

Disguise and imitation of language style in WhatsApp messages

Meike M. de Boer ¹, Susan A. Blackwell², and Fleur van der Houwen³

¹Vrije Universiteit Amsterdam, m.m.de.boer2@vu.nl

²Utrecht University, s.a.blackwell@uu.nl

³Vrije Universiteit Amsterdam, f.vander.houwen@vu.nl

Abstract

The deliberate manipulation of evidence is routinely considered in most branches of forensic science, but in the field of forensic authorship analysis it is a relatively new concern. This paper reports the findings from an experiment which explored what happens when authors are asked to manipulate their language style. We asked 120 Dutch students to write WhatsApp messages in their normal style and in two manipulated conditions: disguise and imitation. Results show that authors are able to introduce major changes in their language use, which are most substantial when they have access to messages written by the person they are trying to impersonate. These findings imply that the possibility of disguise and imitation should be considered when conducting forensic authorship analysis, especially when more than one candidate for authorship could have had access to the same device or account.

Keywords: Authorship analysis, WhatsApp messages, forensic stylistics, disguise, imitation.

Resumo

A manipulação deliberada de provas é uma prática corrente na maioria dos ramos das ciências forenses, mas no domínio da análise de autoria é uma preocupação relativamente recente. Este artigo apresenta os resultados de uma experiência que explorou o que acontece quando se pede aos autores que manipulem o seu estilo linguístico. Pedimos a 120 estudantes holandeses que escrevessem mensagens no WhatsApp no seu estilo normal e em duas condições manipuladas: disfarce e imitação. Os resultados mostram que os autores são capazes de introduzir grandes mudanças no seu uso da linguagem, que são mais proeminentes quando têm acesso a mensagens escritas pela pessoa que estão a tentar imitar. Estes resultados implicam que a possibilidade de disfarce e imitação deve ser considerada aquando da realização de análises forenses de autoria, especialmente quando mais do que um candidato à autoria pode ter tido acesso ao mesmo dispositivo ou conta.

Palavras-chave: Análise de autoria, mensagens do WhatsApp, estilística forense, disfarce, imitação.

1. Introduction

Trying to avoid detection is inherent in criminal behaviour: manipulation of the evidence is therefore a reasonable possibility in virtually any forensic investigation (Ekblom, 1999). Within the context of linguistic evidence, manipulation has so far been mostly considered in the analysis of speech (voice disguise, see e.g. Cambier-Langeveld, 2016; Eriksson, 2005) and threat assessment (Simons & Tunkel, 2021). In the context of authorship analysis, although the possibility of deception has been acknowledged in the literature (Dern, 2008; Marko, 2017), it is mentioned only briefly, if at all, in discussions of cases or procedures. Over the past few years, however, the present authors have been consulted in authorship cases where manipulation could have played a role. Messages had been sent from an alleged murder victim's phone after their disappearance, and the police wanted to know whether the disappeared or deceased persons had sent the messages themselves or whether someone else might have written them¹. According to Grant (2013), in such cases it is possible to indicate the most likely author by comparing the disputed message(s) to the language style of both candidate authors. We were provided with messages written by the victim together with messages written by one or more suspects and compared these to the disputed message(s). In some cases we found no linguistic evidence that the suspect had written the message(s), which were consistent with the language style of the victim. However, we had to insert the caveat that if anyone had imitated the victim's language style, our conclusions would be invalid.

The current study investigates how lay people change their language style in a WhatsApp message when they are asked to disguise their own language style or to imitate that of another person. In order to investigate this, we carried out an experiment in which participants acted as their own controls, producing WhatsApp texts firstly in their own personal style and subsequently under two manipulated conditions (i.e. disguise and imitation).

2. Authorship analysis

For short texts such as WhatsApp messages, a stylistic or qualitative approach is preferred over a stylometric or quantitative approach (Grant, 2013). While stylistic approaches to authorship analysis have been criticized for being subjective or prone to 'cherry picking', since the selection of features tends to be dependent on the analyst (for a discussion see Solan, 2013), the number of WhatsApp messages in forensic cases is often insufficient to adopt a stylometric approach (see e.g. Brennan & Greenstadt, 2009; Juola, 2012). The stylistic approach focuses on the choices made by the author, either deliberately or unconsciously. These have been termed 'style-markers' (McMenamin, 2002: §11.4) or 'possible textual dualities' (Olsson, 2009, p.57) and are features that can be written in more than one way, e.g. abbreviations, symbols, numbers, and (mis)spellings. Together, these are said to form an author's idiolect (Coulthard & Johnson, 2007).

Grant (2013) tried to adopt a more systematic approach, focusing specifically on mobile phone text messages. His method is data-driven, based on a word list contain-

¹For the verdicts of the court and appellate court in one of these cases, see: *Rechtbank Midden-Nederland* (15 December 2015). ECLI:NL:RBMNE:2015:8972
Gerechtshof Arnhem-Leeuwarden (24 February 2017). ECLI:NL:GHARL:2017:1544.

ing all word forms present in the messages of two candidate authors. Word forms are selected as a feature for the analysis if they show a sufficient degree of *consistency* and *distinctiveness*. In order to be consistent, a feature must be present in the messages at least ten times. To be distinctive, the feature must be twice as prevalent in the messages of one candidate author than in the messages of the other. These criteria make it possible to work with relevant features – the word forms that are selected for the analysis are similar to the style-markers that are found without a word list – while objectivity is preserved: the selection does not depend on the analyst and would be the same in replications. However, the approach still requires a decent amount of text per author in order to find enough occurrences of linguistic features, which is often unavailable in forensic casework.

The comparison of style-markers is based on the idea that the choices authors make in their texts are distinctive from other authors' choices. However, it is not possible to distinguish one author from all other possible authors in the world. Grant (2013) instead proposes the notion of 'pairwise distinctiveness', which holds that it is possible to identify the most likely author from a small number of candidates, e.g. the victim and one or more suspects. This, in addition to other evidence, can help the police or judge to discriminate between scenarios (Coulthard & Johnson, 2007). This means that the notion of the 'linguistic fingerprint' or 'stylistic genome', as idiosyncrasy in language has been dubbed (see e.g. Bernhardsson, Correa da Rocha, & Minnhagen, 2009; Luyckx, Daelemans, & Vanhoutte, 2006), is flawed and misleading (cf. Coulthard, 2014).

The implicit assumption of authorship analysis is that the variability of texts written by the same author (within- or *intra*-author variability) is less than the variability of texts written by different authors (between- or *inter*-author variability). There are, however, factors that increase within-author variability: language is continually adapted to fit the circumstances (Nolan, 2001). For example, authors adapt their language style depending on the recipient, topic, context, medium, and genre of the text (e.g. Grant, 2013; Olsson, 2008, 2009; Shapero & Blackwell, 2012). One way in which within-author variability could be artificially increased is by changing features to disguise one's language style (Olsson, 2008), which could lead the analyst to conclude incorrectly that the text is not written by the same author as the reference texts. Authors might also be capable of consciously decreasing between-author variability by imitating someone else's language style (De Vel, Anderson, Corney, & Mohay, 2001), which could lead the analyst to conclude that the text is written by the imitated person instead of the imitator.

3. Disguise and imitation of language style

Scholars within the field of forensic phonetics have long been considering disguise (e.g. Amino, Makinae, & Kamada, 2018; Künzel, 2000; Masthoff, 1996; Perrot, Preteux, Vasseur, & Chollet, 2007). While Eriksson (2005, p.10) describes most voice disguise techniques as rather unsophisticated, he also acknowledges that they can have a 'considerable detrimental effect on speaker identification'. Disguise is estimated to play a role in 15 (Künzel, 2000) to 40 (Clark & Foulkes, 2007) percent of forensic speech investigations, and in the view of Schilling and Marsters (2015) should always be considered as a possibility.

The same would appear to hold true for manipulation of language style in written texts. The literature discussing this issue is very scarce, especially from a stylistic viewpoint (but see Dern, 2008; Marko, 2017). Moreover, each publication seems to create its own terminology to describe the phenomena (e.g. disguise, circumvention, obfuscation, masking, imitation, mimicking), showing a lack of uniformity in the scientific debate. In the current paper, we use the terms disguise and imitation to refer to the two experimental conditions, which are approaches in which an author can attempt to manipulate their language style to deceive the addressee and/or the police – thereby trying to avoid apprehension (cf. Simons and Tunkel 2021, p. 245). The remainder of this section describes some of the few works available on this topic.

Most of the work on disguise and imitation of language style has been done in the stylometric tradition. Computational systems perform at around chance level when attributing disguised texts to authors, and attribute imitated texts mostly to the person who is imitated (see e.g. Brennan, Afroz, & Greenstadt, 2012; Brennan & Greenstadt, 2009). However, although authors might not be individually identified when they manipulate their language style, the deceptive intent may be recognized as a separate style by stylometric systems (Juola, 2012).

Dern (2008, pp. 243-244), discussing manipulation from a stylistic viewpoint, distinguishes four forms, which she based on “common-sense”: 1) arbitrary changes, without a particular pattern or apparent strategy; 2) ‘dumbing down’, i.e., feigning a lower language proficiency, level of education, or social status; 3) feigning non-nativeness; and 4) imitation of the language style of a specific person. According to Dern, arbitrary changes are the easiest way to disguise one’s language style since it does not require much consistency and is not used to evoke a certain image in the reader. Imitation, on the other hand, is described as the most difficult for authors to perform accurately, since it requires a high level of metalinguistic awareness both of one’s own language style and of the style of the person one is imitating. This is corroborated by MacLeod and Grant (2017), who show that for police officers posing as children in undercover operations, imitation is difficult – even after extensive training. A further study by Grant and MacLeod (2020) found that such undercover officers have more success with imitation at the level of lexis, spelling and punctuation than with the pragmatic layer of online messages.

Most of the research so far has focussed on Dern (2008)’s third category, feigned non-nativeness, prompting the participants to adopt a specific disguise strategy. Overall, although authors may be successful in disguising their own language style using this strategy, the act of manipulation has been described as detectable. According to Fobbe (2014), authors who are asked to simulate a non-native speaker identity introduce similar patterns of morphological and syntactic errors, suggesting a recognizable language style in deliberate error production. Dern (2008) states that feigned non-nativeness can be distinguished from genuine low language proficiency, since authors unconsciously show their language proficiency and native skills by their correct use of idioms and choice of words. Moreover, authors who consciously introduce errors into their texts tend to be inconsistent in the type and number of errors they use. Thus, a high occurrence of errors does not necessarily mean that the author cannot be a native speaker, and rather than focussing on individual errors, authorship analysts should examine the pattern of errors throughout the text. Marko (2017) found that whereas people expect

others to use a simulated non-native language style as a disguise strategy, her German and English-speaking participants rarely did so, instead making arbitrary changes such as deliberate spelling errors, changes in register, and changes in the purported identity of the author, such as using 'we' instead of 'I'. Indeed, when not prompted to feign a non-native identity, the changes made by Dern (2008)'s participants were superficial, mostly consisting of random deliberate spelling errors.

A recurring remark in the literature is that there is a lack of knowledge about the manipulation strategies that are used by authors in real life (e.g. Bredthauer, 2013; Dern, 2008; Marko, 2017; Seifert, 2010). Therefore, Bredthauer (2013) studied such strategies in forensic texts investigated by the Bundeskriminalamt (BKA, the German Federal Police Office), which led to a few findings regarding manipulated language styles. First, she found that manipulated texts are typically shorter than authentic texts, possibly because of the higher mental concentration required. Second, 76 per cent of the manipulated texts showed parts written entirely in capitals or entirely without capitals: Bredthauer believed this could be used by the authors to hide their true language proficiency². Thirdly, only 23 per cent of the texts contained punctuation that was 'predominantly correct', whereas 36 per cent of the authors used punctuation in a way that was 'totally incorrect', e.g. by not using any commas. Although the many punctuation errors may indicate that they are easy to manipulate due to high metalinguistic awareness, Bredthauer also suggests that they could equally be a feature of the authors' genuine language styles. Finally, errors in orthography (spelling), morphology, syntax, and lexis were inconsistent, and no pragmatic errors were observed. Overall, Bredthauer (2013) concluded that all manipulation strategies used in the forensic texts have the potential to be deceptive, although the manipulation may be indicated by inconsistencies. It is especially difficult to manipulate several different linguistic parameters simultaneously in a consistent manner over time (Kniffka, 1993): authors tend to revert to their usual style, especially at the end of longer texts (Bredthauer, 2013).

This paper attempts to supplement the meagre research on the manipulation of language style. It describes an experimental study of Dutch messages sent via WhatsApp, which is currently the most popular social media platform in the Netherlands, being used by 80 per cent of the population above the age of 15 (Van der Veer, Boeke, & Peters, 2017). The aim of this study is to compare participants' own, non-manipulated language styles as well as their manipulated styles, rather than merely describing features of manipulated language. In addition, this is the first study on manipulation of language style in Dutch, or as far as we know in any language other than German or English.

²The possible relationship between correct use of capitalization and higher language proficiency might be typical for German, since in German, all nouns start with a capital letter and other word classes do not.

4. Method³

4.1. Participants

The experiment in this study was conducted among 120 students of an introductory course in forensic linguistics at Vrije Universiteit Amsterdam in September 2017. The participants, who came from a variety of academic disciplines, had just commenced the third year of their bachelor's degree and could be regarded as lay people in the field of authorship analysis. Most participants (88.2%) indicated that they had Dutch as their sole native language; eleven (9.2%) were raised bilingually. The participants used their own smartphones to participate in the experiment, using their own preferred keyboard settings. 55 (46.2%) indicated that they used auto-correct (automatic correction of spelling) and 29 (24.4%) used predictive text (automatic suggestion of the next word). Participants had a free choice as to whether to attend the tutorial in which the experiment was conducted and could drop out of the experiment at any time. This led to the exclusion of one participant. There were six different tutorial groups, which were placed in different experimental conditions.

4.2. Materials and procedure

The experiment consisted of two scenarios with short writing tasks. After signing the consent form, the participants received written instructions for the first scenario: they were told to imagine that they had just changed their weekend plans and were asked to communicate this to their mother and a friend via WhatsApp (for each of which we provided a phone number). Scenario 1 was the same for all groups and functioned as a control task, intended to elicit the participants' non-manipulated language styles. After completing this task, the participants received their instructions for scenario 2, in which they had to imagine that a friend (given the pseudonym "Lieke", which we will use henceforth to designate this fictitious person) had died in dubious circumstances while staying with them. They were instructed to send WhatsApp messages from Lieke's phone to her mother and/or friend to tell them that she was going on a camping trip and to cancel their appointments, making it look as if she was still alive. Scenario 2 functioned as the experimental phase and was intended to elicit WhatsApp messages in which the participants manipulated their language style.

While scenario 2 started in the same way for each group, the instructions differed on two parameters: (1) the overall manipulation strategy and (2) the recipient. Participants were instructed either merely to disguise their own language style, or to imitate the language style of Lieke. Some groups were instructed to send messages in both conditions. Regarding the recipient, participants who were placed in both the disguise and the imitation condition were asked to send messages either to Lieke's mother or to her friend "Romy". Participants who were only placed in one condition (either disguise or imitation) were asked to send messages to both recipients, in alternate orders. This design had the purpose both of restricting the number of messages each participant had to send, and of counterbalancing any priming effects due to the sequencing of the tasks (see e.g. Gaito, 1961). Table 1 provides an overview of the different tasks performed by each of the groups.

³A full account of the methodology, along with a second experiment to determine how convincing the manipulations were, is described in De Boer (2018).

group	<i>First manipulation</i>		<i>Second manipulation</i>	
	Strategy	Recipient	Strategy	Recipient
A	disguise	mother	imitation	mother
B	disguise	friend	imitation	friend
C	imitation	mother	imitation	friend
D	disguise	mother	disguise	friend
E	imitation	friend	imitation	mother
F	disguise	friend	disguise	mother

Table 1. Tasks of the experimental phase per group.

Thus, each of the participants sent four messages in total: two in scenario 1 to provide their genuine language style (one to their mother and one to their friend), and two in scenario 2, under different conditions.

Participants who were asked to imitate Lieke's language style received three pages of a WhatsApp conversation between Lieke and her mother and/or 2.5 pages of a WhatsApp conversation between Lieke and Romy. The WhatsApp conversations of Lieke were genuine texts between a real student and her actual mother and friend, although the names had been changed. The participants had 15 minutes to study the conversation before they had to send a message to Lieke's interlocutor in that conversation. They could take notes and keep the conversation and notes with them while sending the imitated message.

After completion of the two writing tasks, the participants were asked to fill out a survey about their demographic details and the strategies they had used in the manipulated messages. These self-reported strategies were used to determine the features for the current analysis, but will not be further discussed in the current paper (but see De Boer, Van der Houwen, & Blackwell, in prep.).

4.3. Analysis of the messages

There were 113 participants who mentioned at least one strategy after completing the disguise and/or imitation task. A coding scheme for possible strategies was drawn up using open coding. In the end, all strategies were coded as present (1) or not present (0) in the participants' answers; the answer of each participant could include multiple strategies. Frequency tables were made to show which strategies were mentioned most often. For a thorough discussion of the self-reports, see De Boer et al. (in prep.). For the ten most frequently reported strategies, excluding those that were considered too vague or subjective to analyse systematically, at least one feature was included in the current analysis of the messages (see Table 2). Each variable was recorded and analysed using SPSS and will be discussed in the remainder of this section, sorted by variable type.

Strategy	Freq.	Feature in the analysis	Type of feature
Emoticons/emojis ^a	62	Number of emojis ^b	scale
Opening	47	Type of first emoji ^b	categorical ^c
		Full opening	qualitative (string)
		Greeting	qualitative (string)
		Addressee	qualitative (string)
Punctuation	45	Punctuation	qualitative (string)
		Number of full stops	scale
		Number of commas	scale
Word use ^d	43	-	-
Writing style ^d	30	-	-
Sentence/message length	21	Number of separate text messages sent	scale
		Number of sentences	scale
		Number of words	scale
		Average sentence length	scale
Formality ^d	21	-	-
Sentence structure	18	Subject deletion	binary
		Article deletion	binary
Exclamation marks ^e	17	Number of excl. marks	scale
Closing	15	Full closing	qualitative (string)

Table 2. Top 10 strategies for manipulating the language style (N = 113).

^aAn emoticon is composed of typed keyboard characters, e.g. :-), whereas an emoji is an image, e.g. 😊 (Pavalanathan & Eisenstein, 2015).

^b This variable includes emoticons; however, most messages only contained emojis.

^c Values: 0 = N.A.; 1 = Facial emoji; 2 = Other emoji; 3 = Emoticon.

^d These features were excluded from the current analysis.

^e ‘Changing the number of exclamation marks’ was often mentioned as an additional strategy besides ‘changing the punctuation’, even in the same responses. Therefore, both were counted separately. In the discussion of the results, however, the features are grouped thematically and the exclamation marks will fall under punctuation.

4.4. Scale variables

The following quantitative variables were counted or calculated in each of the messages:

- number of WhatsApp text messages sent to complete the assignment;
- number of sentences, i.e. strings of words separated from another string of words by a full stop, exclamation mark, question mark, emoji/emoticon, line break, empty line, or the end of the text message;
- number of words, calculated using the ‘word count’ function of Microsoft Word;
- average sentence length, calculated as the total number of words divided by the number of sentences;
- punctuation features: number of full stops; number of commas; number of exclamation marks;
- number of emojis and emoticons.

For each of these variables, means of the different conditions (i.e. disguise and imitation) were compared to the control condition using the non-parametric Wilcoxon Matched-pairs test (due to the data not being normally distributed). In addition to testing whether two means differ significantly, the output of the Wilcoxon Matched-pairs

test shows for how many participants the score changed and in which direction, making it possible to see the adaptations made by the authors in their manipulated messages compared to their non-manipulated language style.

To complement the information regarding the use of emojis and emoticons, the type of the first one used in the messages was categorized as 1) emoticon; 2) facial emoji; or 3) other emoji. Whereas different types of emojis/emoticons were sometimes present within one message, the analysis of the first one was considered the most representative way of systematically analysing all messages.

4.5. Binary variables

To analyse subject deletion and article deletion, these features were coded as present (1) or absent (0). For each of the binary variables, the three conditions were compared with each other one-by-one using a chi-square (χ^2) test. For the significant results, their strength was determined using a post-hoc test which calculated the phi coefficient (ϕ). This coefficient was interpreted according to the recommendations of Huizingh (2010): 0 – 0.3 = weak; 0.3 – 0.7 = moderate; 0.7 – 0.9 = strong; 0.9 – 1.0 = exceptionally strong.

4.6. Qualitative variables

Qualitative variables, i.e., the exact formulations of the openings and closings, were treated as nominal string variables with an unlimited number of options and analysed using frequency tables demonstrating the types and number of formulations used by the participants in each of the conditions (i.e., between-author variability). The openings often consisted of different components that were also analysed separately: a greeting (e.g. 'Hi'), addressee (e.g. 'mom'), and a punctuation mark (e.g. ';'). The majority of the closings only consisted of one component (e.g., an emoji or a word), so they were not divided into subcomponents.

5. Results

5.1. Qualitative features

Of the features that were counted in the messages, only the messages in the imitation condition showed significant differences to the control messages. None of the differences between the control and disguise conditions were significant. Table 3 gives an overview of the quantitative features in the three conditions. For reference, the features of the messages of the "victim" Lieke are presented in the right column. Note, however, that the conversations between Lieke and her mother and friend consisted of semi-live chat sessions in which the interlocutors were involved in turn-taking, whereas the messages created in the experiment consisted of a single message or short unilateral stream of messages with no intention of starting a conversation.

Variable	Control		Disguise		Imitation		Lieke	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
RECIPIENT: MOTHER	(N=87)		(N=53)		(N=55)		(N=49)	
# messages	1.14	0.55	1.04	0.28	1.80	1.08	-	-
# sentences	3.54	1.47	3.66	1.52	4.11	1.75	1.29	0.65
# words	32.60	13.67	32.40	10.87	33.09	16.95	8.14	6.36
Mean sentence length	10.09	4.48	10.53	6.10	8.35	2.97	6.12	3.72
# full stops	1.51	1.31	1.45	1.23	1.18	1.36	0.20	0.46
# commas	1.20	0.86	1.23	0.93	0.87	0.92	0.08	0.28
# excl. marks	1.05	1.03	1.28	1.26	1.11	0.94	0.10	0.31
# emojis/emoticons	0.71	0.82	0.75	1.16	1.96	1.47	0.31	0.55
RECIPIENT: FRIEND	(N=94)		(N=60)		(N=61)		(N=35)	
# messages	1.30	1.17	1.11	0.61	1.54	1.06	-	-
# sentences	4.22	1.72	4.45	1.72	5.08	1.48	1.82	1.12
# words	34.02	12.89	36.33	15.31	37.77	14.49	16.60	13.07
Mean sentence length	8.81	3.86	8.83	3.95	7.62	2.70	9.08	4.90
# full stops	1.94	1.70	1.95	1.75	1.25	1.03	0.34	0.68
# commas	1.19	0.88	1.02	0.95	0.77	0.86	0.26	0.51
# excl. marks	1.62	1.43	1.42	1.23	1.89	1.08	0.37	0.60
# emojis/emoticons	1.05	1.27	0.98	1.64	1.90	0.72	0.57	0.70

Table 3. Means and standard deviations of the quantitative features of the messages. Means in bold were significantly different ($p < .05$) from the means in the control condition (see Table 4 for the statistical results).

As can be seen in table 3, there are a few significant differences between the control and imitation condition. Whereas in the control task, most participants sent one text message to complete the assignment, when imitating the victim, the number of texts increased significantly in the messages to the mother. In addition, more but shorter sentences were used in these messages. In the messages to both recipients, the number of words did not change significantly. In the imitation condition, the use of commas and full stops in the messages to the friend decreased, whereas the use of exclamation marks increased. Table 4 gives an overview of the significant results of the Wilcoxon Matched-Pairs test.

Variable	Messages to the mother				Messages to the friend			
	Z score	↑	↓	=	Z score	↑	↓	=
Number of messages	-4.267***	24	1	30				
Number of sentences	-3.217**	30	12	13	-2.441*	34	15	12
Number of words								
Sentence length	-3.727***	16	37	2				
Number of full stops					-3.248**	14	36	11
Number of commas	-2.462*	13	28	14	-2.802**	15	31	15
Number of excl. marks					-1.997*	28	13	20
Number of emojis/emoticons	-4.832***	34	2	19	-4.306***	42	8	11

Table 4. Significant results of the Wilcoxon Matched-Pairs tests, including details on the number of authors for which the feature increased, decreased, and stayed equal in the imitation condition when compared to the control condition. Note that there were only significant results in the imitation condition. * < .001; ** < .01; * < .05**

Overall, the quantitative features studied were present in the majority of the messages written by the participants (see Table 5). The largest difference between the conditions is the number of emojis: almost all (95 to 98%) of the imitated messages contained at least one emoji or emoticon, whereas in the control condition, this was only 53% (to the mother) and 64% (to the friend). The increased use of emojis in the imitation condition may have been prompted by Lieke's messages, in which emojis were used relatively often (in approximately 25% and 50% of the messages to the mother and friend, respectively).

Grapheme	Messages to the mother				Messages to the friend			
	Cntrl N = 87	Disg. N = 53	Imit. N = 55	Lieke N = 49	Cntrl N = 94	Disg. N = 60	Imit. N = 61	Lieke N = 35
Full stop	74	72	60	18	75	73	74	26
Comma	79	74	55	8	78	63	53	23
Exclamation mark	63	66	76	10	75	75	92	31
Emoji/emoticon	53	42	95	27	64	42	98	46

Table 5. Percentage of messages that included the grapheme listed in the left column (per condition and for the "victim" Lieke).

When looking at the type of symbol (i.e., emojis or emoticons) used, only the type of the first in each message was registered, to allow easy comparison between the messages. This was rarely an emoticon (in a maximum of 10 messages per condition and recipient); in the imitated messages to the mother, no emoticons were found. Of the emojis used, most were facial emojis, e.g. 😊 or 😄. These occurred particularly frequently in the imitation condition (73% to the mother and 87% to the friend). Overall, the emoji use in the imitated messages was very similar to that of Lieke, who used facial emojis in 80-81% of the cases, and 'other' emojis (e.g. a spaghetti emoji) in all other cases.

5.2. Sentence structure

Regarding sentence structure, two strategies were explicitly mentioned by the participants and thus analysed: first-person subject deletion and article deletion. Table 6 gives an overview of the percentage of messages in each condition that contained these strategies at least once.

Feature	Messages to the mother				Messages to the friend			
	Cntrl N = 87	Disg. N = 53	Imit. N = 55	Lieke N = 49	Cntrl N = 94	Disg. N = 60	Imit. N = 61	Lieke N = 35
Subject deletion	47	41	49	16	47	58	64	31
Article deletion	10	9	7	2	13	3	13	9

Table 6. Percentage of the messages in the three conditions that contained first-person subject deletion or article deletion. Results that were significantly different from the control condition are presented in bold. For reference, the percentages are also presented for the victim's messages.

In the control condition, almost half of the participants deleted the first-person subject at least once (see Table 6). In the imitated messages to the friend, this percentage

increased significantly ($\chi^2 = 4.362$, $p < .05$). Lieke used subject deletion nearly twice as often in the messages to her friend than to her mother – which may have caused the feature to have been picked up by the imitators in the messages to the friend but not the mother. The relationship between condition and subject deletion was weak ($\phi = 0.168$). Article deletion was less common in the messages: 10 and 13% in the control condition. The disguised messages to the friend showed a significant decrease in the presence of article deletion ($\chi^2 = 3.943$, $p < .05$). The relationship between condition and article deletion was weak ($\phi = -0.160$).

5.3. Openings

The participants show a high degree of between-author variation in the openings of the manipulated messages. Depending on the condition and recipient, the participants used 38 to 73 different openings. Table 7 gives an overview of the openings that were used by more than one author in the control or disguise condition.

Messages to the mother				Messages to the friend			
Control	Freq.	Disguise	Freq.	Control	Freq.	Disguise	Freq.
Hee mam!	9	Hoi mam,	7	Hee Romy!	6	Hoi Romy,	4
Hoi mam,	7	Hoi mam!	3	Hey Romy,	4	Hee Romy!	3
Hee mam,	6	Hoi mama,	3	Hey Romy!	3	Hee room!	2
[none]	5	Heee mam	2	Hi Room!	3	Hoi Room!	2
Hoi mam!	5	Hey mam!	2	Hee Romy,	2	Hoihoi,	2
Hi mam,	3	Hoi mama!	2	Hee Room!	2	[none]	2
Hey mam,	3	Hoi mama.	2	Heee!	2		
Hoi mam.	2	Mam,	2	Hoi Romy,	2		
Hee mams!	2			[none]	2		
Ma,	2			Romy!	2		
				Room!	2		
				Yo	2		
other	43	other	30	other	58	other	45
<i>Total</i>	<i>87</i>	<i>Total</i>	<i>53</i>	<i>Total</i>	<i>90</i>	<i>Total</i>	<i>60</i>

Table 7. Openings to the mother and friend in the control and disguise condition (sorted on frequency per condition per recipient; openings that occurred only once are not presented). N.B. Other openings were each used by one participant only (per condition and recipient).

Apart from the openings presented in table 7, there were openings that were used by only one author in a condition. In the control messages, 49% (to the mother) and 64% (to the friend) of the messages had such a unique opening; in the disguised messages, this was 57% and 75%, respectively.

In the imitation condition, on the other hand, only 3 of the total number of messages contained a unique opening. Most authors had imitated Lieke's openings: only 4 participants used an opening that was not used by Lieke in the messages provided to them (see Table 8). One participant, who used 'Hey lief!', could be detected in her attempt at imitation, since she imitated the wrong person: Lieke's mother calls Lieke 'lief' ('sweetheart') in all four opening sequences, whereas Lieke did not use it once.

Opening	Mother		Friend	
	Imitation	Lieke	Imitation	Lieke
[none]	1	1	0	2
Hey,	25	2	14	2
Hey!	26	1	45	2
hey!	0	0	1	0
Hey mam,	2	0	0	0
Hey lief!	1	0	0	0
<i>Total</i>	55	4	60	6

Table 8. Openings in the imitation condition compared to the openings used by Lieke in the materials provided to the participants.

A closer inspection of the three possible elements of the openings, i.e., greeting (e.g., hey, hi), naming the addressee (e.g. Romy, Room), and punctuations (e.g. , or !), still shows considerable between-author variation in the control and disguise conditions. Per condition and recipient, there are 15–33 varieties of greetings and 12–20 ways of naming the addressee (see Table 9 for an overview). As there was only minimal variation in the imitation condition, where only two types of greetings were used and the addressee was only rarely mentioned, the verbal part of the openings of imitated messages are not further discussed.

	Opening part 1: Greeting				Opening part 2: Addressee					
	Mother		Friend		Mother			Friend		
	Cont.	Disg.	Cont.	Disg.	Cont.	Disg.		Cont.	Disg.	
[none]	9	9	14	10	[none]	15	3	[none]	25	13
Hallo	0	2	0	0	Ma	2	0	lieverd	2	1
Hee	23	4	16	9	mam	47	27	mop	3	0
Heee	0	2	12	1	Mam	3	4	Romy	29	17
Heeee	1	0	2	2	Mama	3	2	romy	0	2
Heeeee	0	0	2	0	mama	3	8	Room	15	10
Heeey	1	1	3	0	mams	6	3	room	5	2
Heey	2	2	2	0	mamsie	2	0	Rooms	0	2
Hey	13	6	14	7				schat	3	1
Hi	6	2	6	3						
Hoi	21	20	6	12						
Hoihoi	0	0	1	2						
Yo	0	0	2	0						
other	11	5	13	14		6	6		11	12
<i>Total</i>	87	53	93	60		87	53		93	60

Table 9. Overview of the verbal parts of the openings (i.e., greetings and ways to address the addressee) per condition (control, disguise) and recipient. ‘Other’ greetings/addressees were present for only one author.

Despite the overall variation in the verbal parts of the openings, some greetings and ways to address the addressees were clearly most popular: at least half of the messages contained the same few greetings or addressee parts. For instance, in the control condition, 66% of the authors opened the messages to their mother with either ‘Hee’, ‘Hoi’, or ‘Hey’, and 54% addressed their mother with ‘mam’.

When looking at the third part of the openings, the punctuation, there is less variation. 6 different (combinations of) punctuation marks were used in the control con-

dition, 9 types (incl. emojis) in the disguise condition, and 3 types in the imitation condition (see table 10). The majority of the openings (64–84% in the control and disguise condition, and 98–100% in the imitation condition) had either a comma or an exclamation mark as punctuation.

Opening 3: Punctuation	Mother				Friend			
	Cont.	Disg.	Imi.	Lieke	Cont.	Disg.	Imi.	Lieke
!	26	13	27	1	39	24	46	2
,	47	21	27	2	31	18	14	2
no punc.	10	6	1	1	14	13	0	2
.	2	5	0	0	4	4	0	0
!!	0	1	0	0	4	0	0	0
;	1	0	0	0	1	0	0	0
..	0	0	0	0	0	1	0	0
...	1	3	0	0	0	0	0	0
❤	0	2	0	0	0	0	0	0
!:))	0	1	0	0	0	0	0	0
😊	0	1	0	0	0	0	0	0
<i>Total</i>	<i>87</i>	<i>53</i>	<i>55</i>	<i>4</i>	<i>93</i>	<i>60</i>	<i>60</i>	<i>6</i>

Table 10. Overview of the punctuation used in the openings in the different conditions and in Lieke's messages (provided to the participants in the imitation condition).

5.4. Closings

Whereas almost 98% of the messages in the dataset contained an opening of some kind, only about 70% of the messages contained a closing (60.0 – 80.5% per condition and recipient). In fact, the absence of a closing was the most popular ending across conditions and recipients. For the messages that contained a closing of some kind, emojis were quite common. The closings used by at least two participants per condition are presented in table 11.

Closing	Mother				Friend			
	Cont.	Disg.	Imi.	Lieke	Cont.	Disg.	Imi.	Lieke
[none]	17	12	11	2	37	24	24	4
😘 / 😊 *	7	5	17	1	11	3	7	0
X	4	2	0	0	6	4	0	0
❤️ / 🍷 / ❤️ *	4	2	0	0	3	1	2	0
:)	5	2	0	0	2	1	0	0
😊 / 😊 *	4	1	6	0	5	0	16	2
x	3	0	0	0	1	0	0	0
xx	3	0	0	0	2	0	0	0
xxx	2	1	0	0	2	0	0	0
Xxx	2	2	0	0	0	1	0	0
❤️❤️ *	2	0	0	0	1	0	0	0
😘😘 *	2	0	1	0	1	0	0	0
Kus!	2	0	0	0	0	0	0	0
Liefs	1	4	0	0	1	0	0	0
👉	0	0	0	0	3	0	0	0
Miss you! 😘 *	0	0	0	0	0	0	2	1
😊😘 *	0	0	5	1	0	0	0	0
🌲😘	0	0	3	0	0	0	0	0
😁 / 😁 *	0	0	3	0	0	0	1	0
👦	0	0	2	0	0	0	0	0
😊😊	0	0	2	0	0	0	1	0
other	29	22	5	0	22	26	16	0
Total	87	53	55	4	93	60	60	7

Table 11. Closings to the mother and friend in the condition (sorted on frequency per condition per recipient). ‘Other’ closings were used by only one participant in the dataset (per condition and recipient). Closings in the messages by the victim Lieke are provided for reference. *For these emojis, similar variants were treated as the same symbol.

In the control messages, about 30% of the participants used a unique closing (not presented in table 11), and in the disguised messages, about 47% did so. When the participants were instructed to imitate Lieke’s language style, 60% and 70% of the messages to the mother and the friend respectively contained a closing that was also used by Lieke.

As can be seen in Table 11, most closings consisted of one component only (e.g., only an emoji or a word). Hence, the different components of the closings were not analysed separately.

6. Discussion

In this study, we investigated the manipulation of language style in WhatsApp messages. Whereas prior research mainly focussed on the effect of feigned non-nativeness (Dern, 2006, 2008; Fobbe, 2014), in line with the findings of Marko (2017), the participants in the current study did not report to have feigned a non-native identity (cf. De Boer et al., in prep.). Neither did our participants, on the whole, report that they had introduced errors into their manipulated messages, whereas earlier publications often

focused on deliberate error production⁴. This implies that research investigating manipulation of language style should rather focus on strategies adopted spontaneously by authors, as was done in the current paper. The strategies investigated in this study overlapped strongly with features of language use that are described in the literature as ‘textual dualities’ (e.g. Olsson, 2009), showing that stylistic authorship analysis may be vulnerable to manipulation.

The exploitation of textual dualities was most striking in the imitation condition. Whereas people in both experimental conditions introduced changes into their language use, without access to the victim’s messages the changes seem less substantial. For example, most quantitative features of the messages in the disguise condition did not differ significantly from the control messages. Note, however, that this may be partly explained by the participants changing their language use in different directions: some may over-use certain features when compared to their control messages, whereas others may under-use them. This study looked at aggregated features and averages, rather than changes made in the experimental conditions when compared to individual authors’ control messages. It is possible that some individuals were more successful at either disguise or imitation than others: this could be the topic for further research based on our data.

A qualitative investigation of changes across conditions per author could shed more light on the nature of the changes and the extent to which certain linguistic features were manipulated in both experimental conditions. What the current analysis did show is that the messages in the disguise condition showed as much variation as the messages in the control condition, whereas the imitated messages were more similar to each other and were also similar to the victim’s language use. This suggests that it might be easier for lay people to change their language use when they have access to messages written by the person they are impersonating. This finding contradicts Dern (2008)’s assertion that imitation requires the highest level of metalinguistic awareness and is therefore the most difficult task for authors who are trying to manipulate their language use. We suggest that metalinguistic awareness is in fact heightened when comparison material is available and the author is confronted with alternatives to their own language use. Whereas most choices in writing are usually made unconsciously (Bayer, Dal Cin, Campbell, & Panek, 2016; Kniffka, 1993), when other people make different choices, this may draw the writer’s attention. This may be especially true for WhatsApp messages, where authors are less constrained by genre-specific conventions compared to e.g. letters (as used in Dern’s experiment) and show high between-author variability. In addition, WhatsApp messages are typically short and may be largely built up by elements (e.g. opening, closing) that can covertly be copied from earlier messages.

According to prior literature, manipulation of language style can be detected because authors are not able to maintain a consistent language style (as suggested by Bredthauer, 2013; Dern, 2006; Grant & MacLeod, 2018; Juola, 2012; MacLeod & Grant, 2017). A conventional way to detect manipulated language is to consider the consistency of features throughout a text. Since manipulation takes up a lot of effort and energy, it is claimed to be difficult to sustain (Bredthauer, 2013; Fobbe, 2014; Kniffka, 1993). Our data make it hard to assess this claim, since the texts concerned were ex-

⁴Errors were not one of the ten most frequently mentioned manipulation strategies, meaning that it was mentioned in <15 responses in the entire database, and were hence not analysed for this paper.

tremely short: the messages were on average 30 to 40 words in length. Although the beginnings (i.e. openings) of the messages were indeed more often imitated than the endings, the short messages produced by the participants may not be long enough to lose their concentration as found in previous studies. Still, this situation is reflective of a certain type of forensic casework that we have encountered, where there may be only one or two final messages sent from the victim's phone. This implies that imitation in such cases may be easier than so far established from experimental studies.

Overall, we can say that the manipulated messages did not share similarities that could make them stand out as deceptive – at least not on the linguistic features analysed in this study. In the disguise condition, most quantitative features did not display significant differences from the control messages; unlike Bredthauer (2013)'s findings on actual manipulated forensic texts, in our study disguised messages were not shorter (see also Marko, 2018). Of the features that changed, such as the openings and closings, there was still as much between-author variation in the disguised messages as in the messages that were not manipulated. In the imitation condition, the authors did become more similar to each other, but this was always in the direction of the victim's language style. Hence, the shared features between the imitated messages are not signs that the manipulation could be easily detected, but merely that the imitations could have been effective.

Whether the imitated messages are actually effective in misleading a recipient or even forensic linguist is a question that remains unanswered. The comparison between the messages in the imitation condition and the messages of the victim cannot be made one-to-one, as the conversation handed to the participants to study the victim's language consisted of multiple exchanges of messages in which the interlocutors were both involved in the conversation simultaneously. This means that for some features, such as emojis and subject deletion, it appears that the participants may have 'over-done' it – a feature known in the sociolinguistic literature as hypercorrection (e.g. Decamp, 1972), but it may in fact be proportional considering the high information-density of the messages constructed by the participants for the experiment. In addition, the knowledge that the victim may use a certain feature at least once in their authentic messages may be enough for both imitators and analysts to conclude that the feature fits their linguistic repertoire.

The current study should be considered as an initial exploration of the field, with limitations concerning the external validity of the results. Firstly, the diversity and representativeness of the participants in the experiment was limited. The participant group was very homogeneous on demographics such as gender, age, and education level. In addition, they did not represent the relevant population, namely suspects in police investigations for abduction and murder, who are most often men aged between 25 and 44 (Centraal Bureau voor de Statistiek, 2018) with vocational rather than academic education. Metalinguistic awareness may be higher for our participants and it may be easier for the young women in the current study to imitate another woman of a similar age. In fact, the control messages of the participants may already show a high degree of similarity to the messages of our victim. In actual cases, the victims may not be of similar age, gender and background, so perpetrators are expected to show less awareness of how their victims generally talk, act and relate to others. We would expect imitators who did not know the victim to have particular difficulty with the pragmatic level, as

experienced by the undercover police officers described by Grant and MacLeod (2020), and this would be fruitful material for future studies.

In addition, the tasks were different from real-life situations in a number of ways. Firstly, the participants were asked to imagine that a friend had just died, while in fact they were sitting among their fellow students in a classroom. They were not rewarded for their manipulations, while in real-life situations a lot is at stake for the person trying to manipulate their language style. In real cases, authors probably put more effort into their act of deception and spend more time studying the victim's messages and formulating their manipulated messages than the students in this study. In fact, they may have even known their victim and have exchanged messages with them for years, creating familiarity with their language use. On the other hand, criminals in real-life situations may be faced with stressful and unplanned situations, and may not be considering language use suiting their victim's style at all. What could have further affected the results is that the participants were first asked to send a message in a non-deceptive style, which was fresh in their memory and even visible while formulating the manipulated messages. This gave the participants an example of how they would write the message themselves, which is expected to have led to increased metalinguistic awareness of their own style.

7. Conclusion

The goal of this study was to investigate how lay people change their language use in WhatsApp messages when asked to disguise their own language style or to imitate someone else's. The findings show that most authors introduce major changes in their messages on a whole range of linguistic and stylistic features. Their strategies did not involve introducing deliberate errors or feigning non-nativeness; rather, they paid attention to the use of emojis and emoticons, punctuation, and openings and closings in their messages. Moreover, our participants displayed a reasonably accurate degree of awareness of the style of the person they were trying to imitate and impersonate. This study shows that the manipulation of language style is an important topic that should be investigated further. In particular more research is needed on the different strategies pursued by individuals attempting to disguise their own language or imitate that of someone else.

Based on this study, it is strongly recommended that in any authorship case the possibility of manipulation should be considered. The linguistic expert should take into account factors such as the level of education of any possible suspect(s), how well they knew the missing/deceased person and the extent to which they might have had access to texts produced by that person. In some cases this might lead to a report which states that the findings can be interpreted in more than one way, which could be regarded as unhelpful advice. However, acknowledging the possibility of manipulation should still be preferred over presenting an invalid conclusion. Presenting neutral conclusions might in the short term delay police investigations, but hopefully it will lead to greater acceptance of the field as a solid forensic science.

References

- Amino, K., Makinae, H., & Kamada, T. (2018). Auditory discrimination of natural speech and synthetic speech used as voice disguise. *Acoustical Science and Technology*, 39(1), 48–50.
- Bayer, J. B., Dal Cin, S., Campbell, S. W., & Panek, E. (2016). Consciousness and Self-Regulation in Mobile Communication: Consciousness in Mobile Communication. *Human Communication Research*, 42(1), 71–97.
- Bernhardsson, S., Correa da Rocha, L. E., & Minnhagen, P. (2009). The meta book and size-dependent properties of written language. *New Journal of Physics*, 11(12).
- Bredthauer, S. (2013). *Verstellungen in inkriminierten Schreiben: Eine linguistische Analyse verstellten Sprachverhaltens in Erpresserschreiben und anderen inkriminierten Texten*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Brennan, M., Afroz, S., & Greenstadt, R. (2012). Adversarial stylometry: Circumventing authorship recognition to preserve privacy and anonymity. *ACM Transactions on Information and System Security*, 15(3), 1–22.
- Brennan, M., & Greenstadt, R. (2009). Practical attacks against authorship recognition techniques. In *Proceedings of the Twenty-First Innovative Applications of Artificial Intelligence Conference*. (pp. 60–65). California: AAAI Press.
- Cambier-Langeveld, T. (2016). Language analysis in the asylum procedure: a specification of the task in practice. *International Journal of Speech Language and the Law*, 23(1), 25–41.
- Centraal Bureau voor de Statistiek. (2018). *Verdachten; delictgroep, geslacht, leeftijd en herkomstgroepering* (Tech. Rep.). Centraal Bureau voor de Statistiek. Retrieved from <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81947NED/table?ts=1529502620833>
- Clark, J., & Foulkes, P. (2007). Identification of voices in electronically disguised speech. *International Journal of Speech Language and the Law*, 14(2), 3820.
- Coulthard, M. (2014). Whose text is it? On the linguistic investigation of authorship. In S. Sarangi & M. Coulthard (Eds.), *Discourse and Social Life* (1st ed., pp. 270–287). Routledge. doi: 10.4324/9781315838502-15
- Coulthard, M., & Johnson, A. (2007). *An introduction to forensic linguistics: Language in evidence*. London: Routledge.
- De Boer, M. M. (2018). *Disguise and imitation of language style in WhatsApp Messages: An analysis of linguistic and stylistic variation in manipulated texts* (Master's thesis, Vrije Universiteit Amsterdam). Retrieved from https://www.ubvu.vu.nl/pub/fulltext/scripties/13_2121611_0.pdf
- De Boer, M. M., Van der Houwen, F., & Blackwell, S. (in prep.). Metalinguistic awareness of linguistic disguise and imitation in WhatsApp messages.
- Decamp, D. (1972). Hypercorrection and rule generalization. *Language in Society*, 1(1), 87–90. doi: 10.1017/S0047404500006552
- Dern, C. (2006). Bewertung inkriminierter Schreiben: Zum Problem der Verwischung von Spurendurch Verstellung. *Kriminalistik*, 5, 323–327.
- Dern, C. (2008). “Wenzzahle nix dan geht dir schlecht”: Ein Experiment zur sprachlichen Verstellungsstrategien in Erpresserbriefen. *Zeitschrift für Germanistische Linguistik*, 36(2), 240–265.
- De Vel, O., Anderson, A., Corney, M., & Mohay, G. (2001). Mining e-mail content for author identification forensics. *ACM SIGMOD Record*, 30(4), 55–64.

- Ekblom, P. (1999). Can we make crime prevention adaptive by learning from other evolutionary struggles? *Studies on Crime and Crime Prevention*, 8, 27–51.
- Eriksson, A. (2005). Tutorial on forensic speech science. Part I: Forensic phonetics. In *Proceedings of the 9th European conference on speech communication and technology*. Lisbon.
- Fobbe, E. (2014). Fingierte Lernersprachen: Strategien der muttersprachlichen Fehlerproduktion im Dienste der Verstellung. *Zeitschrift für germanistische Linguistik*, 42(2), 196–222.
- Gaito, J. (1961). Repeated measurements designs and counterbalancing. *Psychological Bulletin*, 58(1), 46–54.
- Grant, T. (2013). TXT 4N6: Method, consistency, and distinctiveness in the analysis of SMS text messages. *Journal of Law and Policy*, 58(1), 467–494.
- Grant, T., & MacLeod, N. (2018). Resources and constraints in linguistic identity performance. *Language and Law / Linguagem e Direito*, 5(1), 80–96.
- Grant, T., & MacLeod, N. (2020). *Language and Online Identities: The Undercover Policing of Internet Sexual Crime* (1st ed.). Cambridge University Press.
- Juola, P. (2012). Detecting stylistic deception. In *Proceedings of the EACL Workshop on Computational Approaches to Deception Detection* (pp. 91–96). Avignon, France: European Association for Computational Linguistics.
- Kniffka, H. (1993). Forschungsethik. In H. Goebel, P. H. Nelde, Z. Starý, & W. Wölck (Eds.), *Kontaktlinguistik* (pp. 819–825). Walter de Gruyter.
- Künzel, H. J. (2000). Effects of voice disguise on speaking fundamental frequency. *Forensic Linguistics*, 7(2), 149–179.
- Luyckx, K., Daelemans, W., & Vanhoutte, E. (2006). Stylogenetics: Clustering-based stylistic analysis of literary corpora. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*. Genoa, Italy: European Language Resources Association.
- MacLeod, N., & Grant, T. (2017). “go on cam but dnt be dirty”: Linguistic levels of identity assumption in undercover online operations against child sex abusers. *Language and Law / Linguagem e Direito*, 4(2), 157–175.
- Marko, K. (2017). Strategies for disguise in written threatening communications and ransom demands: an analysis of English and German texts. *International Journal of Speech Language and the Law*, 24(2), 243–247.
- Marko, K. (2018). Underlying motivations for the use of linguistic disguise in written English and German threats and ransom demands in an experimental corpus. *Nordic Journal of Linguistics*, 41(2), 205–226.
- Masthoff, H. (1996). A report on a voice disguise experiment. *International Journal of Speech Language and the Law*, 3(1), 160–167.
- Nolan, F. (2001). Speaker identification evidence: Its forms, limitations, and roles. In *Proceedings of the conference Law and language: Prospect and retrospect* (pp. 1–19). Levi, Finnish Lapland.
- Olsson, J. (2008). *Forensic linguistics* (2nd ed.). New York: Continuum.
- Olsson, J. (2009). *Word crime: Solving crimes through forensic linguistics*. London: Bloomsbury.
- Pavalanathan, U., & Eisenstein, J. (2015). Emoticons vs. Emojis on Twitter: A Causal Inference Approach.

- Perrot, P., Preteux, C., Vasseur, S., & Chollet, G. (2007). Detection and recognition of voice disguise. In *Proceedings, IAFPA*. Plymouth, UK: International Association for Forensic Phonetics and Acoustics.
- Schilling, N., & Marsters, A. (2015). Unmasking Identity: Speaker Profiling for Forensic Linguistic Purposes. *Annual Review of Applied Linguistics*, 35, 195–214.
- Seifert, J. (2010, January). Verstellungs- und Imitationsstrategien in Erpresserschreiben: Empirische Studien zu einem Desiderat der forensisch-linguistischen Textanalyse. *Zeitschrift für angewandte Linguistik*, 2010(52).
- Shapero, J., & Blackwell, S. A. (2012). ‘There are letters for you all on the side-board’: What can linguists learn from multiple suicide-note writers. In S. Tomblin, R. Sousa-Silva, & M. Coulthard (Eds.), *Proceedings of the international association of forensic linguists’ tenth biennial conference* (pp. 225–244). Birmingham: Centre for Forensic Linguistics.
- Simons, A., & Tunkel, R. F. (2021). The Assessment of Anonymous Threatening Communications. In J. R. Meloy & J. Hoffmann (Eds.), *International Handbook of Threat Assessment* (pp. 235–256). Oxford University Press.
- Solan, L. M. (2013). Intuition versus algorithm: The case of forensic authorship attribution. *Journal of Law and Policy*, 21(2), 551–576.
- Van der Veer, N., Boekee, S., & Peters, O. (2017). *Nationale social media onderzoek 2017: Het grootste trendonderzoek van Nederland naar het gebruik en verwachtingen van social media #NSMO*. (Tech. Rep.). Enschede: Newcom Research & Consultancy.