Title
Grammatical Gender and Meaning

Permalink
https://escholarship.org/uc/item/88x2t3h2

Journal

ISSN
1069-7977

Authors
Vigliocco, Gabriella
Vinson, David P.
Paganelli, Federica

Publication Date
2004

Peer reviewed
Grammatical Gender and Meaning

Gabriella Vigliocco (g.Vigliocco@ucl.ac.uk)
Department of Psychology, University College London, Gower Street
London WC1H 6BT, England

David P. Vinson (d.vinson@ucl.ac.uk)
Department of Psychology, University College London, Gower Street
London WC1H 6BT, England

Federica Paganelli (f.paganell@ucl.ac.uk)
Department of Psychology, University College London, Gower Street
London WC1H 6BT, England

Abstract

Two experiments assessed whether grammatical gender of Italian nouns referring to animals and tools affects conceptual representations of the corresponding objects, comparing results from Italian and English. In the first experiment, we elicited semantic substitution errors (e.g., saying “hammer” when “axe” is intended), finding language-specific gender effects (more errors in Italian than English for words sharing gender) for words referring to animals but not for words referring to tools. In the second experiment, words sharing gender were judged as more similar in meaning by Italian speakers than English speakers, again only for animals and not for tools. Moreover, no such gender effect was observed for pictures of the same animals.

Introduction

As Roman Jakobson (1959) put it: "Languages differ essentially in what they must convey and not in what they may convey" (p.236). That is, languages differ in which conceptual or formal properties must be realized in sentential form. For example, in English the word "friend" does not indicate the sex of the friend, while in Italian the corresponding word is differentially inflected for a man (“amico”) or a woman (“amica”). In English, adjectives used as predicates (e.g., "tall" in "The boy is tall" and “The girl is tall”) do not agree in gender with the subject of the sentence, while they must in Italian (e.g., "Il ragazzo e’ alto” or “La ragazza e’ alta”). Such differences in obligatory expression may imply that speakers of different languages pay more or less attention to those dimensions of meaning. For example, Italian speakers may pay more attention to the sex of referents than English speakers. By extension, Italian speakers may tend to think of objects in the world as more male- or female-like on the basis of the words’ grammatical gender (as suggested by the work of, e.g. Boroditsky, Schmidt & Phillips, 2003; Sera, Elieff, Forbes, Burch, Rodriguez, & Dubois, 2002). But how strong and pervasive can these effects be?

Here we present experiments investigating the conditions under which effects of a language-specific property (grammatical gender of Italian nouns) are present, contrasting performance by Italian and English speakers on translation-equivalent nouns. Grammatical gender allows a conservative test of language-specific effects on cognition because it is largely arbitrarily linked to meaning (although see Foundalis, 2002). How could grammatical gender affect conceptual representations for objects? Effects of grammatical gender could arise as a consequence of general language-learning mechanisms based on similarity. According to this hypothesis (to which we will refer as “Similarity and Gender”), words that are similar to each other on any linguistic dimension (including but not limited to grammatical gender) may become more semantically similar as a consequence of the fact that words of the same syntactic class (e.g., same gender, same grammatical class, etc.) appear in the same syntactic contexts. For example, in languages with grammatical gender, nouns are used in sentences along with gender-marked determiners and adjectives, whether the nouns refer to sexuated entities or not. Sensitivity to shared sentence context could allow children to bootstrap properties of similarity in meaning from the syntactic contexts in which the words occur during language acquisition (Landauer & Dumais, 1997). This hypothesis does not require any explicit associations between grammatical gender and sex of human referents; instead it predicts that any effects of grammatical gender on semantic representations should be found in any gendered language (no matter how many gender classes are in the language), and that they should be found for all words (whether the referents are sexuated or not).

However, mechanisms mediating such effects may be more specific and limited. According to this other hypothesis (to which we will refer as “Sex and Gender”), effects of grammatical gender could arise because children would treat all grammatical categories as revealing specific semantic properties (Boroditsky et al., 2003). In the case of grammatical gender, these effects would require linking the grammatical gender of nouns referring to humans to the sex of referents. Across languages there is a core correspondence between grammatical marking of gender and biological sex (Corbett, 1991), although the consistency of this mapping differs across languages. According to this view, children learning a gendered language would first notice the core correspondence between the gender of nouns and male/female semantic properties of human referents (and some animals). They would then generalize this correspondence to other nouns for which there is no clear conceptual foundation of gender,
assigning male or female features to referents in agreement with the grammatical gender of the corresponding words. Thus, words of the same gender would be more similar among themselves than words of different gender because they share male or female-like properties. Such a mechanism could be strongest for languages with the greatest degree of correspondence between the gender of nouns referring to humans and the sex of the referents. This is the case in Romance languages which have only two genders and few exceptions to the consistent mapping between the gender of nouns referring to humans and sex of the referents. It could be weaker (if present at all) in languages with multiple genders and/or in which nouns referring to humans fall into more than two classes. Moreover, any effect of gender could be stronger for words referring to sexuated entities (e.g., animals) than for words referring to objects and abstract concepts, because semantic properties of sex are less relevant in these latter domains. Most of the studies investigating language-specific effects of grammatical gender (e.g., Boroditsky, et al., 2003; Sera, et al., 2002) have tested this hypothesis, either implicitly or explicitly.

In the experiments below we tested some predictions stemming from these views, considering grammatical gender of Italian nouns. As in other Romance languages, all nouns in Italian are marked for gender, either masculine or feminine. For nouns referring to humans and some animals, the gender depends on the sex of the referent (e.g., “ragazzo/ragazza” [boy/girl]; “leone/leonessa” [lion/lioness]), while for other animals gender does not depend upon the sex of the referent (e.g., “lupo” refers to both male and female wolves, although it is possible to mark the gender in some cases). For words referring to objects and abstract entities, instead, there are no such clear semantic correlates (with certain exceptions not addressed here). We investigate two semantic fields, animals and tools, to test the hypotheses outlined above. Both predict language-specific effects of grammatical gender on meaning, such that word pairs sharing Italian gender will show greater semantic similarity effects than the same word pairs in English translation. The two hypotheses make different predictions, however, suggesting that these effects may have different breadth. The Sex and Gender hypothesis predicts that greater language-specific gender effects should be observed for animals than for tools (as animals are sexuated entities), while the Similarity and Gender hypothesis predicts no category difference.

Experiment 1

Here we assessed whether grammatical gender affects online linguistic tasks such as picture naming. We begin with a linguistic task, as finding language-specific gender effects is not only evidence for a “thinking for speaking” view of language-specific effects on cognition (Slobin, 1996), but is also a pre-requisite to testing for broader language specificity in tasks less tightly tied to linguistic encoding. We focus upon semantic substitution errors (i.e., instead of producing a target word, speakers produce another word related in meaning, e.g., saying “hammer” when “saw” is intended) in picture naming. In previous work we have introduced a continuous picture-naming paradigm to elicit such errors (Vigliocco, Vinson, Lewis & Garrett, in press). Here we investigate whether semantic substitution errors in Italian tend to preserve the gender of the target word (i.e., masculine nouns are more likely to substitute for other masculine nouns, and feminine nouns for other feminine nouns). It is generally agreed upon in the language production literature that semantic substitutions arise during the process of retrieving the lexical entry corresponding to a concept, thus tapping into the interface between linguistic and conceptual knowledge (e.g., Garrett, 1984; Levelt, 1989). Moreover, these errors are sensitive to fine-grained semantic similarity: the likelihood of errors increases with greater semantic similarity (Vigliocco et al., in press). Thus, other factors being equal, if grammatical gender has a semantic effect, it should increase the likelihood that words of the same gender substitute for each other in a language such as Italian. However, other factors may not necessarily be equal, particularly non-language-specific factors such as semantic similarity not related to gender in a language-specific manner, or visual similarity among pictorial referents. In order to provide the tightest controls, we selected English as a baseline comparison language, using the same items (translation equivalent words) and the same tasks. This allows us to test for language-specific effects of Italian gender while avoiding concerns related to general semantic or visual similarity among the items used in our experiments. We investigate whether Italian errors tend to preserve gender above the English baseline level (based upon assigning Italian gender to English translations). It should also be noted that substitution errors are sensitive to phonological similarity between target and intruder (Dell & Reich, 1981), and Italian gender does have strong and reliable phonological correlates. To minimize the possibility that any observed language-specific effects of gender are due to phonological similarity, we also conducted analyses in which we excluded all errors with substantial phonological similarity to the target word.

Method

Participants

Participants were 27 native speakers of Italian from the London community and 20 English speakers from the UCL subject pool. The Italian participants had only rudimentary knowledge of English, and none of the English speakers reported moderate or better competence in any Romance language.

Materials

We selected 27 black and white line drawings of animals, avoiding those animals for which the gender of the noun strictly depends upon the sex of the referent. We further selected 50 black and white line drawings of tools. Most pictures came from Snodgrass and Vanderwart (1980), with additional ones prepared for the experiment. Name agreement was established for each of our participants during the experimental session (see also Vigliocco et al., in press); in general there was very strong name agreement in both languages, further ensuring that the words are suitable translation equivalents. Because in previous work we have established that substitution errors in this paradigm do not cross
semantic fields, we used a blocked presentation design, analyzing the data for the animals and the tools separately. We presented 77 blocks of 10 pictures each to every participant. Each block contained only animal or tool pictures (presented in random order within the block), and each picture appeared 10 times in the course of the experiment.

**Procedure**
The experiment began with a name agreement phase in which each picture was presented and participants were asked to name them. This phase allowed us to ensure name agreement across participants and also to identify specific naming preferences by individual participants (which might otherwise have been considered errors). Next, a practice series of blocks were presented in which the speed of presentation of each picture was adjusted for each participant (between 600ms and 1100ms) in order to render the task difficult but manageable for each speaker. After the training, the experiment started. Participants were told that their task was to name each picture as it appeared on a computer screen as quickly as possible.

**Scoring Criteria**
Participants’ responses were transcribed and scored in the following categories: Correct responses: participants uttered the correct target word. Different label: participants used a different word than our intended target (e.g. "stag" for "deer"), but this different label was consistently used by that participant and did not refer to another item in the experiment. Lexical errors: participants produced a word that differed from the target and that did not qualify as a "different label". Lexical errors were further classified as "out of set" (intruding words that are not among the experimental items) and "within set" (those items from within the present response set). Because of the repeated presentation of a limited set of pictures to be named, most lexical errors tended to be other items from the response set. Because of this, and to minimize the possibility of linguistic variability beyond this particular set of items, analyses were performed only upon within-set lexical errors items. Miscellanea: other responses not included above, such as dysfluencies, incomplete utterances, inaudible responses, omissions and self-corrections. Table 1 reports a breakdown of the proportions of responses in the different scoring categories.

<table>
<thead>
<tr>
<th>Response type</th>
<th>IT: A</th>
<th>EN: A</th>
<th>IT: T</th>
<th>EN: T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct &amp; Different Label</td>
<td>.876</td>
<td>.935</td>
<td>.948</td>
<td>.910</td>
</tr>
<tr>
<td>Lexical errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within-set</td>
<td>.021</td>
<td>.024</td>
<td>.013</td>
<td>.018</td>
</tr>
<tr>
<td>Out of set</td>
<td>.004</td>
<td>.001</td>
<td>.002</td>
<td>.002</td>
</tr>
<tr>
<td>Miscellanea</td>
<td>.100</td>
<td>.04</td>
<td>.133</td>
<td>.062</td>
</tr>
</tbody>
</table>

**Results and Discussion**
All analyses were carried out on within-set lexical errors. First, we eliminated all those items for which the average correct performance was not above 75% in both languages or for which the average correct performance differed more than 15% across the two languages in order to exclude additional cross-cultural differences. For each semantic field (animals and tools) we carried out two 2x2 ANOVAs. In all ANOVAs, proportion of errors was the dependent variable with target-error pair as a random factor. Independent variables were language (Italian, English) and Italian gender (shared between target and intruder or not shared). English words were assigned Italian gender for the purpose of the analysis. The first ANOVA was carried out on the within-set errors remaining after we excluded the cases discussed above (for animals: 103 errors in Italian and 73 errors in English; for tools: 90 in Italian, and 117 in English). For animals, this analysis showed a significant interaction between language and Italian gender, such that errors sharing gender with the target were more common in Italian (68%) than in English (41%); $F(1,63) = 8.03, p = .006$. Neither main effect was significant ($F < 1$). The results of the analysis for tools were similar; only the interaction between language and gender was significant: gender preservation was greater in Italian (61%) than in English (36%); $F(1,79) = 4.6, p = .04$; main effect $F_s < 1$).

In the second analysis we excluded all errors for which the target and the intruder shared phonological similarity. Phonological similarity between target and intruder was assessed as in Vigliocco et al. (in press). In this second analysis, only errors for which either of two measures of phonological overlap did not exceed the average + one standard deviation of that measure (in either language) were considered (for animals, leaving 64 errors in Italian and 42 in English; for tools, leaving 39 errors in Italian and 42 in English). This analysis for animals also showed a significant interaction between language and gender; such that even among target-intruder pairs with low phonological similarity, Italian target-intruder pairs tended to share gender (77%) more often than English pairs (43%) (interaction $F(1,37) = 5.88, p = .020$; main effect $F_s < 1$). However, this interaction was not significant in the analysis for tools (all $F_s < 1$). Thus, for the tools, the language x gender effect observed in the complete set of errors may just be a consequence of greater phonological similarity in Italian for words sharing the same gender.

To summarize the results of this experiment, we found language-specific effects of grammatical gender; gender affects the likelihood of producing a lexical error for Italian speakers, compared to the errors produced by English speakers naming exactly the same pictures. This language-specific effect of grammatical gender, however, survives only for words referring to animals once phonological similarity is taken into account. This result suggests that language-specific effects are constrained even in a linguistic encoding task such as picture naming.

**Experiment 2**

The results of the error induction task in Experiment 1 show that language-specific effects of grammatical gender can be observed in an on-line task requiring lexical retrieval. In this
Experiment 2a: Animals

Method
Participants
Participants were 24 native speakers of Italian from the London community, and 24 native English speakers from the University College London participant pool. The Italian participants had only rudimentary knowledge of English, and none of the English speakers reported moderate or better competence in any Romance language.

Materials
Words (and corresponding pictures) referring to 20 animals were selected for the experiment. Words were translation equivalents in Italian and English. The words and the pictures used were a subset of those used in Experiment 1. Triads for the Italian and English conditions were created by first assembling all possible three-word combinations of the 20 items in the experimental set (for a total of 1,140 triads). The order of words in each triad was randomized; then the order of triads was randomized across participants. A separate set of picture triads were then created by replacing each word with its corresponding picture (this set was identical for Italian and English participants). Twelve participants from each language were assigned to the word condition and twelve to the picture condition. In each modality (word or picture) and language (Italian or English) condition, the 1,140 triads were divided into three lists, each containing 380 triads of words or pictures.

Procedure
All participants were told that the experiment concerned participants’ judgments of meaning similarity among groups of words (or pictures), and that their task was to choose the two words (or pictures) out of the three which were more similar in meaning and to delete the odd one out. Instructions emphasized that the decision was to be made on the basis of meaning and not other types of similarity between the items (e.g., phonological similarity among the words or visual similarity among the pictures). After completing the task, participants were asked to describe the strategies they may have used to perform the task, to list the easiest and most difficult decisions, etc. For the purpose of the present study, the most important aspect of these questions was whether any Italian participants mentioned grammatical gender as an overt basis for making their decisions.

Design and Analysis
The dependent variable was similarity ratio: the number of times that a given pair of words/pictures was selected as “similar”, divided by the number of triads in which those two items appeared in the experiment. Four participants completed each list of 380 items; thus each triad (either words or pictures) was judged by four different speakers of a language. Results were analyzed using a three-way mixed ANOVA with item pairs as a random factor. Independent variables were language (English or Italian, manipulated within item pairs), modality (words or pictures, manipulated within item pairs) and Italian gender (same Italian gender; different Italian gender, manipulated between item pairs). As in Experiment 1 this latter factor refers to the gender of the Italian translation or label.

Results and Discussion
No Italian participant indicated that they used grammatical gender in their similarity judgments in the post-experimental questionnaire. Table 2 reports the average similarity proportions for items of same vs. different Italian gender as a functions of language and modality.
Results were analyzed using a three-way mixed ANOVA investigating the effect of language, modality and Italian gender on similarity proportions. Only the three-way interaction (Language x Modality x Gender) reached significance \((F(1,188)=6.539, p=.013)\); no other main effects and interactions were significant (all \(F_s < 1\)). Analysis of simple interaction effects within each modality revealed that the similarity proportion for same-gender items was relatively higher for Italian word judgments than for English word judgments (with the corresponding difference in the opposite direction for words differing in Italian gender), while there was no such language difference for picture judgments by speakers of either language. Thus, in this experiment we found language-specific effects of grammatical gender for words referring to animals, but not for pictures referring to the same animals; Italian speakers’ judgments of meaning similarity seem to be affected by shared grammatical gender. Experiment 2b assessed whether such a gender effect is present for tools.

### Experiment 2b: Tools

#### Method

**Participants**
Participants were 48 native speakers of Italian from the London community, and 48 native English speakers from the University College London participant pool. The Italian participants had only rudimentary knowledge of English, none of the English speakers reported moderate or better competence in any Romance language. None of them had participated in Experiment 2a.

**Materials**
Words (and corresponding pictures) referring to 24 tools were selected for the experiment. Words were translation equivalents in Italian and English. The words and the pictures used were a subset of those used in Experiment 1. Word and picture triads for the Italian and English conditions were created as in Experiment 2a. In this experiment all possible three-word combinations of the 24 items in the experimental set yielded a total of 2,024 triads. These triads were divided into six lists each containing 337 or 338 words or pictures. This experiment was otherwise the same as in Experiment 2a.

#### Results and Discussion

No Italian participant indicated that they used grammatical gender in their similarity judgments in the post-experimental questionnaire. Table 3 reports the average similarity proportions for items of same vs. different (Italian gender as a functions of language and modality.

<table>
<thead>
<tr>
<th>Language</th>
<th>Modality</th>
<th>Grammatical Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Same</td>
</tr>
<tr>
<td>Italian</td>
<td>Words</td>
<td>.318 [.018]</td>
</tr>
<tr>
<td></td>
<td>Pictures</td>
<td>.314 [.020]</td>
</tr>
<tr>
<td>English</td>
<td>Words</td>
<td>.316 [.017]</td>
</tr>
<tr>
<td></td>
<td>Pictures</td>
<td>.308 [.019]</td>
</tr>
</tbody>
</table>

Results were analyzed using a three-way mixed ANOVA investigating the effect of language, modality and Italian gender on similarity proportions. No main effects or interactions were significant (\(F_s < 1\)), with the exception of the main effect of gender which was marginal \((F(1,274) = 2.49, p = .115)\). This indicates an underlying tendency for tools sharing Italian gender to be more similar than items with different Italian gender. Because no interaction between language and Italian gender was observed, this main effect cannot reflect language-specificity. Thus, whereas grammatical gender affected Italian speakers’ judgments of meaning similarity for words referring to land animals, no such effect was observed for either words or pictures referring to tools.

#### General Discussion

In the experiments reported above, we explored language-specific effects of Italian grammatical gender on semantic representations for the corresponding objects. These experiments combined on-line and off-line methodologies, assessing effects for two different semantic fields: one for which associations between grammatical gender and sex can be plausibly built (animals) and one for which they cannot (tools). Moreover, we further explored gender effects within the same task, using both words and pictures as stimuli, in order to establish the generalizability of any effect.

We found that language-specific effects of Italian grammatical gender are present, but highly limited. They are limited to a semantic field (animals) in which entities have biological gender, and in which the gender of some nouns can depend on the sex of the referent. However, this effect does not extend to a field for which distinctions in grammatical gender have no conceptual foundation (tools). These effects are further limited to tasks that recruit linguistic knowledge (picture naming or similarity judgments for words, but not similarity judgments for pictures).
Our results suggest a far more limited role of grammatical gender on semantic representations than it has been suggested in previous studies. For example, Sera et al (2002) showed that grammatical gender of Spanish and French nouns influenced speakers’ assignment of a male or female voice to inanimate (and animate) objects, regardless whether the task was carried out using words or pictures as stimuli. However, speakers might have used gender in a strategic manner in this task. Boroditsky, et al. (2003) report studies suggesting that grammatical gender may have implicit effects. However, these studies are reported without enough methodological detail to address possible reasons for the different results.

Our results suggest that these gender effects are linked to assigning male- or female-like semantic properties to referents in agreement with the gender of the nouns. They provide evidence for a very constrained version of the Sex and Gender hypothesis described in the introduction. According to this view, language-specific grammatical gender effects should be stronger for semantic fields in which there is a conceptual motivation for establishing a link between gender of words and sex of the referent (such as animals), than for fields for which there is not such a clear conceptual motivation (such as tools). This hypothesis also predicts that language-specific effects of gender on meaning should be stronger for languages such as Italian that have strong transparent correspondence between the gender of nouns and sex of referents. No effect of grammatical gender was investigated gender effects on semantic substitution errors for animals (at least when speakers were asked to produce bare nouns). These different lines of investigation converge in suggesting that language-specific effects of gender do not arise as a consequence of a general mechanism sensitive to similarity. Finally, the difference we observe in Experiment 2 between word and picture stimuli suggests that grammatical gender of Italian nouns (referring to animals) affects “thinking for speaking”, but does not affect conceptual representations when language is not required (Slobin, 1996).

Acknowledgments

The work reported here has been supported by a Human Frontier Science Program grant (HFSP148/2000) and an Economic and Social Research Council grant (RES000230038) to Gabriella Vigliocco. We thank Jo Ariciuli and Belen Lopez-Cutrin for help in preparing this manuscript.

References


