

Lisa Hilte\*, Walter Daelemans and Reinhild Vandekerckhove

# Communicating across educational boundaries: accommodation patterns in adolescents' online interactions

<https://doi.org/10.1515/applirev-2020-0140>

Received November 18, 2020; accepted November 22, 2021; published online December 13, 2021

**Abstract:** This paper studies linguistic accommodation patterns in a large corpus of private online conversations produced by Flemish secondary school students. We use Poisson models to examine whether the teenagers adjust their writing style depending on their interlocutor's educational profile, while also taking into account the extent to which these adaptation patterns are influenced by the authors' own educational background or by other aspects of their socio-demographic profiles. The corpus does reveal accommodation patterns, but the adjustments do not always mirror variation patterns related to educational profiles. While salient features like expressive markers seem to lead to pattern-matching, less salient features appear less prone to 'adequate' adjustment. Lack of familiarity with the online behavior of students from other educational tracks is a factor too, since online communication clearly proceeds primarily within 'same-education' networks. The focus on cross-educational communication is quite unique in this respect and highly relevant from a sociological perspective.

**Keywords:** accommodation; adolescents; education; online communication; social media

## 1 Introduction

In the past two decades, a substantial amount of sociolinguistic studies has been devoted to social patterns in online writing (see e.g. Herring and Kapidzic 2015; Tagliamonte and Denis 2008; Varnhagen et al. 2010; Verheijen 2018). Our own previous work focused on the social profiles of adolescent writers and revealed a strong impact of the authors' age, gender, and educational track on their informal

---

**\*Corresponding author: Lisa Hilte**, Department of Linguistics, Universiteit Antwerpen, Prinsstraat 13, 2000 Antwerpen, Belgium, E-mail: [lisa.hilte@uantwerpen.be](mailto:lisa.hilte@uantwerpen.be). <https://orcid.org/0000-0003-0891-319X>

**Walter Daelemans and Reinhild Vandekerckhove**, Department of Linguistics, Universiteit Antwerpen, Prinsstraat 13, 2000 Antwerpen, Belgium

social media writing (Hilte et al. 2018a, 2018b, 2018c, 2020b – see also below). Moreover, we found that apart from the teenage authors' *own* gender, that their conversation partners' gender appeared to be a major determinant too (Hilte et al. 2020c). Obviously, this evokes questions with respect to the impact of their conversation partners' age and educational track as well. The latter is the focus of the present contribution: we investigate whether youths adapt their writing styles depending on their (online) interlocutors' educational background. While linguistic *accommodation* has been widely researched for spoken interactions (see Section 1.1.1), it is still under-researched for written online interactions. Furthermore, the link between conversation partners' educational profiles and linguistic convergence or divergence constitutes a gap in accommodation research that the present study aims to address too.

The paper is structured as follows: an overview of related research can be found in Section 1.1. It leads to the research questions in Section 1.1.3. Section 2 presents the corpus for the present study, while Section 3 describes the linguistic variables and the methodology for the quantitative analyses. Finally, the findings are discussed in Section 4.

## 1.1 Related research

This section presents a selection of relevant studies on linguistic accommodation (Section 1.1.1) and an overview of previous findings that offer a frame for addressing the social variable of educational background (Section 1.1.2). Finally, it describes the research questions that the present contribution aims to answer (Section 1.1.3).

### 1.1.1 Accommodation

*Linguistic accommodation*<sup>1</sup> concerns the adaptation of one's communicative behavior to that of one's conversation partner. Our main point of reference is the sociolinguistic framework 'Communication Accommodation Theory' (CAT), which considers accommodation to be aimed at facilitating interaction as well as regulating social distance among interlocutors (Dragojevic et al. 2015: 10). Common strategies are convergence and divergence, i.e. the adaptation of one's communicative behavior to appear more resp. less similar to others (Giles and Ogay 2007: 295–296). Divergence tends to be evaluated more negatively and

---

<sup>1</sup> This phenomenon is also referred to as e.g. *alignment*, *matching*, or *synchrony*, depending on the scientific field. See Burgoon et al. (2017) for an overview.

convergence more positively, but full convergence is seldom desired, as there appear to exist individually and socio-culturally determined optimal levels of similarity (Burgoon et al. 2017; Dragojevic et al. 2015: pp. 13, 15). Over-accommodation can even be perceived as parody (Jones et al. 2014: 457).

While the inclination to adapt one's communicative style to that of one's addressees may be individually determined (Jones et al. 2014; Xu and Reitter 2015), several accommodative patterns relating to interlocutors' socio-demographic or psychological profiles appear to be quite robust. To our knowledge, educational background, i.e. the variable of interest in the present study, has not yet been studied in this respect, but several studies discuss the impact of the related variable of social power or social position (see Section 1.1.2). So-called 'upward' social convergence has been attested repeatedly, i.e. communicative adaptation towards interlocutors with greater (social) power, and it appears to hold for different kinds of 'traditional'<sup>2</sup> social power (e.g. Dragojevic et al. 2015; Kroll et al. 2018; Muir et al. 2016). This asymmetrical pattern is in line with CAT's predictions with respect to unidirectional shifts towards interlocutors with greater power (Dragojevic et al. 2015: 4), as people with lower social power will have a stronger desire "to gain the approval of the higher-power partner" than vice versa (Muir et al. 2016: 477). However, the role of social power should, at least to some extent, be nuanced. Xu et al. (2018) for instance report that certain low-level linguistic features such as utterance length may have a stronger impact on linguistic alignment than social power relations between interlocutors. Furthermore, Adams et al. (2018) only observe significant upward convergence in an experimental setting and not in naturalistic chat conversations. In addition, de Siqueira and Herring (2009) report on potential 'conflicts of interest' with respect to different aspects of one's socio-demographic profile, while Muir et al. (2016) signal an interaction between social power and certain psychological traits, with specific personality types in lower-power roles converging more strongly than others.

We note that while social power and educational background may be related to each other (see below), we cannot simply equate them. Furthermore, the research on the impact of social power on accommodative behavior generally concerns adult participants, whereas the target group of the present study consists of adolescents. The desire to obtain social approval, while certainly not absent among adults, seems even stronger among teenagers who are often driven by a "need of acceptance" and "fear of rejection" (Taylor 2001: 298). While this makes adolescent age highly interesting for research on accommodative behavior,

---

<sup>2</sup> 'Modern' or 'new media' interpretations of social power, such as people's position and influence in online social networks (e.g. Danescu-Niculescu-Mizil et al. 2011) fall outside the scope of the present paper, since they do not relate to the variable of educational track.

surprisingly, an important aspect of the social profiles of adolescents, i.e. their educational tracks, has hardly been included in research on linguistic accommodation. This may be related to the fact that in many countries, differentiation into educational tracks does not happen as distinctly and as early on in pupils' school trajectory as in Belgium (see Sections 1.1.2 and 2 below).

Another gap that the present study aims to address, concerns the nature of the data. Linguistic accommodation is most often analyzed in spoken face-to-face dialogue. While studies have examined to which extent these findings translate to the context of online communication (e.g. Doyle et al. 2016; Riordan et al. 2013; Scissors et al. 2008, 2009), as of yet, there do not exist any large-scale studies on accommodation in CMC corpora that truly mirror spontaneous face-to-face interactions, as existing studies are either carried out on small corpora (e.g. Adams et al. 2018; de Siqueira and Herring 2009; Kroll et al. 2018; Wolf 2000), on public and/or asynchronous conversations (e.g. Bunz and Campbell 2004; Danescu-Niculescu-Mizil et al. 2011; Dino et al. 2009; Doyle 2016; Pérez Sabater 2017), or on synchronous chat conversations between strangers and/or in lab-based settings (e.g. Gonzales et al. 2010; Kroll et al. 2018; Muir et al. 2017; Niederhoffer and Pennebaker 2002; Scissors et al. 2008, 2009).

### 1.1.2 Educational track

Although several aspects of youths' socio-demographic profiles will be included in the present study, our main focus concerns their educational track. All participants are teenagers who attend one of the three main types of secondary education in Belgium, ranging from highly theory- to highly practice-oriented (see Section 2). While youths' educational track is seldom included in CMC research, it is both an essential aspect of their social profile and an important co-determiner of their online writing style. As mentioned above, a potential explanation why (youths') educational track is seldom included in linguistic studies, is that in many countries, differentiation into educational tracks does not happen as distinctly and as early on in school as in Belgium. Consequently, the distinction between educational tracks (and between practice- and theory-oriented tracks in particular) may not be as present or outspoken in these countries. In this section, we provide more information on the Belgian secondary school system and more particularly on the link between teenagers' educational track, their social class and their (online) language use. For an elaborate description of the educational tracks included in the present paper, we refer to Hilte et al. (2020a).

Teenagers' educational track strongly influences their peer group networks since students in different tracks of Belgian secondary education spend their school days in separate class groups, and often even in different schools or school

buildings. In addition, educational track is also related to social class in multiple ways. First of all, teenagers' educational profile is indicative of their future professional career (de Jager et al. 2009: 253). As today's society has evolved towards a knowledge-based *meritocracy* – i.e. “social stratification based on personal merit” (Macionis 2011: 206) – education and obtained degrees tend to be strong determinants of social status and position (de Jager et al. 2009: 243, 247). Furthermore, youths' educational track also relates to their current social background. Social class origin and family background impact youths' levels of attainment in education, as social differences (and the limitations and opportunities typically faced in different social classes) tend to affect performance at school and decisions within the educational career (Goldthorpe and Breen 2007: 45–47; Vranken et al. 2017: 319–325). Our own previous work revealed that half of our secondary school participants attended an educational track that would most likely lead to the same working-class, middle-class or upper-class profession type as their parents' (Hilte et al. 2018a, 2018b), which points to *social stagnation* or *social immobility*. The other half demonstrated either ‘upward’ or ‘downward’ social mobility (see also de Jager et al. 2009: 254; Vranken et al. 2017: 314–315, 319). However, huge discrepancies between the child's educational profile and the parents' profession type were rare. Finally, we also found a significant correlation between the participating teenagers' educational track and their home language: with Dutch being the official language of education in Flanders, it seemed harder for children from non-Dutch speaking families to get access to more theoretical education systems (see Hilte et al. 2018a, 2018b).

These findings indicate that teenagers' educational track fits into a larger social identity. Furthermore, they show that youths with distinct educational profiles often truly move in different social circles even far beyond the school context. This sets apart education accommodation from gender accommodation since interlocutors with different gender identities may still have highly similar profiles in other respects. We can thus wonder what will happen when teenagers from different educational tracks do interact with each other, and whether the observed patterns will bear resemblance to those attested for gender accommodation (Hilte et al. 2020c).

As for the linguistic effects of teenagers' educational track, we observed significantly distinct online writing styles for adolescents in the three educational tracks included in this paper. The distinction between students on the two ‘ends’ of the educational continuum from theory to practice appeared to be highly consistent: students in the most practice-oriented track tend to deviate most strongly from formal standard writing by incorporating more ‘speech-like’ markers (e.g. regional language features) in their online texts, as well as more typographic markers that typically enhance the expression of social or emotional involvement

(e.g. emoji) (Hilte et al. 2018a, 2020b). Strikingly, students holding an intermediate position on this educational continuum have a much more unpredictable and variable writing style (Hilte et al. 2018a, 2018d, 2020b). Education-related variation that concerns more general text features (e.g. average utterance length) in teenagers' online writing, as well as patterns regarding teenagers' formal school writing, falls outside the scope of the present research, but is discussed in Hilte et al. (2020a), and in Verheijen and Spooren (2017) and Vandekerckhove and Sandra (2016), respectively.

Finally, we examined to what extent youths are aware of attested sociolinguistic patterns in their (peers') social media writing. While the overall intuition on gendered and age-related writing styles appeared to be quite strong, there was a clear lack of awareness with respect to the observed education-related linguistic variation (Hilte et al. 2019). So while gender accommodation in adolescent online writing might be based on actual awareness of gender patterns in online writing, this may not be true for accommodation triggered by the interlocutor's educational track: if significant linguistic adaptation can be observed, this process might be of a more subconscious nature.

### 1.1.3 Research questions

The present paper aims to broaden the scope of accommodation research and examine whether and how youths adapt their online writing style depending on their interlocutor's educational track. More specifically, we will examine whether all youths (non-)accommodate in a similar way irrespective of their own profile, or whether their own educational track and other aspects of their socio-demographic profiles (i.e. their age or gender) play a role too and thus interact with the accommodation pattern.

## 2 Corpus and participants

The corpus consists of a collection of 305,227 private social media messages (>1.7 million tokens) produced by 1,104 Flemish teenagers in Flemish Dutch on Facebook Messenger and WhatsApp between 2015 and 2016. The teenagers were secondary school students between 13 and 20 years old at the time and nearly all lived in the central Flemish (i.e. northern Belgian) province of Antwerp. We personally visited secondary schools and invited the pupils to voluntarily donate (parts of) their chat conversations (produced out of the school context and before our visit). The participants also provided the following metadata: age, gender and educational track. The pupils' (and for minors, also their parents') consent

was asked to store and linguistically analyze their texts after anonymization. We note that this corpus is a subset of a larger collection of social media messages (>2.5 million tokens) (see chapter 1 in Hilde 2019). This subset was selected based on its relevance for the present study.<sup>3</sup>

Our main focus concerns the social variable of educational track. All participants attend one of the three main types of Belgian secondary education, ranging from the theory-oriented general secondary education, where students are prepared for higher education, to the practice-oriented vocational secondary education, where students are prepared for specific, often manual, professions. The technical secondary education holds an intermediate position in terms of theory and practice (FMET 2018: 10). While educational track may strongly impact offline peer networks, the distributions in the dataset offer insight in the corresponding impact on youths’ *online* networks and interactions. With respect to interlocutors’ educational profiles, the conversations in the dataset can be classified into six categories, three of which are ‘same-education’ (i.e. all conversation partners attend the same track in secondary school) and three ‘mixed-education’ (i.e. at least one of the conversation partners attends a different

Table 1: Distributions in the corpus: interlocutors’ educational track.

Same/ mixed education	Interlocutors’ educational tracks	Conversations	Participants	Posts	Tokens
Same	General students only	315 (26%)	367	54,842 (18%)	365,469 (21%)
Same	Technical students only	387 (32%)	223	103,097 (34%)	558,564 (32%)
Same	Vocational students only	226 (18%)	247	49,746 (16%)	277,742 (16%)
Mixed	General and vocational students	51 (4%)	80	7,259 (2%)	44,087 (3%)
Mixed	General and technical students	138 (11%)	205	46,711 (15%)	280,828 (16%)
Mixed	Technical and voca- tional students	110 (9%)	152	43,572 (14%)	219,100 (13%)
Total		1,227	1,104 <sup>a</sup>	305,227	1,745,790

<sup>a</sup>Note that the number of participants in the six subgroups adds up to a higher number than the total number of participants in the dataset, as the same author may occur in multiple conversational settings. For instance, the same vocational student may participate in conversations among vocational students only, as well as in interactions with technical or general pupils.

3 Conversations that fell outside the scope include interlocutors who are not a teenager attending one of the three main Belgian secondary educational tracks: e.g. a parent, a teenager in artistic secondary education, or a youth in tertiary education. Finally, conversations including interlocutors with an unclear educational profile were deleted too.

track). Table 1 presents an overview of these distributions in terms of number of conversations, participants, tokens, and posts.

Although the data processing finally necessitated a further cutdown on this corpus (see Table 2 in Section 3.2.1), we decided to present the ‘initial’ distributions in Table 1, since they offer a realistic view on the actual ratio of same- versus mixed-education conversations that we collected. It can be derived from Table 1 that three quarters (76%) of the conversations in the dataset are same-education, and only a quarter (24%) are mixed-education and thus include interlocutors with different educational profiles.<sup>4</sup> Furthermore, conversations in which the discrepancy between interlocutors’ educational tracks (regarding their position on the continuum from theory to practice) is largest – i.e. chats including vocational and general students – are rarest (4% of the dataset). Presumably the less frequent real-life contact across educational ‘boundaries’ persists in online settings. This might explain why teenagers demonstrate low awareness with respect to the online writing style of peers with different educational profiles (Hilte et al. 2019): they simply seem to interact much less frequently outside of their own educational network.

Since the relationship between conversation partners might have an impact too, it is worth mentioning that the vast majority of interactions includes interlocutors who are (close) friends. Others are lovers or relatives. As for the conversational setting with respect to interlocutors’ gender (i.e. same- versus mixed-gender talks) – which impacts youths’ online writing style too (see Hilte et al. 2020c) – and potential age gaps between interlocutors (younger versus older adolescents), comparable patterns were observed for the six educational groups. Finally, imbalances regarding age, gender and the number of interlocutors in a conversation, will be taken into account systematically as these confounding factors are included in the research design (see Section 3.2). For the distribution of these categories in the final dataset, see Table 3 below.

### 3 Methodology

Below, we present the linguistic variables of this study (Section 3.1) and describe the methodology for the statistic modeling (Section 3.2).

---

<sup>4</sup> Interactions including students from *all three* educational tracks were excluded for two reasons: first, they are highly infrequent, and second, the distinction between one-on-one chats (two interlocutors) and group chats (more than two interlocutors) that will systematically be included in the research design as confounding factor (see below) cannot be made for these specific interactions, since they always include at least three interlocutors.

### 3.1 Linguistic variables

Many prototypical markers of social media writing can be linked to one of three ‘maxims’ or implicit rules of linguistic conduct of informal online writing: the principles of expressive compensation, orality and economy (see e.g. Androutsopoulos 2011: 149). For the present study, we excluded language features relating to the maxim of economy (e.g. acronyms), which consists in strategies aimed at maximizing typing speed and minimizing typing effort, as these are of a highly functional nature and appear not to be subject to social variation to the same extent as other prototypical markers of the genre (De Decker and Vandekerckhove 2017; Hilte et al. 2018a, 2018b). Below, we describe the two maxims that are included in the research design along with their related linguistic features, illustrated with examples from the dataset. The selection of these particular features was based on related research (e.g. Varnhagen et al. 2010; Verheijen 2018) and on our previous work, in order to facilitate systematic comparison between our current and past findings (e.g. Hilte et al. 2020b, 2020c).

The principle of expressive compensation relates to the application of a wide range of (predominantly typographic) strategies to compensate for the absence of certain expressive cues in written communication, such as intonation, volume or facial expressions. We include the following features:

- emoticons/emoji:

e.g. 😂😂😂😂😂, 💕💕💎

- words or phrases rendered in capital letters (‘allcaps’):

e.g. *HOE MOTTIG* (‘HOW UGLY’)

- deliberate repetition of letters or punctuation marks (‘flooding’):

e.g. *Ik oooo!!!!* (‘me toooo!!!!’)

- combinations of question and exclamation marks:

e.g. *Vertel???* (‘Tell me???’)

- typographic rendering of kisses and/or hugs using the letters ‘x’ and ‘xo’:

e.g. *Slaapzacht xxxxx* (‘Sleep tight xxxxx’)

e.g. *veel plezier vanavond xoxo* (‘have fun tonight xoxo’)

- onomatopoeic rendering of laughter:

e.g. *hahahhaa, whahahhaha*

The orality maxim concerns “speech-like” writing: in many forms of informal written online interaction, the register is to a large extent ‘conceptually oral’, reflecting oral communication and typical speech patterns rather than classical written communication. In our dataset of Flemish teenagers’ Dutch instant messages, this maxim results in the insertion of different kinds of non-standard Dutch lexemes and non-standard grammar which render the written utterance more speech-like:

- dialect/regiolect words:

e.g. *tot seffes* (std. Dutch: *tot straks*, ‘see you later’)

- informal/colloquial words or ‘slang’:

e.g. *negeer die gast* (std. Dutch: *negeer die jongen*, ‘ignore that dude’)

- orthographic renderings of non-standard pronunciation or morphology:

e.g. *Ik wist ni da gij kwam* (std. Dutch: *Ik wist niet dat jij kwam*, ‘I didn’t know you were coming’)

Furthermore, Flemish teenagers often use English words or phrases that are part of Dutch adolescent speech:

- English words rendered in their ‘original’ form:

e.g. *Da zou echt lame zijn* (‘That would be really lame’)

- English words adapted to Dutch (in terms of e.g. spelling or morphology):

e.g. *ik zou wel pist zijn* (‘I would be pissed’)

e.g. *Er is geen excuus om te cheaten* (‘There is no excuse for cheating’)

We note that the base language in the dataset is always Dutch, as entire chat conversations in another language were excluded. Furthermore, English loan words that have been integrated in Dutch for a considerable time to the extent that they are generally considered part of standard Dutch vocabulary and included in Dutch dictionaries (e.g. *computer*), are counted as Dutch and not as English in the analyses.

The feature occurrences were detected and counted automatically in the dataset with Python scripts. The scripts’ output was compared to a human annotator’s decisions for a test set of 200 randomly selected posts (1,257 tokens). The software reached an average precision of 91% (i.e. the share of detected feature occurrences that are valid) and an average recall of 88% (i.e. the share of all feature

occurrences in the test set that were detected as such by the software) – and these scores were also sufficiently high for the individual features. So, the software's accuracy is sufficient to find the main trends in the data. Consequently, the scripts' output is reliable and suitable for further linguistic analysis. For an extensive discussion and error analysis of the feature extraction procedure, see Hilte et al. (2020b).

As mentioned above (Section 1.1.2), we previously observed significantly distinct online writing styles for students in more theory- versus more practice-oriented educational tracks for both expressiveness and orality, with vocational students inserting both more oral and more expressive features in their social media discourse than students in the theoretical track, and students in the 'hybrid' technical track demonstrating a much more unpredictable and variable writing style (Hilte et al. 2018a; 2018d, 2020b). Therefore, it is worth investigating whether potential accommodation patterns for these features are different for these three educational groups too.

Finally, we note that our study on gender accommodation (Hilte et al. 2020c) showed significant convergence for (prototypically 'female') expressive markers but not for orality markers. In the present study, we want to investigate which feature set is more susceptible to accommodative change with respect to educational track rather than gender. Are expressive features generally more susceptible to accommodative change or are they only strategically manipulated for gender accommodation?

## 3.2 Methods

We will statistically model the participants' language use with respect to educational accommodation. Our approach to accommodation is of a quantitative nature: we analyze whether the usage frequency for particular features significantly increases or decreases depending on the interlocutor's educational background. Furthermore, we focus on accommodation from a synchronic perspective, comparing youths' writing in different *conversational settings* (depending on interlocutors' profiles) rather than analyzing *the course of* particular interactions. Diachronic analyses, including a temporal dimension, are left for future work.

The present study's methodology is similar to our previous work on gender accommodation (Hilte et al. 2020c) and on sociolinguistic variation in teenagers'

social media writing (Hilte et al. 2020b), which facilitates the systematic comparison of our previous and current findings. Below, we describe the data preprocessing steps (Section 3.2.1) and the model fitting (Section 3.2.2).

### 3.2.1 Data preprocessing

We created a summary of the dataset with each line or ‘observation’ representing one participant in one conversation. Participants can thus occur on multiple lines, i.e. in different conversations. Similarly, conversations can be represented on multiple lines too, as the different conversation partners may each occupy a line. Therefore, we correct for these repeated observations by adding a random effect for participant and conversation (see below). Each line in the dataset contains the participant’s profile information (a unique, anonymous identifier as well as their gender, age and educational track), conversational meta-information (a unique identifier for the interaction, and information on the number of interlocutors: one-on-one/group chat), and finally, the relevant linguistic counts (i.e. the total number of tokens and the number of oral and expressive markers for this participant in this particular conversation).

As a final preprocessing step, we deleted observations which concerned fewer than 10 posts (i.e. a participant who produced fewer than 10 utterances in a particular conversation), since such small samples might not be representative of the authors’ online writing. Table 2 shows the reduced dataset after applying this frequency cutoff. The reduction has a negligible impact on the number of tokens and posts: we lose 1.4% of the tokens and 1.3% of the posts. However, it does affect the number of participants and conversations with a reduction of respectively 21 and 26% (compare Table 2 with Table 1).

Table 3 shows the gender and age distribution in the final dataset in terms of participants and tokens and it also includes the representation of group chats versus one-on-one conversations. Gender is operationalized as the distinction between boys and girls, since a non-binary approach (e.g. operationalizing gender as a continuum) were infeasible with the profile information we had access to. For age, we distinguish between younger teenagers (13–16 years old) and older teenagers or young adults (17–20 years old). Finally, the number of interlocutors was operationalized as a binary variable too: we distinguish one-on-one chats and group chats (including at least two interlocutors).

**Table 2:** Distributions in the final dataset w.r.t. interlocutors’ educational track.

Same/ mixed education	Interlocutors’ educational tracks	Conversations	Participants	Posts	Tokens
Same	General students only	259 (29%)	299	54,047 (18%)	360,414 (21%)
Same	Technical students only	289 (32%)	202	101,917 (34%)	550,313 (32%)
Same	Vocational students only	148 (16%)	172	48,889 (16%)	272,890 (16%)
Mixed	General and vocational students	33 (4%)	54	7,078 (2%)	43,064 (3%)
Mixed	General and technical students	98 (11%)	155	46,132 (15%)	277,202 (16%)
Mixed	Technical and vocational students	80 (9%)	116	43,302 (14%)	217,511 (13%)
Total		907	872 <sup>a</sup>	301,365	1,721,394

<sup>a</sup>Note that the number of participants in the six subgroups again adds up to a higher number than the total number of participants in the dataset, as the same author may occur in multiple conversational settings (see also Table 1).

**Table 3:** Distributions in the final dataset w.r.t. confounding factors.<sup>a</sup>

Variable	Variable levels	Participants	Tokens
Gender	Girls	465 (53%)	1,190,986 (69%)
	Boys	407 (47%)	530,408 (31%)
Age	Younger teenagers (13–16)	543	903,874 (53%)
	Older teenagers (17–20)	442	817,520 (47%)
Number of interlocutors	One-on-one (2 interlocutors)	702	1,249,547 (73%)
	Group chat (>2 interlocutors)	312	471,847 (27%)
Total		872	1,721,394

<sup>a</sup>Percentages are only provided in the participant columns when they make sense: since the participant counts in the ‘number of interlocutors’ column indicate how many participants are included in one-on-one versus group chats, and one participant may occur in both settings, the sum in this column exceeds the total number of participants. Finally, participants can occur in the dataset at different ages (e.g. in early versus later adolescence), so the number of younger and older teenagers exceeds the total participant count too.

3.2.2 Model fitting

We modeled the teenagers’ use of expressive and oral markers with generalized linear mixed models (GLMMs) with a Poisson distribution.<sup>5</sup> Poisson models are

5 As implemented in the ‘lme4’ package for R (Bates et al. 2017).

typically recommended for the analysis of count data (Harrison 2014: 2; Ismail and Jemain 2007: 105), as the Poisson distribution is considered the “simplest distribution for modeling count data” (Zeileis et al. 2008: 5). GLMMs can simultaneously analyze the effect of different predictors or fixed effects, as well as of their potential interaction with each other. We will inspect the impact of authors’ and their conversation partners’ educational track on the response variable, i.e. the counts for expressive and oral markers. In addition, we will analyze the effect of three confounding factors: the authors’ gender and age, and the number of interlocutors in a chat conversation.

As mentioned above, the models take into account the impact of individual chatters and conversations and thus correct for repeated observations, as a random effect for subject and for conversation were included. This way, the models can link observations from one participant in different conversations to each other, thus dealing with individual writing styles, as certain people may *always* write in a more expressive/oral way than others. Similarly, the models can cluster observations from different interlocutors in the same conversation, thus taking into account conversation-specific styles and conventions, as certain people may *always* use many expressive/oral markers when interacting with each other. Furthermore, a random effect for conversation can also take into account the effect of *stylistic cohesion*, meaning that “[messages] belonging to the same conversation are closer stylistically than [messages] that do not” (Danescu-Niculescu-Mizil et al. 2011: 748). In order to deal with troubles of overdispersion (i.e. the variance of the response variable exceeding the mean – see Hilte et al. 2020b), which, if left unaddressed may lead to unreliable outcomes (Harrison 2014: 1, 2, 17–18 and references therein; Ismail and Jemain 2007: 103), we added a third and final random effect for each observation in the dataset (see Harrison 2014: 1). Finally, the models can handle differences in sample size between observations by adding an ‘offset’ for the logarithm of the number of tokens per chatter within a conversation.

In the results section below, we always discuss the best model, i.e. the model with the combination of predictors that resulted in the best fit for the data. After determining the random structure on the full (i.e. most complex) model, we experimentally determined the final best fit through a backwards stepwise procedure, step-by-step deleting irrelevant factors.

## 4 Results and discussion

Previous analyses revealed that Flemish teenagers’ age, gender and educational track significantly determine their online writing style (Hilte et al. 2020b), and that teenagers tend to adapt certain aspects of their writing depending on their

interlocutor’s gender (Hilte et al. 2020c). The present study aims to complement these findings with potential patterns of education accommodation, operationalized as the linguistic effect of the interlocutor’s educational track (and potentially its interaction with the author’s own educational track). In addition, the linguistic impact of confounding factors will be examined too. Below, we discuss the best models for expressiveness (Section 4.1) and for orality (Section 4.2).

### 4.1 Expressiveness

The best model for expressiveness includes all five predictors and one interaction (see Tables 4 and 5 for the fixed effects and the Anova, respectively). The effect of authors’ own educational track on their use of expressive markers is visualized in Figure 1: vocational students use significantly more expressive features than their

Table 4: Expressiveness: Fixed effects.<sup>a</sup>

	Estimate	Std. Error	z-value	Pr(> z )	Signif.
(Intercept)	−2.86315	0.06797	−42.124	<2e-16	***
ageOlder	−0.08305	0.06186	−1.342	0.179471	
genderFemale	0.44017	0.06371	6.909	4.87e-12	***
educationTechnical	−0.06471	0.06138	−1.054	0.291745	
educationVocational	0.14894	0.07255	2.053	0.040089	*
interlocutor_educationTechnical	−0.13039	0.04933	−2.643	0.008210	**
interlocutor_educationVocational	0.22236	0.06152	3.615	0.000301	***
nr_interlocutorsGroup-chat	−0.19844	0.07074	−2.805	0.005027	**
ageOlder:genderFemale	−0.23845	0.07619	−3.130	0.001749	**

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1. <sup>a</sup>Reference category: younger, male authors in general education, interacting one-on-one with another general student.

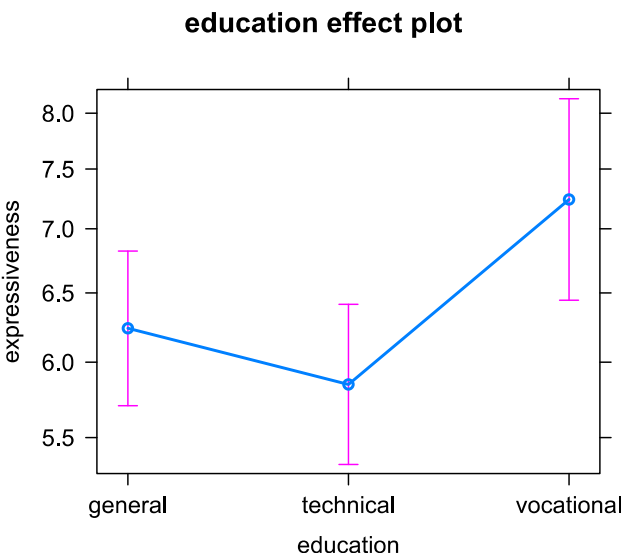
Table 5: Expressiveness: Anova.

	Chisq	Df	Pr(>Chisq)	Signif.
Age	26.7252	1	2.345e-07	***
Gender	38.8825	1	4.501e-10	***
Education	9.2937	2	0.009592	**
Interlocutor education	41.2954	2	1.078e-09	***
Number of interlocutors	7.8697	1	0.005027	**
Age:gender	9.7962	1	0.001749	**

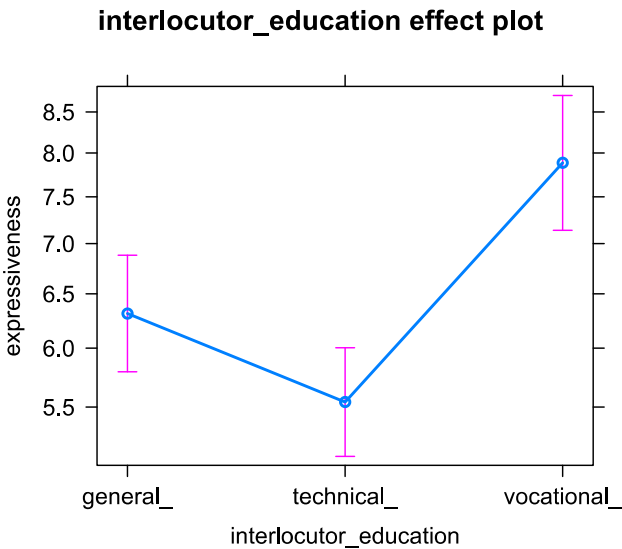
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

peers in technical education, and marginally more (losing significance after Bonferroni correction for multiple testing) than general students. The linguistic difference between general and technical pupils is insignificant. These results are in line with our previous findings for the entire corpus of which the present dataset is a subset (Hilte et al. 2020b). Overall, students in the practice-oriented track displayed a greater tendency towards the use of typographic expressive markers but a weaker tendency towards the verbal expression of sentiment (Hilte et al. 2020a).

The effect of the interlocutor’s education, as visualized in Figure 2, shows a striking resemblance to the pattern for the author’s own education (Figure 1), so teenagers appear to adequately mirror the degree of expressiveness associated with their interlocutor’s educational profile. We recall that such mirroring or convergence, which narrows the linguistic distance and therefore, according to CAT, also the *social* distance between conversation partners, is generally evaluated positively (Burgoon et al. 2017; Dragojevic et al. 2015: 13, 15). The participants’ insertion of expressive markers significantly varies depending on whether their conversation partner attends the general, technical or vocational track. This result is quite remarkable, since Flemish youths’ awareness of educational sociolinguistic variation in their peers’ online writing was found to be very low (Hilte et al. 2019). Consequently, the accommodative adaptation of expressive features by teenagers might be a (predominantly) subconscious process of pattern



**Figure 1:** Effect plot: Expressiveness by the author’s education (counts per 100 tokens).



**Figure 2:** Effect plot: Expressiveness by the interlocutor’s education (counts per 100 tokens).

matching – note that CAT actually points out that convergence often is an unconscious process (Adams et al. 2018: 477; Kroll et al. 2018: 4). However, a subtle difference between ‘actual’ and ‘mirrored’ educational variation can still be observed in the data: the linguistic differences evoked by the interlocutor’s education are more outspoken (i.e. significant for all pairs) than the ones related to the author’s own education (see above: e.g. no significant difference was observed between general and technical students). So the adaptation pattern appears to be somewhat more ‘polarized’ compared to the actual sociolinguistic variation.

Since no significant interaction emerged between the author’s and the interlocutor’s education, we can conclude that the teenagers’ linguistic adaptation towards their interlocutor’s educational profile is not influenced by their own educational track. So pupils in different educational tracks do not adapt their expressive writing to different extents. The adjustive effort made by the different student groups appears to be symmetrical, as opposed to for instance the pattern for gender accommodation, with boys converging much more strongly to a more ‘female’ expressive online standard than vice versa (Hilte et al. 2020c). In studies on accommodation determined by social position, unidirectional patterns of ‘upward’ social convergence have been reported repeatedly too (e.g. Dragojevic et al. 2015; Muir et al. 2016). While educational track and social position/class are interconnected (see Section 1.1.2) and while most of the students in the practice-oriented track have a working-class background whereas those in the general track

do not (Hilte et al. 2018a; 2018b), the model for expressive writing shows no proof of a stronger adaptation by vocational towards general students than vice versa. So, as social class background does not seem to trigger unidirectional accommodation among our participants, peer group solidarity – resulting in convergence from both sides – might predominate over potential social hierarchy for these adolescents.

Finally, some patterns emerge from the model with respect to the confounding factors. Significantly more expressive markers occur in one-on-one conversations than in group chats, thus indicating that these two types of interactions have different conversational dynamics. Note, however, that the number of interlocutors does not interact with the interlocutors' education and thus does not (significantly) influence education-based accommodation (see Section 5 for a discussion). This clearly differs from gender accommodation patterns, that were found to be much more outspoken in one-on-one settings (Hilte et al. 2020c). A potential explanation is that gender-based linguistic adaptation might be of a more personal, intimate nature than education accommodation (for the relation/distinction between gender accommodation and flirting strategies, see Hilte et al. 2020c). A final pattern that is not related to education accommodation concerns the interaction between the authors' age and gender: while all teenagers use fewer expressive markers in their online discourse at an older age, this decrease is much stronger (and only significant) for girls (for a detailed discussion and interpretation, see Hilte et al. 2020b).

**Table 6:** Orality: Fixed effects.<sup>a</sup>

	Estimate	Std. Error	z value	Pr(> z )	Signif.
(Intercept)	−1.707032	0.033981	−50.235	<2e-16	***
ageOlder	0.003430	0.027578	0.124	0.90101	
genderFemale	−0.191622	0.042129	−4.549	5.40e-06	***
educationTechnical	−0.023924	0.043482	−0.550	0.58218	
educationVocational	0.008193	0.052236	0.157	0.87537	
interlocutor_educationTechnical	0.108447	0.022683	4.781	1.74e-06	***
interlocutor_educationVocational	0.049092	0.029294	1.676	0.09377	.
nr_interlocutorsGroup-chat	−0.053203	0.025175	−2.113	0.03457	*
ageOlder:genderFemale	−0.144247	0.035747	−4.035	5.45e-05	***
genderFemale:educationTechnical	0.125367	0.057345	2.186	0.02880	*
genderFemale:educationVocational	0.182605	0.064582	2.828	0.00469	**

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1. <sup>a</sup>Reference group: younger, male authors in general secondary education, interacting one-on-one with another general pupil.

Table 7: Orality: Anova.

	Chisq	Df	Pr(>Chisq)	Signif.
Age	16.6261	1	4.552e-05	***
Gender	46.2951	1	1.017e-11	***
Education	9.8632	2	0.007215	**
Interlocutor education	24.1395	2	5.730e-06	***
Number of interlocutors	4.4662	1	0.034572	*
Age: gender	16.2833	1	5.454e-05	***
Gender: education	9.2576	2	0.009766	**

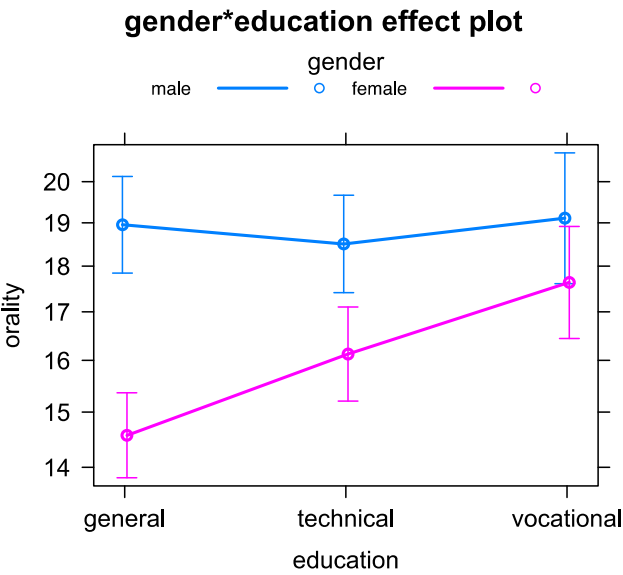
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

## 4.2 Orality

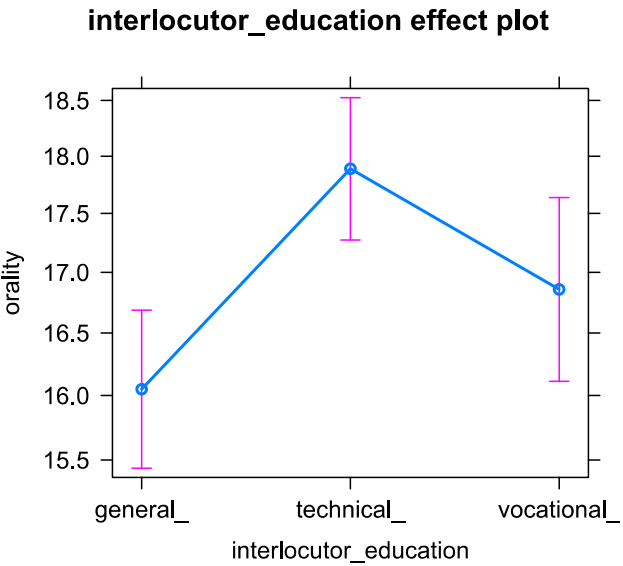
The best model for orality includes all five predictors and two interactions (see Tables 6 and 7 for the fixed effects and the Anova, respectively). Figure 4 shows the effect plot for the interlocutor’s education. When interacting with general education pupils, teenagers use the smallest number of oral markers. The frequency of oral features increases when the conversation partner attends vocational or technical education. However, only conversations including a technical student as interlocutor stand out significantly: the difference in degree of ‘orality’ when interacting with general versus vocational pupils is insignificant. We note that these findings differ from the ones for gender accommodation (Hilte et al. 2020c), as boys nor girls significantly adapted the use of oral markers depending on their interlocutor’s gender. Consequently, we can conclude that the use of expressive markers appears to be adequately adjusted for both gender and education accommodation, whereas the degree of orality is only adapted (for a discussion of the *adequacy* of this adaptation, see below) depending on the interlocutor’s educational profile.

Again, the author’s and the interlocutor’s education do not significantly interact, so the extent to which teenagers adapt their oral writing style to their conversation partner’s educational profile is not significantly impacted by their own educational profile (but see below for an additional analysis).

For orality, as opposed to expressiveness, a direct comparison between the linguistic effect of the interlocutor’s education (Figure 4) and of the author’s own educational track (i.e. actual education-based linguistic variation) cannot be made, since for the use of oral markers, the author’s education and gender significantly interact, and we should thus interpret these two predictors’ effects simultaneously. Figure 3 visualizes how educational differences with respect to orality are different for boys versus girls. Boys clearly insert more oral features in



**Figure 3:** Effect plot: Orality by the author’s education \* the author’s gender (counts per 100 tokens).

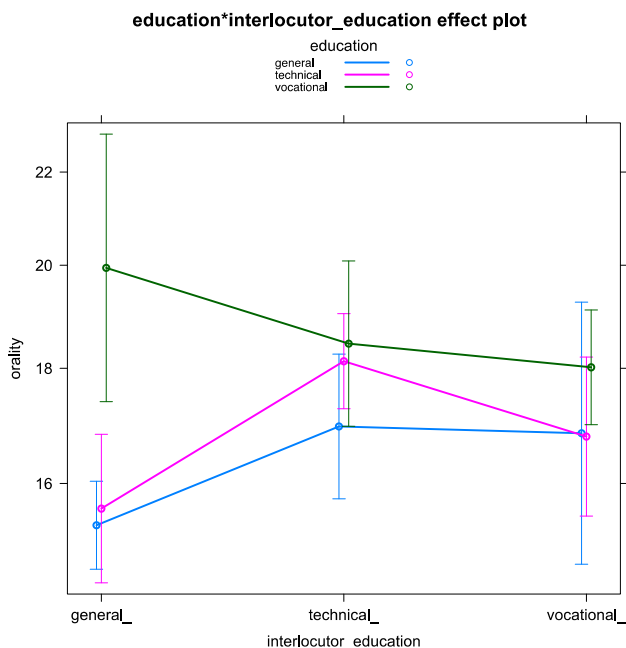


**Figure 4:** Effect plot: Orality by the interlocutor’s education (counts per 100 tokens).

their online discourse than girls do, regardless of their respective educational backgrounds. However, boys' online writing does not show any (significant) educational variation, whereas for girls, a clear educational divide emerges (see also Hilte et al. 2020b). Figure 3 shows how girls in general secondary education use the fewest oral markers, followed by girls in technical and vocational tracks. The difference between each pair is statistically significant (but after Bonferroni correction, only the difference between general and vocational female students remains significant). As for the linguistic adaptation towards one's interlocutor's educational profile, teenagers' less frequent use of oral markers when interacting with general students does reflect the actual educational variation in girls' writing – but then the share of oral markers used in interactions with technical and vocational interlocutors appears to be over- resp. underestimated. However, the observed adaptation towards the interlocutor's educational track clearly diverges from the (absent) educational variation in boys' online discourse. These findings could support the idea of accommodation being caused by 'identity-projection' rather than actual pattern matching (Auer et al. 2005: 201 and references therein). The identity-projection model states that people do not converge towards their interlocutor's actual language use, but rather to a potentially stereotypical image of the social group (resp. role) that the interlocutor belongs to (resp. fulfills) (Auer et al. 2005: 343). So for orality, it is possible that the teenagers' accommodative adaptations rather reflect a stereotyped image of (especially boys') writing rather than their actual style. However, if their accommodative behavior is driven by a stereotypical projection rather than by sociolinguistic reality, it is quite surprising that they primarily seem to associate technical students with the use of oral markers, more so than the most practice-oriented group. For the students in the general track this might be related to the extremely low frequency of interaction with students of the other pole of the educational spectrum, but even then this is still an unexpected pattern.

Two confounding patterns emerge from the model, i.e. patterns that are not related to education accommodation. Just like expressive features, oral markers appear to be inserted significantly more often in one-on-one conversations compared to group-chats, which strengthens our hypothesis that these types of interactions have different conversational dynamics. In addition, a significant interaction between authors' age and gender emerges. Regardless of their age, boys always write in a significantly more 'oral' fashion than girls. But the age dynamics strongly differ for the two gender groups: while girls use significantly fewer oral markers at an older age, boys do not (for a detailed discussion and interpretation, see Hilte et al. 2020b).

We will end this section on orality with a tentative exploratory analysis. As mentioned above, the author's and the interlocutor's education do not interact



**Figure 5:** Effect plot: Orality by the author's education \* the interlocutor's education (counts per 100 tokens).

significantly: so the teenagers' adaptation to their interlocutor's educational profile is not significantly influenced by their own educational track. However, a visual inspection of the raw data seems to suggest that vocational students adapt their orality in a somewhat different way than their peers in more theory-oriented tracks do. Therefore, as an exploratory analysis, we will inspect and (cautiously) interpret the interaction between the author's and the interlocutor's education, which is insignificant at  $p = 0.12$ . The model plot (Figure 5) mirrors the pattern that was observed in the data.

Figure 5 shows how general and technical students accommodate in similar ways, while vocational pupils stand out when they interact with peers from general education by not tempering but rather 'boosting' their use of oral markers. Potential explanatory factors are a weaker familiarity of vocational students with the standard Dutch equivalent of certain oral features, related to differences between educational tracks regarding the focus on formal standard Dutch writing (see Hilte et al. 2020a for an extensive discussion), potentially combined with some sort of reluctance to 'give up' this type of features when confronted with students who have a stronger orientation towards the standard language. Regarding

the latter hypothesis, we note that the observed pattern might be symptomatic of so-called *speech complementarity*, i.e. divergence consistent with (and emphasizing) social roles, which, “if both parties expect and prefer communicative differences, [...] will be positively received” (Muir et al. 2016: 477). This pattern has been observed before in certain mixed-gender interactions and in some interactions including interlocutors with different socio-economic status/power (Burgoon et al. 2017; Dragojevic et al. 2015: 15; Muir et al. 2016: 477).

The slightly different accommodative behavior by vocational students that is suggested by the data and model plot should be investigated further, preferably on a larger dataset, since the ‘deviant’ interactional setting (i.e. vocational students interacting with general students) concerns only 4% of all conversations in the current dataset (see Table 2 above). However, while the collection of additional data for this particular setting might help verify the validity and robustness of the suggested pattern, we note that the limited size of this subset of the corpus is also a (sociological) result in itself, since it appears to be symptomatic of youths’ limited contact and communication across educational boundaries.

## 5 Conclusion and discussion

This study aimed to contribute to research on linguistic accommodation in three ways: by focusing on an under-researched socio-demographic variable (educational track), target group (adolescents) and interactional setting (spontaneous written online conversations).

The analysis of interaction patterns in a large corpus of Flemish secondary school students’ private social media conversations revealed that teenagers’ online communication proceeds primarily within same-education peer group networks and much less frequently across educational boundaries. This could be an explanatory factor for the previously attested low awareness about educational linguistic variation: despite the observation of clearly distinct online writing styles depending on teenagers’ educational track (Hilte et al. 2020b), Flemish teenagers’ insight in these patterns appears to be very weak (Hilte et al. 2019). However, the present study shows that youths do significantly adapt two prototypical aspects of their online writing style depending on their conversation partner’s educational profile.

The way in which teenagers adjust the frequency with which they insert expressive features (such as emoticons) depending on their conversation partner’s educational track strongly reflects the actual frequency pattern for these features in the online discourse of secondary school students in these different tracks. So in that sense the accommodative adaptation of expressive writing appears to be

‘accurate’ and can thus decrease both the linguistic and, following CAT, also the social distance between interlocutors, which is generally perceived as having a positive effect in terms of solidarity and mutual understanding (Dragojevic et al. 2015). Since teenagers’ awareness of educational linguistic variation is very low, we can hypothesize that this convergence is the result of a predominantly subconscious process of pattern matching. However, with respect to the integration of oral markers (e.g. regional or colloquial slang) in online discourse, the accommodative behavior in terms of frequency renders no accurate reflection of the actual educational pattern: rather than suggesting pattern matching, the findings for this linguistic variable seem to support the identity-projection model, with teenagers’ accommodative adaptations reflecting a stereotyped image of their peers’ writing rather than their actual style (Auer et al. 2005). However, we note that the observed ‘projected’ pattern is still somewhat unexpected, as it does not entirely correspond to the common stereotyped image of vocational students’ writing.

A potential explanation for the difference in the ‘accuracy’ of mirroring of the educational variation for oral versus expressive features, concerns the higher ‘visibility’ of many of the (typographic/pictorial) expressive markers. It seems plausible that the insertion of e.g. emoji is more salient and therefore triggers more adequate accommodation than the use of certain low-level colloquial markers. Note that in related work, emoji use in particular has been considered “an important aspect in the [digital] communication that is worth to accommodate” (Kroll et al. 2018: 8). Furthermore, expressive markers may be features that are generally used more deliberately (see also Adams et al. 2018: 475), while teenagers (and especially students in more practice-oriented tracks with a minor focus on formal Dutch writing) might be less in control over their speech-like (e.g. regional, colloquial) rendering of Dutch words. Consequently, it may simply be harder to (adequately) adapt the degree of ‘oral writing’ than that of ‘expressive writing’. Finally, we recall that the actual sociolinguistic variation depending on authors’ educational track is actually more complex for orality (as it interacts with the authors’ gender) than it is for expressiveness, which could also complicate accurate mirroring.

For neither expressive nor oral features a significant interaction emerged between the author’s and the interlocutor’s education, which indicates that the teenagers’ linguistic adaptation towards their interlocutor’s educational profile is not influenced by their own educational track. Or in other words: pupils in one educational track do not adapt their online writing style to a greater extent than their peers in other tracks. These ‘symmetric’ accommodative efforts contrast with the asymmetric gender patterns that were previously found for expressive markers, with boys converging much more strongly to a more ‘female’ expressive online

standard than vice versa (Hilte et al. 2020c; see also Wolf 2000), as well as with the repeatedly attested unidirectional ‘upward’ convergence among interlocutors in distinct social positions (e.g. Dragojevic et al. 2015; Muir et al. 2016). In view of the fact that most vocational students in our dataset have a working-class background and most general students do not (Hilte et al. 2018a, 2018b), this suggests that for these adolescents, peer group solidarity – resulting in convergence from both sides – might predominate over potential social hierarchy. However, the data weakly suggest that vocational students adapt the use of oral markers in a different way than their peers in more theory-oriented tracks, since they appear rather reluctant to ‘give up’ this particular set of features and thus diverge in this respect, especially from their interlocutors at the other pole of the educational spectrum. This pattern certainly needs further investigation, but it is intriguing nonetheless, as it might be symptomatic of *speech complementarity*, i.e. divergence that follows and strengthens social roles – a pattern that reflects certain previous findings with respect to gender- and social status-based accommodation (Burgoon et al. 2017; Dragojevic et al. 2015: 15; Muir et al. 2016: 477).

Finally, an inspection of potential confounding factors revealed that the teenagers’ adaptation towards their interlocutors’ educational profile is not significantly different in dyadic (one-on-one) chats than it is in group chats. That is quite fascinating, since some might expect stronger convergence in the former type of interactions for several reasons. For instance, dyadic talks 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. Group conversations might pose ‘accommodative dilemmas’ (Dragojevic et al. 2015: 17), as there are multiple people that one could adapt to – we note that group talks are, potentially for this reason, more rare in accommodation research. In our previous work on gender accommodation, convergence (for expressive markers) was indeed found to be much more outspoken in one-on-one conversations (Hilte et al. 2020c). This distinction with the present findings could indicate that gender accommodation is of a more personal, intimate nature than education accommodation. Finally, with respect to the comparison between gender- and education-based patterns, we note that the observation of linguistic mirroring in mixed-education interactions is, in a sense, more remarkable than in mixed-gender talks, as teenagers with distinct educational profiles often have quite different social backgrounds/profiles too, and tend to have limited (on- or offline) contact with each other, which does not hold for teenage boys versus girls.

This exploratory study yields promising results in the field of linguistic accommodation with respect to (youths’) educational profiles and opens up several paths for further research. First of all, as a follow-up to this study, we want to zoom

in on the observed macro-scale patterns by analyzing group-bound preferences for some of the most salient oral and expressive markers in order to find out which specific features are most resp. least susceptible to accommodative change and whether accommodative change can be observed for other aspects of these features than their frequency of insertion. For instance, do different groups of secondary school students prefer different types of emoji/emoticons or different kinds of oral tokens? And if so, can any distributional convergence be noted, i.e. do the teenagers adopt a preference pattern that is more similar to that typically associated with their interlocutor's educational profile?

Another path for future work concerns the analysis of additional linguistic variables, such as general textual features (e.g. average sentence length), as well as the inclusion of general linguistic (e.g. lexical) similarity measures.

Finally, we note that the temporal dimension of accommodation falls outside the scope of the present paper. Our current results can be complemented with a study in which the 'diachrony' of accommodative adjustments with respect to interlocutors' educational profiles is analyzed. It will be worth analyzing how and when patterns of convergence (resp. non-adaptation or even divergence) emerge in an interaction and how they evolve, e.g. at what speed or according to which pattern (e.g. linear, fluctuating, ...).

**Acknowledgments:** We thank Ella Roelant for statistical advice.

**Research funding:** This research was funded by the FWO (Research Foundation Flanders) under grant number 12U2620N.

## References

- Adams, Aubrie, Jai Miles, Norah E. Dunbar & Howard Giles. 2018. Communication accommodation in text messages: Exploring liking, power, and sex as predictors of textisms. *The Journal of Social Psychology* 158(4). 474–490.
- Androutsopoulos, Jannis. 2011. Language change and digital media: A review of conceptions and evidence. In Tore Kristiansen & Nikolas Coupland (eds.), *Standard languages and language standards in a changing Europe*, 145–161. Oslo: Novus.
- Auer, Peter, Frans Hinskens & Paul Kerswill (eds.). 2005. *Dialect changes. Convergence and divergence in European languages*. Cambridge: Cambridge University Press.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2017. Package 'lme4'. Available at: <https://cran.r-project.org/web/packages/lme4/lme4.pdf>.
- Bunz, Ulla & Scott W. Campbell. 2004. Politeness accommodation in electronic mail. *Communication Research Reports* 21(1). 11–25.
- Burgoon, Judee K., Norah E. Dunbar & Howard Giles. 2017. Interaction coordination and adaptation. In Alessandro Vinciarelli, Maja Pantic, Nadia Magnenat-Thalmann &

- Judee K. Burgoon (eds.), *Social signal processing*, 78–96. Cambridge: Cambridge University Press.
- Danescu-Niculescu-Mizil, Cristian, Michael Gamon & Susan Dumais. 2011. Mark my words! Linguistic style accommodation in social media. In Sadagopan Srinivasan, Krithi Ramamritham, Arun Kumar, M. P. Ravindra, Elisa Bertino & Ravi Kumar (eds.), *Proceedings of the 20th international conference on world wide web*, 745–754. New York: ACM.
- de Decker, Benny & Reinhild Vandekerckhove. 2017. Global features of online communication in local Flemish: Social and medium-related determinants. *Folia Linguistica* 51. 253–281.
- de Jager, Hugo, Albert Louis Mok & G. Sipkema. 2009. *Grondbeginselen der sociologie*. [Fundamentals of sociology]. Groningen/Houten: Noordhoff.
- de Siqueira, Amaury & Susan C. Herring. 2009. Temporal patterns in student-advisor instant messaging exchanges: Individual variation and accommodation. *Proceedings of the 42nd Hawaii international conference on system sciences*. Big Island: IEEE.
- Dino, Amanda, Stephen Reysen & Nyla R. Branscombe. 2009. Online interactions between group members who differ in status. *Journal of Language and Social Psychology* 28(1). 85–93.
- Doyle, Gabriel. 2016. A robust framework for estimating linguistic alignment in social media conversations. In Jacqueline Bourdeau, Jim A. Hendler, Roger Nkambou, Ian Horrocks & Ben Y. Zhao (eds.), *Proceedings of the 25th international conference on world wide web*, 637–648. Montreal: ACM.
- Dragojevic, Marko, Jessica Gasiolek & Howard Giles. 2015. Communication accommodation theory. In Charles R. Berger & Michael E. Roloff (eds.), *The international encyclopedia of interpersonal communication*, Vol. 1, 1–21. Hoboken: Wiley Blackwell.
- Flemish Ministry of Education and Training [FMET]. 2018. *Statistisch jaarboek van het Vlaams onderwijs. Schooljaar 2016–2017*. Brussels, Belgium: Department of Education and Training.
- Giles, Howard & Tania Ogay. 2007. Communication accommodation theory. In Bryan B. Whaley & Wendy Samter (eds.), *Explaining communication. Contemporary theories and exemplars*, 293–310. Mahwah: Lawrence Erlbaum.
- Goldthorpe, John H. & Richard Breen. 2007. Explaining educational differentials. Towards a formal rational action theory. In John H. Goldthorpe (ed.), *On sociology. Second edition. Volume two: Illustration and retrospect*, 45–72. Stanford: Stanford University Press.
- Gonzales, Amy L., Jeffrey T. Hancock & James W. Pennebaker. 2010. Language style matching as a predictor of social dynamics in small groups. *Communication Research* 37(1). 3–19.
- Harrison, Xavier A. 2014. Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ* 2(1). e616.
- Herring, Susan C. & Sanja Kapidzic. 2015. Teens, gender, and self-presentation in social media. In James D. Wright (ed.), *International encyclopedia of social and behavioral sciences*, 2nd edn. 146–152. Oxford: Elsevier.
- Hilte, Lisa. 2019. *The social in social media writing: The impact of age, gender and social class indicators on adolescents' informal online writing practices*. Antwerp: University of Antwerp Doctoral Dissertation.
- Hilte, Lisa, Reinhild Vandekerckhove & Walter Daelemans. 2018a. Adolescents' social background and non-standard writing in online communication. *Dutch Journal of Applied Linguistics* 7(1). 2–25.
- Hilte, Lisa, Reinhild Vandekerckhove & Walter Daelemans. 2018b. Social media writing and social class: A correlational analysis of adolescent CMC and social background. *International Journal of Society, Culture & Language* 6(2). 73–89.

- Hilte, Lisa, Reinhild Vandekerckhove & Walter Daelemans. 2018c. Expressive markers in online teenage talk: A correlational analysis. *Nederlandse Taalkunde* 23(3). 293–323.
- Hilte, Lisa, Reinhild Vandekerckhove & Walter Daelemans. 2019. Adolescents' perceptions of social media writing: Has non-standard become the new standard? *European Journal of Applied Linguistics* 7(2). 189–224.
- Hilte, Lisa, Reinhild Vandekerckhove & Walter Daelemans. 2020b. Modeling adolescents' online writing practices: The sociolectometry of non-standard writing on social media. *Zeitschrift für Dialektologie und Linguistik* 87(2). 173–201.
- Hilte, Lisa, Reinhild Vandekerckhove & Walter Daelemans. 2020c. Linguistic accommodation in teenagers' social media writing: Convergence patterns in mixed-gender conversations. *Quantitative Linguistics*. <https://doi.org/10.1080/09296174.2020.1807853>.
- Hilte, Lisa, Walter Daelemans & Reinhild Vandekerckhove. 2018d. Predicting adolescents' educational track from chat messages on Dutch social media. In *Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media analysis*, 328–334. Association for Computational Linguistics.
- Hilte, Lisa, Walter Daelemans & Reinhild Vandekerckhove. 2020a. Lexical patterns in adolescents' online writing: The impact of age, gender, and education. *Written Communication* 37(3). 365–400.
- Ismail, Noriszura & Abdul Aziz Jemain. 2007. *Handling overdispersion with negative binomial and generalized Poisson regression models*. *Casualty actuarial society forum*, 103–158. Citeseer.
- Jones, Simon, Rachel Cotterill, Nigel Dewdney, Kate Muir & Adam N. Joinson. 2014. Finding Zelig in text: A measure for normalising linguistic accommodation. In Jun'ichi Tsujii & Jan Hajic (eds.), *Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical papers*, 455–466. Dublin: ACL.
- Kroll, Tobias, Lee-Marie Braun & Stefan Stieglitz. 2018. Accommodated emoji usage: Influence of hierarchy on the adaption of pictogram usage in instant messaging. *ACIS 2018 Proceedings* 82. 1–11.
- Maconis, John J. 2011. *Society. The basics*. Upper Saddle River: Pearson Education.
- Muir, Kate, Adam Joinson, Rachel Cotterill & Nigel Dewdney. 2016. Characterizing the linguistic chameleon: Personal and social correlates of linguistic style accommodation. *Human Communication Research* 42(3). 462–484.
- Muir, Kate, Adam Joinson, Rachel Cotterill & Nigel Dewdney. 2017. Linguistic style accommodation shapes impression formation and rapport in computer-mediated communication. *Journal of Language and Social Psychology* 36(5). 525–548.
- Niederhoffer, Kate G. & James W. Pennebaker. 2002. Linguistic style matching in social interaction. *Journal of Language and Social Psychology* 21(4). 337–360.
- Pérez Sabater, Carmen. 2017. Linguistic accommodation in online communication: The role of language and gender. *Revista Signos. Estudios de Lingüística* 50(94). 265–286.
- Riordan, Monica A., Kris M. Markman & Craig O. Stewart. 2013. Communication accommodation in instant messaging: An examination of temporal convergence. *Journal of Language and Social Psychology* 32(1). 84–95.
- Scissors, Lauren E., Alastair J. Gill & Darren Gergle. 2008. Linguistic mimicry and trust in text-based CMC. In Bo Begole & David W. McDonald (eds.), *Proceedings of the 2008 ACM conference on computer supported cooperative work*, 277–280. New York: ACM.
- Scissors, Lauren E., Alastair J. Gill, Kathleen Geraghty & Darren Gergle. 2009. CMC we trust: The role of similarity. In Dan R. Olsen (ed.), *Proceedings of the SIGCHI conference on human factors in computing systems*, 527–536. New York: ACM.

- Tagliamonte, Sali A. & Derek Denis. 2008. Linguistic ruin? Lol! Instant messaging and teen language. *American Speech* 83(1). 3–34.
- Taylor, Ronald L. 2001. Adolescent peer group language. In Rajend Mesthrie (ed.), *Concise encyclopedia of sociolinguistics*, 297–302. Amsterdam: Elsevier.
- Vandekerckhove, Reinhild & Dominiek Sandra. 2016. De potentiële impact van informele online communicatie op de spellingpraktijk van Vlaamse tieners in school context. *Tijdschrift voor Taalbeheersing* 38(3). 201–234.
- Varnhagen, Connie K., G. Peggy McFall, Nicole Pugh, Lisa Routledge, Heather Sumida-MacDonald & Trudy E. Kwong. 2010. Lol: New language and spelling in instant messaging. *Reading and Writing* 23. 719–733.
- Verheijen, Lieke. 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.
- Verheijen, Lieke & Wilbert Spooren. 2017. The impact of WhatsApp on Dutch youths' school writing. In Egon W. Stemle & Ciara R. Wigham (eds.), *Proceedings of the 5th conference on CMC and social media corpora for the humanities (cmccorpora17)*, 3–4 October 2017, 6–10. Bolzano, Italy: Eurac Research.
- Vranken, Jan, Geert Van Hoetegem, Erik Henderickx & Luc Vanmarcke. 2017. *Het speelveld, de spelregels en de spelers? Handboek sociologie [The field, the rules and the players? Handbook sociology]*. Leuven/The Hague: Acco.
- Wolf, Alecia. 2000. Emotional expression online: Gender differences in emoticon use. *Cyber Psychology and Behavior* 3(5). 827–833.
- Xu, Yang, Jeremy Cole & David Reitter. 2018. Not that much power: Linguistic alignment is influenced more by low-level linguistic feature rather than social power. In Iryna Gurevych & Yusuke Miyao (eds.), *Proceedings of the 56th annual meeting of the Association for Computational Linguistics (volume 1: Long papers)*, 601–610. Melbourne: ACL.
- Xu, Yang & David Reitter. 2015. An evaluation and comparison of linguistic alignment measures. In Tim O'Donnell & Marten van Schijndel (eds.), *Proceedings of the 6th workshop on cognitive modeling and computational linguistics*, 58–67. Denver: ACL.
- Zeileis, Achim, Christian Kleiber & Simon Jackman. 2008. Regression models for count data in R. *Journal of Statistical Software* 27(8). 1–25.