



**International
Journal of Society, Culture & Language
IJSCL**

Journal homepage: www.ij scl.net
ISSN 2323-2210 (online)

Social Media Writing and Social Class: A Correlational Analysis of Adolescent CMC and Social Background

Lisa Hilte^{1a}, Reinhild Vandekerckhove^{2a}, Walter Daelemans^{3a}

ARTICLE HISTORY:

Received February 2018
Received in revised form April 2018
Accepted May 2018
Available online May 2018

KEYWORDS:

Computer-mediated communication
Adolescents
Social class
Social media
Non-standard writing

Abstract

In a large social media corpus (2.9 million tokens), we analyze Flemish adolescents' non-standard writing practices and look for correlations with the teenagers' social class. Three different aspects of adolescents' social background are included: educational track, parental profession, and home language. Since the data reveal that these parameters are highly correlated, we combine them into one social class label. The different linguistic practices emerging from the analyses demonstrate the crucial impact of social class on adolescent online writing practices. Furthermore, our results nuance classical findings on working class adherence to 'old vernacular' by also highlighting working class youth's strong connection to the online writing culture, or 'new vernacular'. Finally, we point out the complexity of the social class variable by demonstrating interactions with gender and age, and by examining groups of teenagers whose social background is ambiguous and therefore hard to operationalize.

© 2018 IJSCL. All rights reserved.

¹ PhD Candidate, Email: lisa.hilte@uantwerpen.be (Corresponding Author)
Tel: 003232654246

² Associate Professor, Email: reinhild.vandekerckhove@uantwerpen.be

³ Professor, Email: walter.daelemans@uantwerpen.be

^a University of Antwerp, Belgium

1. Introduction

In previous studies on informal *computer-mediated communication* (CMC), gender and age have been popular independent variables (e.g., Baron, 2004; Hilte, Vandekerckhove, & Daelemans, in press a; Wolf, 2000). The authors' social class, however, is rarely taken into account. Moreover, certain groups of people systematically tend to be overrepresented in CMC research, as participants are very often middle class people, in most cases middle class youngsters. Consequently, the chat practice of working class teenagers has hardly been studied. Therefore, the present study includes youngsters with this profile and compares their linguistic behavior to that of other social groups.

The study focuses on the correlation between Flemish adolescents' online non-standard writing practices (including typical chatspeak phenomena) and their social background. The paper is structured as follows: First, we discuss the theoretical framework (Section 2) and the methodology (Section 3). Section 4 presents the results of the analyses, and Section 5 is devoted to the discussion of these results and some concluding remarks.

We note that we already conducted a pilot study on this topic (see Hilte, Vandekerckhove, & Daelemans, in press b). The differences with the present study relate to the operationalization of non-standard writing and the overall methodological focus. The present study includes eight more markers of non-standard writing, the combination of different social subfactors has been optimized, and two new sections were added focusing on methodological challenges related to classification of participants with hybrid social profiles and to interactions between social class, age, and gender.

2. Theoretical Framework

In order to obtain a feasible yet accurate operationalization of adolescents' social class, we included criteria from both academic research and Belgian government studies. Taking into account the complexity and multidimensionality of social class, we treat it as a variable consisting of three subvariables

(representing different aspects of class, e.g., cultural, financial, and economic): the teenagers' educational track, their home language, and the profession of their parents. Potential correlations between the three social subfactors will be addressed in Section 4.1.2.

For educational track, we include the three main types of Belgian secondary education (Flemish Ministry of Education and Training – from now on FMET, 2017):

- General Secondary Education: theory-oriented educational track that prepares students for higher education.
- Technical Secondary Education: educational track with a strong practical and theoretical (technical) focus. After graduation, students can either enter higher education or start working.
- Vocational Secondary Education: practice-oriented educational track where students are taught a specific (often manual) profession. Students (can) start working right after graduation. This degree excludes direct access to higher education.

Adolescents' educational track strongly impacts their current and future (adult) social networks and future professional career (de Jager, Mok, & Sijkema, 2009). As today's western societies have evolved toward meritocracies – i.e., “social stratification based on personal merit” (Macionis, 2011, p. 206) – with a strong emphasis on knowledge and skills, education and obtained degrees have become increasingly important determinants of social status and position (de Jager et al., 2009).

Concerning the participants' home language, it is important to note that Dutch is the only official language in Flanders and the only medium of instruction in Flemish education. For the present study, three home language contexts are distinguished:

- The adolescent only speaks Dutch at home.
- The adolescent speaks Dutch and one (or multiple) other language(s) at home.
- The adolescent does not speak Dutch at home, but one (or multiple) other language(s).

In most cases, the “other” language listed by the teenagers appears to be a language which

suggests a recent migration background (e.g., Arabic). Thus home language can be considered an important socio-cultural factor. Furthermore, home language may have an indirect impact on the adolescents' school experience and performance, as it might indicate the presence/absence of a parent who can easily connect with the school context.

The final determinant of minors' social background included in this study is parental profession, as it often has a large impact on the overall family situation (e.g., in financial, economic, and cultural terms). For the classification, we applied the well-known sociological EGP-scheme, which ranks professions based on different criteria, such as degree of autonomy and supervision, and required level of education or skills (Erikson, Goldthorpe, & Portocarero, 1979; Vranken, Van Hootegem, Henderickx, & Vanmarcke, 2017). The requirement of a university degree was added as an extra criterion for distinguishing between upper and middle class professions, so as to fit the current Flemish social landscape more adequately, and the original seven EGP-categories were regrouped into three clusters:

- *Upper class* professions: Non-manual professions for which a university degree is required (e.g., doctor, civil engineer).
- *Middle class* professions: Professions for which a degree of higher education is required, encompassing both non-manual professions for which a non-university degree is required (e.g., secretary, nurse), and manual work for which specific technical degrees are required (e.g., electrician) and that entails a certain degree of autonomy.
- *Working class* professions: unskilled manual professions (e.g., truck driver, cashier).

Whenever the profession of both parents was known, the one that ranked highest served for classification, since the highest ranked profession may have a major impact on the general family situation, e.g., in terms of financial resources and consumption of cultural goods. Finally, we note that we were unable to classify certain social positions which fall outside the scope of the scheme, such as unemployed people or housewives/-men (Marsh, 2000).

In previous research, distinct age and gender patterns were observed in CMC. With respect to gender, women appear to show stronger preferences for expressive markers, such as emoticons (see Section 3.2.1) (e.g., Baron, 2004; Hilde et al., in press a; Parkins, 2012; Varnhagen et al., 2010), which corresponds to older sociolinguistic findings on the strong emotionally and socially connective dimension in women's discourse (e.g., Tannen, 1990).

Concerning age patterns, previous research showed that adolescents tend to use more stylistic chatspeak features than adult chatters (e.g., Argamon, Koppel, Pennebaker, & Schler, 2009; Schwartz et al., 2013). Especially young adolescents appear to favor typical chatspeak features (both expressive markers and unconventional spelling forms) in online interaction (De Decker & Vandekerckhove, 2017; Hilde et al., in press a; Tagliamonte & Denis, 2008; Verheijen, 2015). These age patterns seem indicative of changing linguistic attitudes as adolescents grow older (Verheijen, 2015).

Social class, however, has – to our knowledge – not been operationalized as a linguistic determinant in (adolescent) CMC, and neither have the three social subfactors included in the present study. First of all, parental profession has never been operationalized in CMC research. Educational track and CMC have actually been linked to each other, though from a completely different perspective. Some studies discuss the educational use of CMC (e.g., Heemskerk, Brink, Volman, & Ten Dam, 2005; Yates, 2001). The same holds for home language. Its impact on CMC writing has not been tested, but there has been research on the application of CMC in foreign language teaching (e.g., Warner, 2004) and on the use of English CMC (in a business context) by non-native speakers (Zummo, 2018). Furthermore, some studies examine the impact of CMC on students' writing performance in school contexts (e.g., Vandekerckhove & Sandra, 2016). The latter study points to educational track as a determining variable. Students in Vocational Education seem to have more trouble avoiding chatspeak interferences in formal school writing than their peers in more theory-oriented educational tracks. Still, the question whether home language or educational

track actually influences online writing style remains unanswered.

Although social class has not yet been examined systematically in variationist research on informal CMC, several studies have addressed the visibility of social structures and inequality in the genre. In the early days, digital communication was assumed to be free of inequality, because of the lack of (social) face-to-face cues. However, Yates (2001) concluded that this so-called democratic theory/model of CMC does not hold, because the technology does not “strip away existing social structures” (p. 32), and because “CMC suffers, like all communications media, from the intrusion of existing social relations, including those that are based upon inequalities of access and power” (pp. 32-33).

An important non-linguistic class difference that has been addressed in previous CMC research concerns the access to technology and familiarity with digital writing (Heemskerk et al., 2005; Yates, 2001). Heemskerk et al. (2005) conclude that the use of ICT-tools might actually “increase inequality in education”, because of a “digital divide [...] that follows the

traditional lines of race and social class” (pp. 1-2). This approach falls outside the scope of the present paper, but obviously all teenagers in our corpus have access to the technology and at least some CMC-literacy, since they donated personal CMC-data (see below).

3. Methodology

Below, we discuss the corpus (Section 3.1) and the procedure of the data processing and feature extraction (Section 3.2).

3.1. Corpus

The corpus consists of over 2.8 million tokens (488K posts) produced by 1384 Flemish teenagers in an informal interactive CMC-context. Table 1 shows the distribution of the social variables. All participants’ age, gender, and educational track is known, and for almost all of them, home language could be included too. Parental profession was hard to get access to, as many participants either left this field blank or produced answers which were too vague for classification (e.g., a company name without a job description).

Table 1
Distributions in the Corpus

Variable	Variable levels	Tokens	Participants
Education	General Secondary Education	920114 (32%)	596 (43%)
	Technical Secondary Education	1213483 (42%)	395 (29%)
	Vocational Secondary Education	751487 (26%)	393 (28%)
	Unknown	0 (0%)	0 (0%)
Home language	Dutch only	2563096 (89%)	1154 (83%)
	Dutch + other language	216558 (8%)	87 (6%)
	Other language only	93978 (3%)	105 (8%)
	Unknown	11452 (0.4%)	38 (3%)
Parental profession	‘Upper class’ professions	415965 (14%)	99 (7%)
	‘Middle class’ professions	743952 (26%)	214 (15%)
	‘Working class’ professions	392215 (14%)	87 (6%)
	Unknown	1332952 (46%)	984 (71%)
Total		2885084	1384

As the corpus contains an imbalance for gender (66% of the tokens were produced by girls and 34% by boys) and a slight imbalance for age (younger teenagers (aged 13-16): 55%, older teenagers or young adults (aged 17-20): 45%), we will control for gender and age influences in the linguistic analyses. There is no need to control for other factors such as dialect region or medium, as these are highly constant in the corpus. Almost all tokens (96%) were collected

from participants living in the same dialect region in the center of Flanders, Antwerp-Brabant, which makes region a (quasi-)constant. The same holds for medium and year. Almost all tokens (99%) were extracted from instant messages on Facebook/Messenger or WhatsApp, and the vast majority of the tokens (87%) were produced in 2015-2016 (compared to 10% in 2013-2014 and 2% in 2011-2012).

The data were collected in a school context, but the conversations delivered by the students were produced outside of school and before the time of collection. The participants were instructed to submit conversations with Dutch as the main language. Entire conversations in a language other than Dutch were excluded from this study, but data with some code switching were not. Students were free to participate and donate their chat conversations. Photos were deleted automatically and all data were anonymized in order to guarantee the privacy of the participants.

3.2. Procedure

3.2.1. The Operationalization of Non-Standard Writing

We operationalize adolescents' online non-standard writing as a combination of eleven kinds of deviations from the formal writing standard. These deviations relate to the three "maxims" of informal CMC, i.e., three largely implicit but widely applied rules of linguistic conduct in CMC-contexts: those of orality, brevity (also economy/speed), and expressive compensation (Androutsopoulos, 2011; De Decker & Vandekerckhove, 2017). Below, we discuss the feature sets, define the underlying principles, and provide examples from the corpus.

The largest set consists of *expressive* features: (mostly typographic) linguistic markers which add or enhance the expression of emotional or social involvement in a chat message. They are related to the chatspeak principle of *expressive compensation*, which implies that all kinds of strategies are used to compensate for the absence of certain expressive cues in online communication, such as volume and facial expressions. Seven types of expressive markers were included in the analysis. The selection of these markers is based on related research (Androutsopoulos, 2011; Parkins, 2012; Varnhagen et al., 2010; Verheijen, 2015; Wolf, 2000).

1. Emoticons and emoji: stylized facial expressions and hearts (manually composed with characters or selected as a pictogram from the platform's keyboard interface) and pictograms (representing various objects)

e.g., *zie u graag !* 🥰😘😘😘😘💋❤️ ("love you!")

2. Allcaps: the capitalization of entire words or sentences to convey a feeling (anger, excitement, etc.), to mimic shouting, or to emphasize a particular word

e.g., *IK BEN ECHT BOOS* ("I AM REALLY ANGRY")

e.g., *Dan zijn we om 1u ZEKER thuis* ("Then we will be home by 1 o'clock **FOR SURE**")

3. Deliberate letter repetition: written representation of the oral phenomenon of lengthening a sound to stress a word

e.g., *Een suuuuuuuuuuuuuper gelukkige verjaardag* ("A suuuuuuuuuuuuper happy birthday")

4. Deliberate repetition of question or exclamation marks: to increase their expressive function

e.g., *Ja!!!* ("Yes!!!")

5. Combinations of question and exclamation marks: often used to convey disbelief or shock

e.g., *Serious?!?!* ("Seriously?!?!")

6. The onomatopoeic rendition of laughter

e.g., *Hahahahahahaha*

7. Kisses: the rendition of kisses/hugs through combinations of the letters "x" and "xo"

e.g., *Ik spreek u morgen xxx* ("I will talk to you tomorrow xxx")

The second set of deviations from the formal written standard consists of features related to the *write like you speak* principle. This *orality* principle implies that in spite of the written character of the digital medium, the register in informal CMC is often closer to oral than to written communication. We included:

8. Non-standard Dutch lexemes: dialect words, slang, or written representations of non-standard phonological phenomena (like the deletion of the final "t" in short function words, as shown in the last example below) e.g., *gij hebt niks te vertellen* (std. Dutch "**ji** hebt niks te vertellen", "**you** have got nothing to say")

e.g., *ik ook ni* (std. Dutch “ik ook **niet**”, “me **neither**”)

9. English lexemes (in a Dutch conversation)
e.g., *Die zijn echt heel nice* (“They are really very **nice/cool**”)

The inclusion of English in the orality category may at first sight seem surprising. However, the (abundant) use of English marks Flemish adolescent speech, and most of the English lexemes and utterances have not been integrated into standard Dutch (yet).

The third set of features concerns the principle of *brevity* and covers all kinds of strategies to compress words or utterances and thus maximize typing speed and minimize typing effort. They enable chatters to mimic, to a certain extent, the *flow* of a face-to-face conversation. We included the following cluster of features:

10. typical chatspeak abbreviations and acronyms (none of them standard Dutch)
e.g., *omg das geweldig* (full version: “**oh my god** das geweldig”, “**oh my god** that is awesome”)
e.g., *ja idd* (full version: “ja **inderdaad**”, “yes **indeed**”)

The final set of features included in the research design does not belong to any of the three main categories, but is nevertheless typical of online discourse:

11. Discourse markers: # (“hashtag”, to indicate a topic or express a feeling about it) and @ (“at”, to address one person directly in a group conversation)
e.g., *#verslaafd* (“#addicted”)
e.g., *@robin*

This collection of deviations from the formal written standard consists of both “old vernaculars” and “new vernaculars” (Androutsopoulos, 2011, p. 146), or old and new types of non-standard writing. The typographic expressive features, the prototypical non-standard chatspeak abbreviations and the discourse markers can be considered new vernacular: they cover new ways of deviating from formal written standards that are bound to digital culture (Androutsopoulos, 2011). The non-standard Dutch lexemes can be considered traditional

vernaculars: they represent “locally bound ways of speaking” (Androutsopoulos, 2011, p. 146), or in this context, regional and slang linguistic variants that have marked colloquial speech for ages. The only feature that cannot be classified in terms of old or new vernaculars unambiguously is the use of English lexemes in Dutch chat conversations. As it generally reflects offline colloquial speech practices, it resembles some of the old vernacular features. However, the term “old” is largely inappropriate here, since the increasing impact of “global” English is a relatively recent phenomenon. Moreover, some English practices do not reflect adolescent speech but cover specific terms, acronyms, and memes related to international chat culture.

3.2.2. Feature Extraction

Occurrences of the features were extracted and counted automatically with Python scripts. For a test set of 200 randomly selected posts (1257 tokens), the software’s output was compared to human annotations and judged to be reliable. The average precision score (i.e., the percentage of detected occurrences of a feature that are indeed valid occurrences of that feature) for all eleven features was 0.92. The average recall score (i.e., the percentage of all occurrences of a feature present in the corpus that are detected as such) was 0.88. We note that in the present study, both measures are (equally) important, as we want our software to be precise in its detections without missing relevant occurrences. The average scores as well as the scores for the individual features indicate that the overall feature detection is reliable.

4. Results

This section discusses the impact of (aspects of) adolescents’ social class on their online non-standard writing. First, we analyze the correlation between educational track, home language, and parental profession and evaluate their combined impact (Section 4.1). Next, we broaden up the scope on social class by examining adolescents with hybrid social profiles (Section 4.2) as well as possible interactions between social class, age, and gender (Section 4.3).

4.1. The Impact of Social Class on Non-Standard Writing Practices

We start with a brief discussion of the individual impact of educational track, home language, and parental profession on adolescents' online writing practices (Section 4.1.1). Next, we show how these social subfactors are actually correlated (Section 4.1.2). Finally, we operationalize social class as a combination of educational track and parental profession and examine their combined linguistic impact (Section 4.1.3).

4.1.1. Individual Impact of Educational Track, Home Language, and Parental Profession

Educational track, home language, and parental profession all significantly correlate with the use of non-standard features ($p < 0.0001$ for the three chi-square tests). All of the social patterns remained valid (and equally strong) after correction for age and gender imbalances in the dataset. Students in theory-oriented educational tracks score lower for non-standard features than their peers in practice-oriented tracks, and so do participants with higher class parents compared to their peers with a lower class family background. Finally, teenagers who only speak Dutch at home produce fewer non-standard markers than their peers with a – combined or exclusive – “other language” profile. Interestingly, the “other language” groups' higher rate of non-standardness does not seem to be related to a more frequent use of other languages (e.g., Arabic) in Dutch chat conversations, but instead appears to indicate a

stronger preference for typographic expressive markers (e.g., emoticons).

4.1.2. Correlations Between Educational Track, Home Language, and Parental Profession

We start by examining the potential correlation between the teenagers' educational track and the profession of their parents. The analysis is performed on the profiles (and not on the chat conversations, as no linguistic variable is included here) of participants whose parents' profession is known (400 or 29% out of 1384 participants). Information on the educational track is available for all participants. The data reveal a significant and strikingly strong correlation between educational track and parental profession (chi sq. = 99.638, $p < 0.0001$, Cramer's $V = 0.35$). The mosaic plot (Figure 1) shows that most youngsters of parents with an upper class profession are in General Secondary Education: a theory-oriented educational track in which students are prepared for higher education, through which they may obtain an upper class profession themselves. The majority of adolescents of parents with a working class profession are in the Vocational system: a practice-oriented education type where a specific (often manual) profession is taught and which generally prepares for a working class career. For children of middle class parents, the three education types are balanced.

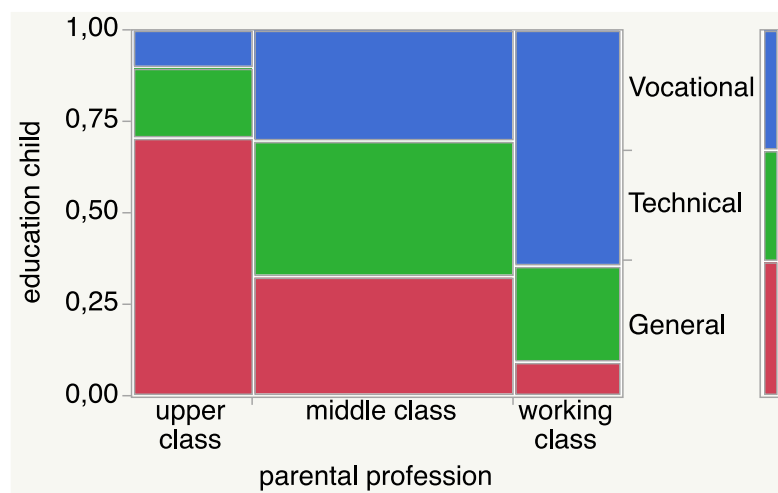


Figure 1

Educational Track by Parental Profession (See also Hilde et al., in press b)

The correlational analysis between adolescents' educational track and their home language was performed for participants whose home language is known (1346 or 97% out of 1384 participants). A significant but not very strong correlation was found ($\chi^2 = 23.249$, $p < 0.0001$, Cramer's $V = 0.09$). The results suggest that it is harder for children from non-Dutch speaking families to get access to more theoretical education systems (see Figure 2). Adolescents with Dutch as their only home language are more likely to attend the theoretical General Education than adolescents who speak another language at home, as 45% of the former category attend General

Education compared to 32% (Dutch + other language) versus 34% (only other language) of the latter group. The data for the Vocational track are even more striking. Only 26% of the students with Dutch as their only home language attend Vocational Education compared to 46% of the students with a combined "Dutch + other language" profile and 39% of the students with an exclusive "other language" profile. The orientation toward Technical Education is comparable for all language groups: 29% of the "Dutch only" teenagers, 22% of the "Dutch + other language" teenagers and 27% of the "other language only" teenagers are students in the Technical track.

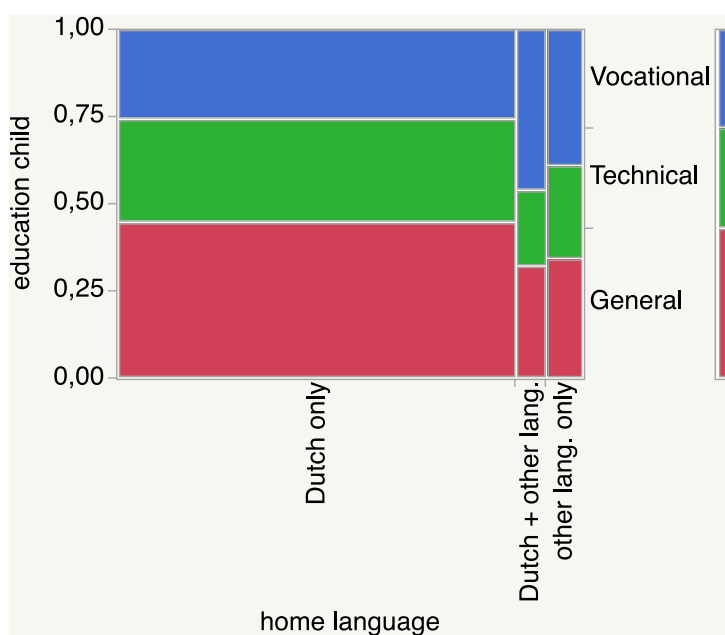


Figure 2
Educational Track by Home Language

The final correlational analysis was performed for participants for whom both parental profession and home language are known (398 or 29% out of 1384 participants). Home language significantly and strongly correlates with parental profession ($\chi^2 = 16.138$, $p = 0.0028$, Cramer's $V = 0.14$). The following pattern emerges (see Figure 3): working class professions seem more common and upper class professions less common in families in which Dutch is not the only home language or is not a home language at all. Most parents in a "Dutch only" home context have a middle class profession (55%), followed by upper class

(27%) and working class (18%) professions. In the families where both Dutch and another language are spoken, middle class professions are still the most common category (52%), but working class professions are far more prominent than in the families where Dutch is the only home language (31%), and upper class professions are less well represented (17%). Finally, in the families where only a language other than Dutch is spoken, half of the parents have a working class profession (50%), followed by middle class (36%), and upper class professions (14%).

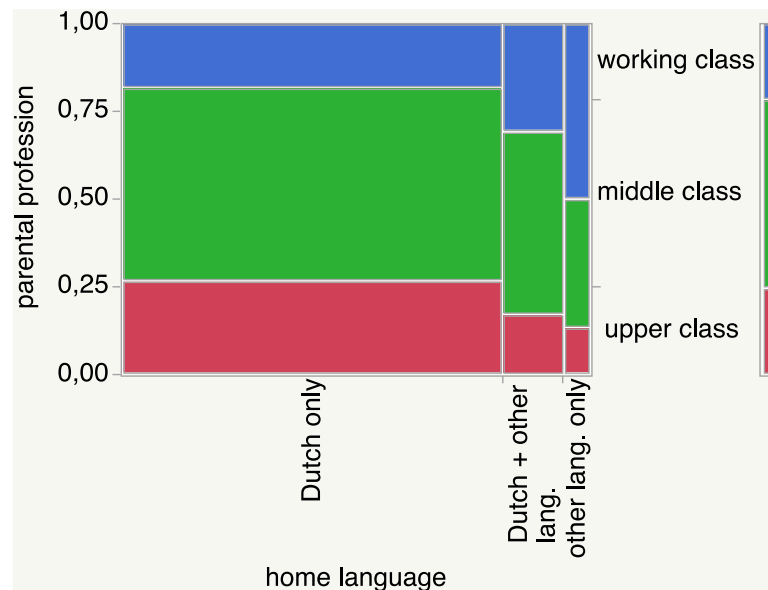


Figure 3

Parental Profession by Home Language

The tendencies visualized in the plots do not only have implications for the processing of the linguistic data (see Section 4.1.3), they clearly have a more general sociological relevance. First of all, Figure 1 shows that both upward and downward social mobility amongst the youngsters is fairly limited (social mobility and status congruence theory will be discussed in Section 4.2). Moreover, while Figure 2 suggests that youngsters with a migration background are relatively overrepresented in the Vocational track, Figure 3 reveals that their parents are overrepresented in working class professions.

4.1.3. Combined Linguistic Impact of Educational Track and Parental Profession

The results of the correlational studies (Section 4.1.2) suggest that the social subfactors representing different aspects of adolescents' social class should not only be examined in isolation, but also in combination. However, the inclusion of home language in the combined analysis had some undesirable consequences (see below). Therefore, three groups of teenagers were distinguished based on the combination of two of the three socio-cultural criteria discussed above, i.e., educational track and parental profession. They were labeled as upper class, middle class, and working class. The upper class group consists of adolescents in General Secondary Education whose parents have an upper class profession. The middle class group contains teenagers in Technical

Education whose parents have a middle class profession. Finally, the working class youngsters are adolescents in Vocational Education whose parents have a working class profession. Table 2 shows an overview of the groups. For two reasons, home language was not included as a criterion for categorization. First, the analyses in the pilot study (Hilte et al., in press b), in which social clusters were created based on all three social subfactors, suggested that home language was too restrictive as a criterion because the dataset for working class youngsters (operationalized in the pilot study as "other language only" students in Vocational Education, with working class parents) became too small. As the large majority of participants speak Dutch at home (either exclusively, or combined with another language), only 8 participants met the three criteria for the working class profile. Additionally, although home language is an important socio-cultural and linguistic factor (see sections 2 and 4.1.1), including it as a criterion implies restricting the analyses to the comparison of the linguistic behavior of "autochthonous" upper class adolescents to that of working class adolescents with a migration background. This implies a questionable simplification of social reality. Obviously many working class families in Flanders are "autochthonous", and needless to say, there are also non-Dutch speaking higher class families, either with or without a recent migration background.

Table 2

Three Prototypical Social Groups

	educational track	parental profession	participants	tokens
'working class' teenagers	Vocational	working class	56	218676
'middle class' teenagers	Technical	middle class	79	387363
'upper class' teenagers	General	upper class	70	221917

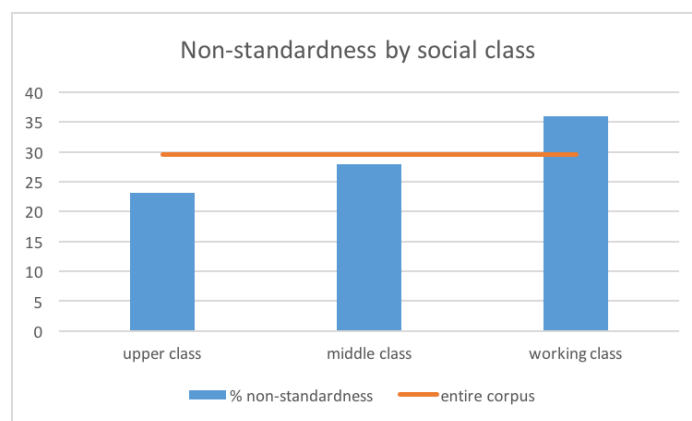
**Figure 4***Non-Standard Writing by Social Class*

Figure 4 shows a gradual pattern for the linguistic variable, with less non-standard writing for adolescents in “higher” social layers. For upper class teenagers, the proportion of non-standard features amounts to 23%, but it rises to 28% and 36% for their middle class and working class peers respectively. The correlation between this construct of social class and non-standardness is statistically significant and also quite strong (chi sq. = 9054.840, $p < 0.0001$, Cramer’s $V = 0.10$, performed on 827956 tokens or 29% out of 2885084). After correcting for age and gender imbalances, the same pattern remains, and the correlation is equally significant and strong.

The differences between the two groups holding extreme positions on the social continuum, i.e., upper and working class youngsters, are very consistent for the different features. Higher frequencies can be found in the working class corpus for eight of the eleven features – for the remaining three (infrequent) features, there are no significant differences. The position of middle class youngsters is quite variable. They hold a middle position for some features (e.g., repetition of punctuation marks), but for other features they have either the lowest frequency scores (e.g., emoticons) or the highest (e.g., kisses). In other words, when it comes to online language practices, middle class adolescents do not just hold an

intermediate position, they have a distinct sociolinguistic profile.

All three social groups deviate from formal writing practices mainly for the sake of orality and expressiveness. Interestingly, the distributions in terms of types of markers also show a gradual difference. The middle class teenagers are strongly oriented toward orality (68% of their non-standard markers are oral features), and much less toward expressiveness (28%). Upper class teenagers show a similar – but less outspoken – preference pattern, with 60% oral features versus 37% expressive markers. For working class adolescents, however, the distribution between expressive and oral features is much more balanced: 53% of their non-standard features serve the purpose of orality, and 44% are used for expressive purposes. In all three groups, chatspeak abbreviations and acronyms score much lower than the other sets of features. They represent 3 to 4% of all non-standard markers.

As working class youngsters use both expressive and oral features significantly more often than their upper class peers, we can conclude that they seem to be attracted more to both “old” vernacular (e.g., dialect words) and “new” vernacular (e.g., typographic chatspeak features such as emoticons). We note that the more frequent use of oral features and of non-

standard Dutch lexemes in particular might also point to a lower proficiency in formal written standard Dutch and/or more carelessness toward standard language norms, which in turn might both be related to a minor focus on standard Dutch proficiency and a stronger focus on skills in practice-oriented education types. The more frequent use of expressive markers, finally, suggests more (typographically) expressive writing by these youngsters.

For brevity-related features, we found no differences between the different groups when dealing with the variables educational track, parental profession, and home language individually and this holds for the combined social profiles. De Decker and Vandekerckhove (2017) already signaled that no gender and hardly any age differences could be attested for the use of acronyms and abbreviations in Flemish CMC, and concluded that these features are the most stable markers of the genre. So, apparently, these features are so useful and functional that they are appreciated by all groups to more or less the same extent.

4.2. Non-Prototypical Social Profiles

The operationalization of adolescents' social class presented in Table 2 leads to three prototypical social groups which we labeled as working class, middle class, and upper class. However, many participants do not fit into one of these categories, but have a more "hybrid" social profile: e.g., teenagers in General Secondary Education whose parents are unskilled manual workers (i.e., working class profession). The online language use of these participants with a hybrid social profile will be examined in this section.

In order to visualize the linguistic behavior for all potential combinations of educational track and parental profession, we adapted the mosaic plot from Figure 1. In Figure 5, the color of the blocks reflects the relative proportion of non-standard features: dark blocks represent higher frequency scores than the paler ones. In every group or block, the participants' profiles in terms of age and gender were checked, and none of the groups were too skewed. Nevertheless, these results should be

interpreted with caution, as some of the smaller blocks consist of few participants. In the bottom left and upper right corners are two of the prototypical groups from Table 2, holding extreme positions on the social continuum. These two groups are youngsters from the upper class and the working class. These two groups' significantly diverging frequency scores for non-standard markers (discussed in the previous section) are now visualized in Figure 5 by extreme color contrasts. The middle block represents the typical middle class youngsters: the orange color shows that their overall frequency score for non-standardness is somewhere in between that of their upper class (pale yellow) and working class (dark red / maroon) peers. The remaining blocks represent youngsters with "hybrid" social profiles. The groups in the upper left and bottom right corner seem to be strikingly deviant concerning their use of non-standard features, as their color stands out. The block in the upper left corner represents adolescents in Vocational Education whose parents have an upper class profession. The pale orange color indicates a relatively low frequency score for non-standard markers. In other words, their language use is fairly standard-oriented. Interestingly, it is more similar to the linguistic profile of their peers with a similar (upper class) family background than to that of their peers in Vocational Education. The opposite pattern can be found for the group in the bottom right corner, which represents adolescents in General Education whose parents have a manual working class profession. The pale yellow reveals that these youngsters produce a relatively small amount of non-standard markers, just like their peers from the same (General) education system and unlike their peers with a similar (working class) family background. Interestingly, the linguistic behavior of these two groups reveals a stronger orientation toward standard writing norms than that of the hybrid groups of Technical students with an upper class family background and General students with a middle class family background. This might point to a tendency of sociolinguistic hypercorrection (see below) amongst youngsters with a strong clash between social family background and educational track.

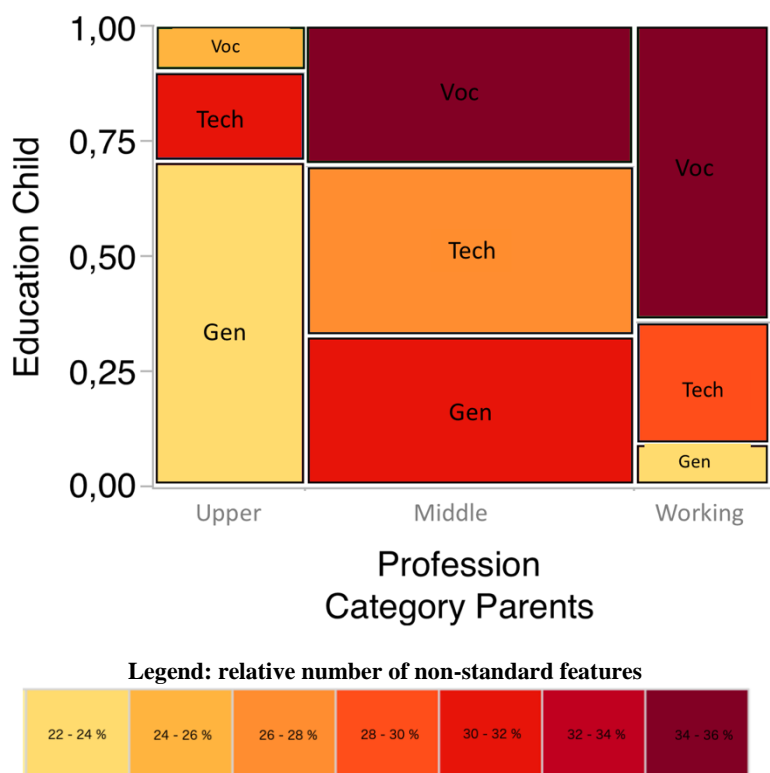


Figure 5
Visualization of Non-Standardness for Different Groups of Adolescents

The, in some respects, ‘deviant’ linguistic practices of particular hybrid groups suggest that some determining factors are still missing in the current operationalization of minors’ social class. The operationalization might be optimized by including attitudinal factors, such as social ambition: Do the youngsters aspire upward social mobility or not? We interpret the adolescents’ social mobility in terms of educational track (assuming this is a reliable predictor for their future professional career) and the professions of their parents. In sociological literature, this type of mobility is called *intergenerational* mobility, as it concerns changes in profession type/class between multiple generations (Vranken et al., 2017). Figure 1, which visualized the number of participants per combination of the different profession and education categories, shows that half of the participants “stagnate” (i.e., no social mobility) (51%): their educational track corresponds to their parents’ profession type. A quarter of the participants move “down” (24%) and a quarter move “up” (25%) the social ladder, since their level of education is likely to lead to a “lower” versus “higher” profession type than that of their parents. We note that these percentages largely correspond to the proportions reported by Vranken et al. (2017)

for father-son intergenerational social mobility in the Netherlands in the 1970s. They report 54% immobility versus 26% upwards and 20% downwards mobility. (Follow-up studies showed a decrease in social immobility in the Netherlands to 45% in the early 2000’s, versus an increase in upwards mobility to 35% and a stagnation of downward mobility, 20%).

In our data, stagnation is clearly most frequent for upper class and working class professions (followed by slight downward or upward mobility respectively), whereas for the middle class professions, the three possibilities (stagnation, upward and downward mobility) are more balanced. The tendencies with respect to social stagnation can be explained by the sociological status congruence theory. *Status congruence* implies that different components of one’s social status are “congruent” or reconcilable, whereas *status incongruence* indicates an imbalance between these components (Vranken et al., 2017). The theory states that status congruence facilitates social interaction and is therefore generally positively reinforced (Vranken et al., 2017). This theory offers a frame for the finding that parents tend to send their children to an education type corresponding to their own status. It predicts

that a lower class background counteracts upward social mobility, while a higher class background counteracts downward mobility. Vranken et al. (2017) therefore conclude that the larger the potential status incongruence, the more mobility will be impeded.

The two groups in the upper left and bottom right corner of Figure 5, whose online language use is most deviant, represent “extreme” social mobility (i.e., they experience the strongest incongruence between family background and future professional career). We see “downward” social mobility for the students in the Vocational track with upper class parents and “upward” social mobility for the students in the General track with working class parents. This type of extreme social mobility appears to be highly infrequent, which seems to confirm the status congruence theory.

Social mobility might affect the teenagers’ language use, making it more dynamic and open to change. While Aitchison (2013) states that lower middle class and upper working class people (i.e., people on the ‘boundaries’ between different social groups) often act as the trendsetters of linguistic change, Labov (1966) already found that the unclear and insecure position of the lower middle class and its aspirations for upward social mobility favor sociolinguistic hypercorrect behavior (see also Labov, 2006). Thus, the dynamic social position of these teenagers might explain the less predictable patterns of non-standard writing practices in their data.

4.3. Interactions between Social Class, Age, and Gender

We focused on how the social class parameters interact, but we did not yet discuss possible interactions between adolescents’ social class and other aspects of their socio-demographic profile, such as their age and gender. These interactions will be examined in this section.

For all linguistic analyses described in this paper, additional ‘weighted’ tests were carried out to correct for possible age and gender imbalances, since both age and gender have proven to impact adolescents’ online writing (e.g., Baron, 2004; De Decker & Vandekerckhove, 2017; Hilde, Vandekerckhove, & Daelemans, 2017; Hilde et al., in press a;

Schwartz et al., 2013; Verheijen, 2015). Moreover, the analyses of the CMC-data for the present case study reveal that age and gender actually interact with social class. In other words, social class does not have the same impact on the online writing practices of boys versus girls, or on those of younger adolescents (aged 13-16) versus older adolescents/young adults (aged 17-20).

The three-way interaction between gender, age, and social class is visualized in Figures 6a, 6b, and 6c. Each figure shows the ‘age*gender’-interaction for one of the three social groups (upper class, middle class, and working class youngsters). In all three plots, the relative number of non-standard features is shown on the y-axis (i.e., the absolute number of features divided by all tokens). The two age categories are shown on the x-axis, and the gender groups are represented by the orange solid lines (girls) and blue dashed lines (boys). Strikingly, different ‘age*gender’-patterns emerge depending on the adolescents’ social class.

For upper class teenagers, a clear interaction can be observed (see the cross pattern in Figure 6a). Age has a different effect on the language use of upper class girls versus boys. Whereas boys tend to use marginally more non-standard markers as they grow older, girls do not, on the contrary: non-standard features decrease as they age. In related research, girls were found to converge more strongly toward the adult standard as they grow older than boys (see Eisikovits, 2006 for adolescents with a working class family background). Eisikovits (2006) ascribed these distinct age patterns to a difference between (working class) boys’ and girls’ attitudes toward society when they graduate from high school; while accepting the responsibilities of adulthood, girls converge toward mainstream societal norms, whereas boys insist on their autonomy more strongly.

Interestingly, and contrary to Eisikovits’ (2006) findings for working class teenagers, we can only find this pattern for the upper class participants. However, the study of Eisikovits (2006) is not perfectly comparable to ours, since she studied spoken language and focused on ‘old vernacular’. For middle class adolescents (Figure 6b), no real interaction seems to emerge between age and gender. Although the figure suggests a marginal

increase for boys and a marginal decrease for girls, the difference between both gender groups essentially stagnates as they grow older.

For working class adolescents, however, Figure 6c does reveal an interaction, but the pattern strongly deviates from that of the upper class group. While girls more or less stagnate, boys clearly use more non-standard markers as they grow older. Strikingly, the girls' frequency scores for non-standard markers consistently exceed those of the boys. Once again, it should be noted that this need not be due to a stronger preference of old vernacular, since the non-standard features include a wide range of

typographic expressive markers and girls tend to use these (much) more frequently than boys do (see Baron, 2004; Hilte et al., in press a; Parkins, 2012; Varnhagen et al., 2010).

Finally, the three plots indicate that gender differences are most outspoken (in both age categories) for working class adolescents. In the middle class group, gender and especially age differences are very small, whereas in the upper class group gender differences are small in early adolescence, but increase toward late adolescence. Summarizing, different patterns of age and gender dynamics emerge depending on the adolescents' social background.

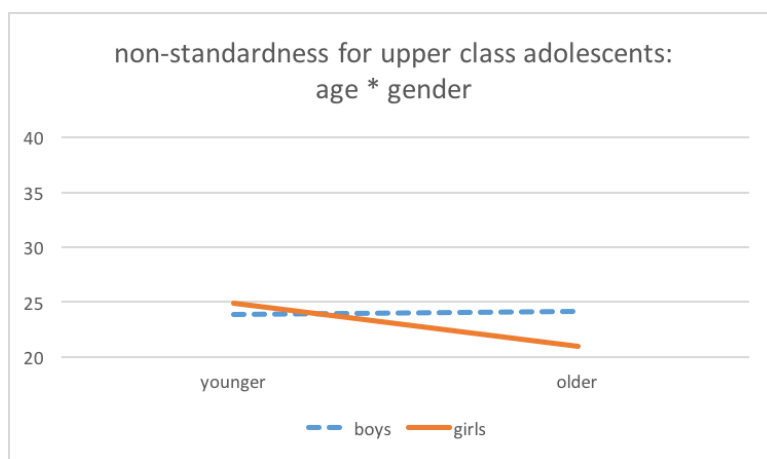


Figure 6a

*'Age*Gender'-Interaction for Upper Class Teenagers*

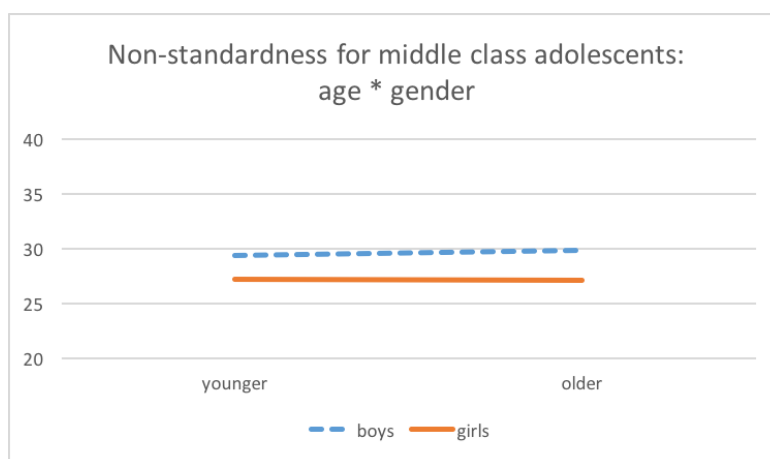
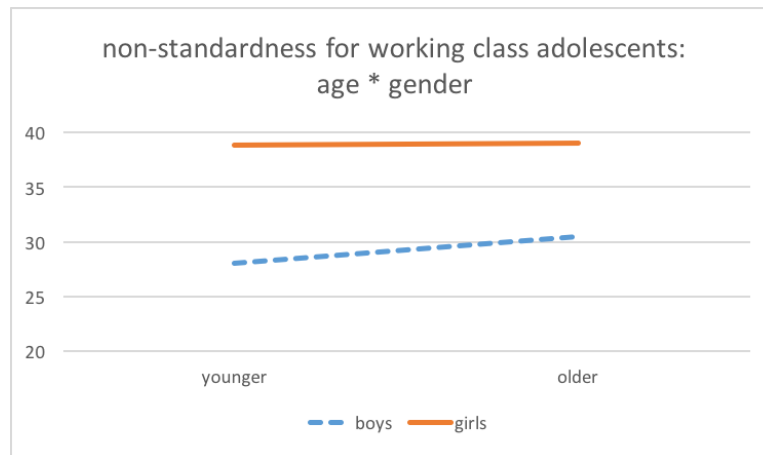


Figure 6b

*'Age*Gender'-Interaction for Middle Class Teenagers*

**Figure 6c**

*'Age*Gender'-Interaction for Working Class Teenagers*

5. Discussion

The present study was devoted to the impact of Flemish adolescents' social class on their informal online writing practices. More specifically, it focused on the occurrence of both old (i.e., traditional regional and slang) and new (i.e., bound to the digital writing culture) vernacular features which generally are no part of formal standard writing and therefore were clustered into a general non-standardness index. The adolescents' social class was operationalized in terms of educational track, home language, and parental profession.

While each of these variables had a significant impact on non-standard writing practices, it was demonstrated that they were correlated rather than independent. For educational track and parents professions, this correlation corroborates previous sociological findings. Therefore, these two factors were clustered so as to create more prototypical social class groups: a working class, middle class, and upper class group. This "clustered" approach revealed more distinct sociolinguistic patterns. Especially upper class and working class youngsters appeared to diverge strongly, with the working class youngsters using much more non-standard markers. The language use of middle class youngsters held an intermediate position when all non-standard features were clustered, but showed a more varied pattern for the individual non-standard markers.

While the distinct online linguistic behavior of the upper versus working class adolescents may

at first sight seem to corroborate classic sociolinguistic findings, the distinction between old and new vernacular features actually changes the perspective to some extent. Ever since Labov (1972), working class people, and especially working class men, have been found to be attracted to the toughness of vernacular speech. The same holds for youngsters (see e.g., Eisikovits, 2006; Trudgill, 1983, and many more). The informal CMC-context offers adolescents a medium for the integration of oral vernacular features in writing and apparently they eagerly exploit this opportunity. In view of previous findings, it is hardly surprising that we attest significantly more of this old vernacular in the CMC-data of working class youngsters. However, they also score much higher for new vernacular features, e.g., they use much more typographic expressive markers that are typical of informal CMC. Thus, these working class adolescents strongly connect to the digital culture too and demonstrate a high chat linguistic dexterity (see Deumert & Lexander, 2013). In other words, by including several sets of features in the category of non-standard markers, it could be demonstrated that working class youngsters certainly do not exclusively exploit classic ways of divergence from standard language norms.

For all groups, the oral vernacular features and the expressive markers largely outnumbered the brevity-related features that are also typical of informal CMC. Interestingly, no social correlations could be found for the latter. This confirms that these features have become stable markers of the genre. Said features are so

functional for all social groups that hardly any social variation emerges (see De Decker & Vandekerckhove, 2017).

An unwanted side-effect of clustering social variables was that more hybrid social groups could no longer be incorporated in the research design. Therefore, we examined the language use of adolescents with non-prototypical social profiles. The, in some respects, “deviant” linguistic behavior of certain groups suggested that more subtle social factors such as aspirations toward social mobility should definitely be included in the operationalization of class in future research on adolescents’ (online or offline) linguistic practices. Furthermore, the operationalization of social class also benefits from including age and gender information, as social class background appeared to interact with both gender and age: different age and gender dynamics were found depending on the youngsters’ social background.

To our knowledge, social class has not been operationalized systematically in variationist sociolinguistic research on youngsters’ informal CMC – and neither have the different aspects of class included in this study. The present paper illustrates both the relevance of the social class variable for this type of CMC-research and the challenges related to the operationalization of such a complex and multidimensional concept which includes several aspects of people’s socio-demographic profile and even of their personality, if we take into account social ambition.

Acknowledgments

This work was supported by the FWO (Research Foundation Flanders) under grant G041115N. We thank the anonymous reviewers for their pertinent feedback.

References

- Aitchison, J. (2013). *Language change: Progress or decay?* Cambridge: Cambridge University Press.
- Androutopoulos, J. (2011). Language change and digital media: A review of conceptions and evidence. In T. Kristiansen & N. Coupland (Eds.), *Standard languages and language standards in a changing Europe* (pp. 145-161). Oslo: Novus Press.
- Argamon, S., Koppel, M., Pennebaker, J. W., & Schler, J. (2009). Automatically profiling the author of an anonymous text. *Communications of the ACM. Inspiring Women in Computing*, 52(2), 119-123.
- Baron, N. S. (2004). See you online: Gender issues in college student use of instant messaging. *Journal of Language and Social Psychology*, 23(4), 397-423.
- De Decker, B., & Vandekerckhove, R. (2017). Global features of online communication in local Flemish: Social and medium-related determinants. *Folia Linguistica*, 51(1), 253-281.
- de Jager, H., Mok, A. L., & Sipkema, G. (2009). *Grondbeginselen der sociologie* [Fundamentals of sociology]. Groningen/Houten: Noordhoff.
- Deumert, A., & Lexander, K. V. (2013). Texting Africa: Writing as performance. *Journal of Sociolinguistics*, 17(4), 522-546.
- Eisikovits, E. (2006). Girl-talk/boy-talk: Sex differences in adolescent speech. In J. Coates (Ed.), *Language and gender: A reader* (pp. 42-54). Oxford: Blackwell.
- Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *The British Journal of Sociology*, 30(4), 415-441.
- [FMET] Flemish ministry of education and training/Vlaams ministerie van onderwijs en vorming. (2017). *Statistisch jaarboek van het Vlaams onderwijs. Schooljaar 2015-2016*. Brussels: Departement onderwijs en vorming.
- Heemskerk, I., Brink, A., Volman, M., & Ten Dam, G. (2005). Inclusiveness and ICT in education: A focus on gender, ethnicity and social class. *Journal of Computer Assisted Learning*, 21(1), 1-16.
- Hilte, L., Vandekerckhove, R., & Daelemans, W. (2017, October). *Modelling non-standard language use in adolescents’ CMC: The impact and interaction of age, gender and education*. Paper presented at the 5th Conference on CMC and Social Media Corpora for the Humanities, Bolzano, Italy.
- Hilte, L., Vandekerckhove, R., & Daelemans,

- W. (in press a). Expressive markers in online teenage talk: A correlational analysis. *Nederlandse Taalkunde*.
- Hilte, L., Vandekerckhove, R., & Daelemans, W. (in press b). Adolescents' social background and non-standard writing in online communication. *Dutch Journal of Applied Linguistics*.
- Labov, W. (1966). Hypercorrection by the lower middle class as a factor in linguistic change. In W. Bright (Ed.), *Sociolinguistics* (pp. 84-113). The Hague: Mouton.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.
- Labov, W. (2006). *The social stratification of English in New York City*. New York, NY: Cambridge University Press.
- Maconis, J. J. (2011). *Society: The basics*. Upper Saddle River, NJ: Pearson Education.
- Marsh, I. (Ed.). (2000). *Sociology: Making sense of society*. Harlow: Prentice Hall.
- Parkins, R. (2012). Gender and emotional expressiveness: An analysis of prosodic features in emotional expression. *Griffith Working Papers in Pragmatics and Intercultural Communication*, 5(1), 46-54.
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS One*, 8(9), e73791.
- Tagliamonte, S. A., & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen language. *American Speech*, 83(1), 3-34.
- Tannen, D. (1990). *You just don't understand. Women and men in conversation*. New York: Ballantine Books.
- Trudgill, P. (1983). *On dialect. Social and geographical perspectives*. Oxford: Blackwell.
- Vandekerckhove, R., & Sandra, D. (2016). De potentiële impact van informele online communicatie op de spellingpraktijk van Vlaamse tieners in schoolcontext. *Tijdschrift voor Taalbeheersing*, 38(3), 201-234.
- Varnhagen, C. K., McFall, G. P., Pugh, N., Routledge, L., Sumida-MacDonald, H., & Kwong, T. E. (2010). LOL: New language and spelling in instant messaging. *Reading and Writing*, 23(6), 719-733.
- Verheijen, L. (2015, September). *Out-of-the-ordinary orthography: The use of textisms in Dutch youngsters' written computer-mediated communication*. Paper presented at the Second Postgraduate and Academic Researchers in Linguistics at York (PARLAY 2014) Conference, University of York, England.
- Vranken, J., Van Hootegeem, G., Henderickx, E., & Vanmarcke, L. (2017). *Het speelveld, de spelregels en de spelers? Handboek sociologie*. Leuven/Den Haag: Acco.
- Warner, C. N. (2004). It's just a game, right? Types of play in foreign language CMC. *Language Learning & Technology*, 8(2), 69-87.
- Wolf, A. (2000). Emotional expression online: Gender differences in emoticon use. *Cyber Psychology & Behavior*, 3(5), 827-833.
- Yates, S. J. (2001). Gender, language and CMC for education. *Learning and Instruction*, 11(1), 21-34.
- Zummo, M. L. (2018). The effect of CMC in business emails in lingua franca: Discourse features and misunderstandings. *International Journal of Society, Culture & Language (IJSCL)*, 6(1), 47-59.