How Do Interlocutors in Instant Messaging Influence Each Other's Writing Style? Three Case Studies on Accommodation in Teenagers' Chat Conversations

Lisa Hilte, Reinhild Vandekerckhove, Walter Daelemans
University of Antwerp
E-mail: lisa.hilte@uantwerpen.be, reinhild.vandekerckhove@uantwerpen.be, walter.daelemans@uantwerpen.be

Abstract
We bring together three complementary case studies on the linguistic phenomenon of accommodation (i.e. people adapting their language use to that of their interlocutor) in Flemish teenagers' instant messages on Facebook Messenger and WhatsApp. Each case study focuses on a different aspect of the interlocutors' socio-demographic profile which may lead to distinct adaptation patterns: their gender, education, and age. Analyses with generalized linear mixed models reveal how teenagers' usage frequency for two sets of prototypical chatspeak features depends on the profile of the conversation partner. Significant adaptation patterns emerge with respect to all three socio-demographic variables. However, the linguistic features that are subject to adaptation and the accuracy of this so-called 'mirroring' differ depending on the interlocutors' profiles, which points to different accommodative behavior by distinct groups of teenagers.

Keywords: accommodation, mirroring, social media, teenagers, gender, education, age

1. Introduction

It has been attested repeatedly that teenagers' socio-demographic profiles and their online writing style correlate: youths with distinct profiles (in terms of e.g. gender or age) tend to favor certain markers of online writing to different extents (De Decker and Vandekerckhove, 2017; Hilte et al., 2020b; Verheijen, 2018). But little is known about the impact of the interlocutor's profile in online contexts such as instant messaging. While the phenomenon of accommodation (i.e. the adaptation of one's communicative behavior to that of one's interlocutor) has been widely investigated in spoken face-to-face interactions (see below), it is under-researched in (spontaneous, synchronous) written online conversations. The present paper aims to fill this gap by bringing together three complementary case studies on accommodation in teenagers' instant messages, with a respective focus on the interlocutors' gender, education, and age[1].

2. Related research

According to the sociolinguistic framework *Communication Accommodation Theory*, accommodation is driven by a desire to facilitate interaction and to regulate social distance between interlocutors (Dragojevic et al., 2015). Adapting one's language to that of others decreases the linguistic but also the social distance between interlocutors.

While the inclination to mirror the communicative style of others differs among individuals, some robust patterns relating to interlocutors' socio-demographic profiles have been found. With respect to gender, asymmetrical convergence has often been established, with women adapting their language use more strongly to men than vice versa (Palomares et al., 2016). But speech complementarity has been attested too, i.e. mutual divergence by men and women in order to consolidate social gender roles (Dragojevic et al., 2015).

In interactions between people of different ages/generations, common patterns are under-accommodation by older interlocutors versus over-accommodation by younger interlocutors, i.e. failing to adjust versus overadjusting one's communicative behavior to others (Giles and Gasiorek, 2011; Williams and Nussbaum, 2001). Overaccommodation to the elderly in particular is called 'patronizing talk', and includes e.g. oversimplified and excessively loud speech (Giles and Gasiorek, 2011; Williams and Nussbaum, 2001).

Interlocutors' educational background, finally, has - to the best of our knowledge - not yet been studied with respect to accommodation. Several papers do discuss the related variable of social power or position. So-called 'upward' social convergence has been attested repeatedly, that is more communicative adaptation towards interlocutors with greater (social) power (Dragojevic et al., 2015).

3. Data and method

3.1. Corpus

The corpus that is investigated in the three case studies contains 456,751 social media messages produced by 1,398 Flemish teenagers (living in Flanders, Dutch-speaking Belgium). The teenagers, aged 13-20, are pupils in secondary education. They attend general, technical, or vocational secondary education, which range from very theory- to mainly practice-oriented. The dataset contains the pupils' spontaneous, private instant messages, produced in Dutch on Facebook Messenger and

[1]Each of the case studies is published or forthcoming (Hilte et al., 2020a; Hilte et al., 2021; Hilte et al., under review). For an extensive discussion of the related work, research design and results per study, we refer to these papers. The present contribution brings these separate articles together for the first time and will focus on comparing and confronting the findings. This synthesis, in which studies on three distinct socio-demographic variables are confronted, offers a more complete and holistic perspective on accommodation.

WhatsApp, mainly between 2015-2016.

The corpus was collected in collaboration with schools. Pupils were invited to voluntarily donate chat conversations that were produced before our school visits. Note that some groups of pupils (e.g. girls) donated more data than others (e.g. boys). The participants also provided the relevant metadata: their own age, gender, and educational track (see Table 1 for the distributions in the corpus) and the age of their interlocutors (in case their data were not part of the teenage corpus). Finally, we asked the pupils' (and for minors also their parents') permission to store and linguistically analyze their anonymized messages.

Variable	Levels	Tokens
Gender	Girls	1,759,067 (66%)
	Boys	894,857 (34%)
Age	Young teens (13-16)	1,385,802 (52%)
	Older teens (17-20)	1,268,122 (48%)
Education	General	747,867 (28%)
	Technical	1,192,595 (45%)
	Vocational	713,462 (27%)
Medium	Facebook Messenger	2,045,396 (77%)
	WhatsApp	579,463 (22%)
Total		2,653,924

Table 1: Distributions in the corpus

3.2. Linguistic Variables

The case studies include two sets of linguistic variables that are prototypical of instant messaging, as they relate to two 'maxims' (implicit rules of linguistic conduct) of informal online writing, the maxims of expressive compensation and orality (Thurlow and Poff, 2013).

The principle of expressive compensation entails several (mostly typographic) strategies that compensate for the absence of certain expressive cues in written communication, such as facial expressions or voice volume. Examples of features are emoticons and emoji (e.g. :D), character repetition (e.g. *niiiice!!!*), and allcaps (e.g. *YES*).

The orality principle concerns speech-like writing: the register in instant messages often reflects typical speech patterns rather than classical written communication. In our corpus of Flemish teenagers' online writing, this results in the use of regional and colloquial features (e.g. *zoekt gij ambras?* for *zoek jij ruzie?*, 'are you picking a fight?'), and the insertion of English words or phrases typical of adolescent talk (e.g. *echt nice*, 'really nice').

All feature occurrences were detected automatically in the corpus with Python scripts. The software's performance was evaluated as reliable on a manually annotated test set (see Hilte et al. (2020b) for an extensive discussion).

3.3. Method

We approach accommodation from a quantitative perspective, i.e. as significant in- or decreases in the authors' usage frequency of linguistic features depending on their interlocutors' profiles. In all three case studies, generalized linear mixed models (Poisson distribution) were used to predict the participants' usage frequency for expressive and oral features, with aspects of their interlocutors' socio-demographic profiles serving as predictors. A random effect for subject and conversation was added to correct for repeated measurements, as both authors and conversations may occur more than once in the corpus (e.g. the same author in different interactions, or one interaction represented by its different participants). An observation-level random effect was added to avoid overdispersion (i.e. the variance of the response exceeding the mean; see Hilte et al. (2020b)). Finally, differences in sample size between observations are dealt with through an offset for (the logarithm of) the number of tokens per observation. Below, we discuss the best models per case study, i.e. the models that best fit the data (experimentally determined through stepwise deletion of insignificant predictors).

4. Results

Below, the results of the case studies are summarized. The present section and the discussion section aim to compare and confront the three separate analyses. For a more detailed discussion of the results per study (including e.g. model summary tables), see Hilte et al. (2020a; 2021; under review).

4.1. Gender

The first case study (Hilte et al., 2020a) compared teenagers' mixed-gender chats (including boys as well as girls; 34% of all conversations in the dataset) to same-gender chats (including only boys or only girls; 66% of all conversations in the dataset). The statistical analyses revealed no significant accommodation for oral markers (e.g. regional and colloquial language features), which are more prototypical 'male' features (Hilte et al., 2020b). But significant adaptation did emerge for expressive markers (e.g. emoji), which are generally inserted much more frequently by girls than boys (Hilte et al., 2020b). Both genders alter their use of expressive markers in mixed-gender chats, making their writing style more similar to that of their interlocutor of the opposite gender (see Figure 1). In one-on-one talks (including two interlocutors), which are shown on the left panel of Figure 1, girls insert significantly fewer ($p = 0.0266$) and boys significantly more expressive markers ($p < .0001$) when interacting with someone of the opposite gender versus when they chat with someone of the same gender. But, as Figure 1 shows, boys converge much more strongly to a (more expressive) 'female' style than vice versa. This contradicts previous work on spoken interactions, in which either a stronger convergence by women has been observed, or mutual divergence (see above). So the phenomenon of gender accommodation does extend from spoken to online written interactions, but the specific gender convergence patterns that have been attested in oral settings do not hold in our corpus of instant messages. Finally, we note that a similar

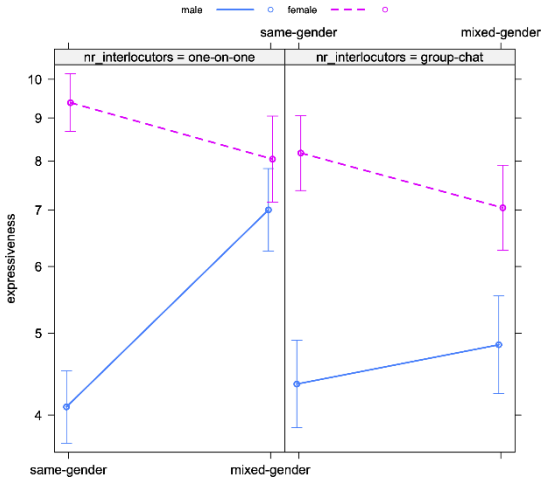

Figure 1: Expressive markers by author's gender, in same- vs mixed-gender conversations that are either one-on-one or group chats (predicted counts per 100 tokens).

Figure 2: Interaction between two female friends.

gender accommodation pattern as in one-on-one chats can be observed in group chats (including more than two interlocutors), as shown on the right panel of Figure 1. However, the accommodative adjustment by girls nor boys appeared significant (p = 0.07 resp. 0.1415). Intuitively, this finding of 'weaker' linguistic mirroring in group conversations makes sense – but we will come back to it in the discussion.

Since significant adaptation was only observed for expressive markers, which can serve as 'tools' for flirting, we examined to which extent gender accommodation can be obscured by flirting. An exploratory qualitative analysis of the corpus revealed that specific expressive markers with a romantic connotation (e.g. heart emoji) need not involve flirting, at least not in all-girls chats: girls use them abundantly in both romantic and non-romantic conversations (e.g. to express close friendship, as illustrated by the example in Figure 2).

Such a non-romantic use of these markers was absent in boys' chats. Consequently, we argue that flirting and gender accommodation (of expressive markers) are related yet distinct phenomena, and that even 'romantic' expressive markers are not solely used for flirting, but are truly part of a general female adolescent online style. Finally, the case study revealed that while the teenage girls in the corpus showed less accommodative behavior than the boys in terms of feature frequency (see above), they seemed to take into account male aversion for particular features (e.g. heart emoji) by avoiding them in non-flirty mixed-gender conversations. The boys, who accommodate more strongly in terms of feature frequency, much less do so in terms of adopting particular female features, except when flirting is involved.

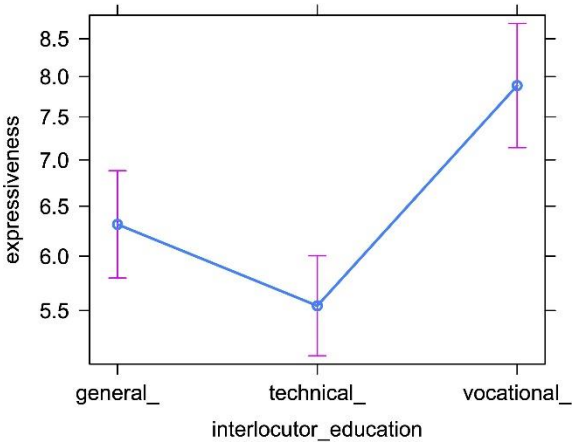

Figure 3: Expressive markers by the interlocutor's educational track (predicted counts per 100 tokens).

4.2. Education

The second case study (Hilte et al., under review) first of all shows that Flemish teenagers' instant messaging primarily proceeds within same-education networks, i.e. between pupils who attend the same educational track in secondary school (76% of the conversations in the dataset). But still, the statistical models reveal that when youths do chat across 'educational boundaries' (24% of the conversations in the dataset), they adapt their online writing style depending on their interlocutor's educational profile: the teenagers' usage frequency of oral and expressive markers significantly differs depending on whether they are interacting with a student from general, technical, or vocational education. However, this adaptation is only an accurate mirroring of the interlocutor's style for expressive features[2] (see Figure 3): i.e. the teenagers actually approach the average usage frequency for emoji, allcaps, ... by general, technical, and vocational pupils. For instance, when interacting with vocational students, who are generally the most ardent users of expressive markers (Hilte et al., 2020b), the teenagers tend to increase their own expressiveness to match that of their interlocutor. The frequency differences for expressive markers that are evoked by the interlocutor's educational profile (i.e. the teenagers' accommodative adjustments with respect to their conversation partner's education) are all significant (difference when chatting with a gen. vs tech. interlocutor p = 0.0082; gen. vs voc. interlocutor p = 0.0003; tech. vs voc. interlocutor p < .0001). The distinction between expressive and oral markers (for which respectively accurate versus inaccurate mirroring was observed) will be addressed in the discussion.

The findings of the second case study also indicate that pupils in different school systems do not adapt their online writing style to different extents (depending on their interlocutor's educational profile), so all teenage

[2]In the present contribution, we do not describe or visualize the pattern for oral markers, since it was not an accurate mirroring of the interlocutor's style (but see Hilte et al. (under review) for a discussion).

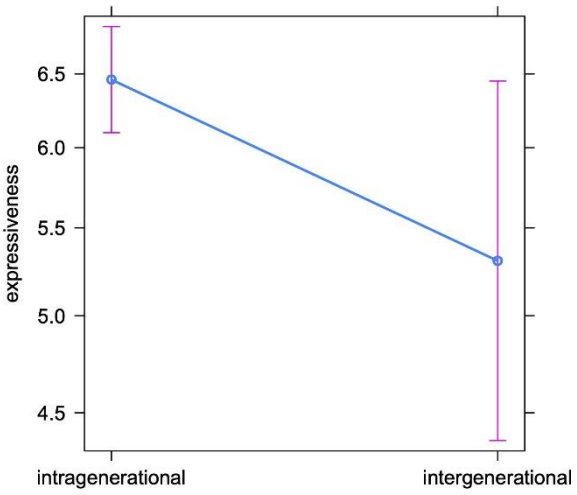

Figure 4: Expressive markers in intra- vs intergenerational conversations (predicted counts per 100 tokens).

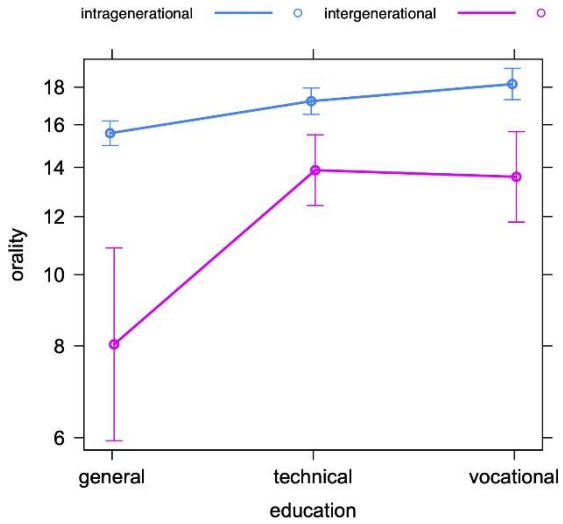

Figure 5: Oral markers by the author's educational track, in intra- vs intergenerational conversations (predicted counts per 100 tokens).

participants, regardless of whether they are general, technical, or vocational pupils, mirror their interlocutor's expressive writing to (more or less) the same extent. These symmetric accommodative efforts contrast with the asymmetric gender patterns that were observed in the first case study, with boys converging much more strongly to a 'female' style than vice versa. We will come back to this in the discussion.

Finally, the teenagers' adaptation to their interlocutors' educational profile is similar in one-on-one chats and group chats. That is quite surprising, since one might expect stronger mirroring in the former type of interactions for multiple reasons – see below.

4.3. Age

The third and final case study (Hilte et al., 2021) demonstrates how teenagers' instant messaging primarily proceeds within peer group networks (i.e. chats among teenagers only; 93% of the conversations in the dataset) and much less frequently across generations (i.e. chats including both teenagers and adults older than twenty; 7% of the conversations in the dataset). Still, the Poisson models show that when the teenagers do chat with older interlocutors, they adapt their own writing style by using significantly fewer expressive and oral markers (see Figures 4 and 5) ($p = 0.0473$ resp. $< .0001$). This decrease can be considered accurate mirroring, as (especially young) teenagers are more ardent users of these features than age groups above theirs (Hilte et al., 2020b; Prada et al., 2018; Verheijen, 2018). Consequently, youths appear to adapt their writing style to (that of) their older conversation partners by inserting fewer expressive markers (e.g. emoji) and fewer oral markers (e.g. colloquialisms) and thus adopting a more 'adult' writing style. This accommodation by the younger interlocutors in intergenerational communication is in line with previous findings on spoken interactions (see above). While the accommodation pattern for expressive compensation is not impacted by the teenagers' own socio-demographic profiles, the observed convergence with respect to oral markers differs depending on the teenagers' own educational track. A stronger adaptation – i.e. a stronger decrease in oral markers when chatting with older interlocutors – is made by pupils in the theory-oriented general secondary education (see Figure 5). So these pupils might have a stronger inclination to linguistic/stylistic mirroring (but further inspection is required). Such mirroring can be considered the product of meta-linguistic skills that are actually part of language teaching in more theory- oriented tracks only (VVKSO (2014)). In addition, a stronger command of written standard Dutch - which can be expected from theory-oriented students in view of their curriculum - might also increase these students' control of their standard versus speech-like rendition of Dutch words and phrases.

Finally, the teenagers' mirroring of older interlocutors' writing style is not any different in one-on-one chats compared to group chats. This echoes our findings for education but not for gender accommodation. The similarities and distinctions between the three case studies in this respect will be discussed below.

5. Discussion

This overview paper brought together three complementary case studies on linguistic accommodation in Flemish teenagers' instant messages, each focusing on a different aspect of the interlocutors' socio-demographic profiles. Convergence patterns emerged in all three studies: the teenagers adapt (aspects of) their online writing style depending on their conversation partner's gender, education, and age. However, differences could be noted concerning (a) the linguistic features subject to adaptation, (b) the accuracy of the mirroring, (c) its (a)symmetry, and (d) the impact of the number of interlocutors in a conversation. With respect to (a), in intergenerational interactions, teenagers converged to older interlocutors with respect to the use of both expressive and oral markers. In mixed-education talks, they adapted both their

expressive and speech-like writing too, though only the former adaptation actually mirrored their interlocutor's style (b). And in mixed-gender interactions, finally, significant convergence was attested for expressive markers only. A potential explanation for these differences concerns the higher 'visibility' of many of the (typographic/pictorial) expressive markers. The use of e.g. emoji may be more salient and therefore trigger more (accurate) convergence than the use of certain low-level colloquial markers. Furthermore, expressive features are generally used more deliberately, making them easier to manipulate, while teenagers (and especially pupils in practice-oriented tracks with a minor focus on formal Dutch writing) might be less in control over their speech-like rendering of Dutch words or the use of non-standard lexemes in general. It may thus simply be harder to (accurately) adapt speech-like writing.

With respect to the convergence pattern (c), symmetric accommodative efforts could be noted concerning interlocutors' educational tracks (i.e. pupils in different tracks did not adapt their online writing style to different extents), but not regarding interlocutors' gender: boys converged much more strongly to a 'female' writing style than vice versa. As for the age pattern, the frequency of oral markers in interactions with older interlocutors was suppressed more by pupils in theory-oriented tracks, which might again be linked to these pupils' stronger command of standard Dutch writing conventions.

We also compared one-on-one chats and group chats (d). Gender convergence was only significant in the former, i.e. instant messaging between one boy and one girl. The teenagers' adaptation towards interlocutors with different age or educational profiles, however, does not differ between these two conversational settings. That is quite fascinating, since stronger convergence might be expected in one-on-one interactions for several reasons. For instance, such interactions tend to be of a more intimate and personal nature than group chats, and trust is said to facilitate communicative convergence (Riordan et al., 2013). Second, linguistic mimicry is naturally more straightforward when there is only one other interlocutor to mirror. The distinction in this respect between the three case studies could indicate that gender accommodation mostly occurs in more intimate settings (while female vs male gender identities are expressed and emphasized more through gendered writing in group chats), in contrast to adaptation based on interlocutors' age or education accommodation.

Finally, recall that convergence narrows the linguistic and therefore also the social distance between interlocutors. The three case studies included in this paper suggest that the desire for social approval and closeness (a driving force behind accommodation) does not only hold among peers with similar socio-demographic profiles, but also across gender boundaries, educational boundaries, and across generations. In addition, note that mirroring in intereducational and intergenerational interactions is, in a sense, less obvious than in mixed-gender settings, as we observed how teenagers' instant messaging primarily proceeds in peer group networks within a same educational track. Still, both the educational background and the age of the interlocutor trigger accommodation.

6. Data availability statement

In order to protect the participants' privacy, and following the guidelines of our university's ethical committee, the collected dataset cannot be made publicly available. For more information on the database, see chapter 1 in Hilte (2019).

7. References

De Decker, B. and Vandekerckhove, R. (2017). Global features of online communication in local Flemish: Social and medium-related determinants. *Folia Linguistica*, 51(1):253–281.

Dragojevic, M., Gasiorek, J., and Giles, H. (2015). Communication accommodation theory. In C. R. Berger et al., editors, *The international encyclopedia of interpersonal communication (vol. 1)*, pages 1–21. Wiley Blackwell, London.

Giles, H. and Gasiorek, J. (2011). Intergenerational communication practices. In K. W. Schaie et al., editors, *The handbooks of aging consisting of three vols. Handbook of the psychology of aging*, pages 233–247. Elsevier Academic Press.

Hilte, L., Vandekerckhove, R., and Daelemans, W. (2020a). Linguistic accommodation in teenagers' social media writing: Convergence patterns in mixed-gender conversations. *Journal of Quantitative Linguistics*.

Hilte, L., Vandekerckhove, R., and Daelemans, W. (2020b). Modeling adolescents' online writing practices: The sociolectometry of non-standard writing on social media. *Zeitschrift fu¨r Dialektologie und Linguistik*, 87(2):173–201.

Hilte, L., Daelemans, W., and Vandekerckhove, R. (2021). Interlocutors' age impacts teenagers' online writing style: Accommodation in intra- and intergenerational online conversations. *Frontiers in Artificial Intelligence*, 4. Hilte, L., Daelemans, W., and Vandekerckhove, R. (under review). Communicating across educational boundaries: Accommodation patterns in adolescents' online interactions.

Hilte, L. (2019). *The social in social media writing: The impact of age, gender and social class indicators on adolescents' informal online writing practices*. Doctoral dissertation, University of Antwerp.

Palomares, N. A., Giles, H., Soliz, J., and Gallois, C. (2016). Intergroup accommodation, social categories, and identities. In H. Giles, editor, *Communication accommodation theory. Negotiating personal relationships and social identities across contexts*, 123–151. Cambridge University Press, Cambridge.

Prada, M., Rodrigues, D. L., Garrido, M. V., Lopes, D., Cavalheiro, B., and Gaspar, R. (2018). Motives, frequency and attitudes toward emoji and emoticon use. *Telematics and Informatics*, 35: 1925–1934.

Riordan, M. A., Markman, K. M., and Stewart, C. O. (2013). Communication accommodation in instant messaging: An examination of temporal convergence. *Journal of Language and Social Psychology*, 32(1): 84–95.

Thurlow, C. and Poff, M. (2013). Text messaging. In S.

Herring, et al., editors, *Pragmatics of computer-mediated communication*, 163–190. Mouton de Gruyter, Berlin and New York.

Verheijen, L. (2018). Orthographic principles in computer-mediated communication: The SUPER-functions of textisms and their interaction with age and medium. *Written Language & Literacy*, 21(1):111–145.

Vlaams Verbond van het Katholiek Secundair Onderwijs (VVKSO). (2014). *Nederlands derde graad aso-kso-tso. Leerplan secundair onderwijs.* VVKSO.

Williams, A. and Nussbaum, J. F. (2001). *Intergenerational communication across the life span*. Lawrence Erlbaum Associates, Mahwah and London.