6 General Discussion

The data collected through the main experiment (SPMT) confirmed an overall lower level of accuracy when participants dealt with the innovative suffix *-ə* compared to the masculine and feminine conditions. While this trend persists to some extent in the follow-up study (TCET) – where the nonbinary condition still shows the lowest accuracy – it nonetheless registers a significant increase in performance relative to the first experiment. The overall improvement observed in the TCET applies across all singular conditions, i.e., nonbinary, masculine, and feminine alike, but is particularly striking in the nonbinary case. This may be attributed to the different task design: the absence of images in the follow-up experiment eliminated visual interference, allowing participants to focus directly on linguistic cues and produce more reliable accuracy scores. The more transparent structure of the TCET thus appears to have enhanced participants' ability to accurately process each linguistic form.

These general trends take on greater significance when anchored to the specific findings of the three analyses conducted on the SPMT. *Analysis 1* (Section 4.2.2.1) indeed showed that the lowest accuracy is found in the nonbinary singular condition, while the masculine

DOI 10.30687/979-12-5742-035-2/006

and feminine conditions obtained a higher, identical score. The main finding revealed by *Analysis 2* (Section 4.2.2.2) points to the actual interpretation of the nonbinary plural condition, i.e., one that serves as a representation of a mixed-gender group that may be composed by male and female members only. Although the exclusively nonbinary group (i.e., NBNBNB) received the highest percentage of selections, the mixed-gender groups lacking a nonbinary member ranked second in frequency. *Analysis 2* indeed proved – through the sharp drop in score recorded for the nonbinary plural condition as a consequence of leaving out MMF and FFM groups – that the groups including at least one nonbinary individual selected by participants were few. Hence, there appears to be a tendency for the innovative suffix *-ə* to be interpreted rather frequently as referring to a mixed group of males and females, without a nonbinary person. It follows that the inclusive use of *-ə* does not necessarily imply nonbinary participation, thus partly redefining the concept of inclusiveness conveyed. Activation of a nonbinary referent in the presence of *-ə* is therefore by no means to be taken for granted.

As far as response times are concerned, what emerges from *Analysis 3* (Section 4.2.2.3) is that higher levels of traditional gender markedness were found to aid in reducing response time, as participants were fastest in dealing with the feminine and masculine conditions. The one exception is the masculine appearing in the plural condition, which recorded a rather slow mean response time. In all likelihood, this is due to the opposition in meaning (i.e., specific *vs* generic) that it generates, posing some sort of conflict to be handled by participants. The fact that the masculine can have a generic meaning might have caused longer response times precisely because it took longer for participants to decide which referents to choose, as opposed to what happened with the feminine, which was predominantly assigned an exclusively-female meaning. The same reasoning can be applied to the neutral condition: being completely deprived of any gender cue, the wide margin of choice prevents participants from an immediate disambiguation, since the selection of potentially any referent is grammatically correct. It is therefore not surprising that the neutral condition is the one with the slowest mean response time in the plural-condition task, and the second-slowest in the singular-condition task. The nonbinary condition was identified as the most irregular condition, as it was attributed the slowest mean response time in the singular-condition task, as opposed to the fastest mean response time in the plural-condition task. This opposite trend might suggest total hesitation in having to identify the corresponding single referent indicated by *-ə* in the singular condition and the maximum freedom to which its plural counterpart alludes to. Diametrically opposite mean response times

thus seem to reflect the different response strategies when dealing with singular and plural *-ə*.

The examination of the internal distribution of the assigned genders through images per condition, carried out in Section 4.2.3, has shed light on the possible conflicts that surfaced from the matching procedure. For the masculine singular condition, a distinct preference for pictures previously rated as 'male' was markedly detected, confirming the activation of the specific reading of the masculine form (cf. Safina 2024 for similar results). The opposition arising with the plural (i.e., specific *vs* generic meaning) seems to be resolved in favor of a specific interpretation of the masculine plural form, as the majority of participants selected a male-only group, rather than a mixed-gender one. This evidence supports the existence of a bias embedded in the masculine form according to which an interpretation of an exclusivity of male referents is naturally triggered. This bias found in Italian is consistent with previous results in other languages (see Section 1.5).

It is worth pointing out that the group obtaining the second-highest score for the masculine plural condition was the FFM group, that is, one with a female majority. This outcome is likely related to the fact that both masculine plural items referred to roles such as "fixed-term type A researchers" and "research fellows", which correspond to lower-ranking academic positions where female representation in Italy remains relatively strong. According to 2023 data from the Italian Ministry of University and Research (MUR), women comprised 50% of research fellows (Grade D) and 46% of junior researchers (Grade C), but their representation drops progressively at higher career levels: 43% among associate professors (Grade B) and only 28% among full professors (Grade A) (MUR 2025, 8). These figures help explain why participants may have associated those entry-level roles with women, contributing to the selection of the FFM group. This steep decline in female presence at higher academic ranks is well-documented and often discussed in public discourse under the label of the 'glass ceiling'. Therefore, the prevalence of female-majority group selections may reflect not only the actual distribution of women in early academic roles but also a broader awareness – partly grounded in public and media narratives – of gender disparities in Italian academia. In this sense, participants' choices align with both statistical data and broader socio-cultural narratives, rather than being merely shaped by collective imagination.

Turning now to the scrutiny of the neutral singular condition, all genders displayed a rather fair representation, based on their frequency of occurrence. Nonetheless, when no gender information was provided, participants defaulted to male. This trend is also reflected in the neutral plural condition, as the male-only group reports the highest frequency of occurrence in the first neutral

plural item. The FFM group was instead selected by the majority of participants in the second neutral item. Given that both items in the neutral plural condition referred to "members of the research team", it is plausible that participants' interpretations aligned with actual demographic distributions in Italian academia, as the recent MUR data provided above confirm. This recurring preference for a female-majority group may thus reflect an implicit awareness of the structural positioning of women in the academic hierarchy – assigning them a central role in research teams, while men remain more strongly associated with top-level positions. Rather than projecting abstract gender balances, participants may have unconsciously mirrored the current gendered division of labor in academia, where women are more visible and numerous in early-stage academic roles. Alternatively, perhaps it is the order in which the images were presented that affected participants' choices in conditions where there is no grammatical gender clue. The first group encountered by participants in the first neutral item was in fact the MMM group. As for the second neutral item, the FFM group was the second group to be introduced. Participants may have therefore settled on the first plausible option without necessarily considering the remaining alternatives.

Looking at the feminine condition, which at first glance appears entirely unambiguous, we observed that singular feminine forms elicited a highly systematic mapping onto female referents, while also triggering nonbinary interpretations to a noticeable extent. This variability becomes even more marked in the plural, where there is a consistent preference for female-only groups alongside a clearly visible secondary assignment to nonbinary-only groups. This reflects a perceptible closeness between female and nonbinary referents, a pattern further corroborated by the data from the nonbinary condition. In the nonbinary singular condition, the vast majority of participants selected the image representing a female referent – a finding interpreted by one reviewer as potentially resulting from visual stimuli that reflected a distorted representation of gender. This observation raised the concern that participants may have interpreted the images intended to depict nonbinary individuals as female, and vice versa. As statistically supported by the significantly lower accuracy observed in the nonbinary condition, this interpretive bias may have undermined the classification of certain responses as nonbinary. While this is a legitimate observation and we acknowledge that some images may not have fully conveyed a clear departure from binary gender norms, we believe several methodological features mitigate this issue.

First, all visual stimuli used to represent nonbinary individuals – as well as male and female individuals – were independently rated by an American sample prior to the original experiment, and subsequently

validated by an additional, independent Italian sample, distinct from the participants involved in the main study. No significant discrepancies emerged between the two groups in how the images were interpreted, suggesting that potential cultural differences in gender perception were minimal, if present at all. Second, and perhaps more importantly, these images were never presented in isolation. Each nonbinary image or group was always displayed alongside – hence shown in contrast with – overtly and stereotypically male and female alternatives. This contrastive structure was designed precisely to help participants disambiguate gender identity through relative comparison. Even if the nonbinary images were not visually optimal or entirely unambiguous, participants encountered them alongside referents that were stereotypically binary, making it more likely that they would interpret the former as belonging to a distinct, nonbinary category. In other words, the contrastive format made it possible for participants to interpret them in relative terms, as different from both male and female options. Importantly, in the nonbinary plural condition marked with the innovative suffix *-ə*, participants most frequently selected the NBNBNB group, further suggesting that they interpreted these figures as neither male nor female, but as a separate gender category. While we acknowledge that visual representations inevitably allow for a degree of interpretive flexibility, the consistency with which participants associated these images with nonbinary linguistic input reinforces the validity of our classification, ultimately supporting the possible exclusive association of the suffix *-ə* with nonbinary referents. Nonetheless, a considerable portion of participants still opted for mixed-gender groups that did not include nonbinary candidates. This finding suggests that although the specific reading of the innovative suffix as denoting nonbinary identity appears to dominate, a more generic or inclusive interpretation may persist to some extent. In this respect, our data point to a more cautious and gradual shift than the one hypothesized by Thornton (2022; see Section 2.2.2), who suggested a possible refunctionalization of *-ə* toward a fully generic plural. Rather than contradicting this perspective outright, our results highlight the coexistence of both readings, with a preference for the specific nonbinary interpretation within our sample.

The follow-up experiment, which removed all visual stimuli and relied exclusively on linguistic input, offers further evidence. A clear majority of participants selected the *nonbinary* label in the condition marked by the neomorpheme; yet, among the remaining responses, *female* was the most frequently chosen alternative, followed by fewer *other* or *male* selections. The persistence of *female* as the second-most frequent choice – even in the absence of visual referents – confirms that the interpretive bias toward the female over the male when processing forms marked with the innovative suffix *-ə* is not primarily

perceptual but likely grammatical in nature and has presumably to do with the morphological structure of the nonbinary forms. The graphic presence of the schwa (*ə*) after the consonant *l* in the determiner, for instance, may be processed as analogous to the suffix *-a* in the feminine determiner *la*. In contrast, the masculine form displays the vowel *i* before the consonant (cf. *il*). In addition, the nonbinary *ə* has a certain level of proximity with the feminine (plural) suffix *-e*, which hardly leaves room for a male reading. The argument that *lə* alone may carry a slight bias toward female reference is, however, weakened by the findings from the nonbinary plural condition, which is also built around the form *lə*: cf. *lə candidatə* 'the.NB.SG/PL candidate.NB.SG/PL'. In this case, nonbinary-only groups were most frequently chosen, demonstrating that the determiner itself does not drive the preference for female referents. Crucially, this proves that the same *lə* form can produce different interpretive outcomes depending on the grammatical context: in the singular, the presence of *l*+*ə* may nudge participants toward a female interpretation, whereas in the plural, the grammatical context (e.g., surrounding nouns) allows the innovative suffix to convey its intended nonbinary meaning, effectively overriding any apparent bias associated with the determiner. Precisely because the same form can lead to such diverse interpretations (i.e., femaleness *vs* nonbinariness), we cannot conclude that the determiner itself has a direct effect on assignment choices.

When considering the role of image removals, it is worth noting – without making it the main focus, for the reasons outlined above – that the feminine condition offers a valuable insight. With feminine items, we found that eliminating visual stimuli led to higher accuracy. The selection of the *nonbinary* label dropped in fact to nearly zero (0-5%). This suggests that the previously observed co-selection of female and nonbinary depicted referents when matching the description *la candidata* 'the.F.SG candidate.F.SG' was possibly driven by the perceptual closeness of some of the images, indicating a more consistent activation of the intended gender when the feminine form is presented alone, thereby assigning a role – or at least some influence – to the dimension represented by the images in the gender-assignment process. Reduced visual interference contributed to higher overall accuracy, including the nonbinary condition, where *female* selections decreased. This suggests that the original visuals may have partially biased participants toward female interpretations. However, it is important to highlight that the nonbinary condition still produced the lowest accuracy. This indicates that despite clearer task design provided by the TCET, the innovative suffix *-ə*, being a non-native and non-established grapheme corresponding to a non-native phoneme, continues to pose cognitive and linguistic challenges for Italian speakers. Its lack of anchoring in the traditional

morphological, phonological, and syntactic system of the language likely contributes to this difficulty. The follow-up experiment, while more explicit in its aims and less 'masked' in design, did not manage to eliminate this challenge.

The responses collected through the questionnaire on the *Sexist/Inclusive Language Use in Italian* (SILUI), reported in Section 4.3 and Section 5.4, revealed that overall, our sample of participants is determined to embrace change, in abandoning the sexist use of the Italian language in favor of an inclusive one, despite a strong awareness of the complexity of the initial circumstances. However, participants seem to hesitate in adopting morphological innovations. With the generic use of the masculine plural introduced in Item 4, for instance, participants are not prepared to drop it. They are tempted to maintain it as it is a linguistically-economic strategy that guarantees straightforwardness. Others still struggle with recognizing sexist language, especially with the adoption of masculine occupational terms to refer to a female individual (Item 16). This finding applies only to the SPMT-sample, given that the TCET-sample was largely able to recognize the degree of sexism in this instance. When asked to make use of the *-ə* suffix to refer to an exclusive-nonbinary plurality, or a single nonbinary referent, or a mixed-gender group (in so doing, introducing a single short form to represent all genders), the majority of participants across both samples welcomed the three different proposals. The only conflicting information deriving from the participants' responses has to do with the use of *-ə* as a general, common word-ending to indicate a group whose gender composition is not known or of a mixed nature, as the majority (from both samples) first did not agree with such a use put forward in Item 10, to then welcome it in Item 19. As much as the object under investigation (i.e., the use of the single inclusive plural suffix *-ə*) is identical in both items, the questions are actually exploring quite different related aspects, in that Item 10 is directly asking, in more absolute terms, to express a judgment as to whether the proposed use is deemed appropriate as well as fair; Item 19, on the other hand, is addressing the participants' willingness to eventually adopt such a strategy. Morality aside, it is testing their being ready to put it in use.

As for the *Questionnaire on Reading Preferences and Habits* (cf. Section 4.4 and Section 5.5), which was included to profile our participants according to their level of expertise in reading and their level of exposure to inclusive language (with a focus on the innovative suffix *-ə*), the majority of participants consider themselves experienced readers. Furthermore, the importance of reading was acknowledged by them all, who also expressed a strong preference for reading on paper. The vast majority of participants claimed to be familiar with inclusive language. As predicted, among the punctuation marks and orthographic markers listed, the innovative

suffix -ə was the most frequently recognized by participants, yet mostly confined to the social and news industry, particularly online. Through the last two questions, we were able to understand to what extent the 'ə' symbol impacts on the process of reading as well as the criticalities arising from producing it phonetically and interpreting it. Although most participants do not see it as an obstacle to reading, they find it quite difficult to pronounce, yet easy to interpret.