

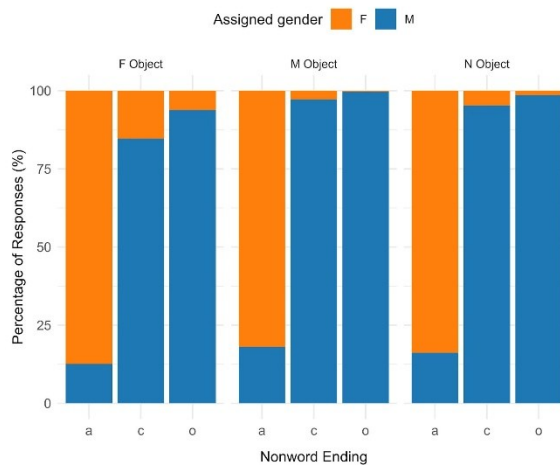
## Ortho-phonological and semantic cues in gender assignment to novel words in Italian

Italian is a language with high gender transparency, meaning that word ending generally represents a reliable cue to grammatical gender: typically, nouns ending in -a are grammatically feminine (F), while those ending in -o are masculine (M), with the latter being the most frequent gender overall (De Martino et al., 2023). This regularity does not apply to the most common types of borrowings in Italian, i.e., anglicisms, often ending in consonant thus not providing any cue to gender. Previous research on gender assignment to novel words in Italian mainly focused on extracting rules from linguistic corpora, showing that borrowings tend to be assigned either the gender of their closest L1 equivalent or a “default” masculine gender, although some degree of within and between speaker variation is often observed in the early stages (Pulcini, 2023). This study investigates gender assignment by monolingual Italian speakers during first exposure to novel words, manipulating word ending (to test an ortho-phonological strategy) and the gender of objects associated with these words to serve as translation equivalents (to test a semantic strategy).

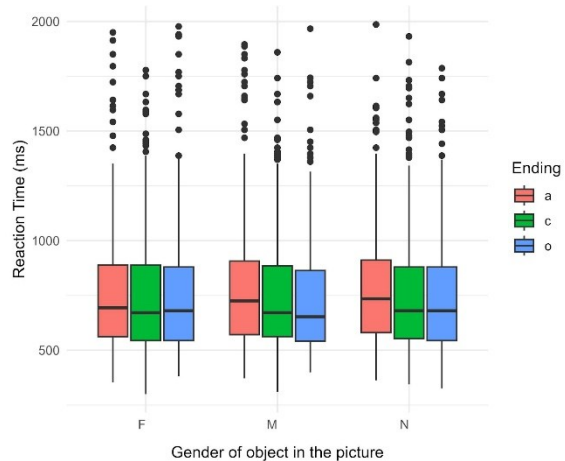
**Our study.** We ran a production study in E-Prime 3 in which a set of images of inanimate objects (N = 180, from the MultiPic database; half feminine in Italian) and non-objects (N = 90, NOUN Database, not associated to any gender), were randomly associated with nonwords ending in -a or -o or consonant, -c (N=270, balanced across the 3 conditions) (Figure 1). Participants (N=12 so far, 50 expected) were asked to produce a sentence starting with “è un/è una...” followed by the novel word (e.g., *dimaba/dimabo/dimab*). We coded the gender assigned to the word on the basis of the article produced (*un/uno*<sub>[M]</sub>, *una*<sub>[F]</sub>); we also recorded speech onset time. Overall, we expect participants to assign M more often across all conditions, being this the default, unmarked option, in Italian. We expect a contribution of the ortho-phonological strategy in assigning gender to nonwords with canonical endings (a/o) and, possibly, a modulation of the gender of the object (semantic strategy), particularly for novel words ending in consonant. We also expect significantly longer reaction times in the case of conflicting cues (e.g., *dimabo*, associated with the picture of boat, *barca*<sub>[F]</sub>).

**Results.** Figure 1 and 2 plot the distribution of assigned gender (M/F) and speech onset latency. Assigned gender (M/F) was entered in a logistic regression model including Ending (c/a/o), and Object (M/F/No-gender) as predictors and their interaction, setting contrasts as (0.5, -0.5, 0) and (0.5, 0, -0.5). The model also included random intercepts for participant, picture and word root. Preliminary results (Figure 3) show a prevalence of the ortho-phonological strategy: -o and -c endings were predominantly assigned M, -a endings F (1st contr.  $z=20.32$ ; 2nd contr:  $z=-10.38$ ). Influence of the object’s gender emerged for words ending in -c, that were more likely to be assigned F when associated with a F object, compared to words ending in -o ( $z=-2.51$ ). We further ran a linear regression model on (log-transformed) speech onset latency with the same random structure as above. Reaction times show longer latencies when a word ending in -a was associated with an M object, compared to words ending in -c ( $z=-2.00$ ), showing a possible interference of the associated object’s gender on gender assignment to the novel word. No other significant interactions emerged.

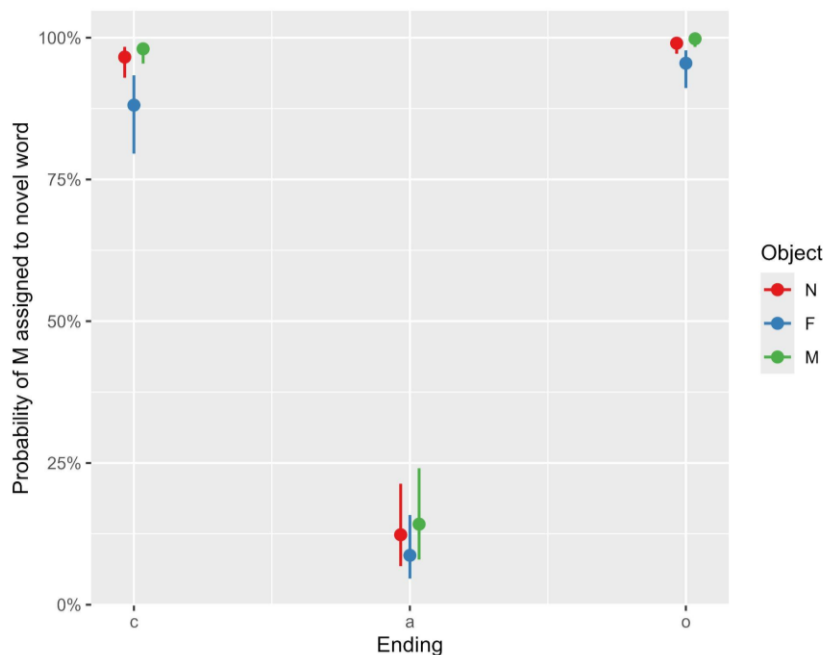
**Discussion.** We show evidence for (i) an orthographic strategy in gender assignment, with M preferably assigned to novel words ending in -o, and F to novel words ending in -a; (ii) evidence for M being the “default” form for novel words ending in -c; (iii) limited but consistent evidence for an influence of the gender of the associated object on gender assignment, as shown in speech latencies for words ending in -a (with M objects) and for gender assigned to words ending in -c (with F objects). Further planned studies with L2-Italian speakers of L1 French (gender-marked language, also exploiting cross-linguistic gender incongruencies), and L1-Dutch (with no gender assigned to inanimate nouns) will allow us to explore the effects of cross-linguistic influence and L2 proficiency on such strategies.



**Figure 1.** Percentage of masculine (M) and feminine (F) articles produced before novel words split by the associated picture's gender (F, M, No-object) and word ending.



**Figure 2.** Response latency between stimulus presentation and speech onset latency by object's gender and nonword ending.



**Figure 3.** Interaction plot of model 1 examining gender assignment to nonwords (M/F).  
(Formula: gender\_assigned ~ Ending\*Object + (1|Participant) + (1|Picture) + (1|Item))

**Selected References** De Martino et al. (2023). Frequency distribution of inflectional properties of nouns: data on written Italian. *CHIMERA: Revista De Corpus De Linguas Romances Y Estudios Lingüísticos*, 10, 121–133; Pulcini, V. (2023). *The Influence of English on Italian: Lexical and Cultural Features*. De Gruyter.