Lost in translation?
Machine translation and gender-fair German

Sabrina Link (Urbino)

Abstract

Efforts to increase the visibility of women in the German language date back to the 1970s. These initiatives have been complemented by more recent attempts to enhance the visibility of non-binary individuals. Consequently, a growing number of language-specific strategies have been proposed to either neutralise the language or to linguistically render individuals of all genders visible. The use of such strategies, rather than the still prevalent masculine generic, is often a highly conscious choice. Therefore, when translating such forms of gender-fair language (GFL), an adequate translation should seek to capture the original intentions expressed in German.

However, in the context of machine translation (MT), on which we are increasingly reliant, previous studies have repeatedly identified a male bias. While most prior research has primarily focused on binary forms and often on translations from natural gender to grammatical gender languages, this study aims to close the existing research gap by analysing MT of neutralisation, binary, and non-binary strategies from a grammatical gender language, German, into another grammatical gender language, Italian, as well as a natural gender language, English. Given the growing presence of large language models (LLMs) in everyday life and their translation capabilities, the study further seeks to determine whether differences exist between commercial translation programmes (DeepL, Google Translate, Bing) and LLMs (ChatGPT, Copilot, Gemini) in translating GFL.

The findings indicate that translations into Italian capture the original GFL intentions slightly better than those into English. Further, it was shown that LLMs provide significantly better and more accurate translations of GFL, with ChatGPT performing best for Italian and Copilot for English translations. However, the primary focus of translations remains on neutralisation and binary forms. Even when specifically prompted to produce translations that include non-binary individuals, the models continue to struggle.

1   Introduction

Being a translator is a demanding profession requiring not only mastery of source and target languages but also a deep understanding of cultural and societal nuances. Translators thus play a crucial role in mediating between cultures and communities (cf. Lardelli/Gromann 2023a: 172). They undergo years of training to translate across languages, cultures, and genres, and their work demands continual adaptation to interlinguistic and extralinguistic change.

A current societal and linguistic shift is the movement towards more gender-fair or gender-neutral language usage in some parts of the world. This change, which is often particularly visible in grammatical gender languages such as German, can pose challenges for translators (cf. ibid.: 166), especially due to the "multiplicity of proposed strategies and gender-specific grammatical differences across languages" (Paolucci et al. 2023: 13). The fact that languages handle and mark gender in different ways, as well as differing in terms of the presence or absence of gender features, implies that a one-size-fits-all solution is unlikely to work. Instead, explication and disambiguation are required (cf. Vanmassenhove 2024: 4). While gender-related issues have always been a subject of interest in translation studies (cf. Von Flotow 1997), and new paradigms such as feminist and queer translation have emerged over the past decades (cf. ibid.; Baer/Kaindl 2017), a lack of awareness beyond the binary remains apparent, both in public discourse and in translations (cf. Misiek 2020).

Furthermore, in daily life, not skilled and trained human translators but machine translation (MT) becomes increasingly prevalent. Over the years, MT has become more sophisticated and refined, establishing itself as a key tool for overcoming language barriers and facilitating cross-cultural communication (cf. García González 2024). In addition to commercial translation systems such as Google Translate and DeepL, large language models (LLMs) have gained mainstream popularity, particularly since the release of ChatGPT's user-friendly interface in late 2022 (cf. Vanmassenhove 2024: 9). Unlike traditional AI, generative AI based on LLMs can respond to prompts and create new content, including text, code, images (cf. Peres et al. 2023), and translations (for more details on the general translation capabilities of LLMs, cf. Hendy et al. 2023; Jiao et al. 2023; Wang et al. 2023).

However, while MT systems may be useful in everyday contexts and are therefore increasingly adopted on a larger scale (cf. Piazzolla/Savoldi/Bentivolgi 2023), concerns persist regarding their tendency to produce gender-biased translations that reinforce stereotypical gender roles (cf. García González 2024). These concerns have led to various studies investigating the relationship between gender and MT across different languages and tools.

Prates/Avela/Lamb (2020), for example, incorporated job positions into sentences such as "He/She is an engineer" and used the Google Translate API to translate these sentences into 12 gender neutral languages, including Hungarian, Chinese, and Yoruba. They reported a strong tendency towards male defaults, particularly in fields typically associated with gender imbalances or stereotypes (cf. ibid.).

Piazzolla/Savoldi/Bentivolgi (2023) studied translations from English into Spanish, Italian, and French in three commercial MT systems – Google Translate, DeepL, and Modern MT – reporting persistent gender bias within these systems. However, they found significant discrepancies in gender translation across the three systems, with DeepL handling feminine gender translation most effectively (cf. ibid.: 222).

Ghosh/Caliskan (2023) examined Bengali and five other low-resource languages, hence languages with a significant lack of data and tools for AI development, and their translations in ChatGPT. They found that ChatGPT perpetuates and even amplifies biases related to gender roles and stereotypes, particularly regarding actions (e. g., *cooking* being classified as a female activity) and occupations (e. g., *doctor* being assumed to be male). Moreover, MT was unable

to correctly interpret and translate the gender-neutral pronoun *they* into an equivalent gender-neutral pronoun in other languages (cf. ibid.).

Vanmassenhove (2024) investigated English-Italian translations in ChatGPT and found that biases persisted even when explicitly prompted to provide alternative translations (cf. Vanmassenhove 2024: 18). These findings suggest that systematic solutions for addressing gender bias in MT remain limited (cf. ibid.: 2). The results contradict OpenAI's claims that extensive bias mitigation measures have been implemented in ChatGPT (cf. Ghosh/Caliskan 2023), as current data indicate that male biases persist across systems (cf. Vanmassenhove 2024: 11f.).

One explanation for these findings is that training datasets for MT systems exhibit underrepresentation and/or misrepresentation of gender identities and gender-related terminology, particularly regarding male and female representation, with male forms being more prevalent (cf. Koehn 2005). Especially LLMs are trained on imbalanced datasets and therefore tend to reflect those imbalances back at us (cf. Kotek/Dockum/Sun 2023). Consequently, the systems tend to default to masculine translations in cases of ambiguity, increasing the likelihood of generating male-biased outputs (cf. Vanmassenhove 2024: 4). This suggests that "the reliance on biased data and statistical patterns in current MT systems can lead to erroneous disambiguation, resulting in the generation of morphologically incorrect alternatives and reinforcing potentially harmful stereotypes" (Vanmassenhove 2024: 18). Furthermore, it calls for tests to "ensure that they treat minoritized individuals and communities equitably" (Kotek/Dockum/Sun 2023: 1)

The studies discussed focus on gender from a binary perspective (male/female). Despite the recent development of queer translation studies (cf. Baer/Kaindl 2017), research in translation studies often neglects non-binary forms (cf. Lardelli/Gromann 2023b), a shortcoming acknowledged by various experts in the field of MT (cf. e. g., Savoldi et al. 2021; Costa-jussà/de Jorge 2020; Stefanovičs et al. 2020; Choubey et al. 2021; Lardelli/Gromann 2023b). Few studies have examined MT and non-binary gender representation, often focusing on media translation (cf. Attig 2022; Misiek 2020; Šincek 2020). Additionally, some studies investigated gender-neutral language. Piergentili et al. (2023), for instance, explored gender-neutral translations from English into Italian, given the significant gender-related linguistic transfer challenges in this language pair. Related studies include Vanmassenhove et al. (2021) and Sun et al. (2021), who worked on developing a gender-neutral rewriter for English.

With regard to German, some previous studies have investigated non-binary forms in MT from English to German (cf. Lardelli/Gromann 2023a, 2023b). A current project aims to develop a new resource for the automatic evaluation and modelling of gender-fair MT from English to German (cf. Lardelli et al. 2024). However, many studies, including those on German, have focused on translations from natural gender languages – mainly English – into grammatical gender languages such as German and Italian.

This paper, by contrast, takes German as the source language. In German, numerous approaches and strategies have been proposed, particularly for person nouns, to avoid the so-called masculine generic and to render the language more neutral or make women and non-binary individuals more visible (cf. Lardelli/Gromann 2023b). Since the masculine generic remains the default in many contexts, some gender-fair forms are still perceived as highly salient, implying that their usage reflects a conscious decision.

Consequently, when using MT, the question arises as to whether MT systems can detect and differentiate forms of GFL and provide translations that convey the intended meaning in German. In this context, it was considered relevant to examine how MT programmes handle GFL when translating from German, a grammatical gender language, into (a) Italian, another grammatical gender language, and (b) English, a natural gender language. This study, therefore, investigates various GFL strategies translated from German into English and Italian using six different MT tools, comprising three commercial translation programmes and three LLMs. The study aims to provide insights into whether the original intentions behind the use of neutral, non-binary, or binary forms are accurately reflected in the translations produced by these systems.

2   An overview of GFL in German, English, and Italian

Since German, English, and Italian differ structurally, their linguistic approaches to GFL also vary. German and Italian are both classified as grammatical gender languages. Hence, each noun has a gender (cf. Corbett 1991: 30), meaning that both personal and inanimate nouns are classified for gender and control the agreement of various other lexical categories, such as determiners, adjectives, or pronouns (cf. Gygax et al. 2019). In contrast, English is a natural gender language and thus does not typically classify nouns according to gender. However, third-person singular pronouns – *he, she,* and *it* – and certain nouns, such as *boy/girl*, are gender-specific (cf. Paolucci et al. 2023). Consequently, personal pronouns distinguish between female and male forms and are used to refer to male or female referents according to their referential sex or gender identity (e. g., *my teacher – she*; *your teacher – he*) (cf. Gygax et al. 2019).

This fundamental difference in language structure has implications for gender fairness. It was therefore considered relevant to examine the approaches employed by MT in translating GFL. In German and Italian, the so-called masculine generic is commonly used when referring to groups of people, as it is presumed to have a universal and neutral value that refers to individuals of any gender (cf. Schmitz/Schneider/Esser 2023: 1; Stahlberg et al. 2007; Nodari 2024). Since the masculine generic and the explicitly masculine form are orthographically and phonologically identical (cf. Schmitz/Schneider/Esser 2023: 1), concerns were raised in German as early as the 1970s regarding a potential male bias activated by the use of the masculine generic (cf. Trömel-Plötz 1978; Pusch 1979).

In response to this and to counteract potential bias while, at the same time, making women more visible in language, the use of the slash was introduced (e. g., *die Lehrer/innen* ('male and female teachers')) as an alternative to the masculine generic (*die Lehrer* ('male teachers/teachers')) (cf. Hensler 2021). In 1981, the capital-I was introduced for the same purpose (*LehrerInnen* ('male and female teachers')) (cf. Hensler 2021). However, the capital-I began to attract criticism for its binary nature, leading to the introduction of the gender gap around the turn of the century (*Lehrer_innen* ('all teachers')), which aims at making non-binary individuals structurally and linguistically visible (cf. ibid.). Since then, various neutralisation, binary, and non-binary strategies have been proposed to avoid the use of the masculine generic in the context of person nouns and to promote gender-fairness in German (cf. Table 1).

Strategy	Form	Examples singular	Examples plural	Translation
Neutralisation strategies	Present participle constructions	*der/die Lehrende*	*die Lehrenden*	'Teaching personnel'
	Gender-neutral nouns	*die Lehrkraft*	*die Lehrkräfte*	
Binary feminisation/ visualisation strategies	Paired form (masc.-fem.) Paired form (fem.-masc.)	*der/die Lehrer und Lehrerin;* *Lehrerin und Lehrer*	*die Lehrer und Lehrerinnnen;* *Lehrerinnnen und Lehrer*	'Male and female teacher(s)'
	Slash	*der/die Lehrer/in*	*die Lehrer/innen*	
	Slash Dash	*der/die Lehrer/-in*	*die Lehrer/-innen*	
	Dash	*der/die Lehrer-in*	*die Lehrer-innen*	
	Capital-I	*der/die LehrerIn*	*die LehrerInnen*	
	Bracket	*der/die Lehrer(in)*	*die Lehrer(innen)*	
Non-binary, queer strategies	Asterisk	*der*die Lehrer*in*	*die Lehrer*innen*	'Male, female, non-binary, queer teacher(s)'
	Colon	*der/die Lehrer:in*	*die Lehrer:innen*	
	Gap	*der/die Lehrer_in*	*die Lehrer_innen*	

Table 1: Generics, neutralisation, binary, and non-binary strategies in German

Albeit originally intended to counteract the male bias evoked by the masculine generic in the plural, gender-fair forms can also be used in the singular. However, it should be noted that while the use of gender-fair noun references in German is relatively unproblematic in the plural – since there is no indefinite article in this form, and the definite plural article is the same for all genders – the singular presents greater challenges concerning article use. For instance, while the present participle in the plural functions well as a neutralisation strategy, the use of the article in the singular leads to a binary interpretation of the form (cf. Table 1).

Not all proposed forms in Table 1 are equally accepted in public discourse. For example, neutralisation strategies, such as gender-neutral nouns, are widely accepted and promoted by some official guidelines (cf. Gesellschaft für Deutsche Sprache 2020). However, the use of special characters – particularly non-binary forms and the binary capital-I – is highly controversial. This is also reflected in the fact that the *Rat für deutsche Rechtschreibung* ('Council for German Orthography') does not recognise them as a core part of German orthography (cf. Rat für deutsche Rechtschreibung 2023). This controversy has sparked intense debate on GFL across various domains, from media and politics to linguistics. Newspapers regularly publish articles discussing the advantages and disadvantages of GFL usage (e. g., Müller-Spitzer 2022; Breuer 2022; Brisant 2025), some political parties and politicians have begun advocating for implementing bans (cf. Schomburg 2023; Straub 2024), and there has even been a court case on GFL

in Germany (cf. Olivares 2023). Furthermore, the reference dictionary *Duden* has been repeatedly criticised by the *Verein Deutsche Sprache* for "follow[ing] an ideology instead of treating the subtleties of the German language with sensitivity" (Verein Deutsche Sprache 2021).

The significant public interest in GFL is further reflected in the growing body of research on GFL in German across different contexts, such as its role in teaching German as a foreign language (cf. Peuschel 2022; Stark 2021; Lipsky 2023; Bachmaier 2021; Stummer 2024; Link 2022; 2024b), its use in administrative contexts (cf. Acke 2019; Müller-Spitzer/Ochs 2023), and its presence in German-speaking media (cf. Sökefeld 2021; Waldendorf 2023; Link 2024a, 2025).

Both the public debate and academic research primarily focus on person nouns. However, syntactical alternatives, such as rephrasing sentences using structures that do not indicate gender (e. g., passive constructions[1]), have also been proposed. Additionally, solutions for gender-fair language across all word classes have been suggested, including neo-systems and approaches that use gender-neutral endings or characters (for a comprehensive overview, cf. Lardelli/ Gromann 2023b: 219–224). However, these approaches are rarely discussed in the public debate.

In Italian suggestions to make the language gender-fairer have also been put forward. Initially, the debate on GFL in Italian primarily focused on the inclusion of women, particularly concerning the use of female forms for professional titles and high-ranking positions (cf. Gheno 2019/2022). However, in recent years, additional strategies have been proposed, such as omitting word endings or using special characters to make Italian genderless and by doing so more inclusive also towards non-binary individuals (cf. Table 2). Since all genderless strategies in Italian replace the gendered suffix with a special character or letter, there is no distinction between singular and plural, apart in case of the schwa, where it has been suggested to use <ə> for singular and <ɜ> for plural, though <ə> is often used for both singular and plural (cf. Gheno 2019/2022).

Similar to German, Italian also has proposals for restructuring sentences to avoid gender-inflected elements (cf. Vitiello 2022). However, no comprehensive publications providing an in-depth overview of the most common strategies for gender-fair language in Italian currently exist (cf. Lardelli/Gromann 2023b: 225). Amongst the suggested strategies in Italian, the asterisk, which was adopted by the LGBTQIA+ community in 2010, is one of the most well-known forms, its origins being described in Marotta/Monaco (2016) and featuring an entry in the grammar section of the *Enciclopedia Treccani* (s. v. *asterisco*). Another widely recognised form is the schwa (written as <ə>/<ɜ>), first suggested by Luca Boschetto in 2015 (cf. Nodari 2024: 244). In 2022, the Italian linguist Massimo Arcangeli launched a petition against the use of the schwa (cf. Arcangeli 2022). Nevertheless, the schwa's popularity has grown in recent years, and it now appears occasionally in some books and journal articles (cf. Lardelli/Gromann 2023b: 226). It is the only strategy that can be pronounced and as stated, distinguishes between singular and plural (cf. Marotta/Monaco 2016). The *Accademia della Crusca*, Italy's official language institute, opposes the use of special characters but recommends various strategies to

---

[1] Example: Masculine generic: *Sonderzahlungen erhalten die Mitarbeiter mit dem Gehalt.* Passive construction: *Sonderzahlungen werden mit dem Gehalt ausgezahlt.*

make women more visible or to avoid gender-marked forms that create gender asymmetry (cf. Zennaro 2023).

Strategy	Form	Examples singular	Examples plural	Translation
Masculine generic	Masculine generic	*il maestro*	*i maestri*	'Male teacher(s)/male and female teacher(s)'
Binary feminisation/visualisation strategies	Pair form	*il maestro e la maestra*	*i maestri e le maestre*	'Male and female teacher(s)'
Genderless strategies	Asterisk	*l* maestr**	*l* maestr**	'Male, female, non-binary, queer teacher(s)'
	At	*l@ maestr@*	*l@ maestr@*	
	Schwa	*Lə maestrə*	*Lɜ maestrɜ*	
	Dash	*l- maestr-*	*l- maestr-*	
	Letter U	*lu maestru*	*lu maestru*	
	Letter X	*lx maestrx*	*lx maestrx*	
	Letter Y	*ly maestry*	*ly maestry*	
	Ending omission	*l maestr*	*l maestr*	

Table 2: Generics, binary, and non-binary strategies in Italian

While both German and Italian use special characters, their strategies differ. In Italian, special characters are used to make the language genderless and thus inclusive of everyone. In German, special characters can be used both binarily and non-binarily, while other strategies are employed to make the language neutral. As a result, German offers various binary, non-binary, and neutralisation strategies, whereas Italian provides only one binary strategy alongside multiple genderless strategies. Another difference is of structural nature. German employs special characters within words, positioning them between the stem and suffix, whereas in Italian, these characters replace gendered suffixes to create a neutral word ending. Consequently, Italian does not typically distinguish between singular and plural – except in the case of the schwa. Furthermore, while German's non-binary special characters can be pronounced using a glottal stop, the schwa is the only non-binary strategy in Italian that has an audible form.

Despite English being a natural gender language and thus largely not distinguishing explicitly between male and female in person nouns, concerns regarding gender bias are not unheard of and have attracted attention since the 1980s (e. g., Sniezek/Jazwinski 1986). In addition to the use of masculine words as generics to describe both men and women in certain cases (cf. ibid.), English also sometimes employs marked forms, creating what is known as gender asymmetry (cf. Hellinger/Bußmann 2001). This is particularly evident in job titles; for instance, in the distinction between *steward* and *stewardess*, where the female form is specifically marked (cf. Prewitt-Freilino et al. 2011). One approach to addressing this issue is the use of gender-neutral terms, such as *flight attendant* instead of *steward/stewardess*. However, in some cases it was found that even the use of gender-neutral terminology (e. g., *member of Congress*) may implicitly convey a gendered interpretation more frequently than gender-symmetrical terminology (e. g., *congressman/congresswoman*) (cf. Prewitt-Freilino et al. 2011). Similar findings apply

to pronouns: replacing masculine generics with gender-symmetrical pronouns, such as *he/she*, has been shown to lead to a greater visualisation of women (cf. Hyde 1984; Switzer 1990).

While these strategies help to make women more visible in English, the language did lack a designated pronoun for individuals who identify outside of the gender binary (cf. Grove 2021). To address this gap, the use of singular *they* was introduced (cf. ibid.). Other non-standard non-binary pronouns include *ze/zir* (cf. ibid.). However, *they* is by far the most widely used non-binary referent (cf. Baron 2020: 126; Hekanaho 2020) and given that speakers are already familiar with singular *they*, the form has a stronger potential for standardisation in English (cf. Baron 2020).

Thus, the difference in linguistic structure underpins the varying strategies the languages employ regarding gender-fair language (GFL), creating difficulties regarding translations. Nevertheless, across all three languages – German, Italian, and English – gender distinctions can be evident in both person nouns and pronouns, albeit to different degrees. Furthermore, the fact that all three languages have seen proposals aimed at increasing the visibility of women and promoting inclusivity for non-binary individuals suggests a growing awareness of gender-related language choices within their respective societies. This implies that the decision to adopt more gender-fair language and to enhance the visibility of all individuals is a conscious choice made by some speakers of these languages.

3 Methodology

3.1 Selection of MT programmes and LLM

The aim of this study is to determine whether commercial machine translation programmes and LLMs are capable of detecting GFL in German person nouns. If so, it further examines how they handle neutralisation, binary, and non-binary strategies, and whether their approaches differ. To this end, freely accessible, well-known web-based translation programmes and LLM-based chatbots, widely used by a large number of people, were selected. The three translation programmes chosen for the analysis are DeepL, Google Translate, and Bing, as they represent leading MT systems. DeepL currently supports 30 languages, including several regional varieties, and is used by more than half a billion individuals (cf. DeepL). Google Translate provides support for over 130 languages, making it the largest existing MT provider (cf. Piazzola et al. 2023: 212). Bing offers translation in 179 languages and language varieties, in addition to text-to-speech functionality in 76 languages (cf. Microsoft 2025).

The LLM-based chatbots selected are ChatGPT, Gemini, and Copilot. In contrast to MT systems, generative AI chatbots based on LLMs are not designed exclusively for translation but are capable of producing text, generating images, and performing translations. Within LLM research, a distinction is commonly drawn between bidirectional Transformer architectures – for example, Bidirectional Encoder Representations from Transformers (BERT) (cf. Devlin et al. 2019) – and unidirectional models such as the GPT series (cf. Vanmassenhove 2024: 9). Unidirectional models process and generate text sequentially, from left to right, producing output token by token (cf. ibid.). By contrast, bidirectional models consider both preceding and succeeding tokens when processing input and generating output, thereby exploiting contextual

information in both directions (cf. ibid.). Although both architectures can be applied to MT, the present study focuses exclusively on unidirectional models.

Among the selected chatbots, GPT-3 was the first version to enable direct user interaction (cf. Marr 2023). Its success generated significant scholarly interest, particularly within translation studies (cf. e. g., Ghosh/Caliskan 2023; Lyu et al. 2023; Vanmassenhove 2024) and prompted the development of competing systems. Microsoft subsequently released Microsoft 365 Copilot, a tool integrated into Microsoft 365 applications and services to enhance workplace productivity (cf. Mauran 2023). In early 2023, Google introduced Bard, an experimental conversational AI system (cf. Comstock 2025). However, shortly after its release, Bard was criticised for producing inaccurate responses. In December 2023, Google launched an updated version powered by the Gemini LLM, which the company claims outperforms GPT-4 (cf. ibid.). In early 2024, Bard was rebranded as Gemini, and the web application was expanded from English-only availability to 40 languages (cf. ibid.).

Unlike commercial MT systems, LLM-based chatbots impose usage restrictions. For example, the free version of ChatGPT limits users to 40 messages per three-hour interval. Gemini's restrictions vary depending on prompt length and complexity, while Copilot imposes a monthly limit of 50 chat messages (cf. GitHub).

3.2   Test sentences

In order to determine whether and how MT systems translate German gender-fair noun references, test sentences were created. As a first step, previous studies on the use of GFL in German-speaking media were consulted to select the most commonly used neutralisation, binary, and non-binary strategies (cf. Link 2024a, 2025; Waldendorf 2023). Consequently, the following strategies from Table 1 were selected: gender-neutral nouns as neutralisation strategies, the pair form and the capital-I as binary strategies and the asterisk and colon as non-binary strategies (cf. Table 3). Despite its frequent use, it was decided to exclude the participle form given the problems that occur in its singular usage (cf. section 2).

The sentences had to be simple to not to create unnecessary complications during the translation process. *Lehrerin/Lehrer* ('teacher') was selected as the subject, as it can also be used as a gender-neutral noun (*Lehrkraft/Lehrperson*) (cf. Lardelli/Gromann 2023b: 218). This choice is particularly relevant for the Italian translation. The most common Italian equivalent for 'teacher' is *insegnante*. In Italian nouns ending in *-e* gender is only indicated by the article. However, since *insegnante* begins with a vowel, the definite singular articles *la* and *il* contract to *l'*. In this case, gender cannot be deduced from either the noun or the article. Instead of *l'insegnante*, alternative terms such as *la maestra/il maestro*, *la professoressa/il professore*, or *la/il docente*, *l'istruttore/l'istruttrice* can be used, all listed as synonyms for *l'insegnante* (cf. *Enciclopedia Treccani,* s. v. *insegnante*). In all these cases, gender is explicitly marked, either by the article alone or by the article in combination with noun suffixes. Thus, Italian offers the possibility of selecting a gender-unmarked translation for the singular definite article while also providing marked alternatives.

After many considerations it was decided to use *Studentin/Student* ('student') as the object, even though no gender-neutral noun equivalent to *Lehrkraft* exists. Hence, it was decided to

use the expression *studierende Person* instead despite the fact that this form is not necessarily well-established (cf. Table 3).

As this study focuses solely on person nouns, the test sentences were designed to exclude pronouns. The sentences were intended to be short, simple, and straightforward. For this reason, the present tense was chosen (cf. Table 3). All sentences were constructed in both singular and plural forms, using both definite and indefinite articles (cf. Table 3). As noted in Section 2, while the use of gender-fair noun references in German is relatively unproblematic in the plural, the singular poses more challenges concerning article use. Although numerous guidelines and recommendations exist for the use of gender-fair person nouns in German, as for instance *Gendern – ganz einfach!* (Diewald/Steinhauer 2019) from the Duden publishing house, the appropriate use of singular articles is rarely discussed. One suggestion for special-character-based forms is the use of both articles, separated by the respective special character (*die*der, die:der, eine*ein, eine:ein*) (cf. Feministisch Sprachhandeln der Humboldt-Universität zu Berlin 2015: 16). This approach was also adopted for the non-binary forms analysed in this study (cf. Table 3).

For the capital-I, articles were separated by a slash (*die/der*, *der/dem*, *eine/ein*, *einer/einem*, etc.). Although an alternative suggestion involves merging *ein* and *eine* into *einE* (cf. Sill 2022), this approach creates grammatical difficulties beyond the nominative and accusative case, as the inflectional endings required for the dative and genitive cannot be merged as seamlessly. For the pair form, constructions using both *oder* and *und* (*der oder die*, *der oder dem*, *eine oder ein*, *einer oder einem*) were tested (cf. Table 3).

Strategy	Form		Definite article	Indefinite article
Neutralisation	Gender-neutral	Sg.	1a) *Die Lehrkraft erklärt der studierenden Person das Thema.*	1c) *Eine Lehrkraft erklärt einer studierenden Person das Thema.*
		Pl.	1b) *Die Lehrkräfte erklären den studierenden Personen das Thema.*	1d) *Lehrkräfte erklären studierenden Personen das Thema.*
Feminisation/ binary	Pair form	Sg.	2a) *Der Lehrer oder/und die Lehrerin erklärt der Studentin oder/und dem Studenten das Thema.*	2c) *Ein Lehrer oder/und eine Lehrerin erklärt einer Studentin oder/und einem Studenten das Thema.*
		Pl.	2b) *Die Lehrer oder/und die Lehrerinnen erklären den Studentinnen oder/und den Studenten das Thema.*	2d) *Lehrer oder/und Lehrerinnen erklären Studentinnen oder/und Studenten das Thema.*
	Capital-I	Sg.	3a) *Die/der LehrerIn erklärt der/dem StudentIn das Thema.*	3c) *Eine/ein LehrerIn erklärt einer/einem StudentIn das Thema.*
		Pl.	3b) *Die LehrerInnen erklären den StudentInnen das Thema.*	3d) *LehrerInnen erklären StudentInnen das Thema.*

Strategy	Form		Definite article	Indefinite article
Non-binary	Asterisk	Sg.	4a) *Die*der Lehrer*in erklärt der*dem Student*in das Thema.*	4c) *Ein*eine Lehrer*in erklärt einer*einem Student*in das Thema.*
		Pl.	4b) *Die Lehrer*innen erklären den Student*innen das Thema.*	4d) *Lehrer*innen erklären Student*innen das Thema.*
	Colon	Sg.	5a) *Die:der Lehrer:in erklärt der:dem Student:in das Thema.*	5c) *Eine:einer Lehrer:in erklärt einer:einem Student:in das Thema.*
		Pl.	5b) *Die Lehrer:innen erklären den Student:innen das Thema.*	5d) *Lehrer:innen erklären Student:innen das Thema.*

Table 3: Sample sentences

3.3 Categorisation of translations

The respective test sentences were entered into commercial translation programmes, while LLM-based chatbots were always given the prompt *Übersetze ins Englische/Italienische* ('Translate into English/Italian') followed by the respective sentences.

The translations were categorised into "complete", "partial", or "no recognition" of gender-fair forms. The categories were determined prior to translating the sentences. The annotation was conducted by one researcher. While it was considered to ask a second, independent annotator to review and classify the translations, ultimately it was found that the annotation process is rather objective, and a second annotation is not needed. To allow for more transparency, some annotation examples will be elaborated in the following.

Google Translate's translation of *Ein Lehrer und eine Lehrerin erklären einer Studentin und einem Studenten das Thema* as 'A teacher explains the topic to a student'*,* for instance, was classified as "no recognition" because it failed to reflect the feminisation through the explicit mention of both masculine and feminine forms in the source sentence. While it could be argued that *teacher* and *student* in English, being a natural gender language, are neutral and can be used for men and women, the original intention of the German sentence has not been considered at all. Similarly, the Italian Google Translate translation for *Ein:eine Lehrer:in erklärt einer:einem Student:in das Thema* as 'Un insegnante spiega l'argomento a uno student', was also classified as "no recognition" since it completely ignored the non-binary strategy employed in the original sentence.

Copilot translated the same sentence as 'Un*una insegnante spiega l'argomento a uno*una student'. Additionally, Copilot provided the following explanatory note: "This translation maintains a gender-neutral perspective, similar to the original German sentence". While this translation exhibits a certain awareness of GFL, it misinterprets the use of the colon (*Lehrer:in*) as a gender-neutral marker rather than a non-binary form. Consequently, this translation was categorised as "partial recognition". Similarly, translations such as 'Un/a insegnante spiega l'argomento a uno/a studente/ssa' for originally non-binary sentences were classified as "partial recognition". Although the translation considers gender-related language decisions, it remains restricted to a binary gender framework rather than a non-binary one. Further, ChatGPT's translation 'The teacher explains the topic to the female student or the male student' for *Der Lehrer oder die Lehrerin erklärt der Studentin oder dem Studenten das Thema* was categorised as "partial recognition", since both genders were explicitly mentioned for *students*, whereas the same

was not done for *teachers*. It should be noted that in Italian, singular translations such as 'L'insegnante spiega l'argomento allo studente' for binary and neutralisation forms were categorised as "partial recognition", given that *l'insegnante* is not gender-marked (cf. section 3.1). While it could be argued that the MT simply used the most common form without the intention of being gender-fair, some translations instead employed explicitly marked forms such as *il/la docente*.

The translation of *Eine/ein LehrerIn erklärt einer/einem StudentIn das Thema* provided by Gemini – 'Un/una docente spiega un argomento a uno/a student' – was categorised as "complete recognition" as the binary nature of the source sentence was accurately reflected in the translation. Likewise, the English translation from ChatGPT for *Die Lehrkräfte erklären den studierenden Personen das Thema* as 'The teaching staff explain the topic to the studying individuals' was classified as "complete recognition", given that the translation captures the neutrality of the original meaning.

4   Results

Overall, the results indicate that translations produced by the LLM-based chatbots are significantly better than those from commercial translation programmes, both in Italian and in English (cf. Table 4). This is partially due to the fact that, while some translation programmes provide alternative translations, generative AI not only allows for multiple translations but also facilitates the creation of additional explanations.

Italian	Singular			Plural			Overall		
	Complete	Partially	Not at all	Complete	Partially	Not at all	Complete	Partially	Not at all
Gemini	4	4	4	4	3	5	8	7	9
ChatGPT	8	4	0	7	5	0	15	9	0
Copilot	6	4	2	4	1	7	10	5	9
DeepL	0	4	8	0	0	12	0	4	20
Bing	0	4	8	0	0	12	0	4	20
Google Translate	0	4	8	0	0	12	0	4	20
English	Singular			Plural			Overall		
	Complete	Partially	Not at all	Complete	Partially	Not at all	Complete	Partially	Not at all
Gemini	4	6	2	1	4	7	5	10	9
ChatGPT	3	1	8	5	1	6	8	2	14
Copilot	4	4	4	4	3	5	8	7	9
DeepL	0	0	12	0	1	11	0	1	23
Bing	0	0	12	0	0	12	0	0	24
Google Translate	0	0	12	0	0	12	0	0	24

Table 4: MT in Italian and English completely, partially, or not reflecting gender-fair German

The three commercial translation programmes generally did not differentiate between the various gender-fair forms and strategies (cf. Table 4). In Italian, only a total of six different translations for the 24 test sentences were produced by all three programmes, depending on whether the German original sentence contained a definite article in the singular (Example 1), an indefinite article in the singular (Example 2), or was written in the plural (Example 3).

(1) 'L'insegnante spiega l'argomento allo studente.'/'Il docente spiega l'argomento allo studente.'
(2) 'Un insegnante spiega l'argomento a uno studente.'/'Un docente spiega l'argomento a uno studente.'
(3) 'Gli insegnanti spiegano l'argomento agli studenti.'/'I docenti spiegano l'argomento agli studenti.'

The object is consistently translated using the masculine generic throughout (cf. Examples 1–3). While in Example 1, *l'insegnante* cannot be clearly assigned as masculine or feminine (cf. Section 3.2), Bing also produced a translation using *il docente*, thereby clearly marking the sentence as masculine generic. The translations with the indefinite article and in the plural exclusively employ masculine forms (Examples 2 and 3).

Similarly, in translations from German to English, Google Translate and Bing produce only three different translations, depending on whether the definite (Example 4), indefinite (Example 5), or plural article (Example 6) was used in the original sentence, ignoring any gender-fair forms. Further, when special characters were used in Google Translate, the programme frequently displayed the message: "Did you mean *Lehrerinnen*?" – thus marking the use of special characters as an error or typo.

(4) 'The teacher explains the topic to the student.'
(5) 'A teacher explains the topic to a student.'
(6) 'The teachers explain the topic to the students.'

Only DeepL provides occasionally alternative sentences to the translations in examples 4–6.[2] Namely, the programme suggests translating *Lehrkräfte* as 'teaching staff'. Hence, while it could be argued that due to the structurally differences gender-fairness is achieved through other means in English (cf. Section 2), the fact that DeepL suggests an alternative translation for the German gender-neutral noun *Lehrkräfte* ws that the programme, in this case, seeks a similarly neutral expression as in the German original. In fact, a recent press release, as well as an external evaluation (cf. Papa/Tavosanis 2020), suggests that the translation quality provided by DeepL for some languages may surpass that of its competitors (cf. Piazzola et al. 2023: 212).

While the LLMs performed considerably better, their performance varied depending on the specific LLM (cf. Table 4). ChatGPT emerged as the best-performing model for Italian and Copilot for English. In Italian, in more than half of the cases (15 out of 24), the translations provided by ChatGPT reflected the original intention in German (cf. Example 7). In nine out of 24 cases, ChatGPT partially accounted for the gender-fair language used in the original sentence (cf. Example 8). In this example, the programme offers a binary differentiation but does

---

[2] It should be noted that in most cases when DeepL suggests alternative translations, the programme provides alternatives for *das Thema* rather than for the subject and object of the sentence.

not reflect the non-binary form used in the original sentence. In none of the translations provided by ChatGPT the gender-fair strategies were completely ignored.

(7) *LehrerInnen erklären StudentInnen das Thema.*
     'I/Le professori/professoresse spiegano il tema agli studenti/alle studentesse.'

(8) *Ein*eine Lehrer*in erklärt einer*einem Student*in das Thema.*
     'Un/a insegnante spiega l'argomento a uno/a studente/ssa.'

Copilot and Gemini performed relatively similarly in their German-Italian translations, with the former producing ten out of 24 translations that reflected the original intention and the latter eight out of 24 (cf. Examples 9 and 10).

(9) *Der/die LehrerIn erklärt der/dem StudentIn das Thema.*
     'L'insegnante spiega l'argomento allo studente o alla studentessa.' (Copilot)

(10) *Ein Lehrer und eine Lehrerin erklärt einer Studentin und einem Studenten das Thema.*
     'Un insegnante e un'insegnante spiegano l'argomento a una studentessa e a uno studente.' (Gemini)

Unlike ChatGPT they provided various alternative translations and explanations. Contrary to the researcher's expectation, this often had a negative effect, as incorrect explanations were given. In case of example 11, Copilot suggests the use of the asterisk in the Italian translation which is usually used to replace the endings and make the word genderless (cf. Table 2). However, in the provided translation the asterisk is used to separate the male and female articles while leaving the nouns unaltered. Thus, in this case the use of the asterisk in Italian does lead to a binary rather than a non-binary translation. Furthermore, the programme incorrectly claims that this translation maintains a gender-neutral perspective from the original sentence, misidentifying the colon as a neutral rather than a non-binary form. Copilot applied this same erroneous explanation to most translations, regardless of the form translated.

(11) *Ein:eine Lehrer:in erklärt einer:einem Student:in das Thema.*
     'Un*una insegnante spiega l'argomento a uno*una studente.' (This translation maintains a gender-neutral perspective, similar to the original German sentence.)

While Gemini gives occasionally more elaborate differentiations, the explanations usually ignored translations beyond the binary. Hence, the programme also provides binary explanations for non-binary source sentences (cf. example 12).

(12) *Die*der Lehrer*in erklärt der*dem Student*in das Thema.*
     Options that reflect inclusivity (with comments):
     'La/il docente spiega l'argomento alla/o student.' (This option explicitly includes both the male and female genders. However, in Italian, this form may seem a bit forced and less natural. It is more common to use the masculine generic, i. e. to use the male term to refer to both genders.).
     'L'insegnante (uomo o donna) spiega l'argomento allo studente (uomo o donna).' (This option is more explicit but may be a bit long and repetitive).'

In English, Copilot slightly outperforms ChatGPT. Both models provide accurate translations in eight out of 24 cases (cf. Example 13 and 14).

(13) *Der Lehrer und die Lehrerin erklärt der Studentin und dem Studenten das Thema.*
     'The male teacher and the female teacher explain the topic to the female student and the male student.' (Copilot)

(14) *Die Lehrkräfte erklären den studierenden Personen das Thema.*
      'The teaching staff explain the topic to the studying individuals.' (ChatGPT)

However, in more than half of the cases (14 out of 14), the translations provided by ChatGPT did not reflect the original gender-fair language adequately and special characters were largely ignored by the model (cf. Example 15).

(15) *Die LehrerInnen erklären den StudentInnen das Thema.*
      'The teachers explain the topic to the students.'

Gemini performed worst across the three LLMs, capturing the original use of GFL in only five out of 24 translations. However, in some instances, Gemini explicitly acknowledged that certain forms, such as the colon, are gender-inclusive (cf. Example 16), suggesting gender-neutral terms as an English equivalent.

(16) The original German sentence uses gender-inclusive language (*Lehrer:in* and *Student:in*). While English doesn't have identical grammatical structures for this, the translations above use gender-neutral terms ('teacher,' 'instructor,' 'student') which are the standard and most inclusive way to express this in English.

It is also noteworthy that, when asked to provide a translation for the colon, Gemini offered various options, mixing different forms in the respective explanations (example 17).

(17) *Ein:eine Lehrer:in erklärt einer:einem Student:in das Thema.*
      1) 'A professor explains the topic to a student.' (*Wenn es sich um einen **Professorin** handelt* ('when it is about a female professor')),
      2) 'An educator explains the topic to a pupil.' (*Wenn es sich um **Schüler*innen** handelt* ('when it is about students of all genders')).

Within the framework of the analysis, it was also examined whether the LLMs performed better when translating original sentences containing definite or indefinite articles. Table 5 illustrates how frequently GFL was fully recognised and translated across the LLMs in sentences with definite and indefinite articles, in both singular and plural. In Italian, gender-fair person nouns with definite articles were slightly more accurately translated (cf. Table 5). The same does not apply for English, however. Further, the LLMs captured GFL relatively equally well between sentences in singular and plural.

	Definite Article		Indefinite Article	
Italian	Singular	Plural	Singular	Plural
Gemini	1	3	2	2
ChatGPT	4	4	4	3
Copilot	4	3	2	2
	Definite Article		Indefinite Article	
English	Singular	Plural	Singular	Plural
Gemini	2	2	1	0
ChatGPT	1	2	3	2
Copilot	2	2	2	2

Table 5: Full recognition of GFL across LLM in singular and plural for the definite and indefinite article

Two further observations were made when annotating the translations. First, various inconsistencies can be found across translations from the same models. For instance, Gemini sometimes provides multiple translations along with explanations. In other cases, only a single translation is provided. Occasionally, translations for individual words are given first and then combined into a single sentence. Secondly, Copilot and Gemini often provide explanations for their respective translations in the target language rather than in German, despite being given a German prompt. Gemini only occasionally provides explanations in German but then switches back to English in subsequent translations.

5   Discussion

The results show that LLMs perform much better than conventional translation programmes when it comes to translating different strategies of GFL from German to Italian and English. Commercial translation programmes hardly ever provided translations reflecting the original gender-fair intention and hence while Google Translate, for instance, claims that they have recently started to provide a male and female output for words where the gender is unclear (cf. Měchura 2023), this cannot be supported by the findings of this study. It is also noteworthy that when inserting special characters in Google Translate, the programme frequently displayed an error message, marking the use of such characters as an error or typo. Hence, none of the three commercial translation programmes could capture the original intentions of the different GFL strategies in German in their translations. Consequently, a significant amount of information contained in the respective original German sentences was lost in translation.

The much better performance from LLMs might not come as a surprise, given how elaborate they are and given that with every search query they are simultaneously trained. However, the three LLMs in question differed in how well they translated the forms, with ChatGPT outperforming the other two in Italian translations and Copilot slightly outperforming ChatGPT in English.

Regarding the two languages, the LLMs perform a bit better in Italian than in English and the discrepancy of translation quality across the LLMs is more pronounced for German-Italian than for German-English. This could be due to the greater structural similarities of the two languages. In Italian, translations based on nouns with the definite article was overall captured slightly more accurately across programmes (cf. Table 5). There are no differences in the quality of translations in singular and plural. The LLMs applied different strategies to translate the respective forms. While ChatGPT provided only one translation per sentence, Copilot and Gemini offered additional explanations and alternative translations. However, these explanations are primarily given in the target language, rendering them largely unhelpful, as translations are typically required when users have limited proficiency in the target language. To receive explanations in the source language, the user must explicitly request them.

Furthermore, the same explanations were often used across different forms, completely disregarding distinctions between neutralisation, binary, and non-binary strategies. In fact, the programmes struggled with any forms beyond the binary. For instance, in Italian, special characters were not used at all, except in some cases in articles. Hence, while binary and neutralisation strategies are sometimes recognised by the programmes and reflected in the translations, non-binary forms have largely been ignored. More generally, there is a lack of differentiation

between strategies, and neutralisation strategies often appear to be treated as equivalent to inclusive strategies. This lack of distinction becomes particularly evident in Example 17 in the results section, where Gemini provides seemingly arbitrary additional explanations, suddenly incorporating the feminine noun *Professorin* and the non-binary asterisk *Schüler*innen*.

Thus, while the results indicate that LLMs perform significantly better in recognising and translating GFL, they struggle with clearly differentiating the various strategies and especially with translating non-binary forms adequately. This raises the question of whether the three LLMs can produce non-binary translations when explicitly prompted to do so, particularly since previous studies suggest that specific prompts might even have an opposite effect (cf. Vanmassenhove 2024: 18). Therefore, based on the results it was decided to test what LLMs do, when explicitly prompted to produce a translation including and addressing also non-binary individuals.

6 Beyond the binary: LLMs and the translation of specific non-binary prompts

The three LLMs investigated in this study were all given the prompt *Übersetze den Satz ins Italienische/Englische, sodass auch nicht-binäre Personen inkludiert sind und sich angesprochen fühlen* ('Translate this sentence into Italian/English so that also non-binary people are included and feel addressed'). The German sentences used were sentences 4a–b and 5a–b (cf. Table 3), given that the programmes performed slightly better in sentences containing the definite article in Italian (cf. Section 4).

The results in Table 6 show that the three LLMs apply different strategies in Italian when given the specific prompt. Copilot argues that the masculine form *lo studente* is gender-neutral and addresses everyone (Table 6). Gemini, similar to the results in Section 4, offers various alternatives and explanations. The main focus of the translations provided is less on actively addressing non-binary individuals and more on making the sentences gender-neutral. The third translation provided ('L'insegnante spiega l'argomento allo studente, alla studentessa o alla persona che studia') can be seen as an attempt to explicitly include non-binary people; however, this is only true for the object, not for the subject of the sentence.

ChatGPT is the only programme that makes use of special characters (cf. Table 6). However, like Gemini, ChatGPT focuses on the object of the sentence. For the subject, the programme chooses *docente* as a translation and, while the suffix -*e* can be used for men and women (cf. Section 2), the translation closely follows the German source, maintaining a binary article usage in Italian. The same applies to the object of the sentence. While in German, the use of a special character between the two articles is suggested to render non-binary individuals visible, in Italian, special characters are usually used to replace the vowels in the article (cf. Section 2). Hence, the lack of integration of special characters into the article in the translations provided means that a binary article system is combined with the use of the non-binary schwa (cf. Table 6).

Italian	Gemini	ChatGPT	Copilot
Singular			
*Die*der Lehrer*in er-klärt der*dem Stu-dent*in das Thema.* *Die:der Lehrer:in er-klärt der:dem Stu-dent:in das Thema.*	'L'insegnante spiega l'argomento alla persona che studia.' This is a neutral option that uses non-gender specific language. 'Chi insegna spiega l'argomento a chi studia.' This is another neutral option with more general wording. 'L'insegnante spiega l'argomento allo studente, alla studentessa o alla persona che studia.' This is a broader option that includes both traditional gender and a non-binary option. 'Il docente spiega l'argomento all'alunno/a.'	'La:il docente spiega l'argomento alla:allo studentə.' (This form makes both male and female forms visible and includes non-binary persons through the <ə>).	'L'insegnante spiega l'argomento allo studente.' (The sentence uses the gender-neutral word *studente* for student and is formulated in such a way that everyone feels addressed.)
Plural			
	Gemini	ChatGPT	Copilot
*Die Lehrer*innen er-klären den Stu-dent*innen das Thema.*	'Le persone che insegnano spiegano l'argomento alle persone che stu-diano.' This is a neutral option that uses non-gender specific language. 'Chi insegna spiega l'argomento a chi studia.' This is another neutral option with more general wording. 'I docenti spiegano l'argomento agli alunni/alle alunne/alle persone che studiano.' This is a broader option that includes both traditional gender and a non-binary option.	'Le persone docenti spiegano l'argo-mento alle persone studentə.' (By *persone docenti* and *persone studentə* a gender-neutral form is used, whereby <ə> is a more inclusive spelling for non-bi-nary persons).	'Ile insegnanti spie-gano l'argomento agli/le studenti/esse.' (This formulation takes into account both non-binary and binary per-sons and uses gender-neutral terms.)
Die Lehrer:innen er-klären den Student:in-nen das Thema.	'Coloro che insegnano spiegano l'argomento a coloro che studiano.' This is a more formal and neutral option.	'I: le docenti spie-gano l'argomento agli:alle studentə.'	

Table 6: Italian translations following prompt to specifically include non-binary individuals

The plural translations are relatively similar to the singular ones. Once more, Gemini offers various options but mainly focuses on neutralisation rather than explicitly non-binary strategies (cf. Table 6). Copilot performs slightly better in the plural, as it does not exclusively use the masculine generic; however, the programme only provides a binary translation, which it claims to be non-binary. ChatGPT also uses the schwa in the plural but only for the German sentences with the asterisk but always in combination with binary articles and only for the direct object. For the translation of the asterisk, ChatGPT further combines neutralisation, by means of the word *persone*, with the non-binary schwa (cf. Table 6).

While in Italian, special characters exist to render the language genderless and include non-binary individuals, in English, other means would be required to indicate that a sentence should

include non-binary people. Table 7 shows that the models do not offer different translations for the various German source sentences. Copilot and ChatGPT both provide only one translation for all sentences in singular and plural, stating that *teacher*(*s*) and *student*(*s*) are both gender-neutral and therefore also include non-binary individuals. Gemini, once more, offers multiple translations and explanations, demonstrating a higher performance than the other two LLMs in this regard. While the model mainly produces gender-neutral translations, it also suggests integrating the pronouns *they* and *their* (cf. Table 7). Additionally, Gemini provides guidance on which of the presented translations are most suitable when aiming to be explicitly gender-neutral (cf. Table 7). The model further advises that the words *educator* and *instructor* are additional neutral alternatives for *teacher*, implying that *teacher* is not inherently neutral. Similarly, *learner* is classified as a neutral term, suggesting that *student* might not be entirely neutral.

English	Gemini	ChatGPT	Copilot
	Singular		
Die*der Lehrer*in erklärt der*dem Student*in das Thema. *Die:der Lehrer:in erklärt der:dem Student:in das Thema.*	'The teacher explains the topic to the student.' (This is the most common and uncomplicated translation. However, it is not explicitly gender-neutral.) 'The teacher explains the topic to the learner.' (This option is slightly more neutral, as *learner* is a gender-neutral word.) 'The teacher explains the topic to the student, they are learning.' (This option uses the gender-neutral pronoun *they* in relation to the student.) 'The educator explains the topic to the student.' (*Educator* is another neutral word for *teacher*.) 'The instructor explains the topic to the student.' (*Instructor* is another neutral word for *teacher*.) 'The teacher explains the topic to their student.' (This is another way of using the gender-neutral pronoun *their*.) 'The educator explains the topic to the learner.' (Two neutral terms are used here.) (The best option depends on the context and the desired emphasis. If you want to be explicitly gender-neutral, the third, sixth or seventh options are the best choice.)	'The teacher explains the topic to the student.' (In English, the words *teacher* and *student* are gender-neutral, so this sentence includes non-binary individuals without the need for further modification.)	'The teacher explains the topic to the student.' (This formulation uses gender-neutral terms and takes into account non-binary persons.)

Plural			
	Gemini	ChatGPT	Copilot
Die Lehrer*innen erklären den Student*innen das Thema. *Die Lehrer:innen erklären den Student:innen das Thema.*	'The teachers explain the topic to the students.' (This is the most common and uncomplicated translation. However, it is not explicitly gender-neutral.) 'The teachers explain the topic to the learners.' (This option is slightly more neutral, as *learners* is a gender-neutral word.) 'The teachers explain the topic to the students, they are learning.' (This option uses the gender-neutral pronoun *they* in relation to the students.) 'The educators explain the topic to the students.' (*Educators* is another neutral word for *teachers*.) 'The instructors explain the topic to the students.' (*Instructors* is another neutral word for *teachers*.) 'The teaching staff explain the topic to the students.' 'Those who teach explain the topic to those who are learning.' 'The teachers explain the topic to their students.' (This is another way of using the gender-neutral pronoun *their*.) (The best option depends on the context and the desired emphasis. If you want to be explicitly gender-neutral, the third or eighth option is the best choice.)	'The teachers explain the topic to the students.' (In English, *teachers* and *students* are gender-neutral terms, so this sentence includes non-binary individuals without needing further modification.)	'The teachers explain the topic to the students.' (This formulation takes into account non-binary persons and uses gender-neutral terms.)

Table 7: English translations following prompt to specifically include non-binary individuals

7   Conclusion

This study aimed to gain a better understanding of whether, and to what extent, MT can recognise GFL and translate it in a way that preserves the original intention. It became clear that LLMs perform significantly better than conventional translation programmes when translating GFL. However, in particular non-binary forms were frequently translated inadequately, resulting in a loss of the original intention.

Even when explicitly prompted to produce translations that address and include non-binary individuals, the LLMs demonstrated only limited improvements. In this context, the translations

provided by Copilot were noticeably weaker than those of the other two LLMs. While in the Italian translations ChatGPT once more outperforms the other two LLMs by using the schwa, Gemini provides more detailed and gender-fair translations in English. However, it becomes evident that all programmes tend to rely predominantly on neutralisation or binary strategies rather than non-binary ones. When genderless markers, such as the schwa, are used in Italian, their application is inconsistent, often combining binary article use with non-binary noun inflection. These results may suggest that the use of these special characters is not sufficiently represented in the training data in order to appear in the translation output. Thus, the study clearly indicates that, in terms of gender-fair noun references, and particularly the implementation of non-binary strategies, MT is not yet adequately equipped, leading to a significant loss of meaning in the translation process.

Despite the new and interesting findings, this study also had some limitations. Firstly, it was restricted to only two target languages. Future research could explore how translation programmes integrate GFL in other languages, including genderless languages. Furthermore, while both target languages are major European languages, it would be valuable to extend the research to less widely spoken languages. Additionally, as with all research involving LLMs, the results of this study represent only a snapshot of the current state of the technology. Since LLMs are continuously trained on user input, the translations they produce evolve daily. In this context, it should also be noted that the output generated by LLMs is highly prompt-dependent. Hence, it is possible that the models would have performed better if given different prompts, such as: "Translate this sentence into Italian/English so that non-binary people are visible". Finally, the relationship between different prompts and output in relation to GFL translation needs to be studied in greater detail in future research.

The study demonstrated that LLMs still have a considerable way to go before they can fully grasp and accurately translate gender-fair German, especially non-binary forms. Once again, it has become apparent that, while MT is an invaluable tool in our daily lives, we cannot blindly rely on it, and it certainly cannot replace qualified translators who are capable of understanding the nuances and intentions of the original text – subtleties that machines often fail to capture.

References

Acke, Hanna (2019): „Sprachwandel durch feministische Sprachkritik". *Zeitschrift Literaturwissenschaftliche Linguistik* 49/2: 303–320.

Arcangeli, Massimo (2022): *La lingua scəma: Contro lo schwa (e altri animali)*. Roma: Castelvecchi.

Attig, Remy (2022): "A Call for Community-Informed Translation: Respecting Queer Self-Determination across Linguistic Lines." *Translation and Interpreting Studies* 18/1: 70–90.

Bachmaier, Elvira (2021): „Gendern im DaF-Unterricht – Einblicke in das Forschungsfeld und Anregungen zur Unterrichtsreflexion". 麗澤大学紀要 = *Reitaku University Journal* 104: 54–60.

Baer, Brian James/Kaindl, Klaus (2017): "Introduction: Queer(ing) translation". In: Baer, Brian James/Klaus Kaindl (eds.): *Queering Translation, Translating the Queer*. London, Routledge: 1–10.

Baron, Dennis (2020): *What's your pronoun?: Beyond he and she*. New York: Liveright Publishing.

Breuer, Rayna (2022): *Ist gendersensible Sprache hilfreich*? dw.com/de/gendersensible-sprache-pro-contra-gendergerecht/a-63677872 [25.03.2025]

Brisant (2025): *Gendern – was spricht dafür, was dagegen*? brisant.de/gendern-argumente-106.html [25.03.2025]

Choubey, Prafulla Kumar et al. (2021): "GFST: Gender-Filtered Self-Training for More Accurate Gender in Translation". In: Moens, Marie-Francine et al. (eds.): *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Punta Cana, Association for Computational Linguistics: 1640–1654.

Comstock, Frannie (2025): *Google Gemini. Generative AI*. britannica.com/technology/Google-Gemini [25.03.2025]

Corbett, Greville G. (1991): *Gender*. Cambridge: Cambridge University Press.

Costa-jussà, Marta R./de Jorge, Adrià (2020): "Fine-tuning Neural Machine Translation on Gender-Balanced Datasets". In: Costa-jussà, Marta R. et al. (eds.): *Proceedings of the Second Workshop on Gender Bias in Natural Language Processing*. Barcelona, Association for Computational Linguistics: 26–34.

DeepL: *DeepL Translator languages*. support.deepl.com/hc/en-us/articles/360019925219-DeepL-Translator-languages [25.03.2025]

Devlin, Jacob et al. (2019): "Bert: Pretraining of deep bidirectional transformers for language understanding". In: Burstein, Jill/Doran, Christy/Solorio, Thamar (eds.): *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (NAACL-HLT 2019). Volume 1 *Long and Short Papers*. Minneapolis, Association for Computational Linguistics: 4171–4186.

Diewald, Gabriele/Steinhauer, Anja (2019): *Gender – Ganz einfach*! Kehl: Duden Verlag.

Enciclopedia Treccani: treccani.it/vocabolario [25.03.2025].

Feministisch Sprachhandeln der Humboldt-Universität zu Berlin (2015): *Was tun? Sprachhandeln – aber wie? W_ortungen statt Tatenlosigkeit*. feministisch-sprachhandeln.org/wp-content/uploads/2015/04/sprachleitfaden_zweite_auflage.pdf [25.03.2025]

García González, Marta (2024): "Gender Bias in Translation Automation. Addressing Bias and Inequality". In: Monzó-Nebot, Esther/Tasa-Fuster, Vicenta (eds.): *The Social Impact of Automating Translation. An Ethics of Care Perspective on Machine Translation*. New York, Routledge: 96–122.

Gesellschaft für Deutsche Sprache (2020): *Leitlinien der GfdS zu den Möglichkeiten des Genderings*. gfds.de/standpunkt-der-gfds-zu-einer-geschlechtergerechten-sprache/ [16.01.2025].

Gheno, Vera (2019/2022): *Femminili Singolari +: Il Femminismo è Nelle Parole*. Florence: Effequ.

Ghosh, Sourojit/Caliskan, Aylin (2023): "Chatgpt perpetuates gender bias in machine translation and ignores non-gendered pronouns: Findings across bengali and five other low-resource languages". doi.org/10.1145/3568813.3600120.

Github: *About individual GitHub Copilot plans and benefits*. docs.github.com/en/copilot/managing-copilot/managing-copilot-as-an-individual-subscriber/about-github-copilot-free [25.03.2025].

Grove, Kenna (2021): ""Where Is They?": The Globalization of English and Gender-Inclusive Language". *The English Languages: History, Diaspora Culture* 7: 10–21.

Gygax, Pascal Mark et al. (2019): "A Language Index of Grammatical Gender Dimensions to Study the Impact of Grammatical Gender on the Way We Perceive Women and Men". *Frontiers in Psychology* 10. doi.org/10.3389/fpsyg.2019.01604.

Hekanaho, Laura (2020): *Generic and Nonbinary Pronouns: Usage, Acceptability and Attitudes*. PhD dissertation, University of Helsinki. helda.helsinki.fi/handle/10138/321581 [31.05.2021]

Hellinger, Marlis/Bußmann, Hadumod (2001): *Gender Across Languages. The linguistic representation of women and men*. Volume 3. London: Benjamin.

Hendy, Amr et al. (2023): *How Good are GPT Models at Machine Translation? A Comprehensive Evaluation.* arxiv.org/abs/2302.09210 [02.09.2025].

Hensler, Sarah (2021): *Eine kurze Geschichte der gendergerechten Sprache*. srf.ch/kultur/gesellschaft-religion/minenfeld-gendersternchen-eine-kurze-geschichte-der-gendergerechten-sprache [13.01.2025].

Hyde, Janet S. (1984): "Children's understanding of sexist language". *Developmental Psychology* 20/4: 697–706. doi.org/10.1037/0012-1649.20.4.697.

Jiao, Wenxiang et al. (2023): *Is ChatGPT a Good Translator? Yes With GPT-4 As The Engine* doi.org/10.48550/arXiv.2301.08745.

Koehn, Philipp (2005): "Europarl: A Parallel Corpus for Statistical Machine Translation". *Proceedings of Machine Translation Summit X: Papers*. Phuket,Thailand: 79–86. aclanthology.org/2005.mtsummit-papers.11/ [15.09.2025].

Kotek, Hadas/Dockum, Rikker/Sun, David (2023): "Gender bias and stereotypes in Large Language Models". In: Bernstein, Michael/Savage, Saiph/Bozzon, Alessandro (eds.): *C '23: Proceedings of The ACM Collective Intelligence Conference.* New York, Association for Computing Machinery: 12–24. doi.org/10.1145/3582269.3615599.

Lardelli, Manuel/Gromann, Dagmar (2023a): "Gender-Fair (Machine) Translation". In: Castilho, Sheila et al. (eds.): *Proceedings of the New Trend in Translation and Technology Conference – NeTTT 2022.* Shoumen, INCOMA: 166–177.

Lardelli, Manuel/Gromann, Dagmar (2023b): "Translating non-binary coming-out reports: Gender-fair language strategies and use in news articles". *The Journal of Specialised Translation* 40: 213–240.

Lardelli, Manuel et al. (2024): "Sparks of Fairness: Preliminary Evidence of Commercial Machine Translation as English-to-German Gender-Fair Dictionaries". In: Savoldi, Beatrice et al. (eds.): *Proceedings of the 2nd International Workshop on Gender-Inclusive Translation Technologies* (collocated with 2024). Sheffield, European Association for Machine Translation (EAMT): 12–21.

Link, Sabrina (2022): „Gendergerechte Sprache: Präsenz in den deutschen Medien und Integration in den DaF-Unterricht". In: Leal Rivas, Natasha (ed.). *Prospettive discorsive e di educazione linguistica internazionale*. Naples, Federico II University Press: 45–65.

Link, Sabrina (2024a): "The use of gender-inclusive language in Austrian, German, and Swiss broadsheets: a contrastive, corpus-based study". *Lingua*. 308: 1–16. doi.org/10.1016/j.lingua.2024.103787.

Link, Sabrina (2024b): „Zum Einsatz von Korpora zur Auseinandersetzung mit genderinklusiver Sprache im universitären DaF-Unterricht: ein Beispiel aus der Praxis". *German as a Foreign Language* 3/24: 108–131.

Link, Sabrina (2025): "Gender-Fair Language (GFL) in Austrian Tabloid and Broadsheet Newspapers: A Corpus-Based Frequency and Context Analysis". *International Journal of Applied Linguistics:* 1–17. doi.org/10.1111/ijal.12717.

Lipsky, Angela (2023): „Was hat Genus mit Geschlecht zu tun? Gendergerechte Sprache als Chance für den DaF-Unterricht in Japan". *Doitsu Bungaku* 22/1: 12–32.

Lyu, Chenyang et al. (2023): "New Trends in Machine Translation using Large Language Models: Case Examples with ChatGPT". doi.org/10.48550/arXiv.2503.10351.

Marotta, Ilaria/Monaco, Salvatore (2016): "Un Linguaggio Più Inclusivo? Rischi e Asterischi Nella Lingua Italiana". *Gender/Sexuality/Italy* 3: 45–57. doi.org/10.15781/ZS16-XH58.

Marr, Bernard (2023): *A Short History Of ChatGPT: How We Got To Where We Are Today*. forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/ [25.03.2025]

Mauran, Cecily (2023): *Meet Copilot, Microsoft's AI tool for work and productivity*. mashable.com/article/microsoft-copilot-ai-assistant-announcement [25.03.2025]

Měchura, Michal (2023): *Google Translate versus gender bias*. fairslator.com/google-translate-versus-gender-bias [25.03.2025].

Microsoft (2025): *Translator language support*. learn.microsoft.com/en-us/azure/ai-services/translator/language-support [25.03.2025].

Misiek, Szymon (2020): "Misgendered in Translation?: Genderqueerness in Polish Translations of English-Language Television Series". *Anglica. An International Journal of English Studies* 29/2: 165–185. doi.org/10.7311/0860-5734.29.2.09.

Müller-Spitzer, Carolin (2022): *Zumutung, Herausforderung, Notwendigkeit? Zum Stand der Forschung zu geschlechtergerechter Sprache*. bpb.de/shop/zeitschriften/apuz/geschlechtergerechte-sprache-2022/346089/zumutung-herausforderung-notwendigkeit/ [25.03.2025]

Müller-Spitzer, Carolin/Ochs, Samira (2023): „Geschlechtergerechte Sprache auf den Webseiten Deutscher, österreichischer, schweizerischer und südtirolerischer Städte". *Sprachreport* 39/2: 1–5. doi.org/10.14618/sr-2-2023_mue.

Nodari, Rosalba (2024): "Gender-inclusive strategies in Italian: stereotypes and attitudes". In: Pfalzgraf, Falco (ed.): *Public attitudes towards gender-inclusive language*. Berlin/Boston, De Gruyter Mouton: 243–286.

Olivares, Daniela (2023): *Klage gegen Audis Gendersprache-Leitfaden endgültig gescheitert*. br.de/nachrichten/bayern/klage-gegen-audis-gendersprache-leitfaden-endgueltig-gescheitert,TkKmcxN [25.03.2025]

OpenAI (2019): *Microsoft invests in and partners with OpenAI to support us building beneficial AGI*. openai.com/index/microsoft-invests-in-and-partners-with-openai/ [02.09.2025].

Papa, Sirio/Tavosanis, Mirco (2020): "Valutazione umana di DeepL a livello di frase per le traduzioni di testi specialistici dall'inglese verso l'italiano". In: Monti, Johanna/Dell'Orletta, Felice/Tamburini, Fabio (eds.) *Proceedings of the Seventh Italian Conference on Computational Linguistics*. Bologna, Accademia University Press: 1–7.

Paolucci, Angela et al. (2023): "Gender-Fair Language in Translation: A Case Study". In: Vanmassenhove, Eva et al. (eds.): *Proceedings of the First Workshop on Gender-Inclusive Translation Technologies*. Tampere, European Association for Machine Translation: 13–23.

Peres, Renana et al. (2023): "On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice". *International Journal of Research in Marketing* 40/2: 269–275. doi.org/10.1016/j.ijresmar.2023.03.001

Peuschel, Kristina (2022): *Keine größere Hürde als...Gendergerechte Sprache im Deutschen aus der Perspektive des Lehrens und Lernens*. bpb.de/shop/zeitschriften/apuz/geschlechter gerechte-sprache-2022/346100/keine-groessere-huerde-als/ [07.12.2024].

Piazzolla, Slivia Alma/Savoldi, Beatrice/Bentivogli, Luisa (2023): "Good, but not always Fair: An Evaluation of Gender Bias for three Commercial Machine Translation Systems". *HERMES – Journal of Language and Communication in Business* 63: 209–225. doi.org/10.7146/hjlcb.vi63.137553.

Piergentili, Andrea et al. (2023): "Gender Neutralization for an Inclusive Machine Translation: from Theoretical Foundations to Open Challenges". In: Vanmassenhove, Eva et al. (eds.): *Proceedings of the First Workshop on Gender-Inclusive Translation Technologies.* Tampere, European Association for Machine Translation: 71–83.

Prates, Marcelo. O. R./Avela, Pedro H./Lamb, Luis C. (2020): "Assessing gender bias in machine translation: a case study with google translate". *Neural Computing and Applications* 32: 6363–6381.

Prewitt-Freilino, Jennifer L. et al. (2012): "The Gendering of Language: A Comparison of Gender Equality in Countries with Gendered, Natural Gender, and Genderless Languages". *Sex Roles* 66: 268–281.

Pusch, Luisa. F. (1979): „Der Mensch ist ein Gewohnheitstier, doch weiter kommt man ohne ihr – Eine Antwort auf Kalverkämpers Kritik an Trömel-Plötz' Artikel über Linguistik und Frauensprache". *Linguistische Berichte* 63: 84–102.

Rat für deutsche Rechtschreibung (2023): *Amtliches Regelwerk der deutschen Rechtschreibung: Ergänzungspassus Sonderzeichen.* rechtschreibrat.com/amtliches-regelwerk-der-deutschen-rechtschreibung-ergaenzungspassus-sonderzeichen/ [16.01.2025].

Savoldi, Beatrice, et al. (2021): "Gender bias in machine translation". TACL 9: 845–874.

Schmitz, Dominic/Schneider, Viktoria/Esser, Janina (2023): "No genericity in sight: An exploration of the semantics of masculine generics in German". *Glossa Psycholinguistics* 2/1: 1–33. dx.doi.org/10.5070/G6011192.

Schomburg, Neele (2023): *Gendern in Schulen und Verwaltung: Welche Bundesländer es verbieten – und welche den Doppelpunkt nutzen*. rnd.de/politik/gendern-in-schule-und-verwaltung-wo-gilt-das-verbot-bundeslaender-uebersicht-UXSOG4QPP5FFJLHZORHZWR4FA4.html [25.03.2025].

Sill, Hannah (2022): *Gendern mit Binnen-I | Alle Regeln leicht verständlich im Überblick.* scribbr.de/richtig-gendern/binnen-i/ [25.03.2025].

Šincek, Marijana (2020): "On, ona, ono: Translating gender neutral pronouns into Croatian" In: Božanić, Petra/Ćurčić, Ana (eds.): *Zbornik Radova Međunarodnog Simpozija Mladih Anglista, Kroatista i Talijanista.* Split, Sveučilište u Splitu, Filozofski fakultet: 92–112. urn.nsk.hr/urn:nbn:hr:172:208617 [02.09.2025].

Sniezek, Janet. A./Jazwinski, Christiane H. (1986): "Gender bias in English: In search of fair language". *Journal of Applied Social Psychology* 16: 642–662.

Sökefeld, Carla (2021): „Gender(un)gerechte Personenbezeichnungen: derzeitiger Sprachgebrauch, Einflussfaktoren auf die Sprachwahl und diachrone Entwicklungen". *Sprachwissenschaft* 46/1: 111–141.

Stahlberg, Dagmar et al. (2007): "Representation of the sexes in language". In: Fiedler, Klaus (ed.): *Social Communication.* New York/London, Psychology Press: 63–187.

Stark, Katharina (2021): *Gendergerechte Sprache im DaZ/DaF-Unterricht: Bewertung und Vermittlung durch DaZ/DaF-Lehrkräfte.* Master's thesis, Universität Augsburg.

Stefanovičs, Artūrs et al (2020): „Mitigating Gender Bias in Machine Translation with Target Gender Annotations". In: Barrault, Loïc et al. (eds.): *Proceedings of the Fifth Conference on Machine Translation.* Stroudsburg, Association for Computational Linguistics: 629–638.

Straub, David (2024): *CDU und AfD wollen Genderverbot in Thüringer Schulen – wie kommt das dort an?* mdr.de/nachrichten/thueringen/gendern-schule-afd-cdu-landtag-100.html [25.03.2025].

Stummer, Nathalie (2024): „Gendersensible Sprache im Fremdsprachenunterricht Deutsch Aktuelle Befunde aus Lehrmaterialien". *Wiener Linguistische Gazette* 96: 1–28.

Sun, Tony et al. (2021): „They, them, theirs: Rewriting with gender-neutral English". doi.org/10.48550/arXiv.2102.06788.

Switzer, Jo Young (1990): "The impact of generic word choices: An empirical investigation of age and sex-related differences". *Sex Roles* 22: 69–82.

Trömel-Plötz, Senta (1978): „Linguistik and Frauensprache". *Linguistische Berichte* 57: 49–68.

Vanmassenhove, Eva (2024): "Gender Bias in Machine Translation and The Era of Large Language Models". doi.org/10.48550/arXiv.2401.10016.

Vanmassenhove, Eva et al. (2021): "Neutral rewriter: A rule-based and neural approach to automatic rewriting into gender neutral alternatives". In: Moens, Marie-Francine et al. (eds.): *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing.* Punta Cana, Association for Computational Linguistics*:* 8940–8948.

Verein Deutsche Sprache (2021): *Duden spaltet die Sprachgemeinschaft.* vds-ev.de/mitteilungen/duden-spaltet-die-sprachgemeinschaft/ [25.03.2025].

Vitiello, Ruben (2022): *Linguaggio inclusivo in italiano: guida pratica per chi scrive per lavoro (e non).* tdm-magazine.it/linguaggio-inclusivo-in-italiano-guida-pratica/ [25.03.2025].

Von Flotow, Luise (1997): *Translation and Gender: Translating in The Era of Feminism.* London: Routledge.

Waldendorf, Anica (2023): "Words of change: The Increase of Gender-Inclusive Language in German Media." *European Sociological Review*: 1–18. doi.org/10.1093/esr/jcad044.

Wang, Longyue et al. (2023): "Document-level machine translation with large language models". doi.org/10.48550/arXiv.2304.02210.

Zennaro, Chiara (2023): *L'Accademia della Crusca ha detto la sua sul femminile negli atti pubblici.* wired.it/article/femminile-schwa-accademia-crusca/ [25.03.2025].