



Experimental evidence for the interpretation of definite plural articles as markers of genericity – How Italian can help

MICHELA REDOLFI 

SERGIO MIGUEL PEREIRA SOARES 

ANNA CZYPIONKA 

TANJA KUPISCH 

**Author affiliations can be found in the back matter of this article*

RESEARCH

]u[ubiquity press

ABSTRACT

In the Romance languages, definite plural articles (e.g., *le rane* ‘the frogs’) are generally ambiguous between a generic and a specific interpretation, and speakers must reconstruct the intended interpretation through the linguistic or extra-linguistic context. Following the “polar bear” paradigm implemented in Czypionka & Kupisch (2019)’s investigation on German, the goal of the present study is to check the suitability of their test on article semantics, by establishing to what extent native speakers of Italian interpret ambiguous definite plural DPs as generic or specific in the presence of a nonlinguistic picture context. We present judgment and reaction time data monitoring the preferred reading of sentences introduced by different kinds of noun phrases (e.g., *Le rane/Queste rane/Le rane di solito sono verdi/gialle* ‘The/These/Usually frogs are green/yellow’), while looking at pictures showing prototypical or non-prototypical properties (e.g., green vs. yellow frogs). Our results show that both possible interpretations of definite plural articles are routinely considered in Italian, despite the presence of a picture with specific referents, validating the “polar bear” paradigm as a suitable test of article semantics.

CORRESPONDING AUTHOR:

Michela Redolfi

University of Verona, Italy

michela.redolfi@univr.it

KEYWORDS:

Definiteness; Ambiguity;
Determiners; Romance
languages; Germanic
languages

TO CITE THIS ARTICLE:

Redolfi, Michela, Sergio Miguel Pereira Soares, Anna Czypionka and Tanja Kupisch. 2021. Experimental evidence for the interpretation of definite plural articles as markers of genericity – How Italian can help. *Glossa: a journal of general linguistics* 6(1): 16. 1–12. DOI: <https://doi.org/10.5334/gjgl.1165>

This paper is concerned with sentences such as (1), which make a generalizing statement (here, being yellow) about members of a particular class (here, bananas). These so-called generic statements illustrate an interesting contrast between Italian (and most other Romance languages) and German (and most other Germanic languages), as they obligatorily require a definite article only in the former type of language.

- (1) a. *Italian*
Le banane sono gialle.
ART.PL banana.PL are yellow
'Bananas are yellow.'
- b. *German*
Bananen sind gelb.
banana.PL are yellow
'Bananas are yellow.'

However, there are claims in the theoretical literature that German allows definite article use with generic plural subjects too, like in (2) (e.g., Brugger 1993; Longobardi 1994; Krifka et al. 1995; Dayal 2004; Oosterhof 2008).

- (2) (Die) Pandabären sind vom Aussterben bedroht (Krifka et al. 1995)
the panda.bears COP.3SG of.the extinction threatened
'Pandas are facing extinction.'

If correct, this implies that German definite articles are ambiguous between specific and generic readings, just like Romance ones, but unlike articles in other Germanic languages.

Until recently, the aforementioned claims about German had not been backed up by corpus or experimental data, and gathering relevant evidence is hard because generic contexts are rare and text specific. Barton et al. (2015) carried out an acceptability judgment task with speakers from 5 different regions in Germany, confirming that, except in the North, speakers of German were inclined to accept definite articles in sentences like (1b) and (2), though accepting more bare nominals. Their study thus allows for the conclusion that Germans tolerate definite articles with generic plural subjects. However, Barton and colleagues worked with contextualized sentences, as is common in acceptability judgment tasks that tap into semantics, where the function of the context is to create a bias towards a specific or a generic interpretation. However, since superficial processing might lead speakers to paying attention primarily to context, it is not entirely clear to what extent the *definite plural article* itself has adopted the feature [+generic] in addition to [+specific].

Other experimental paradigms have used more intuitive judgments, focusing on the interpretation of articles by testing the preferred readings of definite and bare nominals in Romance and Germanic with a truth value judgment task. Typically, the critical sentence was preceded by a context, as in (3) from Pérez-Leroux et al. (2004), and accompanied by a picture (in this case, two spotted zebras).

- (3) Zippy the zebra and Suzy the zebra are spotted. The giraffe wonders why they look different. Now let me ask you some questions Do the zebras have stripes?

A “yes”-answer would be considered generic because zebras generally have stripes, a “no”-answer specific because the zebras on the pictures are spotted. Pérez-Leroux and colleagues found that 6–7 year old English children overaccepted generic readings (see also Gavarró et al. 2006; Kupisch & Pierantozzi 2010; Montrul & Ionin 2010 for studies using similar methodology). However, one problem with this task, as illustrated in (3), is that the referent is explicitly introduced, and using another lexical DP (i.e., article+noun) after the referent has been introduced may be considered somewhat unnatural compared to using the characters' proper names or a pronoun, thus disfavoring a specific reading.

In a recent paper, Czypionka & Kupisch (2019) have provided fresh experimental evidence to tap into the semantics of definite articles in German and specifically the question whether definite article use has spread to generic contexts. To exclude the possibility that ratings

could be influenced or aided by linguistic context, they designed a picture-based experiment (henceforth “the polar bear” paradigm) without a story or any other context that could have cued one specific reading. For example, native speakers of German were shown a picture with pink polar bears, asking whether the sentence *Die Eisbären sind weiß* ‘The polar bears are white’ is correct or not. Speakers’ acceptability judgments reveal the interpretation of definite articles: If they accept *Die Eisbären sind weiß* despite seeing pink polar bears, they allow for a generic interpretation of the definite article, interpreting the sentence as a statement about polar bears in general rather than those in the given context. Their results showed that speakers interpret definite articles as having specific reference, thus compromising the aforementioned claims in the theoretical literature. However, the participants reacted more slowly to sentences containing definite articles than to control items with demonstratives and bare nominals. This could reflect speakers’ making a choice between two possible readings with definite articles and would support the idea that German definite articles are in principle ambiguous (even if the specific reading is preferred). One potential problem in this study is, however, that the presence of pictures might have biased the speakers towards a specific interpretation, thereby obscuring the potential ambiguity of German definites.

Italian constitutes an ideal control case¹ both for the interpretation of definites and for probing the “polar bear” paradigm, given the problems outlined in the previous paragraph. The semantics of plural definites in Italian are less controversial than in German, as Italian allows no bare nominals in subject position and definite articles are required for both specific and generic reference (see, e.g., Longobardi 1994; Chierchia 1998). Articles in sentences like (1a) are completely ambiguous between these two readings. Thus, if Italian speakers perform differently from the German speakers in Czypionka & Kupisch’s study, it would indicate that German definite articles are not ambiguous in the same way that Romance articles are. If, however, Italian and German speakers perform similarly, it would suggest that the picture context in the “polar bear” paradigm biases comprehenders towards a specific reading, and that the paradigm is unsuitable for testing article semantics.

In what follows, we outline aims and research questions. Section 3 presents the experimental design, sections 4 and 5 the results. We discuss and conclude in section 6.

2 AIMS AND RESEARCH QUESTIONS

Following Czypionka & Kupisch (2019), the goal of the present study is to establish to what extent native Italians interpret ambiguous definite plural DPs as generic or specific in the presence of a nonlinguistic picture context, in order to establish whether the “polar bear” paradigm is an appropriate test of article semantics. To this end, we compared the interpretation of plural subject DPs with different determiner types: definite articles (potentially ambiguous), demonstratives (unambiguously specific), plural DPs with a definite article preceded by the adverb *di solito* (‘in general’, unambiguously generic), and partitive articles (fillers). We measured acceptance rates and reaction times for the acceptability judgments in two experiments, closely following the task used by Czypionka & Kupisch (2019). The experiments were designed to answer the following research questions:

RQ1: To what extent do native Italian speakers interpret definite plural DPs as generic or as specific? Does the ambiguity of definite plural DPs surface even when a nonlinguistic context (i.e., a picture creating a potential specific bias) is used?

RQ2: Do the reaction times for stimuli with definites reflect their ambiguity, i.e., are reaction times longer for definites relative to the other conditions?

RQ3: Is there a qualitative difference to the interpretation of German definite plural DPs (supposedly ambiguous between generic and specific readings, but interpreted as specific in the “polar bear” paradigm)?

¹ We are referring to Standard Italian and regional varieties of Standard Italian. We are not making any claims about Italian dialects. To ensure homogeneity in the Italian data, we tested only speakers in the North East of Italy, where the regional Standard variety patterns with Standard Italian in terms of the property in question.

3 MATERIALS AND METHODS

3.1 LANGUAGE MATERIALS

The experimental design was inspired by Gelman & Raman (2003) and Pérez-Leroux et al. (2004). All stimuli consisted of a visual and an auditory stimulus. The visual stimulus provided the nonlinguistic context for the sentence to be rated (a context is essential for making specific reference plausible), and consisted of an image depicting colored fruit, vegetables and animals. The auditory stimulus was a sentence describing the picture in the visual stimulus with different kinds of statements. Visual and auditory stimuli were combined to form 24 conditions, including 12 filler conditions (see below).

Visual stimuli were based on 20 drawings of objects or animals with prototypical and non-prototypical colors or patterns (see [Figure 1](#)). Each drawing showed four identical objects on a white background. There were three different COLOR-VISUAL conditions: **canonical** (prototypical colors, e.g., green frogs), **noncanonical** (unlikely colors, e.g., yellow frogs), and **mixed** (fillers). Most of the stimuli were the same as in Cypionka & Kupisch (2019), but some had to be adapted to fit Italian as target language.

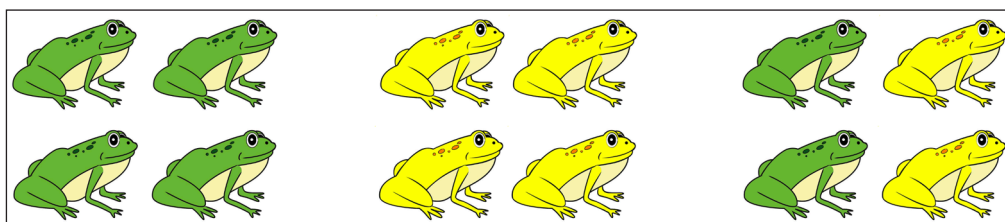


Figure 1 Examples of visual stimuli: *canonical* (left), *noncanonical* (middle), and *mixed* (right) condition.

The auditory stimuli recorded by an Italian native speaker reading with pragmatically neutral intonation (processed in Praat version 6.0.49; Boersma 2016). They totaled 8 conditions ([Table 1](#)), defined by the factors DETERMINATION (determiner type) and COLOR-AUDITORY (canonical or noncanonical color). The factor DETERMINATION had four levels: (1) **definite** (DPs with a definite plural article, e.g., *le rane*, ‘(the)² frogs’), (2) **demonstrative** (DPs with a plural demonstrative, e.g., *queste rane*, ‘these frogs’), (3) **partitive** (DPs with an indefinite plural article, e.g., *delle rane*, ‘some of the frogs’), (4) (explicitly) **generic** (DPs with *di solito*, e.g., *le rane di solito*, ‘(the) frogs usually’). The factor COLOR-AUDITORY had two levels, **canonical** and **noncanonical**. Canonical and noncanonical colors were held constant in the auditory and visual stimuli for each item. In the *definite*, *demonstrative* and *partitive* conditions, statements began with a plural DP (e.g., *Le rane*) followed by the copula (*sono*) and a color or pattern term (e.g., *verdi*). In the *generic* condition, the statements additionally contained *di solito* ‘usually’.

| | COLOR-AUDITORY | |
|----------------------|----------------------------------------------------------------------|------------------------------------------------------------------------|
| DETERMINATION | CANONICAL | NONCANONICAL |
| <i>Definite</i> | <i>Le rane sono verdi.</i> ‘(The) frogs are green.’ | <i>Le rane sono gialle.</i> ‘(The) frogs are yellow.’ |
| <i>Demonstrative</i> | <i>Queste rane sono verdi.</i> ‘These frogs are green.’ | <i>Queste rane sono gialle.</i> ‘These frogs are yellow.’ |
| <i>Generic</i> | <i>Le rane di solito sono verdi.</i> ‘Frogs are generally green.’ | <i>Le rane di solito sono gialle.</i> ‘Frogs are generally yellow.’ |
| <i>Partitive</i> | <i>Delle rane sono verdi.</i> ‘Some of the frogs are green.’ | <i>Delle rane sono gialle.</i> ‘Some of the frogs are yellow.’ |

Table 1 Examples of the eight auditory stimuli conditions.

Visual and auditory stimuli were presented simultaneously. Each of the eight auditory conditions in [Table 1](#) was combined with all three different visual conditions, i.e., *canonical*,

2 Here and in the following we give word-to-word translations to facilitate reading, but since Italian definites can translate into English bare nominals (generic) or definite nominals (specific), we use brackets for the definite article in English.

noncanonical and *mixed* (Figure 1), resulting in 24 conditions per item. Conditions are named for the combination of factor levels in the following order: COLOR-VISUAL (color of the picture), DETERMINATION (determiner type) and COLOR-AUDITORY (color uttered in the predicative position of the auditory stimulus sentence). For example, *Canonical-Definite-Canonical* means that a picture of four green frogs³ was paired with a sentence introduced by the definite article and the canonical (prototypical) color/pattern of the item, i.e., *The frogs are green*.

The general idea behind this design is the following: Under a generic interpretation, participants should accept conditions with COLOR-AUDITORY *canonical*, i.e., sentences like *(The) frogs are green*, and reject conditions with COLOR-AUDITORY *noncanonical*, i.e., *The frogs are yellow*, independently of the color of the visual stimulus (canonical or noncanonical). Under a specific interpretation, participants should accept conditions where COLOR-VISUAL and COLOR-AUDITORY match, irrespective of whether the depicted items have the canonical or the noncanonical color (i.e., they should accept *These frogs are yellow* when seeing yellow frogs and reject it when seeing green frogs). For ambiguous expressions, both types of responses are possible. The participants' responses will reveal which interpretation they chose. More detailed predictions are given in section 3.2.

The stimuli described here were used in both the experiment described below, and the follow-up experiment (see section 5 below). The conditions with mixed visual stimuli and partitive determiners in the auditory stimuli served as fillers and are not included in detailed descriptions and statistical analysis from this point onwards. In Table 2, we illustrate the language material and formulate predictions for the critical conditions as to acceptance/rejection of the sentences in combination with the respective visual stimuli.

| COND. | COLOR-VISUAL | DETERMINATION | COLOR-AUDITORY | EXAMPLE AUDITORY | PREDICTED ACCEPTANCE |
|-------|---------------------|----------------------|---------------------|----------------------------------------------------------------------|-------------------------------------|
| 1 | Canonical | <i>Demonstrative</i> | <i>Canonical</i> | <i>Queste rane sono verdi.</i> 'These frogs are green.' | Yes |
| 2 | | <i>Demonstrative</i> | <i>Noncanonical</i> | <i>Queste rane sono gialle.</i> 'These frogs are yellow.' | No |
| 3 | | <i>Definite</i> | <i>Canonical</i> | <i>Le rane sono verdi.</i> '(The) frogs are green.' | Yes, if generic Yes, if specific |
| 4 | | <i>Definite</i> | <i>Noncanonical</i> | <i>Le rane sono gialle.</i> '(The) frogs are yellow.' | No |
| 5 | | <i>Generic</i> | <i>Canonical</i> | <i>Le rane di solito sono verdi.</i> 'Frogs are usually green.' | Yes |
| 6 | | <i>Generic</i> | <i>Noncanonical</i> | <i>Le rane di solito sono gialle.</i> 'Frogs are usually yellow.' | No |
| 7 | Noncanonical | <i>Demonstrative</i> | <i>Canonical</i> | <i>Queste rane sono verdi.</i> 'These frogs are green.' | No |
| 8 | | <i>Demonstrative</i> | <i>Noncanonical</i> | <i>Queste rane sono gialle.</i> 'These frogs are yellow.' | Yes |
| 9 | | Definite | Canonical | <i>Le rane sono verdi.</i> '(The) frogs are green.' | Yes, if generic No, if specific |
| 10 | | Definite | Noncanonical | <i>Le rane sono gialle.</i> '(The) frogs are yellow.' | No, if generic Yes, if specific |
| 11 | | <i>Generic</i> | <i>Canonical</i> | <i>Le rane di solito sono verdi.</i> 'Frogs are usually green.' | Yes |
| 12 | | <i>Generic</i> | <i>Noncanonical</i> | <i>Le rane di solito sono gialle.</i> 'Frogs are usually yellow.' | No |

Table 2 Stimulus conditions with examples for visual and auditory stimuli (critical conditions in bold).

³ The conditions with partitive determiners were added to offer a third determiner type. To make this third determiner felicitous in some conditions, the *mixed* visual stimuli were implemented.

3.2 METHODS

The language material consisted of the stimuli outlined in the preceding section. The critical conditions are **Noncanonical-Definite-Canonical** and **Noncanonical-Definite-Noncanonical** (i.e., *The frogs are green/yellow* paired with a picture of yellow frogs). If participants accept sentences with canonical COLOR-AUDITORY, and reject sentences with noncanonical COLOR-AUDITORY, their interpretation of definite plural DPs is generic. If their acceptance pattern is the opposite, then their interpretation of definite plural DPs is specific. To avoid too many repetitions of the same items we split the stimuli into 4 lists; each item occurred in 8 conditions per list. Each list contained 120 randomized trials, with 5 different items per condition.

Participants 24 participants were tested. All were recruited at the University of Verona (Italy) and spoke Italian as their only native language. Participants were aged 20–28 (mean age = 22.48 years, SD = 2.43), 20 were female. All participants gave written and informed consent and were paid 4 euros for participation.

Procedure Participants sat in front of a computer screen, with their fingers resting on a keyboard with response keys marked with green and red stickers. They were instructed in oral and written form about the procedure. They were told that they would see pictures and a sentence and they were supposed to say whether the sentences were correct or not. The presentation of the visual and the auditory stimulus began simultaneously. The visual stimulus revealed the COLOR-VISUAL condition, the first word of the auditory stimulus revealed the DETERMINATION condition, and the last word of the auditory stimulus revealed the COLOR-AUDITORY condition. 50 ms after the offset of the auditory stimulus and 150 ms after the offset of the sentence-final color term, the picture was replaced by the Italian question: *È corretta questa frase?* ('Is this sentence correct?'), presented with green letters on a black screen. Once the question appeared, participants could answer by pressing the green or red button for 'yes' or 'no', respectively. For half of the participants the green button was on the right side of the keyboard and the red button on the left; for the other half it was the opposite. The question remained on the screen until participants responded. A blank screen of 800 ms separated the trials. To avoid response bias no feedback was provided.

Two dependent variables were recorded: (i) acceptance rates and (ii) reaction times (RTs) starting from the appearance of the question on the screen. A training session of 4 trials was performed before the experiment. The experiment lasted 10–15 minutes. Stimulus presentation and recordings were performed with the software PsychoPy3 (Version 3.0.5).

PREDICTIONS

Acceptance rates: We are interested in the six planned comparisons listed below. To better understand the conditions analyzed in the comparisons and their associated predictions, see [Table 2](#). The element of contrast is marked in bold.

- i. *Noncanonical-**Demonstrative**-Noncanonical vs. Noncanonical-**Definite**-Noncanonical*: If participants routinely consider both possible readings of definites, acceptance rates should be lower for definites than for demonstratives. If participants only consider the reading of definites that makes most sense in the given context, i.e., the specific reading, acceptance rates for definites and demonstratives should be similar.
- ii. *Noncanonical-**Demonstrative**-Canonical vs. Noncanonical-**Definite**-Canonical*: If participants routinely consider both possible readings of definites, acceptance rates should be higher for definites than for demonstratives. If participants only consider the specific reading of definites, acceptance rates for definites and demonstratives should be similar.
- iii. *Noncanonical-**Definite**-Canonical vs. Noncanonical-**Generic**-Canonical*: If participants routinely consider both possible readings of definites, acceptance rates should be lower for the definite than for the generic condition. If participants only consider the generic reading of definites, acceptance rates for definites and the generic condition should be similar.
- iv. *Noncanonical-**Definite**-Noncanonical vs. Noncanonical-**Generic**-Noncanonical*: If participants routinely consider both possible readings of definites (or only the specific reading),

acceptance rates should be higher for definites than for the generic condition. If participants only consider the generic reading of definites, then acceptance rates for the definite and the generic condition should be similar.

- v. *Noncanonical-Definite-Canonical* vs. *Noncanonical-Definite-Noncanonical*: If participants always consider the specific reading of the definite article, the acceptance rates for the former condition should be lower than for the latter one. If they always consider the generic reading, the acceptance rates for the former condition should be higher than for the latter. If they always consider both possible readings, acceptance rates should be similar for the two conditions.
- vi. *Canonical-Definite-Noncanonical* vs. *Noncanonical-Definite-Noncanonical*: If participants always consider the specific reading of the definite article, the acceptance rates should be lower in the former condition than in the latter one. We expect the opposite if they always interpret the definite article as generic. If they always consider both possible readings, then acceptance rates in the two conditions should not differ.

Reaction times: Participants can judge the felicity of the sentences either against the visual context or against their world knowledge. We expect participants to do the former with unambiguously specific conditions, and the latter with unambiguously generic conditions. In ambiguous conditions, participants need to choose between these two options. We therefore expect longer reaction times in ambiguous than unambiguous conditions, reflecting the additional workload associated with the former.⁴

4 RESULTS

Reaction times and acceptance rates from all participants were analyzed for the conditions relevant for the experiment (see [Table 2](#)). [Table 3](#) gives an overview of the results for both acceptance rates and reaction times. Data were prepared for statistical analysis in R (R Development Core Team, 2005), using core functions and the packages *reshape* (Wickham 2007), *plyr* (Wickham 2011), and *car* (Fox & Weissberg 2011). Data were analyzed using the packages *lme4* (Bates et al. 2015, *glmer* function for acceptance rates and *lmer* function for reaction times) and *LMERConvenienceFunctions* (Tremblay & Ransijn 2015, *summary* function).

| COND. | COLOR-VISUAL | DETERMINATION | COLOR-AUDITORY | % ACCEPTANCE | MEAN RTS/MS (SD IN PARENTHESES) |
|-------|---------------------|----------------------|---------------------|--------------|---------------------------------|
| 1 | Canonical | <i>Demonstrative</i> | <i>Canonical</i> | 92.00 | 627 (399) |
| 2 | | <i>Demonstrative</i> | <i>Noncanonical</i> | 0.00 | 534 (328) |
| 3 | | <i>Definite</i> | <i>Canonical</i> | 100.00 | 648 (414) |
| 4 | | <i>Definite</i> | <i>Noncanonical</i> | 1.87 | 625 (398) |
| 5 | | <i>Generic</i> | <i>Canonical</i> | 100.00 | 557 (396) |
| 6 | | <i>Generic</i> | <i>Noncanonical</i> | 0.96 | 572 (372) |
| 7 | Noncanonical | <i>Demonstrative</i> | <i>Canonical</i> | 1.00 | 587 (381) |
| 8 | | <i>Demonstrative</i> | <i>Noncanonical</i> | 99.00 | 628 (400) |
| 9 | | <i>Definite</i> | <i>Canonical</i> | 63.48 | 826 (507) |
| 10 | | <i>Definite</i> | <i>Noncanonical</i> | 44.55 | 845 (516) |
| 11 | | <i>Generic</i> | <i>Canonical</i> | 92.73 | 611 (372) |
| 12 | | <i>Generic</i> | <i>Noncanonical</i> | 7.55 | 653 (439) |

Table 3 General mean acceptance rates and reaction times per condition over participants.

4.1 ACCEPTANCE RATES

Mean acceptance rates per condition are summarized in [Table 3](#). Only conditions with noncanonical visual stimuli are informative for the interpretation of determiners (see [Table 2](#)). Therefore, we will limit a detailed presentation and discussion of the results to these conditions.

⁴ See Hino et al. (2006) and Szmałec et al. (2008) for discussions of the link between decision making, response conflict and increased reaction times.

- *Demonstrative* conditions: participants accepted sentences with colors/patterns matching the property shown in the picture (99% acceptance rate) and rejected sentences with colors/patterns not matching the picture (1% acceptance rate). Thus, as expected, they interpreted sentences like *These frogs are yellow* as specific.
- *Generic* conditions: Participants accepted sentences with *di solito* only when the property described in the sentence was the prototypical property of the animal/fruit in the picture (92.73% acceptance) and rejected sentences with the non-prototypical property (7.55% acceptance). Thus, sentences with *di solito* always triggered the generic interpretation, as expected.
- *Definite* conditions: When presented with pictures showing noncanonical colors/patterns (e.g., yellow frogs) and sentences with definite articles, participants seemed to consider both the generic reading (leading them to accept *The frogs are green*, 63.48% acceptance) and the specific reading (leading them to accept *The frogs are yellow*, 44.55% acceptance). After a closer look at individual differences in both *Noncanonical-Definite* conditions, we observed that 36% of the participants were consistent in their responses. Of these participants, 67% consistently gave generic responses and 33% consistently gave specific responses.

4.1.1 Statistical analysis of acceptance rates

The results for acceptance rates are very clear descriptively. We nevertheless pursued six planned comparisons (see above), which were analyzed with binomial generalized linear mixed models. In comparisons I-IV, DETERMINATION was specified as fixed effect, participant and item as random intercepts, and DETERMINATION as random slope for participants and items. In comparison V, we specified COLOR-AUDITORY as fixed effect, participant and item as random intercepts, and COLOR-AUDITORY as random slope for participants. For comparison VI, we specified COLOR-VISUAL as fixed effect, participants and items as random intercepts, and COLOR-VISUAL as random slope for participants and items. Only statistically significant comparisons are reported in detail below (alpha = .05).

- Noncanonical-**Demonstrative**-Noncanonical vs. Noncanonical-**Definite**-Noncanonical* (99% vs. 44.55% acceptance): The difference between the conditions is significant ($z = 5.172, p < 0.001$).
- Noncanonical-**Demonstrative**-Canonical vs. Noncanonical-**Definite**-Canonical* (1% vs. 63.48% acceptance): The difference between the conditions is significant ($z = -2.893, p < 0.01$).
- Noncanonical-**Definite**-Canonical vs. Noncanonical-**Generic**-Canonical* (44.55% vs. 92.73% acceptance): The difference is significant ($z = 2.364, p = 0.05$).
- Noncanonical-**Definite**-Noncanonical vs. Noncanonical-**Generic**-Noncanonical* (44.55% vs. 7.55% acceptance): The difference is significant ($z = -5.525, p < 0.001$).
- Noncanonical-**Definite**-**Canonical** vs. Noncanonical-**Definite**-**Noncanonical*** (63.48% acceptance vs. 44.55% acceptance): The difference is not significant.
- Canonical**-**Definite**-Noncanonical vs. **Noncanonical**-**Definite**-Noncanonical* (1.87% acceptance vs. 63.48% acceptance): The difference is significant ($z = -3.942, p < 0.001$).

In summary, the statistical analysis of planned comparisons revealed that descriptive differences were statistically significant, except for Comparison V. Participants accepted sentences with definite determiners and canonical or noncanonical properties at similar rates (63.48% acceptance for *The frogs are green*, indicating a generic interpretation, and 44.55% acceptance for *The frogs are yellow*, indicating a specific interpretation; no significant difference). This matches the general assumption that definites are truly ambiguous, and that both possible interpretations are routinely considered in the present paradigm.

4.2 REACTION TIMES

Only the conditions with noncanonical visual stimuli were analyzed statistically, as these are the only ones allowing us to distinguish between generic and specific interpretation. Before data analysis, we removed reaction times shorter than 200 ms and longer than 6000 ms, leading to the removal of 16.15% of the data. Reaction times per condition were analyzed for

the responses matching the expected interpretation (for unambiguous demonstratives and generics). This led to the removal of 1.91% of the data. Moreover, we calculated two standard deviations from the mean as a cut-off, leading us to discard another 4.5% of the data.

4.2.1 Statistical analysis of reaction times

Log-transformed RTs were analyzed using linear mixed effects models. We specified DETERMINATION and COLOR-AUDITORY as fixed effects with full interactions, participant and item as random effects, and DETERMINATION and COLOR-AUDITORY as random slopes. Only statistically significant effects are reported, unless stated otherwise. We found a significant main effect of DETERMINATION (*definite* vs. *demonstrative* $t = -4.285$, $p < .001$; *definite* vs. *generic* $t = -3.632$, $p < .001$, but no significant difference between *demonstrative* and *generic* $p > .4$). Reaction times for definites were significantly longer than for generics and for demonstratives. Thus, the statistical analysis revealed that participants took longer to rate the sentences in ambiguous (i.e., definite) conditions than in the unambiguously specific or generic conditions. Reaction times between the two unambiguous conditions, i.e., demonstrative and generic, did not differ significantly. We will discuss the implications of our findings below.

4.2.2 Interim-Discussion

Our findings show similar acceptance rates for definites with generic and specific interpretation, and longer reaction times for definites compared to unambiguously specific and generic conditions. This suggests that Italian definites are truly ambiguous between both readings. In contrast, German definites showed a clear preference for specific interpretation in a similar experiment (Czypionka & Kupisch 2019). One point of concern is the extent to which the Italian and German experiments are directly comparable. In our experiment four levels of the factor DETERMINATION were used, namely *definite*, *demonstrative*, *partitive*, and *generic*, whereas Czypionka & Kupisch (2019)'s only had three.⁵ A follow-up experiment was performed to alleviate this concern.

5 FOLLOW-UP EXPERIMENT: THREE DETERMINERS

To allow for a direct comparison to the German experiments in Czypionka & Kupisch (2019), which had only three determiner types (*bare*, *definite*, *demonstrative*), we conducted a follow-up study without the *generic* DETERMINATION condition. Stimuli were the same as in the previous experiment, excluding the conditions with DETERMINATION *generic* (leaving the levels *definite*, *demonstrative* and *partitive*), since in Italian there is no unambiguously generic condition based on article choice alone. To avoid too many repetitions of the same items, we split the stimuli into 4 lists with each item occurring in 8 conditions per list. Each list contained 90 randomized trials, with 5 different items per condition.

25 participants were tested. All were recruited at the University of Verona (Italy) and had Italian as their only native language. Participants were aged 20–26 (mean age = 22.68 years, $SD = 1.86$), 22 were female. All gave written and informed consent and were paid 4 euros for their participation. The procedure and data analysis were run in parallel with the previous experiment; data from all participants were included.

5.1 RESULTS

Results are summarized in [Table 4](#). Since only conditions with *noncanonical* COLOR-VISUAL are informative for the interpretation of articles, we will limit our summary to these conditions. Details of the statistical analyses can be found in the Appendix ([LINK](#)).

5.1.1 Summary of acceptance rates

Demonstrative: Participants accepted sentences with the noncanonical COLOR-AUDITORY (*These frogs are yellow*) and rejected them with the canonical COLOR-AUDITORY (*These frogs are green*; 93.55% acceptance vs. 5.32% acceptance). This matches an (expected) specific interpretation of demonstratives.

⁵ We had four determination conditions for two reasons: (i) we needed an unambiguous generic condition in Italian (as exists in German); (ii) we needed three determiner types based on article alone to parallel the German experiment.

| COND. | COLOR-VISUAL | DETERMINATION | COLOR-AUDITORY | % ACCEPTANCE | MEAN RTS/MS (SD IN PARENTHESES) |
|-------|---------------------|----------------------|---------------------|--------------|---------------------------------|
| 1 | Canonical | <i>Demonstrative</i> | <i>Canonical</i> | 96.16 | 542 (339) |
| 2 | | <i>Demonstrative</i> | <i>Noncanonical</i> | 0.00 | 583 (352) |
| 3 | | <i>Definite</i> | <i>Canonical</i> | 98.96 | 548 (324) |
| 4 | | <i>Definite</i> | <i>Noncanonical</i> | 0.00 | 611 (393) |
| 5 | Noncanonical | <i>Demonstrative</i> | <i>Canonical</i> | 5.32 | 641 (351) |
| 6 | | <i>Demonstrative</i> | <i>Noncanonical</i> | 93.55 | 533 (266) |
| 7 | | Definite | Canonical | 42.20 | 715 (425) |
| 8 | | Definite | Noncanonical | 66.67 | 732 (418) |

Table 4 Mean acceptance rates, reaction times and standard deviation per condition over participant for the follow-up experiment.

Definite: Acceptance rates were close to 50%, both for sentences with the canonical and the noncanonical property. When presented with a picture of yellow frogs, participants accepted 42.2% of the sentences with the *canonical* auditory stimulus (*The frogs are green*) and 66.67% of the sentences with the *noncanonical* auditory stimulus (*The frogs are yellow*). This similarity of acceptance rates is in line with an ambiguous interpretation of definites. The small descriptive difference between both conditions did not reach significance. The findings of both experiments show that the number conditions did not influence participants’ preferred interpretation of definites.

5.1.2 Summary of reaction times

The statistical analysis revealed that there is an effect of DETERMINATION on conditions with *noncanonical* sentences: Reaction times are shorter for unambiguously specific sentences than for sentences ambiguous between specific and generic readings. Conditions with *canonical* sentences show no such effect, possibly because these sentences are usually rejected in the *demonstrative* conditions, but accepted or rejected in the *definite* conditions. Taken together, these findings are in line with the results of the acceptability ratings and of the main study, suggesting that definites are ambiguous. We will discuss the implications below.

6 GENERAL DISCUSSION AND CONCLUSION

The present study investigated to what extent Italian speakers interpret definite plural DPs (in subject position) as generic or as specific, and how this translates to acceptance rates in the “polar bear” paradigm. In our experiments, we provided a nonlinguistic context in the form of pictures. This made a specific interpretation of the sentences possible, while not ruling out a generic interpretation.

Acceptance rates show that Italian speakers do not have a clear preference for a generic or specific interpretation of definite plural DPs, and that definite plural DPs are truly ambiguous with respect to specificity (however, some of the participants developed strategies, choosing one and the same reading in all their sentence interpretations; both options were chosen, suggesting no general preference for a particular reading across participants). Reaction times for the acceptability judgments lend further support to the ambiguity of definite plurals. Reaction times were longer for definite conditions than for the unambiguously specific demonstrative or the unambiguously generic conditions. We interpret this as reflecting a choice that is necessary when judging definite conditions, but not necessary in the other (unambiguous) conditions.

The main goal of the present study was to provide a qualitative comparison to earlier findings from a parallel experiment on the interpretation of plural definite DPs in German (Czypionka & Kupisch 2019), in order to establish whether the “polar bear” paradigm is an appropriate test of article semantics. As outlined in the introduction, German definites were claimed to be ambiguous between generic and specific interpretations, similar to definites in Italian, but unlike definites in (most) other Germanic languages. However, the acceptance rates reported in Czypionka & Kupisch (2019) show that in the current experimental paradigm, German definite plurals are reliably interpreted as specific, not as generic. The only hint at a potential ambiguity of German definites was found in longer reaction times for definites compared to other conditions. In contrast, the interpretation of Italian definite plurals in the same paradigm

is close to 50% for each of the possible readings. This shows that truly ambiguous readings surface in the experimental task employed in these studies, and that the findings reported in Czypionka & Kupisch (2019) cannot be reduced to bias towards specific interpretation introduced by the nonlinguistic picture context. The “polar bear” paradigm can therefore be considered a suitable tool for researching article semantics.⁶

Taken together, the findings of both studies show that Italian definite plural DPs are completely ambiguous, as specific and generic readings are routinely considered, while German definite plural DPs are preferentially interpreted as specific. While reaction times for German definites suggest that their interpretation might involve making a choice, akin to that necessary for the interpretation of Italian definites, this choice is consistently made in favor of the specific interpretation in the current paradigm. Hence, our study lays important groundwork for quantitative comparison, as has been done in research on German-Italian bilinguals (e.g., Kupisch 2012), corroborating with experimental evidence that German and Italian definite articles are indeed semantically different.

ABBREVIATIONS

ART = article; COP = copula; PL = plural; SG = singular.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendix:** Detailed statistical analysis of the follow-up experiment. DOI: <https://doi.org/10.5334/gjgl.1165.s1>

FUNDING INFORMATION

This article was supported by generous funding to Sergio Miguel Pereira Soares by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska Curie grant agreement No 765556.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Michela Redolfi  orcid.org/0000-0001-7725-2300

University of Verona, Italy;
University of Konstanz, Germany

Sergio Miguel Pereira Soares  orcid.org/0000-0002-1709-6059

University of Konstanz, Germany

Anna Czypionka  orcid.org/0000-0003-2247-0110

University of Konstanz, Germany

Tanja Kupisch  orcid.org/0000-0002-2653-2692

University of Konstanz, Germany
UiT, The Arctic University of Norway, Norway

REFERENCES

- Barton, Dagmar, Nadine Kolb & Tanja Kupisch. 2015. Definite article use with generic reference in German: an empirical study. *Zeitschrift für Sprachwissenschaft* 34(2). 147–173. DOI: <https://doi.org/10.1515/zfs-2015-0009>
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using

⁶ We would like to thank an anonymous reviewer for pointing out that a possible task effect might be mirrored in longer RTs in generic responses vs. specific responses in the ambiguous conditions. We checked the results in both ambiguous conditions in both experiments and no relevant patterns were detected. Thus, this potential effect did not affect our overall conclusions.

- Ime4. *Journal of Statistical Software* 67(1). 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Boersma, Paul & David Weenink. 2016. Praat: doing phonetics by computer [Computer program]. Version 6.0.19, retrieved 13 June 2016 from <http://www.praat.org/>.
- Brugger, Gerhard. 1993. Generic interpretations and expletive determiners. *University of Venice Working Papers in Linguistics* 3. 1–30.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405. DOI: <https://doi.org/10.1023/A:1008324218506>
- Czypionka, Anna & Tanja Kupisch. 2019. (The) polar bears are pink. How (the) Germans interpret (the) definite articles in plural subject DPs. *The Journal of Comparative Germanic Linguistics* 22(3). 247–291. DOI: <https://doi.org/10.1007/s10828-019-09111-y>
- Dayal, Veneeta. 2004. Number marking and (in) definiteness in kind terms. *Linguistics and Philosophy* 27(4). 393–450. DOI: <https://doi.org/10.1023/B:LING.0000024420.80324.67>
- Fox, John & Sanford Weisberg. 2011. An R companion to applied regression. Thousand Oaks, California. Online: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
- Gavarró, Anna, Ana Teresa Pérez-Leroux & Thomas Roeper. 2006. Definite and bare noun contrasts in child Catalan. *Language Acquisition and Language Disorders* 41. 51. DOI: <https://doi.org/10.1075/lald.41.04gav>
- Gelman, Susan & Lakshmi Raman. 2003. Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development* 74(1). 308–325. DOI: <https://doi.org/10.1111/1467-8624.00537>
- Hino, Yasushi, Penny Pexman & Stephen Lupker. 2006. Ambiguity and relatedness effects in semantic tasks: are they due to semantic coding? *Journal of Memory and Language* 55(2). 247–273. DOI: <https://doi.org/10.1016/j.jml.2006.04.001>
- Krifka, Manfred, Francis Jeffrey Pelletier, Gregory Carlson, Alice ter Meulen, Gennaro Chierchia & Godehard Link. 1995. Genericity: An introduction. In Gregory Carlson & Francis Jeffrey Pelletier (eds.), *The Generic Book*, 1–124. University of Chicago Press.
- Kupisch, Tanja. 2012. Specific and generic subjects in the Italian of German–Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition* 15(4). 227–244. DOI: <https://doi.org/10.1017/S1366728911000691>
- Kupisch, Tanja & Cristina Pierantozzi. 2010. Interpreting definite plural subjects: A comparison of German and Italian monolingual and bilingual children. In Katie Franich, Kate Iserman & Lauren Keil (eds.), *Proceedings of the 34th Boston University Conference on Language Development*, 245–254. Cascadilla Press.
- Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4). 609–665.
- Montrul, Silvina & Tania Ionin. 2010. Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition* 13(4). 449–473. DOI: <https://doi.org/10.1017/S1366728910000040>
- Oosterhof, Albert. 2008. *The semantics of generics in Dutch and related languages*. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/la.122>
- Pérez-Leroux, Ana Teresa, Alan Munn, Cristina Schmitt & Michelle DeIrish. 2004. Learning definite determiners: Genericity and definiteness in English and Spanish. In Alejna Brugos, Linnea Micciulla & Christine E. Smith (eds.), *28th Boston University Conference on Language Development Proceedings – BUCLD 28, Online Proceedings Supplement*, (<http://www.bu.edu/buclid/files/2011/05/28-perez-leroux.pdf> 29 April 2015).
- R Development Core Team. 2005. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Szmałec, Arnaud, Frederick Verbruggen, André Vandierendonck, Wouter De Baene, Tom Verguts & Wim Notebaert. 2008. Stimulus ambiguity elicits response conflict. *Neuroscience Letters* 435. 158–162. DOI: <https://doi.org/10.1016/j.neulet.2008.02.023>
- Tremblay, A., & J. Ransijn. (2015). LMER Convenience Functions: Model selection and post-hoc analysis for (G) LMER models. *R package version*, 2(10).
- Wickham, Hadley. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21(12). 1–20. DOI: <https://doi.org/10.18637/jss.v021.i12>
- Wickham, Hadley. 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software* 40(1). 1–29. DOI: <https://doi.org/10.18637/jss.v040.i01>

TO CITE THIS ARTICLE:

Redolfi, Michela, Sergio Miguel Pereira Soares, Anna Czypionka and Tanja Kupisch. 2021. Experimental evidence for the interpretation of definite plural articles as markers of genericity – How Italian can help. *Glossa: a journal of general linguistics* 6(1): 16. 1–12. DOI: <https://doi.org/10.5334/gjgl.1165>

Submitted: 25 November 2019

Accepted: 04 December 2020

Published: 05 February 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.