in cases in which the inference was logically indeterminable, in addition to the cases in which the deduction would be clearly false.

The *generic animal condition* was designed to test whether German speakers were more likely to draw an erroneous inference when the sex specified for the biological

property given in the premise and the grammatical gender class of the target animal's basic-level name were consistent (e.g., female—feminine) than when they were inconsistent (e.g., female—masculine). For example, the participants saw the premise "All and only female animals have property *X* inside" and then the target (e.g., "Die Maus"). Here, the deductive reasoning is logically indeterminable, as the biological sex of the target animal is unknown. Thus, "No" is the correct answer. Nevertheless, German speakers may falsely deduce that the mouse has the female-specific property because the grammatical gender of the target animal agreed with the biological sex specified in the premise; even if they responded correctly, their judgments may be delayed by the influence of the gender–sex agreement. Note that, unlike in English, the *(gender-marked) definite article + noun phrase* (e.g., *die Maus, der Elefant*) is a construction commonly used to express a generic meaning in German (Behrens, 2005). In contrast, no such asymmetry should be found in Japanese speakers. It should be easy for them to reject the inference in this logically indeterminable case, regardless of whether the (German) gender of the target noun is consistent or inconsistent with the sex of the animal.

To balance the numbers of "No" and "Yes" responses and to test for possible baseline differences in deductive reasoning across the two language groups, we included the *generic animal control condition*. Here, participants were to judge the correctness of the reasoning about a property true for all animals regardless of their sex (e.g., "All and only animals have *X* inside"), while the targets were exactly the same as in the *generic animal condition*. In this condition, we should not see difference between German and Japanese speakers.

The *sex-specified animal condition* was set up to test whether grammatical gender affects reasoning in German speakers even when the sex of the animal is explicitly specified in the conclusion. For example, if the premise stated that the property was shared by all and only female animals, then the target was either *die weibliche Giraffe* (consistent: "the$_{[FEM]}$ female giraffe") or *die männliche Giraffe* (inconsistent: "the$_{[FEM]}$ male giraffe"). Here, the property inference should of course be made based on the agreement between the sex specified in the premise and the target animal's sex as indicated by the adjective. It is interesting to see whether consistency between grammatical gender and sex affects German speakers' judgments in this obvious case. If this is indeed the case, the difference in error rates and/or the speed of responses between consistent and inconsistent trials should be seen in German speakers but not in Japanese speakers. Note that, unlike the case with the *generic animal* condition, it is in the inconsistent condition that German speakers need to suppress the semantic motivation of the gender-marking article. We thus expect a higher rate of error responses and/or slower response times for the inconsistent condition than in the consistent condition in German speakers. When compared cross-linguistically, we expect poorer performance in German speakers as compared to Japanese speakers in the gender–sex inconsistent case.

We also included the *sex-specified animal control condition* to rule out an alternative explanation in the event that an effect was found for German speakers in the *sex-specified animal condition*. Provided that the expected effect was obtained, it may also have arisen simply from the difference in the difficulty in processing of the two types of noun phrases (i.e., grammatical gender–sex-specifying adjective matching and mismatching) rather than

the difficulty in the deductive reasoning. To disambiguate the two possibilities, the targets in this condition were exactly the same as those in the *sex-specified animal condition* (e.g., [*die weibliche Giraffe*] or [*die männliche Giraffe*]), but the property in the premise was shared by all animals. Finding a gender effect in German speakers in this control condition would indicate that the effect arises during local phrase processing and is thus not directly linked to reasoning about sex-specific properties.

We additionally included the *artifact condition* to examine whether German speakers' property inferences for non-animate entities was affected by grammatical gender. The target object was an artifact whose grammatical gender was either consistent or inconsistent with the sex specified in the premise. For example, for the "All and only female animals have *X* inside" premise, the consistent and inconsistent target would be "die[FEM] Säge" (saw) and "der[MASC] Hammer" (hammer), respectively. The conclusion was logically determinable and should always be "No" because the premise states that the sex-specific property was shared only by female or male *animals*. This condition allows us to evaluate the scope of the influence of grammatical gender on inferences about sex-specific biological properties: If the motivated sex–gender mapping is applied even in the realm of entities without sex, this would suggest that the influence of grammatical gender is truly overarching in German speakers.

Provided that we found the expected results in Experiment 1, at least two possible explanations should be considered. One is that the effect arises at the interface of syntax and semantics but not in the conceptual representation of animal kinds (Bender, Beller, & Klauer, 2011; Vigliocco et al., 2005). In other words, the speakers may project sex-relevant properties onto gender articles (but not onto names) and hence the effect may only be detected when a speaker is processing the gender-marking article or pronoun. An alternative possibility is that the overgeneralized grammar–concept mapping penetrates into the conceptual representation of generic-level animal kinds. If this is the case, the effect should be seen even when generic-level animal names are presented without the gender-marking article.

To separate the two possibilities, in the second experiment, we presented the target words in plural forms without any marking of gender class. As stated earlier, if grammatical gender affects the representation of objects, we should see an effect due to gender–sex congruency or incongruency in German speakers (but not in Japanese speakers) both in Experiments 1 and 2. If we find the effect only in Experiment 1 but not in Experiment 2, this would suggest that the characteristics of biological sex are projected onto the categories of grammatical gender, but grammatical gender does not penetrate into the conceptual representation of objects (including both animals and inanimate things).

3. Experiment 1

In this experiment, we tested whether grammatical gender affects speakers' deductive inference about a sex-specific biological property when the grammatical gender of the target object was explicitly invoked by the gender-marking article.

3.1. Method

3.1.1. Participants

Twenty-one native German-speaking undergraduates from Zurich and 17 native Japanese-speaking undergraduates from Tokyo participated for payment. In both language groups, most of the participants were undergraduates majoring in engineering or social sciences (including psychology). Many of the German speakers in the final sample knew language(s) other than German, but German was the language they used in everyday settings. Bilinguals who fluently spoke a language with gender grammar other than German were not included in the sample. The Japanese participants knew English and sometimes an additional language, but none were fluent bilinguals who could use the second language comfortably and spontaneously.

Four German participants and one Japanese participant were removed from the analyses because their error rate exceeded 50% in the control conditions. Thus, the final analysis included data of 17 German and 16 Japanese participants.

3.1.2. Design and materials

As described earlier, there were five within-subjects conditions: *generic animal*, *generic animal control*, *sex-specified animal*, *sex-specified animal control*, and *artifact* (see Table 1 for sample German stimuli).

In the *generic animal condition,* 36 animal names (half grammatically feminine, half masculine in German) that were commonly known to speakers of both languages were used as targets. Each animal appeared once in the sex–gender consistent trials and once in the inconsistent trials, yielding a total of 72 trials in this condition. Here, the premise stated either "All and only female animals have *X* inside" or "All and only male animals have *X* inside." Following the premise, a basic-level animal name was presented. German speakers saw the animal name with the gender article, whereas Japanese speakers saw just the bare noun.

Nouns may be accompanied by a numeral classifier in Japanese. These classifiers carry semantic information unrelated to gender (e.g., shape, animacy, functionality, etc.). Japanese classifiers are used only with a number, which is specified only when it is pragmatically important. Thus, in most contexts, it would not be natural to present a noun, for example, *inu* ("dog"), together with a numeral and classifier, as in *i*(one)-*ppiki* (classifier for small animals) *no* (genitive particle). More important, the *classifier + noun* construction does not allow for a generic reading of the noun (Mizuguchi, 2004). For this reason, we chose to present the noun in the bare form (see Appendix for sample Japanese stimuli). As described earlier, the correct response was "No" for all trials in this condition, as the deduction was not logically determinable.

The same 36 animal names (with the gender-marking determiner) were used in the *generic animal control condition,* in which the property given in the premise sentence was general to all animals (i.e., "All and only animals have *X* inside"). Here, of course, the correct response was "Yes" for all trials.

Table 1
Examples of the items used for German participants in the five different deductive inference conditions of Experiments 1 and 2

Condition	Premise	Sex–Gender Consistent	Sex–Gender Inconsistent
Generic animal	Alle weiblichen und nur weibliche Tiere haben *X* innen drin. (All and only female animals have property *X* inside.)	Exp 1: die$_{[FEM]}$ Giraffe (the giraffe) Exp 2: Giraffen (giraffes)	Exp 1: der$_{[MASC]}$ Elefant (the elephant) Exp 2: Elefanten (elephants)
	Alle männlichen und nur männliche Tiere haben *X* innen drin. (All and only male animals have property *X* inside.)	Exp 1: der$_{[MASC]}$ Elefant (the elephant) Exp 2: Elefanten (elephants)	Exp 1: die$_{[FEM]}$ Giraffe (the giraffe) Exp 2: Giraffen (giraffes)
Generic animal control	Alle Tiere und nur Tiere haben *X* innen drin. (All and only animals have property *X* inside.)	Exp 1: der$_{[MASC]}$ Elefant Exp 2: Elefanten Exp 1: die$_{[FEM]}$ Giraffe Exp 2: Giraffen	
Sex-specified animal	Alle weiblichen und nur weibliche Tiere haben *X* innen drin. (All and only female animals have property *X* inside.)	Exp 1: die$_{[FEM]}$ weibliche Maus (female mouse) Exp 2: weibliche Mäuse (female mice)	Exp 1: die$_{[FEM]}$ männliche Maus (male mouse) Exp 2: männliche Mäuse (male mice)
	Alle männlichen und nur männliche Tiere haben *X* innen drin. (All and only male animals have property *X* inside.)	Exp 1: der$_{[MASC]}$ männliche Hamster (male hamster) Exp 2: männliche Hamster (male hamsters)	Exp 1: der$_{[MASC]}$ weibliche Hamster (female hamster) Exp 2: weibliche Hamster (female hamsters)
Sex-specified animal control	Alle Tiere und nur Tiere haben *X* innen drin. (All and only animals have property *X* inside.)	Exp 1: die$_{[FEM]}$ weibliche Maus (female mouse) Exp 2: weibliche Mäuse (female mice)	Exp 1: die$_{[FEM]}$ männliche Maus (male mouse) Exp 2: männliche Mäuse (male mice)
Artifact	Alle weiblichen und nur weibliche Tiere haben *X* innen drin. (All and only female animals have property *X* inside.)	Exp 1: die$_{[FEM]}$ Säge (saw) Exp 2: Sägen (saws)	Exp 1: der$_{[MASC]}$ Hammer (hammer) Exp 2: Hämmer (hammers)
	Alle männlichen und nur männliche Tiere haben *X* innen drin. (All and only male animals have property *X* inside.)	Exp 1: der$_{[MASC]}$ Hammer (hammer) Exp 2: Hämmer (hammers)	Exp 1: die$_{[FEM]}$ Säge (saw) Exp 2: Sägen (saws)

In the *sex-specified animal condition,* 18 animal names (half grammatically feminine, half masculine) that were not used in the *generic animal condition* were presented twice, once in a consistent and once in an inconsistent trial. If the premise stated that the

property was shared by all and only female animals, then the target was either *die weibliche Giraffe* (consistent: "the$_{[FEM]}$ female giraffe") or *die männliche Giraffe* (inconsistent: "the$_{[FEM]}$ male giraffe"). If the premise stated that the property was shared by all and only male animals, then the target was either *der männliche Hamster* (consistent: "the$_{[MASC]}$ male hamster") or *der weibliche Hamster* (inconsistent: "the$_{[MASC]}$ female hamster"). It is interesting to see whether consistency between grammatical gender and sex affects German speakers' judgments in this obvious case. If this is indeed the case, the difference in error rates and/or the speed of responses between consistent and inconsistent trials should be seen in German speakers but not in Japanese speakers. Here, the correct response was "Yes" for consistent and "No" for inconsistent trials. The same target nouns were used for the *sex-specified animal control condition*, but here, the property in the premise was not sex specific (e.g., "All and only animals have *X* inside"). In this case, the correct response was "Yes" for all trials.

In the *artifact condition,* the premise was sex specific ("All and only female/male animals have *X* inside"), and 14 artifact names (half grammatically feminine, half masculine) served as targets. All artifact names appeared once in a sex–gender consistent and once in an inconsistent trial. The correct answer of course was "No" for all trials.

3.1.3. Procedure

Prior to the experiment, participants were told that they would have to indicate whether the property specified in the premise is inherited by the target by pressing a designated key for "Yes" or "No." They were informed that the premise always includes "Property *X*" and that *X* is a property that was internal and important for the object specified in the premise. They were also instructed to give a "No" response if the conclusion is logically indeterminable. Before proceeding to the main experiment, the participants went through 24 practice trials covering all the five conditions with the same procedure as the experimental trials but using target nouns that were not used in there. There were 12 trials for which the "Yes" response was correct and 12 trials for which the "No" response was correct. During the training, incorrect responses were indicated by a beep.

In the experimental session, there were 208 trials including 90 trials for which the "Yes" response was correct and 118 trials for which the "No" response was correct. The presentation order of the 208 trials of all conditions was completely randomized within and across participants. During the experimental trials, no feedback was given to any response.

In each trial, a fixation cross appeared on the screen for 1 s. The premise statement was then shown for 1.5 s, followed by a blank screen for 0.5 s. After the response, the screen remained blank for 1.5 s and the next trial was then started.

3.2. Results and discussion

We report the results separately for each condition with error rates and response latencies as dependent measures. For our research questions, comparisons across different conditions were not relevant. In other words, we were not interested in higher order

interactions involving the five conditions. Furthermore, as task difficulty is expected to differ substantially across the conditions, an overall ANOVA including all the conditions would make the variance of the error term unnecessarily large and hence the sensitivity to detect the effect of interest, that is, the influence of grammatical gender, would be seriously affected. We thus conducted separate ANOVA models for each condition.

3.2.1. Generic animal condition

As expected, there was a significant Language (German vs. Japanese) *X* Consistency (sex–gender consistent vs. inconsistent) interaction effect both for the subject analysis and the item analysis, $F_1(1, 31) = 8.9$, $\eta_p^2 = 0.22$, $F_2(1, 35) = 86.190$, $\eta_p^2 = 0.71$, both ps $< .01$. Paired t tests were performed on subject (t_1) and item means (t_2) contrasting the performance in consistent and inconsistent trials. German speakers were much more likely to erroneously deduce that the target animal had the property when the sex in the premise and the grammatical gender of the target were consistent (54.7%) than when they were inconsistent (29.9%), $t_1(16) = 3.133$, $d = 0.626$, $p < .01$, $t_2(35) = 13.447$, $d = 2.898$, $p < .01$. No such difference was found in Japanese participants (17.0% vs. 17.2%), $t_1(15) = -0.102$, $p = .920$, $t_2(35) = -0.067$, $p = .95$ (see Fig. 1).

In contrast, the performance in German speakers in the *generic animal control condition* was slightly better than that of Japanese speakers (German: 92.5%; Japanese: 83.0%), $t_1(31) = 1.821$, $d = 0.630$, $p = .078$; $t_2(35) = 4.645$, $d = 1.071$, $p < .01$, which indicates that the difficulty the German speakers showed in the *generic animal* condition could not be due to their generally poor deductive reasoning ability (see Fig. 1).

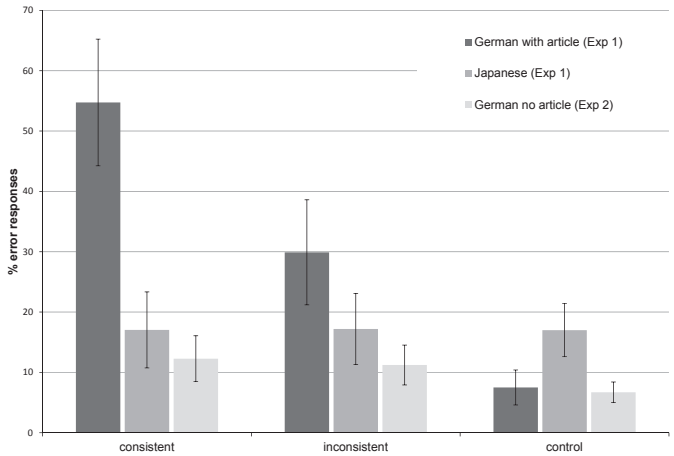

Fig. 1. Percentages of error responses in the *generic animal condition* (with sex-specific premises) and in the *generic animal control condition* (with sex-general premises) in Experiments 1 and 2. In Experiment 1, the target noun appeared with the gender article. In Experiment 2, the target noun was presented in the plural form without the article. In consistent trials, the sex specified in the premise and grammatical gender of the target agreed. In inconsistent trials, the sex specified in the premise and grammatical gender of the target animal disagreed.

One may wonder why the error rate is still relatively high (29.9% error rate) in German speakers when the grammatical gender and the sex in the premise were incongruent. In this condition, regardless of the gender article, participants had to judge that the deduction was indeterminable to correctly make a "No" response. Thus, German speakers had to inhibit the sex that was (automatically) invoked when gender article is processed, and this extra inhibitory process may have affected the reasoning process.

We examined whether the general difficulty experienced by German speakers in this task was reflected in response times. Here, because German speakers' error rates were too high to conduct ANOVA analysis when error responses were excluded (only 49% correct, averaged across the consistent and inconsistent trials), we included both correct and incorrect responses in the analyses. The means appeared that German speakers were slower (consistent: 777 ms; inconsistent: 797 ms) than Japanese speakers (consistent 742 ms; inconsistent 721 ms) to respond. However, this difference was not statistically supported, $F(1, 31) = 0.847$, nor was there a Language *X* Consistency interaction $F(1, 31) = 3.357$, both ps $> .05$.

3.2.2. Sex-specified animal condition

Here, both error rates and response times were considered, as error rates were sufficiently low for both groups to conduct a response latency analysis. On error rates, again, a significant Language *X* Consistency interaction effect was found, $F_1(1, 31) = 8.5$, $\eta_p^2 = 0.21$, $p < .01$, $F_2(1, 17) = 8.224$, $\eta_p^2 = 0.33$, $p = .01$. Here, different from the *generic animal* condition, the inconsistent trials (e.g., *die männliche Giraffe*) were expected to be more difficult than the consistent trials (*die weibliche Giraffe*). Indeed, German speakers made more erroneous inferences when the sex specified in the premise and the grammatical gender of the animal were incongruent, even when the animal's sex was clearly indicated by the adjective (inconsistent trials: 11.1%) as compared to when the grammatical gender of the animal and the sex indicated by the adjective agreed (consistent trials: 3.9%), $t_1(16) = 2.917$, $d = 0.878$, $p = .01$, $t_2(17) = 2.735$, $d = .994$, $p = .01$ (Fig. 2). No such difference was found in Japanese participants (consistent = 11.8%; inconsistent = 8.3%), $t_1(15) = -1.274$, $p = .22$, $t_2(17) = -1.294$, $p = .21$. When we compared German speakers' response to that of Japanese speakers, we found a significant difference in the consistent condition, $t_1(31) = -2.150$, $p = .039$, $t_2(17) = 2.364$, $p < .01$, but not in the inconsistent condition, t_1 (31) = 0.776, $p = .443$, t_2 (17) = 1.039, $p = .313$. These results suggest that grammatical gender may facilitate the reasoning when grammatical gender and the biological sex were congruent.

The results of the response time analyses tended to converge onto those of the error rate analysis, although the critical Language *X* Consistency interaction effect was not significant here $F_1(1, 31) = 1.889$, $p = .124$, and $F_2(1, 17) = 1.744$, $p = .204$. However, when the consistent and inconsistent trials were compared within each language, German speakers took more time to draw deductive inferences in the inconsistent (819 ms) than in the consistent trials (736 ms), $t_1(16) = 3.414$, $d = 0.436$, $p < .01$, $t_2(17) = 2.517$, $d = 0.851$, $p = .022$, whereas no such difference was found in Japanese responses

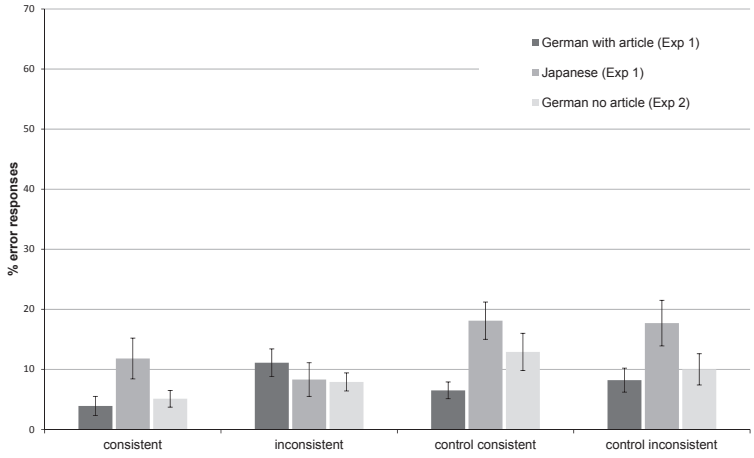

Fig. 2. Percentages of error responses in the *sex-specified animal condition* (with sex-specific premises) and the *sex-specified animal control condition* (with sex-general premises) in Experiments 1 and 2.

(consistent: 747 ms, inconsistent: 772 ms), $t_1(15) = 0.717$, $t_2(17) = 0.468$, both *ps* > .05 (Fig. 3).

In contrast to the sex-specified animal condition, there was no Language *X* Consistency effect at all in the *sex-specified animal control condition* on either error rate, F_1 (1, 31) = 0.394, $F_2(1, 17) = 0.601$ (see Fig. 2), or response time, $F_1(1, 31) = 0.002$, $F_2(1, 17) = 0.165$ (see Fig. 3).

Thus, even when the biological sex of an animal was explicitly indicated, grammatical gender affected German speakers' inferences about sex-specific animal properties. The results of the error rates and the response time analyses both showed that grammatical gender could both disturb and facilitate the reasoning depending on the task context. Note that German speakers' performance did not differ from that of Japanese speakers in the *sex-specified animal control condition* (where the premise was not sex specific), indicating that the sex–gender consistency effect in the *sex-specified animal condition* emerged through the process of reasoning rather than from mere disturbance due to gender–sex mismatch during the local-level processing of the target noun phrase.

3.2.3. Artifact condition

In the *artifact condition*, in neither language group did sex–gender consistent and inconsistent trials differ with respect to error rates (German: 3.8% vs. 1.3%; Japanese: 0.4% vs. 0.4%) nor response times (German: 751 ms vs. 738 ms; Japanese: 628 ms vs. 639 ms). There was no Language *X* Consistency effect on error rates, $F_1(1, 31) = 1.230$, $F_2(1, 13) = 2.131$, both *p*s > . 1, or response times, F_1 (1, 31) = 0.659, $F_2(1, 13) = 0.108$, both *p*s > .1. Thus, the influence of grammatical gender on sex-specific biological properties found in the animal domain did not extend to the artifact domain.

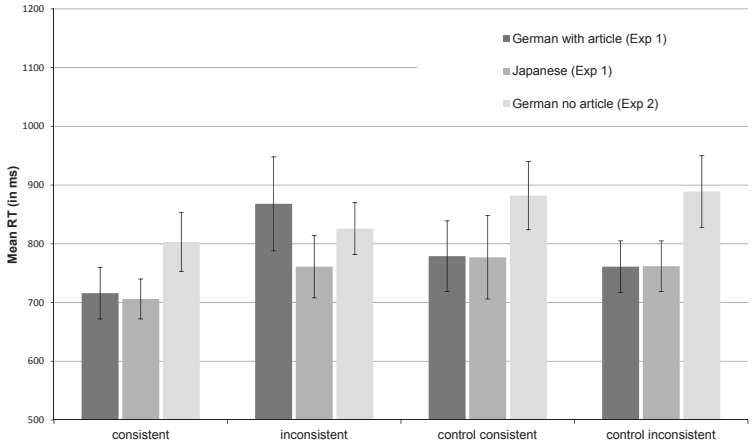

Fig. 3. Response times (in milliseconds) for correct responses in the *sex-specified animal condition* (with sex-specific premises) in Experiment 1 and the *sex-specified animal control condition* (with sex-general premises) in Experiments 1 and 2. In Experiment 1, the target noun appeared with the gender article. In Experiment 2, the target noun was presented in the plural form without the article. In consistent trials, the sex specified in the premise, the grammatical gender of the target, and the sex of the target all agreed. In inconsistent trials, the sex specified in the premise and the grammatical gender of the target agreed, but the sex of the target animal disagreed with the former two. Incorrect responses as well as responses exceeding 1,500 ms were excluded from the analysis.

The results of Experiment 1 showed that German speakers were indeed influenced by the motivated but logically orthogonal gender–sex mapping: When the biological sex specified in the premise agreed with grammatical gender of the target animal, they often erroneously reasoned that a sex-specific biological property was possessed by the target animal, even though its biological sex was unspecified and hence the deduction was indeterminable. Furthermore, German speakers experienced difficulty in rejecting the deductive conclusion when the biological sex specified for the property and grammatical gender of the target animal agreed, even though the target animal's sex was explicitly indicated as otherwise by a sex-specifying adjective, but their reasoning was facilitated when grammatical gender and the sex specified by the adjective were congruent.

These results naturally lead to a question addressed in Experiment 2: Are the same effects obtained when the target animal name is presented without the gender-marking article? If German speakers' representation of animals per se is affected by grammatical gender, the same effects should be observed without explicit invocation of grammatical gender. Alternatively, the gender effects in Experiment 1 may vanish when the animal name is presented without the gender article. If so, this would indicate that the gender effect arises because biological sex is projected onto the gender-marking article but not onto the animal names. Experiment 2 was conducted to disambiguate these two possibilities.

4. Experiment 2

4.1. Method

4.1.1. Participants

Twenty-nine German-speaking undergraduates from Zurich who had not participated in Experiment 1 took part in this study. Their demographic properties were identical to the participants of Experiment 1. Five participants were removed from further analyses because accuracy rate did not differ significantly from chance in the control conditions.

4.1.2. Design, materials, and procedure

The design, materials, and procedure of Experiment 2 were identical to those in Experiment 1 with one exception: All target words were presented in plural form without articles. In the *generic animal condition,* for example, the target *die* [FEM] *Maus* (mouse) was now presented as *Mäuse* (mice) and in the sex-specified animal condition, *die männliche* (male) *Maus* was now presented as *männliche Mäuse*. The plural form was chosen because a singular noun with no article is a rarely used construction in German (Behrens, 2005) and could thus be felt unnatural by native German speakers.

4.2. Results and discussion

In stark contrast to Experiment 1, we found no significant difference between the gender–sex consistent and inconsistent trials in any of the conditions on either error rates or response times (for t_1 and t_2: all ps > .05; see Figs. 1–3). When the performance of German speakers in this experiment was compared to that of Japanese speakers in Experiment 1, in no condition was there any Language *X* Consistency effect.

We also conducted a series of ANOVA analyses comparing German participants' performance in Experiments 1 and 2 for each condition. We found a significant interaction for the proportion of error responses in the generic animal condition (Experiment 1 consistent: 54.7%, inconsistent: 29.9%; Experiment 2 consistent: 12.3%, inconsistent: 11.2%), $F(1, 39) = 12.199$, $\eta_p^2 = 0.238$. For the sex-specified animal condition, although the mean difference between the consistent and inconsistent trials was larger in Experiment 1 than in Experiment 2 on both error rates (Experiment 1 consistent: 3.9%, inconsistent: 11.1%; Experiment 2 consistent: 2.8%, inconsistent: 7.9%) and the response times (Experiment 1 consistent: 736 ms, inconsistent: 819 ms; Experiment 2 consistent: 797 ms, inconsistent: 833 ms), on neither measures (error rates: $F(1, 39) = 2.364$ and $F_2(1, 17) = 2.481$, $ps > .1$; response times, $F_1(1, 39) = 2.240$, $F_2(1,17) = 0.309$, ps > .1) did the interaction reach the level of statistical significance.

In other conditions (i.e., generic animal control, sex-specified animal control, artifact), on neither error rates nor response times was there an Experiment *X* Consistency interaction effect (all ps > .1).

The results of Experiment 2 indicate that the grammatical gender effects found in Experiment 1 arise only when the speakers see the target animal name together with the

gender-marking article. This suggests that German speakers project sex on the gender article rather than on nouns; that is, the effect of grammatical gender in the inference of sex-specific properties did not arise because German speakers' representation of animals per se was shaped by gender grammar. However, this conclusion should be treated with care because the Consistency *X* Experiment interaction did not reach the level of significance for the sex-specified animal condition. The lack of the interaction effect was because the inconsistent trials induced slightly more errors and longer response times than in the consistent trials in Experiment 2 as well. It is possible that grammatical gender is weakly activated when German speakers see animal nouns even without the presence of the gender-marking article, but the level of activation is usually not strong enough to affect the performance in behavioral tasks (also see Vigliocco et al., 2004, 2005).

5. General discussion

Researchers investigating the relation between the speakers' conceptual representation of objects and gender grammar have traditionally approached the question in light of whether masculine or feminine images or attributes were projected on objects according to the grammatical gender of the name (e.g., Boroditsky et al., 2003; Konishi, 1993; Ramos & Roberson, 2010; Sera et al., 2002), how grammatical gender affects (implicit or explicit) judgments of similarity of objects by virtue of belonging to the same grammatical gender category (e.g., Vigliocco et al., 2005), or whether grammatical gender class is preserved in speakers' semantic substitution errors during rapid naming (e.g., Kousta et al., 2008; Vigliocco et al., 2004, 2005). In this research, we examined the relation between gender grammar and cognitive processes more directly: We investigated how grammatical gender affected inference about sex-specific biological properties.

We identified two contexts in which German speakers overgeneralize this grammar–concept mapping and, as a consequence, make erroneous deductive inferences. First, German speakers tended to erroneously reason that the sex-specific property given in the premise is inherited by the target animal when the sex specified in the premise and the grammatical gender of the basic-level animal name agreed (with the explicit presence of the gender-marking article). Second, the sex–gender agreement affected the inference even when the sex of the target animal was explicitly indicated: German speakers experienced difficulty in rejecting the deduction when asked, for example, to judge whether a female-specific property would be true for *die*$_{[\mathrm{FEM}]}$ *männliche Maus* (male mouse), while agreement of the grammatical gender and biological sex facilitated for accepting the deduction. However, these effects were apparent only when the gender-marking article was processed. Thus, German speakers seem to project biological sex-related properties onto gender-marking articles but not onto the conceptual representation of animals per se. Furthermore, this mapping does not go so far as to affect inferences when the targets do not have biological sex, which in turn suggests that grammatical gender categories do not affect core conceptual representations of animals and inanimate objects (also see Bender et al., 2011).

Our results were somewhat different from the results from Vigliocco et al. (2005) study, in which an effect of grammatical gender was found for Italian speakers but not for German speakers. Vigliocco et al. argued that the semantic–grammar link is less transparent in German than Italian because German has three gender classes and that this accounted for the lack of the gender effect on construed similarity among objects belonging to the same gender class. Our results showed that, despite irregular mappings between grammatical gender classes and the semantic categories of biological sex, German speakers are sensitive to the grammar–concept link and tend to overgeneralize the grammatical categories of masculine and feminine when drawing inferences about properties related to the biological sex of animals.

However, our results are consistent with Vigliocco and colleagues' work in that the influence of grammatical gender does not penetrate into inferences about objects without biological sex (Vigliocco et al., 2005) and that gender effects only occurred when nouns were marked for gender (e.g., when used with determiners) (Vigliocco et al., 2004, 2005).

5.1. Can German speakers' poor performance in Experiment 1 be explained by unnaturalness of the task?

Studies reporting cross-linguistic differences that are attributed to linguistic categorization systems are often criticized in that the effect may have arisen because the speakers were artificially led to use the target grammatical category as a cue in performing a task because no other cue was available (e.g., Bender et al., 2011). However, this criticism does not apply to our study. In our task, all the German-speaking participants must have been aware that the choice of gender-marking determiner is independent of the biological sex of the animal being talked about. This was especially apparent in the sex-specified animal condition, where the target animal's sex was explicitly specified by the adjective. The three experimental conditions (i.e., generic animal condition, sex-specified animal condition, and artifact condition) and two control conditions were presented in a completely random and mixed fashion throughout the experiment. Thus, German participants must have been constantly reminded that determiners should *not* be used for making decisions. Furthermore, they went through practice trials, in which they were explicitly alerted when they made an incorrect response. Thus, the experiment was designed so that German speakers would form a conscious strategy *not* to use the grammatical gender as a cue. This makes the result of the generic animal condition—that German speakers erroneously reasoned that the premise about a sex-specific property was applicable to category members of the target animal in general when the grammatical gender agreed with the sex in the premise—all the more striking.

Non-speakers of German may wonder whether the determiner phrase (e.g., *der Elephant*) can really be read in the generic meaning, and thus the poor performance by German speakers could have been due to confusion. However, languages differ in terms of the constructions used to express generic meaning; the constructions used in Experiments 1 (*definite article + singular noun*) and 2 (*no article + plural noun*) are both very

commonly used in German to refer to basic-level kinds (Behrens, 2005). One may also be concerned that the poor performance in German speakers in the generic animal condition in Experiment 1 was due to the unnaturalness of the task with logically indeterminate questions. However, this account is also unlikely, given that German speakers performed well in this logically indeterminable case when they did not see the gender-marking articles in Experiment 2.

5.2. Implications for research on the relation between language and thought

In the literature, researchers have often debated whether a cross-linguistic difference, if found, should be interpreted as evidence for a strong or a weak version of linguistic relativity, or this effect should be interpreted as indication of "thinking for speaking" instead of linguistic relativity (e.g., Boroditsky et al., 2003; Gao & Malt, 2009; Gilbert, Regier, Kay, & Ivry, 2006; Huettig, Chen, Bowerman, & Majid, 2010; Lucy, 1992; Vigliocco et al., 2005; Winawer et al., 2007). In this tradition, our results are probably best interpreted as supporting a weak version of linguistic relativity. In light of the traditional criteria for distinguishing the weak versus strong version of linguistic relativity, it is difficult to interpret our findings as support for a strong version for two reasons: First, we did not find that grammatical gender categories penetrate into the representation of objects. Second, the effect found in Experiment 1 (nouns presented with the gender article) was limited to objects with biological sex (i.e., animals) but was not found with objects without it (i.e., artifacts).

Some researchers may further argue that the gender effect here is only support for "thinking for speaking" (Slobin, 1996) rather than for the influence of language on thought per se because the effect was obtained in a task using language (cf. Vigliocco et al., 2005). Nevertheless, the influence of grammatical gender we found in this research is important, considering the overall role of language on thought. For speakers of languages with grammatical gender, explicit gender marking by articles or pronouns is the *norm* rather than the exception in everyday discourse. If speakers of these languages unconsciously link the grammatical gender of an animal's name to its biological sex (even though the two are orthogonal) and apply this link when making inferences about sex-specific properties of animals (even though they consciously know that they should not do so) from childhood (Saalbach et al., 2012), then we should conclude that grammatical gender has non-trivial cognitive consequences for these speakers, which goes beyond mere attention to language-specific patterns for speaking.

In fact, we may argue that our approach is more meaningful and ecologically valid than the traditional approach asking whether grammatical gender affects object representation in purely non-linguistic contexts. Informal inferences about animals' sex-related properties take place frequently in everyday contexts, and they are made using language, after all. Framing the investigation of the relation between language and thought only around the traditional approach to linguistic relativity—that is, focusing only on *non-linguistic* cognitive differences between speakers of different languages—may cloud our understanding of the role of language on thought, or even on the nature of human cognition (see Bender et al., 2011; for a similar view). Recent neurological studies in

color perception suggest that language is automatically accessed and creates cross-linguistic differences in the brain activity even in seemingly "pure" perceptual tasks (Tan et al., 2008; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009). Furthermore, researchers have demonstrated that the effect color category perception is obtained when colors are seen in the right visual field but not in the left (Gilbert et al., 2006; see also Roberson, Pak, & Hanley, 2008). In this light, perhaps it is not at all simple—perhaps not possible—to define what "purely non-linguistic" thought would be.

The finding from this research is striking in that even small grammatical particles such as gender articles led German speakers to false reasoning about 50% of the time (German speakers in the *generic animal condition* in Experiment 1) while they were able to perform almost perfectly when they did not see the articles in Experiment 2. This difference within German speakers across the two linguistic contexts—whether nouns are presented with or without the gender-marking articles—is as striking as the cross-linguistic difference across the speakers of German and Japanese.

It has been repeatedly demonstrated that language could greatly affect speakers' online cognitive processes, including attention, memory, construal of entities, reasoning, and decision making. For example, in a classic study on eyewitness memory, English-speaking participants were shown a video of an auto accident and were later asked whether they witnessed a broken headlight of the car (Loftus, 1979). Even though the video did not show a broken light of the car, the participants were more likely to say "yes" when the experimenter asked, "Did you see *the* broken light?" than when she said, "Did you see *a* broken light." As reviewed, the framing effect demonstrated by Tversky and Kahneman (1981) is an excellent example of language affecting people's reasoning and decision making. Perhaps the results of the present research—that German speakers' deductive reasoning about sex-specific properties of animals is largely biased by gender-marking articles—could be seen as an extension of this context effect on reasoning.

To conclude, simply demonstrating a cross-linguistic difference (or lack of it) due to a given linguistic category and arguing for (or against) the Whorfian hypothesis is no longer satisfactory to fully understand the relation between language and thought (Imai & Masuda, 2013; Imai & Mazuka, 2007; Imai & Saalbach, 2010; Saalbach & Imai, 2007, 2012). The present research not only showed that grammatical gender affects German speakers' reasoning about sex-specific properties but also specified how the influence arises and the scope of the influence. We maintain that, in investigating the relation between language and thought, it is important to finely specify the kind of role language plays in representation and cognitive processes in meaningful cognitive contexts regardless of whether they involve conscious access to language, rather than to just demonstrate the existence of a cross-linguistic difference in purely non-linguistic tasks. Lastly, however, the present research should be extended to other languages, especially those with only two gender classes (e.g., Spanish, Italian, or French) and those with more than four classes (e.g., Zande). Given a weak, statistically non-significant tendency shown by German speakers in Experiment 2, it is possible that speakers of languages with two grammatical gender classes may show the influence of grammatical gender even when target nouns were presented without gender-marking articles. In any case, it is important for

future research to broadly examine the generalizability of the effect found in this research and how the effect interacts with language-specific semantic and structural properties of a given grammatical gender system.

Acknowledgments

This research was supported by MEXT KAKENHI (15300088, 22243043, 23120003) for Imai and by DAAD Post-Doc Fellowship to Saalbach. We thank Yuichiro Anzai and Mitsuhiro Okada for insightful discussions and comments for an earlier version of this manuscript.

References

Aikhenvald, A. (2000). *Classifiers: A typology of noun categorization devices*. New York: Oxford University Press.

Behrens, L. (2005). Genericity from a cross-linguistic perspective. *Linguistics*, *43*(2), 275–344.

Bender, A., Beller, S., & Klauer, K. C. (2011). Grammatical gender in German: A case for linguistic relativity? *The Quarterly Journal of Experimental Psychology*, *64*(9), 1821–1835.

Boroditsky, L., & Schmidt, L. A. (2000). Sex, syntax, and semantics. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 42–47).

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 61–80). Cambridge, MA: MIT Press.

Corbett, G. G., & Fraser, M. N. (2000). Gender assignment: A typology and a model. In G. Senft (Ed.), *Systems of nominal classification* (pp. 293–325). Cambridge, MA: Cambridge University Press.

Flaherty, M. (2001). How a language gender system creeps into perception. *Journal of Cross-Cultural Psychology*, *32*(1), 18–31.

Fox, A. (1990). *The structure of German*. New York: Oxford University Press.

Gao, Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes*, *24*, 1124–1179.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Science*, *103*, 489–494.

Huettig, F., Chen, J., Bowerman, M., & Majid, A. (2010). Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye-gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture*, *10*, 39–58.

Imai, M., & Masuda, T. (2013). The role of language and culture in universality and diversity of human concepts. In M. Gelfand, C. Y. Chiu & Y. Hong (Eds.), *Advances in culture and psychology, vol. 3* (pp. 1–61). New York: University Oxford Press.

Imai, M., & Mazuka, R. (2007). Revisiting language universals and linguistic relativity: Language-relative construal of individuation constrained by universal ontology. *Cognitive Science*, *31*, 385–414.

Imai, M., & Saalbach, H. (2010). Categories in mind and categories in language: Do classifier categories influence conceptual structures? In B. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 138–164). New York: Oxford University Press.

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research*, *22*(5), 519–534.

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*, 843–858.

Loftus, E. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Lucy, J. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, MA: Cambridge University Press.

Mizuguchi, S. (2004). *Individuation in numeral classifier languages: A case of Japanese classifiers and plurals*. Tokyo, Japan: Shohakusha.

Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 928–933).

Ramos, S., & Roberson, D. (2010). What constrains grammatical gender effects on semantic judgements? Evidence from Portugues. *European Journal of Cognitive Psychology*, *23*, 102–111.

Roberson, D., Pak, H. S., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, *107*, 752–762.

Saalbach, H., & Imai, M. (2007). The scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, *136*, 485–501.

Saalbach, H., & Imai, M. (2012). The relation between linguistic categories and cognition: The case of numeral classifiers. *Language and Cognitive Processes*, *27*, 381–428.

Saalbach, H., Imai, M., & Schalk, L. (2012). Grammatical gender and inferences about biological properties in German-speaking children. *Cognitive Science*, *36*, 1251–1267.

Sera, M. D., Berge, C. A. H., & del Castillo Pintado, J. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development*, *9*, 261–292.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodriguez, W., & Dubois, D. P. (2002). When language affects cognition and when does it not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General*, *131*, 377–397.

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking". In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, MA: Cambridge University Press.

Tan, L. H., Chan, A. H. D., Kay, P., Khong, P. L., Yip, L. K. C., & Luke, K. K. (2008). Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences of the USA*, *105*, 4004–4009.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the USA*, *106*, 4567–4570.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language*, *45*, 368–390.

Vigliocco, G., Vinson, D., Indefrey, P., Levelt, W., & Hellwig, F. (2004). Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *30*, 483–497.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implication for language learning and language use. *Journal of Experimental Psychology: General*, *134*, 501–520.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the USA*, *104*, 7780–7785.

Zubin, D., & Köpcke, K.-M. (1986). Gender and folk taxonomy: The indexical relation between grammatical gender and lexical categorization. In C. Craig (Ed.), *Noun classes and categorization* (pp. 139–180). Philadelphia: Benjamins.

Appendix:
Examples of the items used for Japanese participants in the five different deductive inference conditions of Experiments 1 and 2

Condition	Premise	Sex–Gender Consistent	Sex–Gender Inconsistent
Generic animal	すべてのメス動物は、そしてメス動物のみが、X を持っている. Subete (all) -no (GEN.) **mesu** (female) doubutsu (animal) -wa (TOP.) soshite (and) mesu (female) doubutsu (animal) nomi (only) ga (SUBJ.) X-o (ACC.) motteiru (have). (All and only female animals have property X inside.)	キリン Kirin (giraffe)	ゾウ Zou (elephant)
	すべてのオス動物は、そしてオス動物のみが、X を持っている. Subete (all) -no (GEN.) **osu** (male) doubutsu (animal) -wa (TOP.) soshite (and) osu (male) doubutsu (animal) nomi (only) ga (SUBJ.) X-o (ACC.) motteiru (have). (All and only male animals have property X inside.)	ゾウ Zou	キリン Kirin
Generic animal control	すべての動物は、そして動物のみが、X を持っている. Subete (all) -no (GEN.) doubutsu (animal) -wa (TOP.) soshite (and) doubutsu (animal) nomi (only) ga (SUBJ.) X –o (ACC.) motteiru (have). (All and only animals have property X inside.)	ゾウ Zou キリン Kirin	

(continued)

Table (continued)

Condition	Premise	Sex–Gender Consistent	Sex–Gender Inconsistent
Sex-specified animal	すべてのメス動物は、そしてメス動物のみが、X を持っている.	メスのネズミ(female mouse) **Mesu**-no nezumi	オスのネズミ (male mouse) **Osu**-no nezumi
	すべてのオス動物は、そしてオス動物のみが、X を持っている.	オスのハムスター(male hammster) **Osu**-no hamusutah	メスのハムスター (female hammster) **Mesu**-no hamusutah
Sex-specified animal control	すべての動物は、そして動物のみが、X を持っている.	メスのネズミ Mesu-no nezumi	オスのネズミ Osu-no nezumi
Artifact	すべてのメス動物は、そしてメス動物のみが、X を持っている.	ノコギリ(saw) Nokogiri	ハンマー(hammer) Hammah
	すべてのオス動物は、そしてオス動物のみが、X を持っている.	ハンマー Hammah	ノコギリ Nokogiri