

Carolina Flinz, Eva Gredel, and Laura Herzberg

The negotiation of pronominal address on talk pages of the German, French, and Italian Wikipedia

Abstract: The adequate use of social deixis is highly dependent on the situation and context and has therefore always been at the center of linguistic pragmatics. So far, principles of pronominal address have mainly been modelled with a focus on oral, co-present interaction. The use of pronominal address in computer-mediated communication (CMC) with its translocal and partially anonymous contexts is still a research gap.

This paper asks, from a contrastive perspective, how the appropriate use of address pronouns is negotiated on talk pages of the German, French, and Italian Wikipedia. The talk pages of Wikipedia share features of CMC genres such as a dialogic structure and an informal writing style with non-standard language. There are two types of Wikipedia talk pages, whose data are considered in this study based on the multilingual corpora by the Leibniz Institute for the German Language: article talk pages, where authors negotiate online encyclopedic content, and user talk pages, where the contributions of individual authors are discussed. These two types of talk pages will be analysed for the study.

Based on corpus data, it can be shown that the unidirectionality of this transition from the formal form (in German: *Sie*, in French: *vous*, in Italian: *Lei*) to the informal form (in German: *du*, in French: *tu*, in Italian: *tu*) in CMC is not always given. In both analysed Wikipedia subcorpora, i.e., the Wikipedia article talk pages on the one hand and the article talk pages on the other hand, a greater deal of discussions about addressing styles takes place on the user talk pages, with the informal *you* variant being discussed more frequently than the formal *you* variant. Aspects of pronominal address among speakers of German, French and Italian are characterized by instability and uncertainty – especially in CMC. Moreover, it can be shown that Wikipedia authors discuss, among others, the reasoning behind their preference for a certain form as well as the notion of “distance” in which informal variants show familiarity which is not perceived as desirable by all users.

Carolina Flinz, Università degli Studi di Milano, e-mail: carolina.flinz@unimi.it

Eva Gredel, University of Duisburg-Essen, e-mail: eva.gredel@uni-due.de

Laura Herzberg, Leibniz Institute for the German Language, e-mail: herzberg@ids-mannheim.de

Keywords: social deixis, corpus linguistics, crosslingual study, Wikipedia, linguistic pragmatics, computer-mediated communication, pronominal address

1 Introduction

One of the central tasks that interlocutors have to solve in interaction is to make others feel addressed and thus obliged to respond.¹ Hence, “[a]ddress forms are an indispensable part of the communicative process” (De Oliveira 2013: 291). While in face-to-face communication, addressing can be achieved multimodally through various resources (in addition to language, such resources include bodily orientation, facial expressions, gestures as well as gaze), interlocutors in many CMC genres are highly dependent on linguistic means of address. In this context, pronouns are among the most important means by which linguistic addressing can be realized.² In many languages, the selection of certain pronouns from a pronominal paradigm allows the speaker or writer to encode social relations or social distance between the interlocutors linguistically. This has been studied since Fillmore (1975: 76) under the label “social deixis”.

The appropriate use of socio-deictic signs is highly dependent on the situation and context and has always been at the center of linguistic pragmatics (cf. Nübling et al. 2017: 205). However, principles of pronominal address have so far been mainly modeled with a focus on oral interaction where speakers are co-present (cf. Kretzenbacher 2010). The use of pronominal address in CMC with its translocal and (partially) anonymous contexts poses special challenges for writers and has been considered in only a few initial studies (see Gredel 2023 for German and Rabelos and Strambi 2009 for Italian). This paper aims to fill this research gap by analyzing meta-discourses on pronominal address in the CMC genre of Wikipedia talk pages. With the multilingual Wikipedia corpora of the Leibniz Institute for the German Language, digital language resources are used that allow a contrastive approach to this object of investigation.

The languages German, French, and Italian, which are considered in this paper, each have a binary system of pronominal address comprising a T form (GER: *du*,

¹ The three authors have written the paper jointly. Carolina Flinz is responsible for the data and analyses of Italian, Eva Gredel for German, Laura Herzberg for French. Introduction (§1) and Conclusion (§5) were written jointly.

² Other means of address are lexical forms, such as first names, familiarizers and terms of endearments (cf. Helmbrecht 2006; Formentellii and Hajek 2015).

FR: *tu*, IT: *tu*) and a V form (GER: *Sie*, FR: *vous*, IT: *Lei*).³ In oral face-to-face interaction, the selection of the appropriate pronoun in each communicative dyad is generally linked to variables such as social status, age, gender, and conversation situation of the interaction partners (cf. Nübling et al. 2017: 205). In CMC, these variables are not always apparent to writers, so the selection of appropriate pronouns must follow other principles. This paper focuses on this aspect through the analysis of meta-discourses in which interlocutors controversially discuss forms of pronominal address in metapragmatic comments on Wikipedia talk pages and thereby adopt certain stances with the aim of positioning themselves and others.

Regarding CMC, it is interesting that different customs or netiquettes for the use of the appropriate address pronouns have developed on various digital platforms (cf. Gredel 2023). On the multilingual platform Wikipedia, there are differences between the netiquettes of the considered language versions. However, there is no consensus on these netiquettes, and they are subject to controversial discussions. Based on the Wikipedia corpora of the Leibniz Institute for the German Language (IDS), in this paper we explore whether and how writers negotiate the use of address pronouns on Wikipedia talk pages. We also analyze which aspects of the use of pronominal address are being discussed on talk pages of the German, French, and Italian Wikipedia. Specifically, we would like to find answers to the following questions:

- Do writers negotiate the use of pronominal address on German, French and Italian Wikipedia talk pages?
- How frequent do writers negotiate the use of pronominal address on German, French, and Italian Wikipedia talk pages?
- Which aspects of the use of pronominal address are discussed on Wikipedia talk pages of the German, French, and Italian language versions?

First, we describe the theoretical background to pronominal address in general and in CMC in particular for the three languages studied (section 2). Then we discuss the data and methods (section 3), followed by the empirical analysis of the corpus data (section 4). Section 5 summarizes the results.

³ For the sake of simplicity, we use the abbreviations T and V here in the tradition of Brown and Gilman (1960), because they seem adequate for this case study of German, French and Italian pronominal address. However, we are well aware of the criticism of this approach (Simon 2003: 7) that the reduction to a binary system is not appropriate for all languages and for all stages of a language (In Italian for example in the past there were more forms, but after a process of simplification, some of them are no more used, see among others Formentelli and Hajek 2015).

2 Pronominal address

The following section provides a theoretical outline of studies that examine pronominal forms of address in general for the three languages under investigation: German, French, and Italian (section 2.1). This is followed by a description of the current state of research on pronominal forms of address in CMC (section 2.2). Finally, central concepts of the discussion of pronominal address in sociolinguistics (including ‘power’, ‘solidarity’ and ‘social distance’) are specified and metalinguistic utterances on address are theoretically framed as stancetaking (section 2.3).

2.1 Pronominal address in German, French, and Italian

Many languages in the world make it possible to encode social relationships with interlocutors linguistically by selecting pronouns from a pronominal paradigm (cf. Simon 2003), that is specific to each individual language.⁴ Which elements or grammatical forms of the pronominal paradigm are used as honorific pronouns⁵ varies from language to language. There are languages that have developed the honorific form of address from the 2nd person plural (e.g., in French *vous*). This can be explained from a cognitive-linguistic and diachronic perspective with the metaphor “plural is power”: Whoever is addressed in the plural is attributed power (Nübling et al. 2017: 2008). According to Brown and Levinson (1987: 23), this represents an act of positive politeness and has developed over time from a conversational implicature to a conventional one (Nübling et al. 2017: 2008). Other languages, such as Italian, use the 3rd person singular (e.g., in Italian *Lei*),⁶ whereby the indirect address using the 3rd person instead of the 2nd person is intended to mitigate face-threatening acts (FTAs) in the sense of Brown and Levinson (1987). From a typological perspective, German is a special case here: It combines both described strategies for linguistic politeness (i.e., the plural and the 3rd person), as the 3rd person plural is used as honorific pronoun of address (Duden 2016: §361 on the pronoun *Sie*).

4 Depending on the type of language, verbal morphology may play a central role too. For Italian for example subject pronouns are inherent in the verb conjugation: the suffix is pertinent to the grammatical person (cf. Renzi 1995).

5 Honorifics are linguistic forms (e.g., in many languages certain pronouns) that can be used to signal politeness (cf. Brown and Levinson 1987: 102; 185).

6 This applies to modern Italian. There is in fact another V form, i.e., the third person plural (*Loro*), which is considered archaic and limited to very formal and ritual social situations (Rebelos and Strambi 2009).

At the end of the 20th century, the *Grammatik der deutschen Sprache* (GDS) described the prototypical contexts for using the T form (cf. Zifonun et al. 1997: 317) in German, including interactions with relatives, friends, and children as well as with peers, colleagues, and party members. As in many other languages, the choice of the appropriate form of address in German depends on factors such as age, social status, and gender. However, Kretzenbacher points out that both T form contexts and V form contexts have “fuzzy edges”, such as the pronominal form of address for parents-in-law of one’s own children (Kretzenbacher 2010: 7).

In addition to such “fuzzy edges” regarding pronominal address, the transition from the T form to the V form is also associated with uncertainty among the interlocutors and must be brought about explicitly. The transition from the V form to the T form in German takes place consensually on the initiative of the older, higher-ranking or female interlocutor (Zifonun et al. 1997: 317). The one-time transition from the V form to the T form is also common when a change in the interactants’ social relationship has taken place (e.g., a longer acquaintance); this transition is essentially irreversible (Simon 2003: 124). Changing between different pronouns within a communicative dyad is basically not possible (Simon 2003: 124). Moreover, the transition is only common in one direction – namely from the V form to the T form (Zifonun et al. 1997: 317).⁷

In French, the address pronoun system consists of the T forms singular *tu*/plural *vous* and the V forms singular *vous*/plural *vous* (cf. Clyne 2004: 1). The pronominal usage in France, from the late Middle Ages to the early eighteenth century, remained essentially the same. It was based on class status: T forms were used to address inferiors, and V forms to address superiors (cf. Maley 1972: 1002). This pattern of pronoun usage remained dominant in France until the French Revolution, when the Committee for Public Safety ordered everyone to use T forms on all occasions. Nonetheless, V forms did not completely diminish. More so, T forms were applied in areas in which they have not been used previously, for example, husbands and wives are using T forms with each other as well as children when addressing their parents. Further, Maley points out the majority of grammarians from the preceding centuries made few detailed comments on the usages of the T and V forms. This continued to the twentieth century as well (cf. Maley 1972: 1002).

7 The homonymy of the honorific form of address with *Sie* and the form of address of a group in the 3rd person plural (also with *sie*) can lead to ambiguities in German. These ambiguities could lead to misunderstandings in digital interaction, especially if capitalization plays a subordinate role in interaction-oriented writing. However, no such hits were found in our corpus samples, which is why we did not consider this aspect further in the qualitative analysis.

V forms are the default form for neutral interactions but indicate respect or subservience when used in opposition to the T forms that indicate closeness and intimacy (cf. Schoch 1978: 57; Bouissac 2019: 140). Bouissac suggests that the use of titles commands the V form. The use of the first name is compatible with both the T and V forms. The nickname belongs to the realm of the T form (cf. Bouissac 2019: 143). Students of the French language are taught that the forms of the second person singular are *tu*, *te*, and *toi*, but that the plural form *vous* is also used as the proper way of address to address a single interlocutor formally (cf. Bouissac 2019: 140). Helmbrecht also points out the importance of the singular V form *vous* as politeness marker:

As an honorific pronoun it stands in a paradigmatic opposition to *tu* [...]; the driving force for the development of this usage is politeness, i.e., the avoidance of direct reference to the socially superior hearer/addressee (cf. Helmbrecht 2015: 181).

A greater use of T forms correlates with the younger age group with some decline over time as people grow older. In addition, T forms are the norm for relations between people of equal status and who have known each other for a certain length of time, for example, coworkers. Reciprocal V forms have an important place, and are still the pronoun of choice in initial encounters between strangers and between people who want to avoid familiarity (cf. Clyne 2004: 5; Helmbrecht 2006: 428). They are also used between people who know and see each other on a regular basis but want to show respect or deference, or to keep a certain distance one with respect to another, for example, doctors and patients, or parents and their children's teachers (cf. Morford 1997: 12). Additionally, Morford describes the existence of two basic dyadic forms: symmetrical *vous* and asymmetrical *tu/vous* or *vous/tu* usages. Symmetrical *vous* is in fact still the normal starting point for public interactions between adults who have no prior relation. It is commonly used between strangers or people who see each other rarely. People who otherwise address one another as *tu* may also adopt a symmetrical *vous* to mark the formality of certain circumstances, such as professional evaluations (cf. Morford 1997: 12).

Social deixis can be expressed in Italian by the pronouns of the second person singular *tu*, *ti*, and plural *voi*, *vi*, the pronouns of third person singular *Lei*, *Le* and plural pronoun *Loro* (cf. Da Milano 2015: 70).

Subject pronouns are not always expressed phonetically but are inherent in the verb conjugation (cf. Renzi 1995). They are considered as an optional rule in Italian, so when they are used, they strategically add pragmatic meanings to a speaker's utterance (Davidson 1996), because they are a marked choice in discourse (see among others Duranti 1984; Dal Negro and Pani 2019). Stewart (2003) argues,

in fact that, when adding a non-obligatory subject pronoun, speakers flout Grice's "Maxim of Quantity" and, by doing so, they generate pragmatic implicatures.

Italian deploys personal pronouns and related verbal morphology depending on the two parameters of symmetry/asymmetry and familiarity/distance coupled with context formality (Brown and Gilman 1960; Renzi 1995, 2001; Molinelli 2002; Formentelli and Hajek 2015). There is a complexity of norms regulating address pronoun choice. The current standard of address system in Italian is a bipartite T-V-system with the personal pronouns *tu* as the T form and *Lei* as the V form (Maeder and Werner 2019). The binary distinction of address strategies is also codified by lexical forms as honorifics (*signore, signora*), titles (*professore, professor-essa*), "titles + last names" as V forms; first names, familiarizers (*bello, bella*), and terms of endearments (*tesoro*) are used instead as T forms (see also Helmbrecht 2006; Formentelli and Hajek 2015). Address pronouns and nominal forms are often found in the same utterance in Italian, where they take on similar pragmatic values (cf. Formentelli and Pavesi 2022).

The reciprocal use is preferred to index familiarity or to signal social distance and/or mutual respect (cf. Formentelli and Hajek 2015: 122). Reciprocal V forms are for example, considered the default option in academic interactions. The non-reciprocal use (*Lei/tu*) occurs to a lesser extent when there is an asymmetrical distribution of power (age, job rank, and social status, see also Renzi 1993). In particular, it is well established in primary and secondary education, with an increase in the use of "tu + first names" by teaching staff (cf. Formentelli and Hajek 2013: 88–90). Also, in the family, there can still be a non-reciprocal use, for example, the interaction with the mother-in-law (the relationship change comes from the superior person and happens through a ritual). It is important to say that this is not unchangeable as roles and identities are continuously negotiated depending on different factors as levels of formality of the setting, degree of familiarity, and individual preferences (cf. Clyne, Norrby, and Warren 2009).

2.2 Pronominal address in German, French, and Italian CMC

Pronominal forms of address in CMC have been mentioned in linguistic publications since the 1990s and early 2000s, which De Oliveira explains as follows:

Computer-mediated communication (CMC) offers a different investigative milieu in which to study interaction than does face-to-face communication, as both the researcher and the interaction itself are place-independent and, in asynchronous CMC, time-independent, as well. A researcher can gain access to enormous quantities of data that are neither 'contaminated' by her presence [...] nor dependent on it. (De Oliveira 2013: 292)

The first mentions of pronominal address in CMC can be found in Schulze (1999: 80–81), Bader (2002: 52 and 127) and Hess-Lüttich and Wilde (2003: 167), who see the T form as the unmarked and predominant form of address in CMC.

More comprehensive work focusing on pronominal forms of address in CMC has been done for German mainly in the context of the Melbourne Address Project. In an explorative case study based on Usenet data (news groups and Internet Relay Chat) groups, Kretzenbacher takes a differentiated look at the question of the appropriate form of pronominal address in CMC:

Based on the research we have done with focus groups and Network interviews, we can hypothesize that in the case of the Usenet, we find a parallel to off-line communication in the coexistence of two systems, one tending towards unmarked *du*, the other towards unmarked *Sie*. (Kretzenbacher 2005: 6)

This thesis was examined in the course of the project on the basis of a broader database, which also contains data from German speaking online fora: Kretzenbacher (2011) and Kretzenbacher and Schüpbach (2015) came to the conclusion that the T form is not the exclusive and not always the dominant form of pronominal address in German-speaking CMC genres. When the T form is used, Kretzenbacher (2010: 6) interprets this as a symptom of a shared “perceived commonalities” (in German: *geteilte virtuelle Lebenswelt*).

A more recent study on pronominal forms of address is that of Truan (2022), who analyzed data from the Twitter account of *Deutsche Bahn*. She also comes to the conclusion that pronominal forms of address are a controversial topic in CMC, with both supporters and opponents of the T form. Gredel (2023) is the first study on pronominal forms of address in Wikipedia. She shows that beyond the pure question of the form of address (T form versus V form), aspects such as reciprocity of the pronominal form or irreversibility of the transition from the V form to the T form are also important aspects of the analysis with regard to CMC.

For French and Italian, the analysis of pronominal address in CMC has so far largely been a research gap. Williams and van Compernelle (2007, 2009a) studied pronominal address in French CMC, in particular in IRC (Internet Relay Chat) messages. They pointed out factors that “favor” either using T or V forms such as syntactic frame, discursive effect, transitivity, ambiguity, linguistic conservatism and ease with address (cf. Williams and van Compernelle 2009a: 417). When investigating the chat messages, the authors found an overwhelming preference for *tu* as T form (98.02%) compared to V forms (1.08%). These results confirmed their findings from a study two years prior, in which the authors matched the reported *tu* use as high as 99.02% (Williams and van Compernelle 2007). The authors illustrated that the uttered V forms happened to be produced relatively

soon after the participants had entered the room, at the same time when people were still using V forms to address individuals in the room. All made the switch to *tu* as a form of address rather quickly (cf. Williams and van Compernelle 2009a: 420). In addition to chat, Williams and van Compernelle (2009b) analyzed pronominal address in discussion fora. They concluded that two primary factors influence the T and V usage in discussion fora:

- 1) the medium itself (i.e., the technological affordances and constraints of discussion fora)
- 2) each participant's preference to maintain a traditional, off-line paradigm instead of the on-line system of address pronoun use that has emerged in synchronous chat and discussion fora (Williams and van Compernelle 2009b: 364).

The results are similar to the usage of T forms in chat with an overall usage rate of 84.5% when compared to V forms. Similarly, the number of a specific posting plays a role as well. Since the first turn in almost all new threads is a message to anyone and everyone who might happen to read the posting, almost all subject pronouns in Turn 1 were either V forms or a non-second-person pronouns, like *on* (Williams and van Compernelle 2009b: 371).

Lastly, a study on blogs by Douglass (2009) underlines the online preference for T forms. The author compared different genres of blogs: personal blogs, current events blogs, and sports and entertainment blogs. Overall, *tu* was the most frequently used form (68.2%), followed by *vous* (31.8%). Importantly, this result averages all *tu* and *vous* usages across all blog genres. When looking at each of the genres individually, *tu* was overwhelmingly preferred on the personal blogs (93.3%) whereas *vous* was preferred by 71.9% on the current events blogs. Besides the aforementioned factors by Williams and van Compernelle (2007, 2009a), the computer-mediated format has to be taken into consideration as well.

For Italian, Pistolesi (1998), considering the reduced availability of visual and auditory cues, highlights the more egalitarian character of the CMC communication, being age, gender, social, and racial background unavailable; social hierarchy is therefore weakened and simplified. Gastaldi (2002) investigating the linguistic features of chat exchanges between native Italian speakers, finds out a high incidence of the familiar *tu* pronoun between interlocutors. Studies on IRC make evident typical colloquial and informal features (cf. Pistolesi 1997). Rebelos and Strambi (2009), in their study which focuses on the potential role of CMC in promoting learners' understanding of norms regulating address pronoun choice, analyses address pronouns across different forms, asynchronous and synchronous, concluding that in chat rooms participants address each other exclusively with the informal pronouns (*tu*, fewer *voi*). There is a complete absence of a formal address. In discussion boards, there is a higher frequency of explicit forms of address, with a

higher incidence of singular informal pronouns. The only formal pronoun used, *Lei*, is observed on three occasions to signal opposition and social distance in conflict. The identified use is considered a formal demonstration of respect but also a way for the interlocutors to signal the superiority of one of them (see also Dewaele 2004: 384). The two authors show that the most common form of address used across all types of CMC is the informal *tu* pronoun, confirming their initial hypothesis that the observed informality of CMC would manifest itself in the data through a prevalent use of informal pronouns. Another interesting point is whether users perceived the exchange as one-to-one or one-to-many. All of the observed interactions took place in a public space accessible by the online community, with the possibility of participants to observe the interaction and join the conversation. Thus the communication can be between an infinite number of interlocutors, and often the plural address pronoun *voi* is used to reflect this. In the case of a direct and public interaction with a well-known figure of high social standing, the perceived formality of the exchange leads to a more frequent use of the formal pronoun *Lei*. It was also found an instance of explicit reference to a discussion of address form: a participant explained his choice of using an informal pronoun in addressing an interlocutor whom he did not know personally. Maeder and Werner (2019) also cite the study of Rebelos and Strambi (2009), highlighting the dominance of T forms, considering vicinity, familiarity and solidarity the norm.

The contrastive analysis for the languages German, French, and Italian is also still a desideratum, which we would like to address in this case study. The Wikipedia data are interesting and adequate for this project because although the Wikipedia talk pages of the three language versions can be assigned to one and the same CMC genre, the institutional dimension of the language versions (e.g., via wiki-quotas and other metalinguistic negotiations) has developed relatively differently in the more than 20 years of Wikipedia's existence. The data is therefore particularly suitable for comparing language and culture in the context of pronominal forms of address.

2.3 Further theoretical framing of pronominal address

This paper is interested in what meta-linguistic utterances Wikipedia authors make about the use of pronominal forms of address and the associated social deixis. As such meta-linguistic utterances are used by authors to reveal their position on the use of the T or V form and associated values, they can be linked to the sociolinguistic concepts of 'social positioning' (Deppermann 2015) and 'stancetaking' in the tradition of Ochs (1996) and Du Bois (2007), as Truan (2022) does for pronominal forms of address in Twitter. Du Bois states:

One of the most important things we do with words is take a stance. Stance has the power to assign value to objects of interest, to position social actors with respect to those objects, to calibrate alignment between stance takers, and to invoke presupposed systems of sociocultural value. (Du Bois 2007: 139)

In the context of the theory of social positioning, linguists are concerned with the fact that identities are not simply “brought into” interactions as stable entities, but must be negotiated again and again in social interaction. In doing so, the interlocutors permanently position themselves and others in the social space (Spitzmüller 2022: 272).

Stancetaking is then specifically about the positioning of subject 1 in relation to an object of evaluation in the interaction with subject 2. In interaction, the viewpoints of subject 1 and subject 2 considering the object are then also compared and a (dis-)alignment takes place (see figure 1). Du Bois (2007) conceptualizes this as a triadic process, which he reduces to the following formula: “I evaluate something, and thereby position myself, and thereby align with you” (Du Bois 2007: 163).

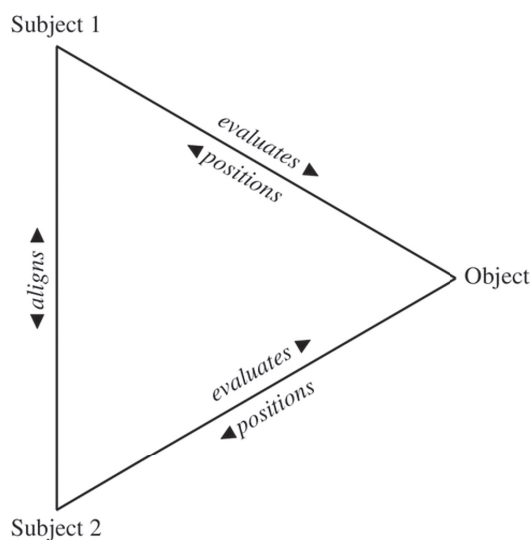


Figure 1: The stance triangle (Du Bois 2007: 164).

Contexts in which the objects of evaluation are linguistic forms, linguistic patterns and language ideologies are particularly interesting for sociolinguistic work (Spitzmüller 2022: 275). It is precisely such explicit meta-linguistic comments that can be found in the corpus data analyzed for this case study. In this context, it is also interesting to note which values and aspects are associated with the use of the T and V

form (see section 4). Very often, aspects of power are addressed by the authors (cf. Gredel 2023) because Wikipedia was launched as a project on the social web with egalitarian and grassroots democratic aspirations (Stegbauer 2009). However, the reality of the platform is that, over the years, roles and hierarchies have gradually emerged in the ad hoc meritocracy of the online encyclopedia for the purposes of quality assurance.

A theoretical framework for thinking about hierarchies in the context of addressing in linguistics was provided by Brown and Gilman (1960) in their widely received paper “The pronouns of power and solidarity”. In this paper, they introduce the vertical scale of the power dimension and the horizontal scale of the solidarity dimension. The power dimension is concerned with hierarchical structures, which can lead to non-reciprocal forms of address (the higher-ranking interlocutor is addressed with the V form, while the lower-ranking interlocutor only receives the T form). The horizontal dimension of solidarity is about the fact that interlocutors with a higher intimacy or familiarity address each other (i.e., reciprocally) with the T form, whereas the use of the V form indicates a lack of familiarity and intimacy. Later, the two authors use the concept ‘distance’ instead of the politically connotated concept of ‘solidarity’ (Kretzenbach 2010: 4). In the meantime, this approach by Brown and Gilman (1960) has been widely criticized. One reason for this is that the use of the abbreviations V and T for all languages suggest a binary system of pronominal address for all languages, although there are systems of pronominal address with more than two elements (Simon 2003: 7). Nevertheless, even today many linguists continue to build on this theoretical background and expand the theory by introducing new categories and concepts or further differentiating the concepts of Brown and Gilman (cf. Leech 1983: 126; Svennevig 1999: 34; Molinelli 2002: 283). Molinelli (2002) for example argues that the dichotomy solidarity (parity of communication; every participant uses the same instruments and there is reciprocity) and power (who has more power determinates the interaction and when there is difference of power there is an asymmetric communication, in which the power uses *tu*, see also Orletti 2000) and their interplay, are not sufficient. Respect and distance have to be added. Respect is in fact different from power; so there can be respect between persons at the same power level. Distance is more changeable than the other parameters; it depends on the culture and is determinant in the culture in which it is socially codified (Molinelli 2002: 294). So all four dimensions, expressed in parameters [$\pm$ power], [$\pm$ solidarity], [$\pm$ distance], [$\pm$ respect] have to be considered, because they are present in the speaker and in the hearer (codification/decodification), cf. Figure 2:

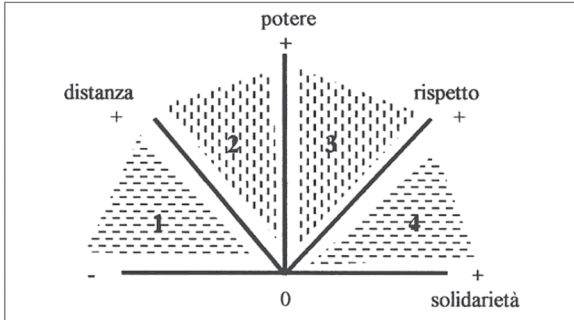


Figure 2: Four address parameters by Molinelli (2002: 296): Distance - Power - Respect - Solidarity.

Molinelli (2002: 296) explains that in the figure, which represents a scheme not necessarily realised in every language and at every stage of a language, each parameter is conceived as a continuum from presence (+) to absence (-) of the trait.

3 Method & data

Wikipedia consists of many structural elements, referred to as namespaces. A namespace is “a virtual container for different types of content on the wiki; namespaces are defined by different prefixes, such as *talk:* or *Wikipedia:* which appear before page names; articles are in the main namespace” (Ayers et al. 2008: 473). Figure 3 presents a schematic overview of the Wikipedia data and namespaces.

In addition to the well-known articles as the core content of Wikipedia, the talk pages also have a central function, because the online encyclopaedic content is negotiated there in digital interaction. It is often a matter of controversy in which terms events or facts are described in Wikipedia. Additionally, the revision history documents in detail the development of articles over the years and allows one to understand the dynamics of the content. Many long-standing authors introduce themselves to the community on user pages and the corresponding user talk pages offer the opportunity to thank or criticize authors. Finally, there are many wiki pages with policies and guidelines as well as links to external sources. From these different areas of Wikipedia, different digital language resources have been compiled in the form of linguistic corpora at the IDS.

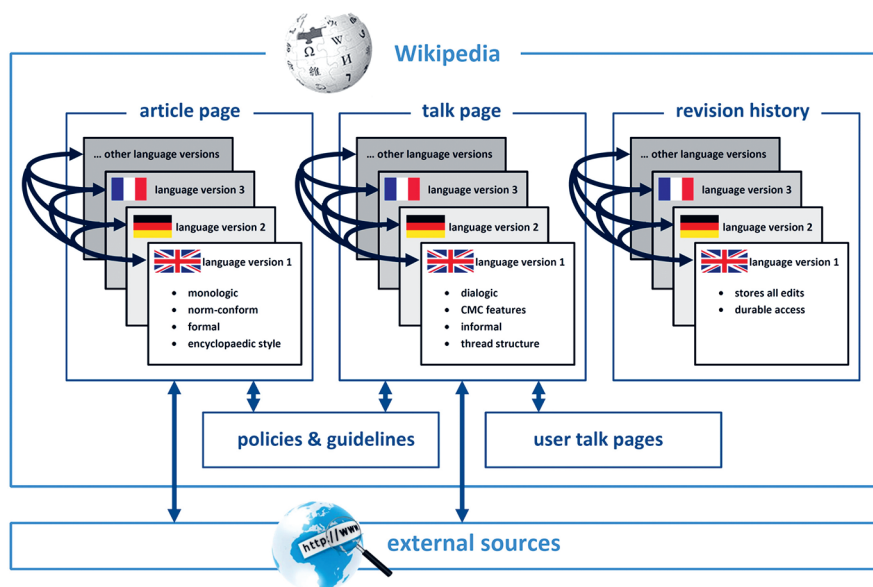


Figure 3: Schema of data and namespaces in Wikipedia (Gredel 2017).

The Wikipedia terms of use permit sharing and reusing of Wikipedia content under free and open licenses, which makes the language data it contains usable for research. IDS has been offering access to Wikipedia data via its corpus infrastructure since 2005. Wikipedia language data is converted into linguistically processed corpora and thus made accessible for research purposes. This is done in a structured way, e.g., by namespace, i.e., there are Wikipedia sub-corpora for article, talk, and user talk pages. In addition to different namespaces, several language versions are also available. Since 2011, corpora of all German-language encyclopedia articles as well as all associated talk pages have been created every two years for this purpose from a Wikipedia dump published by the Wikimedia Foundation. This dump contains a “snapshot” of the database content of an entire single-language Wikipedia at a specific point in time (Beißwenger and Lungen 2022: 439). The Wikipedia corpora collection is thus constantly being expanded and extended. The Wikipedia data forms an archive of the German Reference Corpus (DeReKo), which is the most comprehensive collection of written contemporary language (cf. Lungen and Kupietz 2020). The Wikipedia corpora are available via the IDS corpus research systems COSMAS II_{web} and KorAP. COSMAS II_{web} is an operating system-independent WWW application that enables corpus research in a conventional WWW browser. In COSMAS II_{web}, the corpora are managed in three archives,

whereby the German and English Wikipedia content is organized in one archive each and a further eight language versions are available in one archive.

The talk pages of Wikipedia share features of CMC genres such as a dialogic structure and an informal writing style with non-standard language (cf. Storrer 2017). As we look at the Wikipedia talk pages, we would like to make one more linguistic differentiation: There is text-oriented writing that can be found on the article pages with monologic structure, standard language and elaboration of conventional textual patterns. On Wikipedia talk pages other different specifics can be identified such as dialogic structure, informal writing style and the tendency for the speed of an answer to outweigh an elaborate wording. This study focuses on the interaction-oriented language on Talk pages. There are two types of Wikipedia talk pages: article talk pages, where authors negotiate online encyclopedic content and user talk pages, where the contributions of individual authors are discussed. The metadata for the corpora used are as follows, cf. Table 1:

Table 1: Size of the corpora in tokens and corpus abbreviations⁸ (DeReKo 2022 in COSMAS II_{web} 2024).

Language	Article talk pages	User talk pages
German	373,161,686 (wdd17)	309,390,966 (wud17)
French	138,068,162 (wdf15)	374,390,445 (wuf15)
Italian	52,070,465 (wdi15)	130,067,969 (wui15)

To be able to investigate meta-discourses and thus the negotiation of appropriate address pronouns, we use the following search strings when conducting queries in COSMAS II_{web}:

- GER: &siezen and &duzen
- FR: *vouvoyer* and *tutoyer*⁹
- IT: *dare del Lei* and *dare del tu*¹⁰

⁸ The corpus abbreviations read as follows, **wdd17** is the **W**ikipedia corpus of German (**d**eutsch) article talk (**D**iskussion) pages created from a 2017 Wikipedia dump; **wud17** represents the **u**ser discussion pages.

⁹ All inflected forms were queried in a rather complex REG# (regular expression) search string: #REG(^tuto(ie(nt|r(a(i(s|(en)?t)|s)?|i?(ez|ons)|ont)?|s)?|y(a(i(s|(en)?t)|nt|s(s(e(nt|s)?|i(ez|ons)))?)?|â(mes|t(es)?|é(es)?|er|èrènt|i?(ez|ons)))\$) oder #REG(^vouvo(ie(nt|r(a(i(s|(en)?t)|s)?|i?(ez|ons)|ont)?|s)?|y(a(i(s|(en)?t)|nt|s(s(e(nt|s)?|i(ez|ons)))?)?|â(mes|t(es)?|é(es)?|er|èrènt|i?(ez|ons)))\$).

¹⁰ All inflected forms were queried in a rather complex REG# (regular expression) search string: #REG(^d(à(nno)?|a(i|n((d|n)o|te)|r(à|a(i|nno)|e(bbe(ro)?|i|mm?o|st(e|i|i)|te)?|d)|t(a|e|i|i|o)|v(

The German search string uses the base form operator „&-ampersand“, also called *lemmatization operator* or short: *lemmatizer* that allows searching not only for inflectional forms, but also for word formation forms for a base form.¹¹ The lemmatization of COSMAS II_{web} is tailored to German. Therefore, the usage of the basic form operator ‘&’ in queries to the foreign language Wikipedia corpora cannot be applied. For English and Italian, a workaround, i.e., placeholder operators, can be used (e.g., search operator *) or the specific forms can be explicitly listed in the search expression. We used the latter option and implemented queries consisting of regular expressions to gain results. A total of 12 queries, four for each language version with two for each Wikipedia subcorpus, i.e., T form and V form on article talk and user talk pages were conducted via the COSMAS II_{web} interface.¹²

4 Data analysis

In this section, the results of the quantitative analysis as well as the language-specific qualitative analyses are presented. First, the results of the queries that were conducted using the Wikipedia corpora within the COSMAS II_{web} research interface are described in section 4.1. In section 4.2, examples from the German, French and Italian Wikipedia article talk and user talk pages are used to show the crosslingual similarities and differences when Wikipedia authors make meta-linguistic utterances about the use of pronominal forms of address.

4.1 Quantitative analysis

In the following, it will be quantitatively demonstrated to what extent corpus hits referring to a meta-discourse on social deixis can be found in the three languages under consideration.

a((m|n)o|te|i)?|i|o))?)|e(mmo|s(s(e(ro)?|i(mo)?))|st(e|i)|tt(e(ro)?|i))|i(a((m|n)o|te)?|e(d(e(ro)?|i)))|o|ð)\$)/+w1 del /+w1 (tu oder lei oder Lei).

11 A *base form* describes an uninflected word or word formation morpheme.

12 The results of these queries are described in detail in section 4.

Table 2: Results of the search queries in the Wikipedia corpora (DeReKo 2024 in COSMAS II 2024).

Language	Wikipedia namespace	Query	Occurrences	pMW ¹³	Texts
German	Talk page	wdd17: V form (<i>siezen</i>)	322	0.86	208
German	Talk page	wdd17: T form (<i>duzen</i>)	993	2.66	682
German	User talk page	wud17: V form (<i>siezen</i>)	395	1.21	290
German	User talk page	wud17: T form (<i>duzen</i>)	2,052	6.29	1,659
French	Talk page	wdf15: V form (<i>vouvoyer</i>)	103	0.75	95
French	Talk page	wdf15: T form (<i>tutoyer</i>)	449	3.25	426
French	User talk page	wuf15: V form (<i>vouvoyer</i>)	200	0.53	181
French	User talk page	wuf15: T form (<i>tutoyer</i>)	1,655	4.42	885
Italian	Talk page	wdi15: V form (<i>dare del L/lei</i>)	29	0.56	9
Italian	Talk page	wdi15: T form (<i>dare del tu</i>)	61	1.17	61
Italian	User talk page	wui15: V form (<i>dare del L/lei</i>)	84	1.61	82
Italian	User talk page	wui15: T form (<i>dare del tu</i>)	372	7.14	308

For the German language version, it can be shown that both corpora (wdd17 and wud17) contain hits for both search strings (V form *siezen* and T form *duzen*), cf. Table 2, with occurrences of *duzen* being more frequent in both name spaces, i.e., the article talk as well as the user talk pages. For the French language version, Table 2 shows that both search strings (V form *vouvoyer* and T form *tutoyer*) yield results which, however, differ in their frequency: For both Wikipedia name spaces, i.e. the article talk pages as well as the user talk pages, inflected forms of the T form *tutoyer* are more frequently discussed than forms of the V form *vouvoyer*. French Wikipedia authors debate the means of addressing with each other; in sum more often on their own talk pages than on the article talk pages. The Italian language version contains hits for both search strings in both corpora, cf. Table 2. In particular the T form *dare del tu* is more discussed than the V form *dare del Lei/lei*.

Across all three investigated languages, meta-linguistic utterances about the use of pronominal forms of address are more frequent on the user talk pages than the article talk pages.

¹³ The abbreviation *pMW* stands for *occurrences per million words*. It is a measure of relative occurrence frequencies that are also normalized to a common base (one million current word forms). This allows for comparing frequencies in corpora of different sizes. To calculate pMW values, we need to divide the raw frequency by the total number of words in the corpus and multiply the result by one million.

We used a qualitative corpus-based approach in our study. For hit lists with more than 100 hits, random samples of N=100 were drawn, which were then used to deductively test the hypotheses mentioned at the beginning of the paper.¹⁴

4.2 Qualitative analysis

This section presents the qualitative results of the pronominal address analyses in German (section 4.2.1), French (section 4.2.2) and Italian (section 4.2.3) using language specific examples from the Wikipedia talk pages.

4.2.1 Pronominal address on German Wikipedia talk pages

Wikipedia's policies and guidelines in the German language version contain a detailed Wikiquote. There it says on pronominal address:

Viele der Umgangsformen der Wikipedia stammen noch aus der Zeit, als sich eine überschaubare Gruppe von Enthusiasten die Aufgabe gestellt hatte, eine Enzyklopädie zu schreiben, und nicht ahnen konnte, welchen gesellschaftlichen Stellenwert die Wikipedia eines Tages einnehmen würde. Aus dieser Zeit stammt auch das hier übliche vertrauliche „Du“ im Miteinander der Benutzer.

Many of Wikipedia's codes of conduct date back to the time when a small group of enthusiasts set themselves the task of writing an encyclopaedia and had no idea of the social status Wikipedia would one day have. The customary confidential "Du" form of address among users also dates from this time (Wikipedia 2024).

However, this seems to be the view of those Wikipedia authors who have been part of the online community for a long time (= high level of seniority) and came up with this rule in the early Wikipedia. The reason is that these authors have shared "perceived commonalities" (in German: *geteilte virtuelle Lebenswelt*, Kretzenbacher 2006: 10) for many years and are familiar with each other, even if they have never met face-to-face.

This is also shown by the example (1) from 2008. A long-time author demands the *Du* with reference to the Wikiquote: He not only refers linguistically to the Wikiquote (*nachlesbar auch unter WP:DU – also available at WP:DU*), but even

¹⁴ Our aim was to conduct a qualitative cross-linguistic study of forms of address. The examples in the following sub sections are taken from these lists and grouped thematically.

adds a corresponding hyperlink to the abbreviation WP:DU. In doing so, he suspends the factors (age and level of education) that are usually decisive in face-to-face communication for the choice of the appropriate pronominal form of address. Another author in (2) states on a talk page that *Sie* – the German V form – is impolite in Wikipedia. In (3) an author characterizes the V form of address as inappropriate (*unangebracht*) in Wikipedia:

- (1) Also vielleicht auch zum Einstieg die Information, dass sich in der **WP alle duzen**, unabhängig von Alter und Bildungsstand (nachlesbar auch unter WP:DU) (WUD17/K37.41593)
So maybe to start with the information that everyone in the WP uses du, regardless of age and level of education (also available at WP:DU)
- (2) Ich habe mit dir nichts zu diskutieren, und **sieze mich** nicht, das ist hier sehr unhöflich.
I have nothing to discuss with you, and don't use Sie, that's very impolite here (WDD17/C81.68097)
- (3) Es steht dir selbstverständlich frei zu **Siezen**, es wirkt aber **unangebracht**.
You are of course free to use the pronoun Sie, but it seems inappropriate. (WDD17/B09.74909)

The long-standing Wikipedia authors therefore adopt a pro T form stance and a contra V form stance and propagate the T form as the only correct form of address in Wikipedia: In their view, deserving authors correctly use the familiar T form with the aim to suspend differences due to various educational levels, for example. They thus suggest, at least superficially, an egalitarian claim. The V form, on the other hand, is characterized by them as an index of limited experience with Wikipedia. Through this stancetaking, they position themselves as superior to other interlocutors who use the V form.

Examples (4) to (6) are three examples of authors rejecting this wikiquote and also shared perceived commonalities:

- (4) Wir leben nicht mehr in den 90er Jahren. Wie Sie vielleicht festgestellt haben, wird sich mittlerweile auch im Internet auf seriösen, ‚erwachsenen‘ Seiten **zunehmend gesiezt**. Der Wikipedia würde ein seriöserer Anspruch auch unter seinen Mitarbeitern sicherlich nicht schaden. (WUD17/B72.41245)
We no longer live in the 1990s. As you may have noticed, on serious, ‘adult’ sites on the Internet authors now increasingly use the V form. Wikipedia

would certainly not be harmed by a more serious standard, even among its contributors.

- (5) ich sieze, weil sie mich auch Siezen würden, **stände ich vor Ihnen** – ebenso würde ich Sie siezen. (WDD17/E85.13333)
I use the V form because if I were standing in front of you, you would also use the V form – and vice versa.
- (6) Da ich Sie (Herr Hob) nicht kenne und ich Ihnen **das ‚Du‘ nicht angeboten** habe bitte ich, mich zu Siezen. (WDD17/B60.35170)
Since I don't know you (Mr. Hob) and I haven't offered you the T form, I ask that you address me with the V form.

From the broader context of the examples (4) to (6), it can be concluded that these are authors who have not been contributing to Wikipedia for that long. In (4), the informal form of address is dismissed as a 1990's bad habit. In (5), the comparison is drawn to the face-to-face situation in which strangers would be formally addressed. In (6), the author points out that there was no offer to change from V to T. Thus, he sees no legitimate basis for using the informal form. Authors with a low level of seniority in the ad hoc meritocracy therefore reject the use of the T form by default on the talk pages. Therefore, they adopt a pro V form stance and a contra T form stance.

Interesting insights also arise from corpus hits in which not only pronominal forms of address, but also nominal forms of address and nominal reference play a crucial role, as in examples (7) to (9):

- (7) Du/Sie?: Bei Wikipedia duzen wir uns alle – egal, ob 14 oder 140, **Schüler oder Professor** (oder sonst was). (WUD17/B58.97203)
T/V: On Wikipedia, we all use T - no matter if you're 14 or 140, a student or a professor (or whatever).
- (8) P.S. Das Duzen stellt tatsächlich in der Wikipedia die Höflichkeitsform dar, Siezen gilt allgemein als Zeichen von Distanzierung bzw. von Unkenntnis (**,Newbie‘**) (WUD17/P51.87518)
P.S. Duzen actually represents the polite form in Wikipedia, Siezen is generally considered a sign of distancing or of ignorance (,newbie‘)
- (9) Letzten Endes möchte ich noch darauf hinweisen, dass man sich in der Wikipedia allgemein duzt. Ich habe Sie bis zu diesem Punkt gesiezt, um Sie als **Wikipedia-Neuling** nicht vor den Kopf zu stoßen, gehe zukünftig aber zum ‚Du‘ über, wie es in der Wikipedia-Etikette Standard ist. (WUD17/K25.79885)

Lastly, I would like to point out that on Wikipedia we generally say Sie. I have been using Du with you up to this point so as not to offend you as a Wikipedia newcomer; but will switch to 'you' in the future, as is standard in Wikipedia etiquette.

Long-time authors pretend to abolish hierarchies of institutional offline contexts and to suspend professional roles (example 7). They do this with the alleged aim of achieving maximum equality between all authors. At the same time, however, long-time authors introduce a new hierarchy in the ad hoc meritocracy Wikipedia, which focuses on seniority in the online encyclopedic project. New authors are explicitly addressed nominally as *newbie* (8) or Wikipedia newcomer (*Wikipedia-Neuling*, 9) and thus degraded in a certain respect. Via the analysis of metalinguistic comments in the German language version, it thus becomes clear that lines of conflict do not run along classical hierarchies, but along newly created hierarchies that are oriented towards the level of seniority in the Wikipedia project.

4.2.2 Pronominal address on French Wikipedia talk pages

In French, the usage of the V form *vous* form of address is greatly reflected upon. There are for example specific user boxes which can be implemented on a user page that indicate how a user wishes to be addressed. Although some users prefer the informal *tu*, the formal *vous* form of address still plays a rather important role in Wikipedia user addressing.

The following examples show a preference for the formal form of address, i.e., a pro V form and contra T form stance:

- (10) Serait-il possible d'éviter les familiarités, **je ne tutoie personne**, merci de faire de même, nous ne sommes pas des amis ou ennemis de Facebook ! (WDF15/A73.66675)
Would it be possible to avoid familiarities, I'm not addressing anyone with „tu“, please do the same, we're not Facebook friends or foes!
- (11) Merci de **ne me pas me tutoyer**, on n'a pas gardé les cochons ensemble. (WDF15/B68.70940)
Thanks for not addressing me with „tu“, “we didn't keep pigs together“.
- (12) Je vous ai vouvoyé, je **n'accepte donc pas être tutoyée** (WDF15/M14.64419)
I addressed you with the formal „you“, I won't accept being addressed with the informal „you“

In (10) the user prefers being addressed with *vous* and makes a clear distinction between Wikipedia and other digital platforms, such as Facebook.

In (11), the French idiom *ne pas avoir garder les cochons ensemble* (*don't get so familiar with me!* or *don't get so pally with me!*) is generally used if somebody expresses an inappropriate or unwanted level of familiarity. The term *familiarity* plays a keyword role in the discussion about pronominal address, as it is also being explicitly mentioned in Example 1.

In example (12), the formal addressing given to the receiver is equally demanded by the sender.

- (13) @Guil2027 Le vouvoiement, vois-tu, permet de **maintenir une certaine distance** avec des contributeurs dont on souhaite justement éviter la proximité (WDF15/C18.89546)

You see, being formal allows you to maintain a certain distance from contributors whose proximity you wish to avoid.

- (14) Donc **merci** d'éviter de me tutoyer comme si j'étais un copain et **de garder vos distances** ; je garderai les miennes à votre égard comme je l'ai toujours fait (WDF15/M66.99763)

So please don't address me with "tu" as if I were a friend, and keep your distance; I'll keep mine from you as I've always done.

- (15) Je ne tutoie pas, c'est **une distance nécessaire** : nous ne sommes pas une bande de copains qui causent au coin d'un comptoir de troquet (WDF15/D18.63069)

I'm not on familiar terms, it's a necessary distance: we're not a bunch of buddies chatting at the corner of a bar.

The notion and importance of 'distance' as another key concept is explicitly addressed in postings. Again, the authors take a pro V form stance by underlining the importance of social distance between each other, as in (13) and (14). In (15) the user draws a clear distinction between work in Wikipedia and, for example, communication to close people, such as friends. Looking at this from another perspective, it could also exemplify a general hesitation in terms of committing to a more informal T form. Once the authors are exchanging terms of familiarity and closeness, it might be difficult to return to the initial situation. Peeters summarizes this accordingly:

Le tutoiement signale souvent un point de non-retour, un degré d'intimité auquel il est difficile de renoncer, alors que le vouvoiement constitue un comportement moins engagé. (Peeters 2004: 7)

Being on familiar terms often signals a point of no return, a degree of intimacy that is difficult to relinquish, whereas being formal is a less committed behavior.

Lastly, personal preference as a factor, closely related to linguistic conservatism and ease with address, is also mentioned as a reasoning in terms of opting for a T form:

- (16) Bonjour, vous **pouvez me tutoyer**, depuis le temps que nous nous croisons sur le projet ! (WDF15/F62.64673)
Hello, you can address me with “tu” for as long as we’re working on this project together!
- (17) Certes, le **tutoiement est d’usage sur Wikipédia** comme sur beaucoup de sites collaboratifs, mais n’est en rien une obligation. (WDF15/C18.89546).
Of course, as on many collaborative sites, it’s customary to use informal forms of address on Wikipedia, but it’s by no means compulsory.
- (18) Je suis fatigué de faire des listes de personnes à tutoyer ou à vouvoyer: **sur WP le tutoiement est quasi l’usage général** et je tutoie toute personne qui a un compte (pour les IP c’est selon mon humeur...) (WDF15/H09.03179)
I’m tired of making lists of people to be addressed with „vous“ or „tu“: on WP, „tu“ is almost the general rule, and I use „tu“ with anyone who has an account (for IPs, it’s up to me...)

In (16), the joint project work qualifies for using the informal *tu*. (17) and (18) underline the defacto-standard, or customary usage of informal forms of address on Wikipedia. Moreover, in (18) the user expresses some level of frustration regarding the topic of addressing by stating that they are *tired of making lists of people to be addressed with „vous“ or „tu“*: on WP, „tu“ is almost the general rule, and I use „tu“ with anyone who has an account – again, one’s own, personal preference is highlighted.

In several user surveys no consensus regarding a standardized form of address in the French Wikipedia could be generated, so both forms of addressing continue to be used depending on a user’s preference.¹⁵ Such a preference can be publicly stated on a user page by implementing the respective *Wikipedia:Userbox*.¹⁶ There are boxes dedicated to the preferred form of pronominal address. A userbox stating

¹⁵ https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Le_Bistro/24_novembre_2014#Tutoiement (last accessed 14 February 2025).

¹⁶ Note that userboxes are not mandatory therefore their frequency should not be used in terms of generalization.

the preference for a V form is called *Modèle:Utilisateur vouvoiment* and for a T form it is *Modèle:Utilisateur Tutoiement*. It is possible to assess the numbers of each userbox by consulting the respective category pages. A total of 112 users has the userbox *Tu ou vous*¹⁷ shown on their page, indicating they do not mind using V and T forms interchangeably, depending on the preference of their counterpart. On 83 user pages the box *Utilisateur vouvoiment*¹⁸ is placed, showing that these users would like to be addressed with a V form. A total of 411 user pages has the userbox *Utilisateur Tutoiement*¹⁹ implemented, i.e., being in favor of T forms of address.

4.2.3 Pronominal address on Italian Wikipedia talk pages

The shared “perceived commonalities” and the familiarity due to being part of a community, even if they have never met face-to-face, is also a characteristic of the Italian Wikipedia talk pages. The use of ,tu’ is considered a convention (examples 19 and 20) and used so by default. The fact of belonging to a community of Wikipedians and being so among *wikifili* is highlighted in (21), while in (22) the focus is on being part of a group. The use of the deictic „here“ is often present to emphasize this use in a specific context: namely Wikipedia. It can be deduced from the examples that new authors tend probably at first to use the V form (19 and 20) (it is an index of limited experience with Wikipedia). Longtime authors are the ones who propose the T form, propagating the T form as the only correct form of address in Wikipedia, enforcing the wikiquette and suspending, superficially, differences (22). Through this stancetaking, they position themselves as superior to other interlocutors who use the V form.

- (19) [...] qui ci si da del **TU** per **convenzione** [...] (WUI15/A28.07711)
[...] *here we adress each other with ,TU’ by convention* [...]
- (20) [...] Per **convenzione**, wikipedia usa “di default” il **tu** [...] (WUI15/A06.20577)
[...] *By convention, wikipedia uses ‘by default’ ,tu’* [...]
- (21) Credo che una distinzione più appropriata potrebbe essere fatta sulla committenza e sul pubblico cui l’opera si rivolge: volendo passare al tuo esempio musicale (**ti do del tu**, visto che siamo tra **wikifili**), ti invito a riflettere come

17 https://fr.wikipedia.org/wiki/Cat%C3%A9gorie:Utilisateur_Tu_ou_vous (last accessed February 2025).

18 https://fr.wikipedia.org/wiki/Cat%C3%A9gorie:Utilisateur_vouvoiment (last accessed February 2025).

19 https://fr.wikipedia.org/wiki/Cat%C3%A9gorie:Utilisateur_Tutoiement (last accessed February 2025).

canzoni di musica leggera hanno valore artistico riconosciuto e successo di pubblico (vd. il caso De André). (WDI15/B24.05597)

I think a more appropriate distinction could be made on the patronage and the audience to which the work is addressed: turning to your musical example (I'm addressing you with ,tu', since we are among wikifili), I invite you to reflect on how pop music songs have recognised artistic value and public success (see the De André case).

- (22) Caro Alberto. Non ti do del tu perché sono d'origine siciliana o perché sono anch'io un arabista (in realtà sono più storico e islamista, ma con 40 anni di studio dell'arabo, una laurea e un perfezionamento con Francesco Gabrieli e Alessandro Bausani). **Ti do del tu** perché qui **da noi si usa darsi del tu**, a meno che non ci siano particolarissime idiosincrasia in merito. (WUI15/A08.57762)

Dear Alberto. I do not use 'tu' because I am of Sicilian origin or because I too am an Arabist (actually, I am more of a historian and Islamist, but with 40 years of Arabic studies, a degree and further education with Francesco Gabrieli and Alessandro Bausani). I use 'tu' because it is customary to address each other 'tu' here, unless there is a particular idiosyncrasy about it.

Social status, hierarchies, age are no more a factor in choosing the appropriate form of address pronoun. Factors such as titles in (23), age in (24 and 25) and sympathy (25) are suspended. Other aspects like equality of all Wikipedians are applied (26): there are no differences between administrators and users.

- (23) Ciao, meglio essere meno formali, di solito **su wikipedia ci si dà del tu e si evita l'uso dei titoli**. (WUI15/A26.79745)

Hi, better to be less formal, usually on wikipedia we address each other using „tu“ and avoid the use of titles.

- (24) [...] Un cordiale benvenuto anche da parte mia! (ovviamente sai già che qui **ci si dà del TU** [...] **senza tener conto di alcun limite**; io ad es. sono del **1945**) [...] (WUI15/A28.07711)

A warm welcome from me too! (Of course, you already know that we address us with ,tu' here by convention, regardless of any limits; I, for example, am from 1945) [...]

- (25) **Non conta l'età o la simpatia**. Siamo tutti coinvolti in un progetto (**noi wikipediani** amiamo scrivere Progetto, con la maiuscola) [...] (WUI15/A08.57762)

Age or liking does not count. We are all involved in a project (we Wikipedians like to write Project, with a capital)

- (26) [...] e anche per un discorso che, su wikipedia, **non ci sono distinzioni di sorta**, cioè un **amministratore è importante tanto quanto un utente** [...] (WUI15/A06.20577)
and also for a discourse that, on Wikipedia, there are no distinctions whatsoever; i.e., an administrator is just as important as a user

Uncertainty can be attributed to new authors, but in the corpus, we can also find examples of rejection to the wikiquette and rejection to the shared virtual reality. In (27) and (28) authors realize that they have been addressing with 'tu' from the start and ask for confirmation (27) or even apologize (28). In (29) the author alternates between 'Lei' and 'tu', emphasizing that it is a matter of new habits. In (30) there is a mixing of the two forms, with a tone of criticism and muted with irony (the character from a comedy film (Fantozzi) who tended to mix forms in speech is quoted). In (31) and (32) the use of the form is simply an adaptation to the chosen form of the interaction partner.

- (27) Grazie (**ti ho dato del tu** fin dal primo momento ... posso, vero?) (WUI15/A04.67431)
Thank you. I've been on familiar terms with you from the first moment ... I can, can't I?
- (28) Salve Angelo, perdonami se **ti do del tu** [...] (WUI15, 18.03.2009)
Hello Angelo, forgive me for addressing you with ,tu' [...]
- (29) **Guardi**, ovvero **guarda** bisogna che mi abitui agli usi correnti dove qui **tutti si danno del TU** (WUI15/A15.97376)
Look, I mean look I have to get used to the current customs where everyone here calls each other TU
- (30) LOL, pure la lezione di grammatica. Il problema è che l'hai usato eccome in altre occasioni **il terzo pronome singolare per rivolgerti a me** (stile: «Fantozzi che fa? **Mi dà del tu?** No, **no batti lei!**» :-PPP). Pace e bene, fratello :-P“ (WUI15/A17.53481)
LOL, even the grammar lesson. The problem is that you've used the third singular pronoun to address me on other occasions (like: 'Fantozzi, what are you doing? Are you calling me 'you'? No, no, you're beating me!' :-PPP). Peace and good, brother :-P“

- (31) Ciao. Visto che **mi ha dato del Lei**, glielo darò anch'io. (WUI15/A19.40028)
Hi. Since you addressed me with ,Lei', I'll address you like that too.
- (32) Visto che **mi dai del tu** te lo do anche io. (WUI15/A24.27639)
,Since you adress me with ,tu', I'll address you like that too.

The analysis of metalinguistic comments in the Italian language version shows again that lines of conflict do not run along classical hierarchies; there are new hierarchies, in which the level of seniority in the Wikipedia project plays a central role.

5 Conclusion

This paper analyzed pronominal forms of address in the CMC genre of Wikipedia talk pages for three language versions (German, French, and Italian). When comparing all three languages, the frequencies of discussing pronominal address meta-linguistically are significantly different between the German, French, and Italian Wikipedia language versions.²⁰ In both analysed Wikipedia subcorpora of the three language versions, i.e., the Wikipedia article talk pages on the one hand and the user talk pages on the other hand, a greater deal of discussions about addressing styles takes place on the user talk pages, with the T form being discussed more frequently than the V form.

In German, people who meet for the first time in face-to-face encounters often use the V form (*Sie*) for pronominal address. Only in special cases (e.g., in the case of certain shared leisure activities such as football) do interlocutors switch directly to the T form (*du*). In the case of the CMC genre of Wikipedia talk pages considered here, the wikiquote explicitly stipulates that interlocutors also address each other directly using the T form – even if they have never met before. This wikiquote was

²⁰ This holds for testing between the three languages, with the chi-square statistic being 87.5197. The p-value is < 0.00001 . The result is significant at $p < .05$ for comparing together the frequencies of the formal *you* variant as well as the informal *you* variant between the three languages, with the chi-square statistic being 61.361. The p-value is < 0.00001 . The result is significant at $p < .05$, cf. <https://www.socscistatistics.com/tests/chisquare2/default2.aspx>. For each language, the differences in frequencies between the two analysed corpus types, i.e., Wikipedia article talk pages and user talk pages, are significant for the formal *you* variant in German and French, e.g., for the formal *you* variant in German, *Sie*, the difference between the name spaces is significant with the chi-square statistic being 27.5725. The p-value is $< .00001$. The result is significant at $p < .05$; French: The chi-square statistic is 7.6534. The p-value is .005667. The result is significant at $p < .05$; not for Italian: The chi-square statistic is 0.4735. The p-value is .491403. The result is not significant at $p < .05$.

established by authors with a high level of seniority. This preference of longstanding authors for the T form goes hand in hand with the fact that they explicitly suspend the usual factors and hierarchies that are decisive for the choice of the appropriate pronominal form of address in face-to-face communication (e.g., age, level of education, professional roles) in their utterances on the talk pages. With this strategy, they suggest – at least superficially – that all authors of the Wikipedia online community have egalitarian rights.

However, new Wikipedia authors do not necessarily share this preference for the T form, as they lack the “perceived commonalities”. The qualitative analysis also showed that the discussions on the pronominal forms of address on the talk pages are particularly controversial when long-time Wikipedia authors linguistically construct new Wikipedia-specific hierarchies (*newbie*) in order to discredit new authors. Basically, the use and negotiation of pronominal forms of address in Wikipedia is an indication that Wikipedia is not comprehensively a platform with egalitarian aspirations, but that long-standing authors constitute linguistically an ad-hoc-meritocracy.

In French, symmetrical *vous* is in fact still the normal starting point for public interactions between adults who have no prior relation. It is commonly used between strangers or people who rarely interact with each other. In several user surveys no consensus regarding a standardized form of address in the French Wikipedia could be generated, so both forms of addressing continue to be used depending on a user’s preference. When investigating Wikipedia talk pages, there are users who prefer the informal *tu*; nonetheless, the formal *vous* form of address plays an important role in Wikipedia user addressing. The authors adopt a pro-V form stance, emphasizing the value of social distance between themselves and others. They make a clear separation between work on Wikipedia and, say, communicating with friends or family. These findings point to a certain degree of reluctance to adopt a less formal T form. Once the writers express terms of familiarity and intimacy, it might be challenging for them to get back to a formal stage of using V forms. In cases where Wikipedia authors are explicitly opting for a T form, personal preference as a factor closely related to one’s own overall ease with address are mentioned as reasonings by Wikipedia authors.

Subject pronouns are considered as an optional choice in Italian, because they are inherent to the verbal morphology. So when they are used, sometimes also together with nominal forms, they strategically add pragmatic meanings to a speaker’s utterance. Symmetry/asymmetry and familiarity/distance coupled with context formality regulate their use. The reciprocal use is preferred to index familiarity or to signal social distance and/or mutual respect: reciprocal V forms are considered the default option in academic interactions; the non-reciprocal use when there is an asymmetrical distribution of power (age, job rank, and social status). Different

factors as levels of formality of the setting, degree of familiarity, and individual preferences play an important role.

The dominance of T forms, considering vicinity, familiarity and solidarity, the norm in CMC communication is confirmed also in Wikipedia, where social hierarchy is weakened and simplified. The shared “perceived commonalities” and the familiarity due to being part of a community, even if they have never met face-to-face, is also a typical characteristic of the Italian Wikipedia. The use of the T form is considered a convention and used by default from longtime authors as the only correct form of address in Wikipedia, enforcing the wikiquote and suspending, superficially, differences. Through this stancetaking, they position themselves as superior to other interlocutors (new authors with limited experience with Wikipedia), who use the V form. The analysis of metalinguistic comments in the Italian language version shows again that lines of conflict run along new hierarchies, in which the level of seniority in the Wikipedia project plays a central role.

Aspects of pronominal address among speakers of German, French, and Italian are characterized by instability and uncertainty. The use of T forms (GER: *du*, FR: *tu*, IT: *tu*) is controversial among Wikipedia authors. They discuss aspects of hierarchies, seniority, levels of proximity and distance, as well as preference. The overall preference for either a V or a T form is highly individual and greatly influenced by notions of social status and familiarity between the users. Pronominal address in this CMC genre is a topic that has received little attention so far, but is very informative for linguistic investigation: This is the case because in CMC with its (partially) anonymous and translocal contexts, pronominal address poses different challenges for interlocutors than in face-to-face communication. In addition, the CMC genre of Wikipedia talk pages is particularly suitable for cross-linguistic comparisons because comparable corpus data is available in large quantities in numerous languages. In further studies, it would certainly also be interesting to examine the difference between pronominal forms of address on the various talk pages in more detail: The central question would then be whether the types of pronominal form of address on user talk pages differ from that on article talk pages.

References

- Ayers, Phoebe, Charles Matthews & Ben Yates. 2008. *How Wikipedia works: And how you can be a part of it*. San Francisco: No Starch Press.
- Bader, Jennifer. 2002. *Schriftlichkeit und Mündlichkeit in der Chat-Kommunikation*, Networx Nr. 29.
- Beißwenger, Michael & Harald Lungen. 2022. Korpora internetbasierter Kommunikation. In Michael Beißwenger, Lothar Lemnitzer & Carolin Müller-Spitzer (eds.), *Forschen in der Linguistik. Eine Methodeneinführung für das Germanistik-Studium*, 431–448. Paderborn: Fink.

- Bouissac, Paul. 2019. Forms and functions of the French personal pronouns in social interactions and literary texts. In *The Social Dynamics of Pronominal Systems*. 133–150. Amsterdam: John Benjamins.
- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Brown, Roger & Albert Gilman. 1960. The pronouns of power and solidarity. In Thomas A. Sebeok (ed.), *Style in Language*, 253–276. Cambridge: MIT Press.
- Clyne, Micheal, Heinz L. Kretzenbacher, Catrin Norrby & Jane Warren. 2004. Address in some Western European languages. In Christo Moskovsky (ed.), *Proceedings of the 2003 Conference of the Australian Linguistic Society*.
- Clyne, Michael, Catrin Norrby & Jane Warren. 2009. *Language and Human Relations. Styles of Address in Contemporary Language*. Cambridge: Cambridge University Press.
- Da Milano, Federica. 2015. Italian. In Konstanze Jungbluth & Federica Da Milano, Federica (eds.), *Manual of Deixis in Romance Languages*, 59–74. Berlin & Boston: De Gruyter.
- Dal Negro, Silvia & Giuseppina Pani. 2019. Tu ed io nel discorso: deissi, allocuzione e accordo come problema di ricerca e di didattica. In Elena Nuzzo & Ineke Vedder (eds.), *Lingua in Contesto: La Prospettiva Pragmatica*. Studi AItLA, 9, 47–63.
- Davidson, Brad. 1996. 'Pragmatic weight' and Spanish subject pronouns: the pragmatic and discourse uses of 'tú' and 'yo' in spoken Madrid Spanish. *Journal of Pragmatics*. 26, 543–565.
- De Oliveira, Sandi Michele. 2013. 12. Address in computer-mediated communication. In Susan Herring, Dieter Stein & Tuija Virtanen (eds.), *Pragmatics of Computer-Mediated Communication*, 291–314. Berlin & Boston: De Gruyter Mouton.
- Deppermann, Arnulf. 2015. Positioning. In Anna de Fina & Alexandra Georgakopoulou (eds.), *The Handbook of Narrative Analysis*, 369–387. Oxford: Wiley-Blackwell.
- Dewaele, Jean-Marc. 2004. "Vous or tu? Native and non-native speakers of French on a sociolinguistic tightrope". *IRAI* 42, 383–402.
- Douglass, Kate. 2009. Second-person pronoun use in French-language blogs: Developing L2 sociopragmatic competence. In Abraham, Lee B. & Lawrence Williams (eds.), *Electronic Discourse in Language Learning and Language Teaching*, 213–240. Amsterdam: John Benjamins.
- Du Bois, John W.. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, 139–182. Amsterdam: John Benjamins.
- Duden-Grammatik = Wöllstein, Angelika. 2016. *Die Grammatik*. Duden, Band 4. Berlin: Dudenverlag.
- Duranti, Alessandro. 1984. The social meaning of subject pronouns in Italian conversation. *Text* 4, 277–311.
- Formentelli, Maicol & John Hajek. 2013. Italian L2 address strategies in an Australian university setting: a comparison with L1 Italian and L1 English practice. In Bert Peeters, Kerry Mullan & Cristine Béal (eds.), *Cross-culturally Speaking, Speaking Cross-culturally*. 77–106. Newcastle: CSP.
- Formentelli, Maicol & John Hajek. 2015. Address in Italian academic interactions: the power of distance and (non-)reciprocity. In Catrin Norrby & Camilla Wide (eds.), *Address Practice as Social Action. European Perspectives*, 119–140. Basingstoke: Palgrave Macmillan.
- Formentelli, Maicol & Maria Pavesi. 2022. The pragmatic functions of tu and lei in films: Converging patterns of address across translated and original Italian dialogue. *Journal of Pragmatics* 201, 15–31.
- Fillmore, Charles. 1975. *Santa Cruz lectures on deixis 1971*. Bloomington: Indiana University Linguistics Club.
- Gastaldi, Erika. 2002. Italiano digitato. *Italiano e oltre* 17, 3, 134–138.
- Gredel, Eva. 2017. Digital discourse analysis and Wikipedia: Bridging the gap between Foucauldian discourse analysis and digital conversation analysis. *Journal of Pragmatics* 115, 99–114.

- Gredel, Eva. 2023. Siezt du noch oder duzt du schon? Korpusstudie zum Gebrauch und zur Aushandlung sozialdeiktischer Zeichen auf digitalen Plattformen. In Simon Meier-Vieracker, Lars Bülow, Konstanze Marx & Robert Mroczynski (eds.), *Digitale Pragmatik*, 39–57. Berlin & Heidelberg: Metzler.
- Helmbrecht, Johannes. 2006. Typologie und Diffusion von Höflichkeitspronomina in Europa, *Folia Linguistica* 39, 3–4, 417–452.
- Helmbrecht, Johannes. 2015. A typology of non-prototypical uses of personal pronouns: synchrony and diachrony, *Journal of Pragmatics* 88, 176–189.
- Hess-Lüttich, Ernest W.B. & Eva Wilde. 2003. Der Chat als Textsorte und/oder als Dialogsorte?, *Linguistik online* 13, 1, 161–180.
- Kretzenbacher, Heinz & Doris Schüpbach. 2015. Communities of Addressing Practice? Address in Internet forums Based in German-Speaking Countries. In Catrin Norrby & Camilla Wide (eds.), *Address Practice As Social Action: European Perspectives*, 33–53. New York: Palgrave.
- Kretzenbacher, Heinz & Wulf Segebrecht. 1991. *Vom Sie zum Du – mehr als nur eine Konvention?* Hamburg: Wallstein Verlag.
- Kretzenbacher, Heinz. 2010. Man ordnet ja bestimmte Leute irgendwo ein für sich ...“. Anrede und soziale Deixis. *Deutsche Sprache* 38, 1, 1–18.
- Kretzenbacher, Heinz. 2011. Addressing Policy on the Web: Netiquettes and Emerging Policies of Language Use in German Internet Forums. In Catrin Norrby & John Hajek (eds.), *Uniformity and Diversity in Language Policy: Global Perspectives*, 226–241. Bristol: De Gruyter, Blue Ridge Summit: Multilingual Matters.
- Kretzenbacher, Heinz. 2005. ‚hier im großen internetz, wo sich alle dududuzen‘ *Internet discourse politeness and German address*. Paper given at the 3rd International Conference on Language Variation in Europe (ICLaVE).
- Leech, Geoffrey. 1983. *Principles of pragmatics*. London & New York: Longman.
- Lüngen, Harald & Marc Kupietz. 2020. IBK- und Social Media-Korpora am Leibniz-Institut für Deutsche Sprache. In Konstanze Marx, Henning Lobin & Axel Schmidt (eds.), *Deutsch in Sozialen Medien: Interaktiv – multimodal – vielfältig*, 319–342. Berlin: De Gruyter.
- Maeder, Costantino & Romane Werner. 2019. T-V address practices in Italian: diachronic, diatopic, and diastratic analyses. In Paul Bouissac (ed.), *The Social Dynamics of Pronominal Systems: A Comparative Approach*. 99–131, Amsterdam: John Benjamins.
- Maley, Catherine. 1972. “Historically Speaking, Tu or Vous?” *The French Review* 45, 5, 999–1006.
- Molinelli, Piera. 2002. “Lei non sa chi sono io!”: potere, solidarietà, rispetto e distanza nella comunicazione. In *Linguistica e fonologia* 14, 283–302.
- Morford, Janet. 1997. Social Indexicality in French Pronominal Address. *Journal of Linguistic Anthropology* 7, 1, 3–37.
- Nübling, Damaris, Antje Dammel, Janet Duke & Renata Szczepaniak. 2017. *Historische Sprachwissenschaft des Deutschen. Eine Einführung in die Prinzipien des Sprachwandels*. Tübingen: Narr.
- Ochs, Elinor 1996. Linguistic resources for socializing humanity. In John J. Gumperz & Stephen Levinson (eds.), *Rethinking Linguistic Relativity*. 407–437, New York: Cambridge University Press.
- Orletti, Franca. 2000. *La conversazione diseguale. Potere e interazione*. Roma: Carocci.
- Peeters, Bert. 2004. TU OU VOUS? *Zeitschrift Für Französische Sprache Und Literatur* 114, 1, 1–17.
- Pistolesi, Elena. 1997. “Il visibile parlare di IRC (Internet Relay Chat)”. In *Quaderni del Dipartimento di Linguistica* 8, 213–246.
- Pistolesi, Elena. 1998. IRC (Internet Chat Relay): Una nuova tecnologia Guida storica, linguistica e tecnica. *Web Italica della RAI*.
- Renzi, Lorenzo. 1993. La deixis personale e il suo uso sociale. *Studi digrammatica italiana* 15, 347–390.

- Renzi, Lorenzo. 1995. La deissi personale e il suo uso sociale. In Lorenzo Renzi & Anna Cardinaletti (eds.), *Grande grammatica italiana di consultazione*, 350–375, Bologna: il Mulino.
- Rebelos, Margareta & Antonella Strambi. 2009. Address Pronouns in Italian CMC Exchanges: A 'Good Example' for L2 Learners?. *Italica* 86, 1, 59–79.
- Schoch, Marianne. 1978. Problème sociolinguistique des pronoms d'allocution < tu > et < vous >. Enquête à Lausanne, *La linguistique* 14, 1, 55–73.
- Schulze, Markus. 1999. Substitution of paraverbal and nonverbal cues in the written medium of IRC. In Bernd Naumann (ed.) *Dialogue analysis and the mass media*. Proceedings of the international conference in Erlangen, April 2-3, 1998, 65–82, Tübingen: Niemeyer.
- Simon, Horst. 2003. *Für eine grammatische Kategorie ›Respekt‹ im Deutschen: Synchronie, Diachronie und Typologie der deutschen Anredepronomen*. Tübingen: Niemeyer.
- Spitzmüller, Jürgen. 2022. *Soziolinguistik. Eine Einführung*. Berlin: Springer.
- Stegbauer, Christian. 2009. *Wikipedia: Das Rätsel der Kooperation*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Stewart, Miranda. 2003. 'Pragmatic weight' and face: pronominal presence and the case of the Spanish second person singular subject pronoun tú. *Pragmat* 35, 2, 191–206.
- Svennevig, Jan. 1999. *Getting acquainted in conversation: A study of initial interactions*. Amsterdam: John Benjamins.
- Truan, Naomi. 2022. (When) Can I say *Du* to You? The metapragmatics of forms of address on German-Speaking Twitter. *Journal of Pragmatics* 191, 227–239.
- Williams, Lawrence & Rémi A. van Compernelle. 2007. Second-Person Pronoun Use in On-Line French-Language Chat Environments. *The French Review* 80, 4, 804–20.
- Williams, Lawrence & Rémi A. van Compernelle. 2009a. On versus tu and vous: Pronouns with indefinite reference in synchronous electronic French discourse, *Language Sciences* 31, 4, 409–427.
- Williams, Lawrence & Rémi A. van Compernelle. 2009b. Second-person pronoun use in French language discussion fora. *Journal of French Language Studies* 19, 3, 363–380.
- Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Berlin & New York: De Gruyter.