The Impact of WhatsApp on Dutch Youths' School Writing

Lieke Verheijen, Wilbert Spooren
Radboud University (Nijmegen, the Netherlands)
Email: lieke.verheijen@let.ru.nl; w.spooren@let.ru.nl

Abstract

Today's youths are continuously engaged with social media. The informal language they use in computer-mediated communication (CMC) often deviates from spelling and grammar rules of the standard language. Therefore, parents and teachers fear that social media have a negative impact on youths' literacy skills. This paper examines whether such worries are justifiable. An experimental study was conducted with 500 Dutch youths of different educational levels and age groups, to find out if social media affect their productive or perceptive writing skills. We measured whether chatting via WhatsApp directly impacts the writing quality of Dutch youths' narratives or their ability to detect 'spelling errors' (deviations from Standard Dutch) in grammaticality judgement tasks. The use of WhatsApp turned out to have no short-term effects on participants' performances on either of the writing tasks. Thus, the present study gives no cause for great concern about any impact of WhatsApp on youths' school writing.

Keywords: computer-mediated communication, social media, WhatsApp, writing, literacy

1. Introduction

Youths are nowadays constantly using computer-mediated communication such as WhatsApp, Facebook chat, Snapchat, and Twitter. Examples (1)–(3) present chat messages by Dutch youths:

(1) **OMG!** Had je mijn mijn verhaal gezien
Hahahahhaahhaaha kwam ik pas vanochtend achter
k kan me **nie** eens herinneren **da** ik die gemaakt heb
Miss in **mn** slaap **ofzo hagahagagaa**
('OMG! Did you see my my story
Hahahahhaahhaaha only found out this morning
i cant even remember making dat
Mayb in me sleep or somethin hagahagagaa')

(2) Beetje te vroeg ik val echt in slaap maar alvast **happy birthdayyyyyy toooooooo youuuuuuuuuuu**! 💓💓👯👯🕺🕺❤❤🎉🎉 **loveyouuuuuu xxxxxxxxxx**
('Bit too early I'm really falling asleep but anyway…')

(3) **Liefie**❤ gaat ie weer met jou? Wat het je😷😟 bel me **weneer** je online bent😎✅ **ly**❤❤❤❤ zie je morgen😊 **BEL ME** 🙂 **chatt**😛😚
('Luv❤ you doin okay again? What hare you😷😟 call me whn you are online😎✅ ly❤❤❤❤ see you tomorrow😊 CALL ME 🙂 honeyy😛😚')

All the words in bold deviate from the Dutch standard language norms. They contain non-standard abbreviations, letter repetitions, phonetic respellings, overuse of capitalisation, and emoji. Such deviations from Standard Dutch lead to fears that informal written CMC may negatively interfere with writing in more formal settings (Spooren, 2009). These fears have existed for decades now;[1] it is just the medium under critique that changes every few years. Yet there are also scholars who point out that youths' literacy skills may benefit from social media use, via creativity with language, greater exposure to written texts, more engagement in writing, and greater metalinguistic awareness (Wood, Kemp, & Plester, 2013).

[1] Or, for that matter, for centuries. See Deutscher (2005) for a historical note on the concerns about language deterioration.

2. Research Goals and Hypotheses

Since prior research does not provide a conclusive answer about the impact of CMC on literacy (Verheijen, 2013), our research aims to contribute to this debate. The goal of this study is to determine whether Dutch youngsters' constant use of social media affects the way they write at school. We conducted an experiment to find out whether engaging in synchronous written CMC directly impacts their productive or perceptive school writing skills. Our focus was on the chat program WhatsApp, because this is currently a very popular medium among Dutch youths.

Furthermore, we aim to establish whether the demographic variables of age, education, and gender have an impact on youths' writing skills – specifically, if they have a mediating effect on the possible impact of WhatsApp on those writing skills. Therefore, the analysis will include four independent variables: not only condition (exposure vs. non-exposure to WhatsApp), but also educational level (lower, intermediate, higher), age group (adolescents vs. young adults), and gender (boys vs. girls). We hypothesize that a greater impact of CMC use on school writing skills may be displayed by youths of a younger age group or lower educational level, as it may be more difficult for them to keep these registers separate.

3. Methodology

3.1 Participants

Participants were 500 youths from secondary and tertiary educational institutions, from different educational levels and age groups, in Nijmegen and surroundings. The data collection period lasted from October to December 2016. Testing took place in an educational setting. Most participants were tested in class; only a small number in higher tertiary education voluntarily took part outside of class. The latter were reimbursed for their participation with gift certificates of € 5. Students from higher tertiary education belonged to different faculties and studies, including communication and information sciences, biology, and literary and cultural studies. The adolescents (N = 300) were around 14 years old ($\bar{x}$ age = 14.2 yrs, range 13-16; 151 male, 149 female), all in the third grade. The young adults (N = 200) were around 20 years old ($\bar{x}$ age = 20.4 yrs, range 18-27; 72 male, 128 female). Table 1 shows an overview of the participants. Afterwards, underage participants were given a document with more information about the study and the researchers' contact details, to take home to their parents or caretakers.

		Educational level		
		lower	intermediate	higher
Age group	adolescents: secondary education	101	92	107
	young adults: tertiary education	102	-	98

Table 1: Overview of participants.

3.2 Data Collection

3.2.1. Priming: WhatsApp vs. Colouring

All classes that were tested were divided into two groups. The experimental groups were primed with CMC via social media: they were instructed to chat via WhatsApp on their own smartphones for fifteen minutes, in small groups of three or four students. They could chat about whatever they preferred; no specific conversation topics were provided, in order to generate as natural chat conversations as possible. During that time, the control groups performed a non-CMC-related control task, namely colouring mandalas. These tasks were chosen because, in a pilot study, they proved to be effective in revealing differences with respect to orthography and language correctness (Riemens, 2016).

3.2.2. Measuring Productive Writing Skills: Stories

To test their productive writing skills, all participants wrote a story in class, starting with the following sentence: "I was alone in a dark room. My hand groped for the light switch, but suddenly…" [translated from Dutch]. The formal writing genre that was tested was that of narrative storytelling. Since not all classes had easy access to computers and laptops, all stories were hand-written for consistency's sake.

3.2.3. Measuring Perceptive Writing Skills: GJTs

Participants also completed grammaticality judgement tasks (GJTs), to test their receptive grammar and spelling skills. These consisted of twenty sentences in which they had to spot and correct 'language errors'. These were orthographic deviations typical of CMC: various types of textisms (phonetic respelling, reduplication of letter, shortening, single letter homophone, initialism); missing capitalisation, diacritics, and punctuation; spelling 'errors' that are heavily frowned upon by Dutch language prescriptivists (*is/eens*, *d/t*, *jou/jouw*); emoticons; omissions; English borrowings; and extra spacing. Five sentences contained no orthographic deviations, so participants could spot and correct fifteen 'errors'.

3.3 Data Analysis

3.3.1. GJT Scores

For the grammaticality judgement tasks, two scores were computed for each participant. First, the choice score: whether they correctly identified the sentence as containing an 'error' or not (max. 20 points). Second, the correction score: whether they correctly managed to correct that 'error' (max. 15).

3.3.2. T-Scan Analysis

The stories were automatically analysed with T-Scan, software for conducting complexity analyses of Dutch texts (Pander Maat et al., 2014). T-Scan provided us with a staggering 411 variables for each text, out of which a theory-based selection of 27 relevant variables was made:

1) Zin_per_doc: number of sentences per essay
2) Word_per_doc: number of words per essay
3) Let_per_wrd: number of letters per word
4) Wrd_per_zin: number of words per sentence
5) Bijzin_per_zin: number of subordinate clauses per sentence
6) Pv_Frog_d: density of finite verbs
7) D_level: D-level
8) Nom_d: density of nominalisations
9) Lijdv_d: density of passive forms
10) AL_gem: average of all dependency lengths per sentence
11) AL_max: maximal dependency length per sentence
12) Bijw_bep_d: density of adverbials
13) TTR_wrd: type-token ratio (for words)
14) MTLD_wrd: measure of textual lexical diversity (for words)
15) Inhwrd_d: density of content words
16) Pers_vnw_d: density of personal and possessive pronouns
17) Ww_mod_d: density of modal verbs
18) Huww_tijd_d: density of auxiliary verbs of time
19) Koppelww_d: density of copula verbs
20) Imp_ellips_d: density of imperatives and elliptical constructions
21) Vg_d: density of conjunctions
22) Lidw_d: density of articles
23) Nw_d: density of nouns
24) Tuss_d: density of interjections
25) Spec_d: density of names and special words
26) Interp_d: density of punctuation
27) Afk_d: density of abbreviations

3.3.3. Exploratory Factor Analysis

Because the twenty-seven variables selected from T-Scan were still too many to put into a regression analysis, we used an exploratory factor analysis (with the extraction method of principal component analysis, PCA), to further reduce these to a set of writing components indicative of the writing quality of stories.

An orthogonal rotation method was chosen, namely varimax with Kaiser normalization: this method, which does not allow correlations between factors, facilitated the interpretation of results, since it maximizes the spread of loadings for a variable across all factors. There was no multicollinearity, because none of the correlation coefficients were $r \geq .84$. Missing values were replaced with the mean, because listwise deletion would result in a loss of participants in the analysis, and pairwise deletion would lead to a non-positive definite matrix. The Keyser-Meyer-Olkin measure was well above .5 (*KMO* = .644), which verified the sampling adequacy for the analysis. Bartlett's test of sphericity showed that correlations between items were sufficiently large for PCA: χ^2 (351) = 6267.569, $p < .001$. The proportion of residuals with an absolute value greater than 0.05 was 50%. An initial analysis yielded eigenvalues for each component in the data. The large sample size of this study (500 participants) allowed us to use a scree plot with eigenvalues over 1 for deciding how many components to extract. The inflexion of the scree plot justified retaining three components.

Table 2 shows the results of the PCA after rotation. The items that cluster on the same components suggest that component 1 represents syntactic complexity, 2 lexical richness, and 3 writing productivity. The total variance explained by the three factors is 38.08%. The resulting factor scores were saved as Anderson-Rubin variables, so they did not correlate.

Rotated Component Matrix			
Writing variable	Rotated factor loadings		
	1	2	3
AL_max	**.868**	.141	.106
D_level	**.818**	-.193	.016
Bijzin_per_zin	**.792**	-.089	.077
AL_gem	**.764**	.221	.232
Wrd_per_zin	**.720**	.004	-.087
Interp_d	**-.718**	-.077	.065
Vg_d	**.556**	-.287	.072
Tuss_d	-.240	-.117	.186
Bijw_bep_d	.221	-.004	.107
Spec_d	-.147	.016	-.057
Pv_Frog_d	-.167	**-.762**	-.037
Nw_d	-.059	**.698**	-.166
Pers_vnw_d	-.102	**-.680**	.066
Let_per_wrd	-.002	**.624**	-.153
Inhwrd_d	-.045	**.554**	.054
Lidw_d	-.047	**.520**	-.232
MTLD_wrd	-.060	**.450**	.004
Nom_d	-.034	**.423**	.028
Koppelww_d	-.127	-.189	.020
Ww_mod_d	.057	-.145	.136
Imp_ellips_d	.039	-.099	.066
Word_per_doc	.023	.004	**.917**
TTR_wrd	-.141	.281	**-.800**
Zin_per_doc	**-.529**	-.038	**.782**
Huww_tijd_d	-.078	-.043	-.378
Lijdv_d	-.038	.028	-.139
Afk_d	-.095	.053	-.095
Eigenvalues	4.496	3.246	2.539
% of variance	16.652	12.021	9.405

Note: loadings > .40 appear in bold and colour.

Table 2: PCA rotated factor loadings for the story analysis.

3.3.4. Linear Multiple Regression

The next step of the statistical analysis was linear multiple regression. The outcome variables were the three A-R factor scores resulting from the exploratory factor analysis of the stories and the two GJT scores. The predictor variables were condition (colouring versus WhatsApp), the three demographic variables educational level, age group, and gender, plus all interactions between condition and the demographic variables. As we had no preconceived ideas about which variables would be significant predictors, they were all entered with the forced entry method. The first block of the regression only contained the main effects. The interactions were entered in subsequent blocks.[2]

4. Results and Discussion

Table 3 shows the means and standard deviations of participants' performances on the writing tasks:

	Dependent variables				
	Stories			GJTs	
Independent variables	syntactic complexity: $\bar{x}$ (*SD*)	lexical richness: $\bar{x}$ (*SD*)	writing productivity: $\bar{x}$ (*SD*)	choice score: $\bar{x}$ (*SD*)	correction score: $\bar{x}$ (*SD*)
Condition:					
Colouring, *N* = 207	-0.02 (1.00)	0.06 (1.04)	0.00 (1.09)	14.44 (3.16)	13.82 (1.03)
WhatsApp, *N* = 201	0.07 (1.05)	0.09 (0.96)	0.03 (0.99)	14.71 (3.04)	13.74 (0.99)
Educational level:					
Lower, *N* = 203	0.19 (1.06)	-0.03 (0.93)	-0.17 (1.04)	12.68 (2.87)	13.53 (0.99)
Higher, *N* = 205	-0.13 (0.96)	0.18 (1.06)	0.20 (1.00)	16.44 (2.00)	14.03 (0.96)
Age group:					
Adolescents, *N* = 208	0.01 (1.14)	-0.24 (0.92)	0.00 (1.08)	14.10 (3.09)	13.75 (1.01)
Young adults, *N* = 200	0.05 (0.88)	0.40 (0.98)	0.03 (1.00)	15.06 (3.04)	13.81 (1.01)
Gender:					
Male, *N* = 179	0.16 (1.16)	0.09 (1.00)	-0.14 (1.11)	14.25 (3.14)	13.60 (1.02)
Female, *N* = 229	-0.08 (0.89)	0.06 (1.0)	0.14 (0.97)	14.82 (3.06)	13.92 (0.98)
TOTAL, *N* = 408	0.03 (1.02)	0.07 (1.00)	0.01 (1.04)	14.57 (3.10)	13.78 (1.01)

Table 3: Descriptive statistics.

4.1 Syntactic Complexity

One writing component was syntactic complexity, presented in Table 2 in column '1'. Educational level was a significant negative predictor: higher educated youths wrote syntactically less complex stories. At a first glance, this may seem surprising. However, this rather fits the genre of narrative storytelling, which does not require complex, long sentences – as opposed to, for example, expository discussion as in essays. So the higher educated youths showed more mastery of the genre of stories. Gender was a significant negative predictor: male participants wrote syntactically more complex stories.

[2] Participants of the intermediate secondary educational level were eventually omitted, as they were not part of the original research plan and would decrease the reliability of the analyses because of an empty cell in the research design: no youths of intermediate tertiary education were tested (see Table 1).

Independent variables	B	SE B	β
	Dependent variable: syntactic complexity		
Condition	0.09	0.10	0.04
Educational level	**-0.32**	**0.10**	**-0.16****
Age group	0.06	0.10	0.03
Gender	**-0.24**	**0.10**	**-0.12***
R^2 / Adjusted R^2	.04 / .03		
ANOVA	*F* (4, 403) = 4.24 (*p* < .01)		

Table 4: Regression results for syntactic complexity.[3]

4.2 Lexical Richness

Another writing component was lexical richness. Table 2 shows the variables that loaded onto this component, in the column labelled '2'. Lexical richness was positively predicted by educational level and age group: the stories of higher educated and of older participants were lexically richer. In addition, there was a significant interaction between gender and condition. For boys, WhatsApp had a small significant positive effect on their stories' lexis; for girls, the effect was negative but non-significant.

Independent variables	B	SE B	β
	Dependent variable: lexical richness		
Condition	0.32	0.19	0.16
Educational level	**0.31**	**0.13**	**0.16***
Age group	**0.64**	**0.13**	**0.32*****
Gender	0.11	0.13	0.06
Educational level × condition	-0.14	0.19	-0.06
Age group × condition	0.08	0.19	0.03
Gender × condition	**-0.50**	**0.19**	**-0.22****
R^2 / Adjusted R^2	.13 / .12		
ANOVA	*F* (7, 400) = 8.90 (*p* < .001)		

Table 5: Regression results for lexical richness.

4.3 Writing Productivity

The third component of the stories, writing productivity, is presented in column '3' in Table 2. It was positively predicted by educational level: youths with a higher educational level produced significantly longer stories. Gender was a significant positive predictor of writing productivity too: female participants wrote longer stories.

Independent variables	B	SE B	β
	Dependent variable: writing productivity		
Condition	0.03	0.10	0.01
Educational level	**0.37**	**0.10**	**0.18*****
Age group	-0.01	0.10	0.00
Gender	**0.27**	**0.10**	**0.13****
R^2 / Adjusted R^2	.05 / .04		
ANOVA	*F* (4, 403) = 5.19 (*p* < .001)		

Table 6: Regression results for writing productivity.

4.4 GJT Choice Score

For the grammaticality judgement tasks, educational level and age group were significant positive predictors of the choice score, so higher educated youths and older youths were more successful in spotting 'language errors'.

Independent variables	B	SE B	β
	Dependent variable: GJT choice score		
Condition	0.21	0.24	0.03
Educational level	**3.77**	**0.24**	**0.61*****
Age group	**0.99**	**0.24**	**0.16*****
Gender	0.35	0.24	0.06
R^2 / Adjusted R^2	.40 / .39		
ANOVA	*F* (4, 403) = 67.21 (*p* < .001)		

Table 7: Regression results GJT choice score.

4.5 GJT Correction Score

The correction score was significantly positively predicted by educational level and gender: both higher educated and female participants were more successful in correcting 'language errors'. The interaction between gender and condition was also significant. For girls, WhatsApp had a small significant negative effect on their correction score; for boys, the effect was positive but non-significant.

Independent variables	B	SE B	β
	Dependent variable: GJT correction score		
Condition	-0.09	0.19	-0.05
Educational level	**0.39**	**0.13**	**0.20****
Age group	-0.11	0.14	-0.06
Gender	**0.51**	**0.14**	**0.25*****
Educational level × condition	0.22	0.19	0.10
Age group × condition	0.31	0.19	0.13
Gender × condition	**-0.44**	**0.20**	**-0.20***
R^2 / Adjusted R^2	.11 / .09		
ANOVA	*F* (7, 400) = 6.71 (*p* < .001)		

Table 8: Regression results for GJT correction score.

An overview of the results of all the linear multiple regressions is presented in Table 9 below:

	Dependent variables				
	Stories			GJTs	
Independent variables	syntactic complexity	lexical richness	writing productivity	choice score	correction score
Main variables:					
Condition					
Educational level	−	+	+	+	+
Age group		+		+	
Gender	−		+		+
Interactions:					
EL × C					
AG × C					
G × C		−			−
EL × AG × C					
EL × G × C					
AG × G × C					
EL × AG × G × C					

Note: + = positive predictor, − = negative predictor, C = condition, EL = educational level, AG = age group, G = gender.

Table 9: Overview of regression results.

[3] For all tables with results, *p < .05, **p < .01, ***p < .001.

5. Conclusion

This paper reports on an experimental study measuring whether the use of WhatsApp has a direct impact on the writing quality of Dutch youths' stories or on their ability to detect 'spelling errors' in grammaticality judgement tasks. Educational level was a significant positive predictor for four writing variables, and age group for two. Gender predicted three writing variables. Condition did not affect the writing variables. We can thus conclude that WhatsApp does not appear to impact Dutch youths' productive or perceptive writing skills. Only two minor interactions between condition and gender were found, which suggests that perhaps there might be a slight impact of WhatsApp, moderated by gender, in which boys' lexical richness might benefit from CMC and girls' ability to correct language errors might be affected by it.

Two objections to this conclusion may be raised. One might doubt whether our measuring instrument was sensitive enough to detect differences in writing quality. However, the effectiveness of our testing method is confirmed by finding main effects for three demographic variables: these show that analysing the stories with T-Scan, as well as the GJT scores, are successful ways to detect differences in youths' writing skills. One might also argue that our experimental manipulation, the use of WhatsApp for fifteen minutes, was not strong enough to generate any effects. That cannot be the case either, because we found some interactions with gender; moreover, the prime already proved to yield significant results in a pilot study conducted in advance. All in all, the present study gives no cause for concern about the impact of WhatsApp on school writing.

6. Future Work

We hypothesized that particularly writers of a younger age group and lower educational level could experience possible interference of social media on their school writings. Prior research also suggests that youths of a lower educational track have more trouble distinguishing informal online writing (CMC) from more formal offline writing repertoires (Vandekerckhove & Sandra, 2016). Further research could explore other ways to test for such interference. Perhaps effects of social media crop up in minor orthographic details of their school writings, such as non-standard punctuation, capitalisation, spacing, or diacritics, because in the pilot study, these were the items on which WhatsApp use had the greatest impact. The frequent omission of punctuation and capitalization (sentence-initial or with proper names) in school writings was also noted by Vandekerckhove and Sandra (2016). The stories written for the present experiment could thus be analysed for the occurrence of such non-standard orthographic details.

In addition, the WhatsApp chats produced by roughly half of the participants during the priming phase were nearly all collected afterwards, of course with their consent (sent to the first author via email), but were not analysed. If properly formatted and annotated, the CMC data thus compiled could be a valuable corpus for further analysis. We could study the nature of these WhatsApp interactions, e.g. for the use of textisms, to find out to which extent these chats actually differ from Standard Dutch in terms of orthography and grammar and whether the amount of deviations affected the direct impact of CMC use on the writing tasks.

7. Acknowledgements

This study is part of a research project funded by the Dutch Organisation for Scientific Research (NWO), project number 322-70-006.

8. References

Deutscher, G. (2005). *The Unfolding of Language*. London: Random House.

Pander Maat, H., et al. (2014). T-Scan: A new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, pp. 53--74.

Riemens, N. (2016). De directe invloed van WhatsApp op schrijfvaardigheid. BA thesis, Radboud University.

Spooren, W. (2009). Bezorgde ouders? De relatie tussen chat en schrijfkwaliteit. In W. Spooren, M. Onrust, & J. Sanders (Eds.), *Studies in Taalbeheersing* 3. Assen: Van Gorcum, pp. 331--342.

Vandekerckhove, R., and Sandra, D. (2016). De potentiële impact van informele online communicatie op de spellingpraktijk van Vlaamse tieners in schoolcontext. *Tijdschrift voor Taalbeheersing*, 38(3), pp. 201--234.

Verheijen, L. (2013). The effects of text messaging and instant messaging on literacy. *English Studies*, 94(5), pp. 582--602.

Wood, C., Kemp, N. and Plester, B. (2013). *Text Messaging and Literacy: The Evidence*. New York, NY: Routledge.