



Effects of Relationship Goal on Linguistic Behavior in Online Dating Profiles: A Multi-Method Approach

Tess van der Zanden*, Alexander P. Schouten, Maria Mos, Chris van der Lee and Emiel Kraemer

Department of Communication and Cognition, Tilburg University, Tilburg, Netherlands

OPEN ACCESS

Edited by:

Sidarta Ribeiro,
Federal University of Rio Grande do
Norte, Brazil

Reviewed by:

Debra Titone,
McGill University, Canada
Peter Boot,
Huygens Institute for the History of the
Netherlands (KNAW), Netherlands

*Correspondence:

Tess van der Zanden
T.vdrZanden@tilburguniversity.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 27 November 2018

Accepted: 01 May 2019

Published: 28 May 2019

Citation:

van der Zanden T, Schouten AP,
Mos M, van der Lee C and Kraemer E
(2019) Effects of Relationship Goal on
Linguistic Behavior in Online Dating
Profiles: A Multi-Method Approach.
Front. Commun. 4:22.
doi: 10.3389/fcomm.2019.00022

This study uses two methods to examine whether online daters looking for a long-term relationship behave linguistically different in their profile texts compared to daters seeking casual relationships. To investigate these linguistic differences, 12,310 existing Dutch dating profiles were analyzed using the Linguistic Inquiry and Word Count (LIWC) program and a word-based classifier. Results of both methods suggest there are reliable differences in the linguistic behavior long-term and casual relationship seekers employ in their dating profiles: long-term relationship seekers mention more topics that are relevant when looking for a long-term relationship, such as internal personality traits and qualities. Additionally, long-term relationship seekers seem to self-disclose more in their profile texts by providing more personal information and using more I-references. Profile texts of casual relationship seekers are more diffuse and harder to classify. Moreover, the study demonstrates that using a multi-method approach, with LIWC and a data-driven word-based classifier, provides a deeper understanding of linguistic differences between the two relationship seeking groups.

Keywords: online dating, relationship goals, language use, text analysis, LIWC, machine learning

INTRODUCTION

According to some studies, online dating has now surpassed more traditionally popular ways of meeting partners (Ross, 2017). In online dating, a person's dating profile is the key element; it is the gatekeeper to further interaction and ultimately even to the establishment of the intended relationship goal (Ellison et al., 2012). Dating profiles typically consist of pictures, basic demographic information, and an open-ended component in which online daters can create a textual self-description (Rosen et al., 2008). In this description, profile owners can express their interests and hobbies, characteristics sought in a potential partner, and relay their intentions and goals to others.

Despite the importance of profiles during the online dating process, little attention has been paid to the textual component of dating profiles. Most studies on the textual component in dating profiles focused on deceptive behavior and the profiles' accuracy (e.g., Toma and Hancock, 2012; Lo et al., 2013). Moreover, some studies have investigated the extent to which personality (Weidman et al., 2015) and other stable characteristics such as gender (Groom and Pennebaker, 2005; Van Berlo and Ranzini, 2018) and age (Davis and Fingerman, 2016) are expressed in textual self-descriptions. This usually ties in with other existing research in a variety of writing domains including collected work of writers (e.g., essays, novels and poems; (Pennebaker and King, 1999);

Pennebaker and Stone, 2003) and social media language use (e.g., Twitter, blogs and Facebook; Yarkoni, 2010; Nguyen et al., 2013; Schwartz et al., 2013).

Studies show that linguistic behavior and the use of particular linguistic characteristics are not only affected by a writer's personality, gender, and other stable traits, but may also be affected by more dynamic characteristics of the writer. A writer's emotions may be one such factor that influences language use: positive and joyful writers may behave linguistically different from negative and angry writers, for instance, with respect to affective language use, negations and punctuation use (Hancock et al., 2007; Gill et al., 2008). For example, positive emotion expressers used approximately six times more the number of exclamation marks and five times less negative affect words than negative emotion expressers (Hancock et al., 2007). Moreover, a volatile factor like a writer's goals (Russell and Schober, 1999) and intentions (Toma and Hancock, 2012) may also guide how people linguistically behave.

Research on how these more dynamic characteristics affect linguistic behavior is scarce, though, and no research has investigated how people's goals in terms of their desired romantic relationship may affect language use. It is likely that people's relationship desires influence how they behave and express themselves, because relationships are fundamental in our lives. This (linguistic) behavior could be especially apparent in the online dating domain where most people adapt to the minimal cue environment by developing strategies—consciously and unconsciously—about how to present themselves. As such, online daters' textual self-presentations may differ depending on profile owners' motives and intended relationship goals (Ranzini and Lutz, 2017). The first goal of this study is therefore to examine to what extent dating intentions affect linguistic behavior in online dating profile texts, and what particular linguistic elements are important when trying to distinguish between profile texts written by people with different relationship goals. To do so, the language use in existing dating profiles of online daters who aim to find a long-term relationship partner is compared with that of daters who search for a casual, less-involved relationship.

In order to investigate whether linguistic behavior is affected by online daters' relationship goals, two different computer-based text analysis methods are employed in the present study. First, by using a theoretical motivated approach, we identify linguistic characteristics that may differ among profile owners looking for a long-term or casual relationship. The Linguistic Inquiry and Word Count program (LIWC; Pennebaker et al., 2015) is subsequently used to automatically analyze 12,310 profiles derived from a Dutch dating site. In addition, we use a data-driven approach to further extract content-specific features that can be used to classify texts of the two relationship seeking groups. In this way, the theoretically postulated linguistic differences between the two groups of profile texts we investigate can then hopefully be confirmed, as well as explore whether there are additional linguistic differences not captured by the first approach. Accordingly, the second goal of our study is to determine to what extent a data-driven word-based classifier adds to a method like LIWC when textually analyzing online dating profile texts. Using both methods may be effective for obtaining

a finer-grained and more comprehensive picture of linguistic differences between profile texts written by people with long-term and casual relationship goals.

BACKGROUND

Online Dating Intentions and Relationship Goals

People's online dating intentions and relationship goals may differ: some aim for a traditional life-lasting relationship, where intended intimacy goals may eventually develop into long-term commitment (Sternberg, 1986; Eastwick et al., 2018). These people start dating someone with the intention, but not certainty, of a high-involved relationship (Buunk et al., 2002). Others seek casual, potentially sexual, dates which may involve personal contact without the intention to become high-involved, intimate relationship partners (Peter and Valkenburg, 2007; Clemens et al., 2015; Chan, 2017). For them, meeting a dating partner or gaining dating experience can already be indicative of success (Gibbs et al., 2006; Sharabi and Dykstra-DeVette, 2019).

On many dating sites, users can indicate their own relationship goal and can see the relationship goals of others while searching or scrolling through profiles. The relationship goal is then mentioned explicitly as a basic characteristic on a dating profile, together with self-reported information about a profile owner's age, level of education, place of residence, etc. In addition to being mentioned explicitly as a basic characteristic, a dater's relationship goal can also be reflected in other aspects of the dating profile. For instance, online daters looking for a long-term or short-term relationship differ in both the presence as well as the content of the picture(s) on their profiles. Long-term relationship seekers are overall more likely to display a profile picture, while casual relationship seekers tend to wear less clothes on the pictures they post (Gallant et al., 2011). In addition, serious-minded daters posted more profile pictures that were deemed realistic by others than daters looking for casual relationships as for the first group future interactions are more likely to occur (Toma and Hancock, 2011). Both relationship seeking groups emphasize specific attributes and as such convey information about their intentions: whereas long-term relationship seekers self-disclose more by being more inclined to provide a (realistic) profile picture, short-term relationship seekers emphasize physical attractiveness and sexuality.

While Gallant et al. (2011) and Toma and Hancock (2011) have shown that relationship goals can influence daters' selection of profile pictures, this has not been investigated yet with respect to language use in dating profiles. It is conceivable that dating intentions are also reflected in the linguistic behavior that online daters employ in their profile texts (Toma and D'Angelo, 2017).

Language Use in Online Dating Profiles

Fairly stable characteristics of profile owners affect linguistic behavior consciously and unconsciously at the same time: whereas men and younger adults are more likely to consciously and strategically write words related to income or status, women use more words related to sexuality and physical appearance. On the other hand, in their dating profiles, men and older adults

tend to use more first-person plural pronouns (e.g., “we”) without being aware, and women and younger adults write more first-person singular pronouns (e.g., “I”) (Groom and Pennebaker, 2005; Davis and Fingerman, 2016; Van Berlo and Ranzini, 2018).

While profile owners’ language use is structurally affected by these stable characteristics, it is well-established that language use may also vary according to the specific setting or the goal that a writer has in mind when writing the text. While producing language, people are known to adapt their style and register to their intended audience (e.g., Giles et al., 1991; Clark, 1996). This language adaptation can improve communication, for instance, when a teacher or parent uses language in such a way that the addressee (pupil or child) is more likely to understand the message. But adaptation can also be strategic (Gallois et al., 2005; Toma, 2014). In particular, it seems likely that online daters use specific words and phrases to illustrate or clarify their interpersonal relationship goals, in an attempt to find a partner online with comparable romantic relationship intentions. Based on the attributes that are considered important for a particular relationship goal, online daters can choose what information to cover and disclose about themselves, by strategically mentioning or avoiding certain topics in profiles (Whitty, 2008; Toma and Hancock, 2012). In this way, profile owners can highlight their own relationship goals, but can also increase the chances of getting replies from other site users with similar intentions. Although language adaptation may thus be a strategic and conscious choice, having a specific relationship goal in mind while writing the profile text may also unconsciously affect what linguistic features are used in the text. This unconscious effect on linguistic behavior can, for instance, be reflected in the use of certain function words; those are used to signal grammatical relationships among words in a sentence, but have themselves little lexical meaning (e.g., auxiliary verbs, pronouns, articles). Such function words are often produced without the writer being aware, but can reveal information about a writer’s social and psychological processes, for example, about the focus of attention, emotional state, social status, and (dis)honesty (Tausczik and Pennebaker, 2010; Pennebaker, 2011; Toma, 2014). In short, relationship goals may both consciously and unconsciously affect linguistic behavior of online daters when constructing their profile texts.

The text analysis program LIWC has been widely used for the study of consciously and unconsciously produced linguistic features in all sorts of texts (Tausczik and Pennebaker, 2010), and was also used for the aforementioned analyses of age, gender, and (dis)honesty in dating profiles (e.g., Groom and Pennebaker, 2005; Toma and Hancock, 2012; Davis and Fingerman, 2016). LIWC is based on the assumption that while writing, the use of particular words provides insights into writers’ linguistic and psychological processes as well as their mental states. The words in the profile texts presumably also carry information that reveals more about intentions and goals of daters, as is investigated here. As LIWC provides a large number of (mostly thematic) linguistic categories on which texts are analyzed, below specific hypotheses are formulated regarding the expected differences in language use of long-term and casual relationship seekers on the most relevant of those categories.

Hypotheses

Body & Sexuality and Status

Casual relationship seekers have a higher focus on external characteristics, such as sexual desirability and physical attractiveness, that are considered important to find in a lower involved relationship partner (e.g., Regan et al., 2000; Li and Kenrick, 2006). The higher the level of relationship involvement, the more individuals become attentive to internal qualities that become more important over the long-term, such as particular personality traits and resource acquisition (Buunk et al., 2002; Eastwick and Finkel, 2008). Following the assumption that casual relationship seekers focus more on external characteristics (e.g., “good-looking,” “fit”) and long-term relationship seekers more on internal characteristics (e.g., “income,” “job”), it is expected that based on their own relationship goal, profile owners consciously decide to present and emphasize such characteristics in their dating profiles (Toma and D’Angelo, 2017).

H1. Online daters looking for a casual relationship use more words related to the body and sexuality than online daters looking for a long-term relationship.

H2. Online daters looking for a long-term relationship use more words related to status than online daters looking for a casual relationship.

Positive Emotion Words

Emotional and psychological processes unconsciously affect the use of particular emotion words (Pennebaker and King, 1999). Long-term relationship seekers are more ready to emotionally involve, to commit, and to bond with a romantic partner, whereas casual relationship seekers look more often for contact with a lower focus on intimacy and emotional involvement (Gibbs et al., 2006; Stephure et al., 2009). Moreover, online daters aiming for long-term relational goals are more inclined to put effort in creating a profile that is deemed positive by others (Gibbs et al., 2006). It has earlier been shown that positive text messages affect romantic satisfaction positively (Luo and Tuney, 2015). Consequently, we hypothesize that long-term relationship seekers use more positive emotion words, including words that express emotional closeness (e.g., “love,” “loyal”) and emphasize positive personality traits (e.g., “intelligent,” “self-confident”).

H3. Online daters looking for a long-term relationship use more positive emotion words than online daters looking for a casual relationship.

Personal Pronouns

How often automatically produced personal pronouns are used can give away information about an individual’s attitudes, goals, and roles within relationships (Pennebaker, 2011), as well as about immediacy and involvement (Walther, 2007). It has been argued that people who engage in self-relevant goals tend to refer more to the self than those who are high involved in other people’s lives (Slatcher et al., 2008; Pennebaker, 2011). We hypothesize that online daters who look for a casual, lower-involved relationship are more self-focused than long-term relationship seekers, and consequently use more first-person singular pronouns that refer to the self (e.g., “I,” “me”).

Online daters with a desire of high relationship involvement and high intimacy are more likely to adopt a listening role (McAdams et al., 1984). The use of pronouns that refer to others (e.g., “you,” “your”) is indicative of other-focused attention (Pennebaker, 2011). The expectation is that long-term relationship seekers unconsciously use more you-references to explicitly acknowledge the importance of the other’s presence.

Furthermore, using first-person plural pronouns (e.g., “we,” “our”) is often correlated with a measure of social integration (Sanderson et al., 2007; Pennebaker, 2011) and highly committed partnership (Agnew et al., 1998; Slatcher et al., 2008). Since long-term relationship seekers have a stronger focus on committing in long-term intimate relationships, they are expected to emphasize connection and interdependence without being aware, by using more we-references.

H4. Online daters looking for a casual relationship use more I-references than online daters looking for a long-term relationship.

H5. Online daters looking for a long-term relationship use more you-references than online daters looking for a casual relationship.

H6. Online daters looking for a long-term relationship use more we-references than online daters looking for a casual relationship.

LIWC and the Word-Based Classifier

To test the hypotheses, we use the program LIWC, perhaps the most commonly used automated text analysis program in the field of language and social psychology (Tausczik and Pennebaker, 2010). The program counts occurrences of specific words in a text, which are taken from list of words that are assigned to different predefined categories related to thought processes, emotional states, and intentions (Pennebaker et al., 2015). The word and category validations by human judgments make LIWC appealing for hypothesis-testing research (Tausczik and Pennebaker, 2010). At the same time, these validated categories are also sometimes perceived as one of the method’s limitations (e.g., Schwartz et al., 2013; Boyd, 2017). LIWC works with a closed-vocabulary approach and hence not all words in a text necessarily occur on this fixed word list. As a result, words that are characteristic of a particular phenomenon under investigation (like a dating profile text) may be missed. In addition, LIWC categories can be rather broad, with some categories containing over a thousand words (e.g., positive emotion words category). As such, differences between groups of texts on the use of words from these broader categories provide information about underlying psychological constructs of writers, but lack information about the specific words within these categories that discriminate the two text groups, which may complicate interpretation of observed differences by LIWC.

Although studies focusing on language use in dating profile texts have used LIWC (e.g., Toma and Hancock, 2012; Davis and Fingerman, 2016), the lexicon’s suitability to the domain has received little regard. It may be the case that LIWC does not cover the content-specific features that are specifically relevant and common among online dating profile texts. Dating profiles can be rather noisy; they are of informal nature and written

by a very diverse population, who do not necessarily adhere to standard language conventions (Van der Zanden et al., 2018). Furthermore, LIWC can provide a wide perspective on structure, tone and the extent to which psychological processes of intentions are reflected in dating profiles, but reveals little about which content-specific features are particularly distinctive for profile texts of long-term and casual relationship seekers. Perhaps, differences in language use between the two relationship seeking groups manifest themselves not so much in the words captured by the broader LIWC categories, but rather in the use of particular content-specific features, where profile owners may use particular words or word combinations that are considered to be important in a profile text.

In contrast to LIWC, word-based machine learning methods do not rely on a priori word or category judgments but use the texts as linguistic input. Such open-vocabulary methods offer finer-grained methods for profile text analysis, yielding additional insights and more information that leverages findings of closed-vocabulary LIWC analyses (Schwartz et al., 2013). It thereby captures the content-specific features that are specifically relevant within the online dating domain. On the other hand, compared to a text analysis method like LIWC, the output of (word-based) computational approaches is harder to interpret. These computational approaches tend not to explicitly show how underlying psychological constructs, such as personal (relationship) goals, are reflected in language use.

Prior online dating studies using machine learning methods primarily aimed at building recommendation systems, also those that focused on (natural) language in online dating (e.g., Diaz et al., 2010; Akehurst et al., 2011; Tay et al., 2018). One exception is the study of Van Berlo and Ranzini (2018) who used a data-driven word-based classifier approach to investigate how male and female users of Tinder differ from each other in their textual self-presentations, by focusing on their usage of pronouns, nouns, adjectives, and verbs. For instance, they found nouns like “music,” “film,” “friend,” and “student” to occur relatively frequently: men were more likely to mention “film” and women used “student” more often, while neither men nor women were distinctive in their use of “music” and “friend.” The frequencies in which words occur give an indication of what Tinder profile owners prioritize in their self-presentation.

Both LIWC and a data-driven word-based classifier are methods that come with advantages and disadvantages. For that reason, it is beneficial to combine the two computer-based text analysis methods. Previous studies that combined LIWC and a machine learning approach to investigate complex, natural language in online environments have shown that using both methods results in a better explanation for linguistic behavior than when only of the two is used (e.g., Gill et al., 2008; Paltoglou and Thelwall, 2012; Schwartz et al., 2013). By employing this multi-method approach, the second research goal is addressed: to investigate the extent to which it is valuable to use an open-vocabulary word-based classifier with a closed-vocabulary approach as LIWC for dating profile text analysis. By doing so, it can be investigated which additional content-specific features can be uncovered that differentiate between profile texts written by long-term and casual relationship seekers.

METHODS

Corpus

Our sample included 12,310 dating profiles from a popular Dutch dating site, which presents itself as “the dating site for everyone.” The site has more than 75,000 active members of different ages and education levels, as the site explicitly mentions that it is open for everyone. The profile texts for our sample were extracted automatically from the site by means of the free online tool Web Scraper. For anonymity reasons, pictures and user names of profile owners were not collected. The main analyses were performed on an aggregated level, which means that only differences between long-term and casual relationship seekers were examined. Ethical clearance for data collection and text analysis was obtained in 2017 by the Ethics Committee (ETC) of the Tilburg School of Humanities and Digital Sciences.

When creating a profile on this dating site, members are asked to write a short piece of text in a section called “about myself,” consisting of information about who the profile owner is and the type of partner and relationship they look for. Together with the profile text, the profile owner’s self-reported gender, age, education level, and desired relationship goal were extracted. These were the standard profile characteristics that were directly visible when scrolling through profiles of other site members. Only the first hundred words of each profile text were analyzed as this is what other site users see when initially searching for potential dates or partners. Furthermore, only profiles were included in the sample that had a word count of more than fifty words and were written in Dutch by someone who indicated to live in the Netherlands. Each profile text therefore contained between fifty and hundred words ($M = 80.71$, $SD = 12.99$). Below, an example of an anonymized, translated version of a profile text is presented.

My name is [name] and I live in [place]. I look for a woman with both IQ and EQ. For me it is important that you feel comfortable in your skin and that you know what you want in life. I am a no-nonsense type myself. Also I am a passionate cyclist and I relax the most in my classic car. Moreover, I have my heart on my sleeve but I like it if someone pushes back a bit. Everything with respect for each other. Show who you are by being yourself. That’s what I do too.

Site users can select one of six relationship preferences provided by the dating site when setting up their profile (“I don’t care,” “mail contact,” “friendship,” “date,” “LAT relationship” (i.e., living-apart-together) or “long-term relationship”). For this

study, the group of long-term relationship seekers consisted of profile owners who selected “long-term relationship” as the preferred relationship goal ($n = 10,696$). Our casual relationship seeking group contained profiles of those who selected the option “date” as the preferred relationship outcome ($n = 1,614$).

The sample contained only profiles of heterosexuals because the focus of the dating site is primarily on heterosexual singles. The profile owners’ mean age was 42 years and 8 months ($SD = 11.7$) and 64.2% ($n = 7,907$) were men. Profile texts were written by people with different educational backgrounds (lower education: 42.3%, higher education: 57.7%). Lower educated people were the profile owners with a vocational education background and higher educated those who completed higher levels of high school or had a higher professional education or university background. **Table 1** shows that the long-term and casual relationship seekers were comparable in terms of their self-reported gender, mean age, and level of education.

LIWC

All profile texts from the sample were analyzed by means of LIWC, which calculates the proportions of words related to the different predefined linguistic categories. The Dutch LIWC2015 vocabulary was used, containing more than 13,000 words, in which each word is assigned to one or more categories (Boot et al., 2017; Van Wissen and Boot, 2017). By default, this vocabulary analyzes each text file on 73 established linguistic categories. In this study, the focus is on three sets of categories, with a total of six categories. The first set of categories looked into the use of words related to body and sexuality, and status. Here, the words in the predefined LIWC categories Body and Sexuality were grouped to form one umbrella category. From LIWC’s main category Personal Concerns, the categories Work and Money were merged into one category Status, as potential partners’ careers and incomes are often associated with status in the dating environment. Other sub categories within this main category (i.e., Leisure, Home, Religion, and Death) did not cover word related to status. In these umbrella categories, words that occurred in more than one of the original LIWC categories were listed only once. The second set gauged the use of positive emotion words and only included the words from the LIWC category Positive Emotions. The third set looked into the use of personal pronouns, in which our focus was on the proportion of I-, You-, and We-references in the profile texts.

Finally, an additional LIWC category was defined by the authors consisting of a list with 60 words related to long-term relational involvement (e.g., “to settle,” “long-term”)

TABLE 1 | Demographic composition of the sample for both relationship seeking groups.

	Gender		Mean age	Education level	
	Men	Women		Low	High
Long-Term	6,773 (63.3%)	3,923 (36.7%)	42.8 (11.83)	4,644 (43.3%)	6,052 (56.6%)
Casual	1,134 (70.3%)	480 (29.7%)	42.2 (11.08)	567 (35.1%)	1,047 (64.9%)
Total	7,907	4,403	42.7 (11.74)	5,211	7,099

This concerns profile owner’s self-reported demographic information.

(see the **Supplementary Material** for the complete list). This category Long-Term Relational Involvement was compiled as a manipulation check to examine whether, compared to casual relationship seekers, profile owners looking for a long-term relationship mention more words related to long-term commitment and involvement.

To test our hypotheses with LIWC, a multivariate analysis of variance (MANOVA) was conducted, with relationship goal as independent variable and six dependent variables, being the proportion of words matching with the words within the six linguistic categories, viz. Body & Sexuality, Status, Positive Emotion Words, I-references, You-references, and We-references.

In order to obtain more information about differences in frequencies of words within the six linguistic categories, a log-likelihood ratio analysis was conducted (Dunning, 1993). This analysis tests whether a word's relative frequency differs significantly for profiles written by people looking for a long-term or casual relationship partner.

Word-Based Classifier

The word-based classifier is based on the classifier approach of Van der Lee and Van den Bosch (2017) (see also Aggarwal and Zhai, 2012). Six different machine learning methods are used: linear SVM (support vector machine), Naive Bayes, and four variants of tree-based algorithms (decision tree, random forest, AdaBoost, and XGBoost). In contrast with LIWC, this open-vocabulary approach does not deal with any preassembled word list but uses aspects from the profile texts as direct input and extracts content-specific features (word n-grams) from the texts that are distinctive for either of the two relationship seeking groups.

Two steps were applied to the texts in a preprocessing stage. Most of the stop words from the regular list of Dutch stop words from the Natural Language Toolkit (NLTK), a module for natural language processing, were not considered as content-specific features. Exceptions are the personal pronouns that are part of this list (e.g., "I," "my," and "you"), because these function words are assumed to play an important role in the context of dating profile texts (see the **Supplementary Material** for the materials used). The classifier operates on the level of the lemma, meaning that it converts the texts into distinctive lemmas. Lemmatization was performed with Frog (Van den Bosch et al., 2007).

To maximize the chances that the classifier assigned a relationship type to a text based on the investigated content-specific features rather than on the statistical chance that a text is written by a long-term or casual relationship seeker, two similarly sized samples of profile texts were needed. To do so, 1,614 texts of each relationship group were used: the entire subset of the group of casual relationship seekers' texts and an equally large subset of the 10,696 texts for the long-term relationship seekers. This subset of long-term texts was randomly stratified on gender, age and level of education based on the distribution of the casual relationship group. Consequently, the baseline chance of the word-based classifier to classify a profile text in the correct relationship group was 50%.

A ten-fold cross validation method was used, meaning that the classifier uses ten times 90 percent of the data to classify the other 10 percent. To obtain a more robust output, it was decided to run this ten-fold cross validation ten times using ten different seeds. To control for text length effects, the word-based classifier used ratio scores to calculate feature importance scores rather than absolute values. These importance scores are also known as Gini importance (Breiman et al., 1984), and are normalized scores that together add up to one. The higher the feature importance score, the more distinctive that feature is for texts of long-term or casual relationship seekers.

RESULTS

LIWC

Overall, LIWC recognized 80.9% of the words in the profiles ($SD = 6.52$). Profile texts of long-term relationship seekers were on average longer ($M = 81.0$, $SD = 12.9$) than those of casual relationship seekers ($M = 79.2$, $SD = 13.5$), $F_{(1,12309)} = 26.8$, $p < 0.001$, $\eta^2 = 0.002$. Other results were not influenced by this word count difference because LIWC operates with proportion scores. In the **Supplementary Material**, more detailed information about other text characteristics of the two relationship seeking groups can be found. Moreover, it was found that long-term relationship seekers use more words related to long-term relational involvement ($M = 1.05$, $SD = 1.43$) than casual relationship seekers ($M = 0.78$, $SD = 1.18$), $F_{(1,12309)} = 52.5$, $p < 0.001$, $\eta^2 = 0.004$.

Hypothesis 1 stated that casual relationship seekers would use more words related to the body and sexuality than long-term relationship seekers because of a higher focus on external characteristics and sexual desirability in lower involved relationships. Hypothesis 2 concerned the use of words related to status, where we expected that long-term relationship seekers would use these words more than casual relationship seekers. In contrast with both hypotheses, neither the long-term nor the casual relationship seekers use more words related to the body and sexuality, or status. The data did support Hypothesis 3 that posed that online daters who indicated to look for a long-term relationship partner use more positive emotion words in the profile texts they write than online daters who seek for a casual relationship ($\eta^2 = 0.001$). Hypothesis 4 stated casual relationship seekers use more I-references. It is, however, not the casual but the long-term relationship seeking group that use more I-references in their profile texts ($\eta^2 = 0.002$). Furthermore, the results are not in line with the hypotheses stating that long-term relationship seekers use more you-references because of a higher focus on others (H5) and more we-references to emphasize connection and interdependence (H6): the groups use you- and we-references equally often. Means and standard deviations for the linguistic categories included in the MANOVA are presented in **Table 2**.

Interaction effects of relationship goal and gender were only significant for text length, $F_{(1,12306)} = 6.49$, $p = 0.011$, $\eta^2 = 0.001$, and the use of I-references, $F_{(1,12306)} = 7.83$, $p = 0.005$, $\eta^2 = 0.001$. Simple effects analyses showed that

TABLE 2 | Average proportion score (SD) for each linguistic category and relationship goal.

Linguistic category	Long-Term	Casual	F ^a	p
Body & sexuality	0.30 (0.70)	0.30 (0.75)	0.040	0.854
Status	1.53 (1.29)	1.53 (1.24)	0.027	0.869
Positive emotions	6.26 (3.36)	5.70 (3.19)	39.70	<0.001
I-references	7.33 (3.35)	6.88 (3.36)	25.20	<0.001
You-references	1.76 (1.89)	1.82 (1.99)	1.50	0.200
We-references	0.23 (0.33)	0.21 (0.25)	1.64	0.220

^aWith for all measures of relationship goal $F_{(1, 12308)}$.

TABLE 3 | Top 10 most important content-specific features with the importance score (IS) per relationship seeking group.

Long-term			Casual		
Dutch	English	IS	Dutch	English	IS
betrouwbaar	trustworthy	0.032	date	date	0.087
samen	together	0.028	zin	feel like	0.017
mijn profiel	my profile	0.027	weten	to know	0.016
rustig	calm	0.026	vrouw	woman	0.011
eerlijk	honest	0.026	gek	crazy	0.010
ik	I	0.024	geen	no	0.009
dag	day	0.023	even	for a while	0.009
serieus	serious	0.023	eten	to eat	0.005
mijn	my	0.022	komen	to come	0.005
genieten	to enjoy	0.022	sturen	to send	0.005

men wrote significantly longer texts when looking for a long-term relationship, $F_{(1, 12306)} = 33.33, p < 0.001$, while for women relationship type sought did not affect profile length, $F_{(1, 12306)} = 0.597, p = 0.440$. Men and women both use more I-references when seeking a long-term than a casual relationship, but this difference was more pronounced for women than for men (men: $F_{(1, 12306)} = 6.47, p = 0.011$; women: $F_{(1, 12306)} = 25.46, p < 0.001$). No significant interaction effect of relationship goal and level of education was found, and results with and without age as covariate were similar.

Word-Based Classifier

XGBoost appeared to be the most effective algorithm for the word-based classifier, with an accuracy score of 59.6%, improving accuracy above chance with 9.6%. For both relationship seeking groups, the ten most important content-specific features are provided in **Table 3**, together with their English translations and the importance score for each feature. These features are obtained based on the importance scores given by XGBoost, the best performing tree-based algorithm. Together, all the features given by the classifier add up to one. Only those features with a total frequency score of hundred or more were selected.

Table 4 reports the percentages of texts being classified correctly and incorrectly in the group to which it belongs. Profile texts of long-term relationship seekers were more often

TABLE 4 | Percentages of correctly and incorrectly classified texts based on the distinctive content-specific features.

True text type	Correctly classified	Incorrectly classified
Long-term	61.5%	38.5%
Casual	57.7%	42.3%

TABLE 5 | Top 10 most distinctive words used within the LIWC category Positive Emotion Words with the log-likelihood ratio score (LLR).

Positive emotion		
Dutch	English	LLR
graag	would like	24.07**
genieten	to enjoy	20.10**
eerlijk	honest	12.25**
lief	sweet	7.43*
delen	to share	25.97**
heerlijk	lovely	7.14*
rustig	calm	16.63**
zorgzaam	careful	16.80**
betrouwbaar	trustworthy	20.08**
liefde	love	9.46*

Each of these words were most frequently used by long-term relationship seekers. * $p < 0.01$, ** $p < 0.001$.

classified correctly than those of casual relationship seekers, $\chi^2_{(1)} = 4.79, p = 0.03$. A smaller proportion of long-term texts were incorrectly classified as casual text (38.5%) than vice versa (42.3%).

Distinctive LIWC Category Words

An additional log-likelihood ratio analysis was conducted to examine which words that were part of the six LIWC categories were most distinctive for long-term and casual relationship seekers. This log-likelihood ratio analysis calculated whether the relative frequency of a word within a LIWC category was significantly higher in profile texts of long-term or casual relationship seekers, focusing on the three broader LIWC categories Body & Sexuality, Status, and Positive Emotion, with a particular interest for the Positive Emotion category as for this category a significant different LIWC score between the two relationship groups was found. **Table 5** presents the ten most distinctive words that are part of the Positive Emotion LIWC category in which profile texts of long-term and casual relationship seekers were distinctive. Considering that long-term relationship seekers were more prolific users of positive emotion words, it may not be surprising that the top ten of most distinctive words within this category were all words used more frequently by long-term relationship seekers.

For the other two broader LIWC categories, for which no significant difference in use was found (i.e., Body & Sexuality and Status), relatively few words in that category were found to be distinctive for either of the relationship groups. The few words within the Status word category that were used relatively frequent

and which were used significantly more by long-term than by casual relationship seekers were “work,” “job,” and “company.” However, those words in the category Body & Sexuality that were used significantly more often by long-term or casual relationship seekers occurred so infrequently which makes it difficult to interpret (e.g., “sexy” was used 11 times in the total of 10,969 long-term texts and 9 times in the 1,614 casual texts).

DISCUSSION

The main aim of the current study was to examine whether dating intentions of online daters are reflected in the language use they employ in their dating profiles and, if so, which linguistic elements are important when trying to distinguish between profile texts written by people with two different relationship goals: on the one hand, online daters aiming for a high-involved, long-term relationship, on the other hand, online daters who indicated to look for a casual, lower-involved relationship. This first research goal was addressed by textually analyzing 12,310 existing dating profiles from a dating site.

Two types of analyses were applied to the texts, both operating on word-count basis but pursuing different approaches: LIWC has a top-down approach, in which predefined categories have been manually compiled and validated by scholars, enabling us to set up and test predefined hypotheses. The word-based classifier, on the other hand, uses a data-driven approach, where word n-grams from the profile texts are taken as the linguistic input and provides distinctive content-specific features for both text groups. The secondary goal of this study was to determine to what extent it is effective to add a word-based classifier to LIWC, so far the most commonly used computerized program for profile text analysis.

Linguistic Behavior of Long-Term Relationship Seekers

Results of LIWC indicated that long-term and casual relationship seekers behave linguistically different in their profile texts on some of the linguistic categories investigated: long-term relationship seekers seem to use more positive emotion words and references to the self. Likewise, the word-based classifier provided distinctive content-specific features for the two relationship seeking groups. These content-specific features were relatable to the findings examined by LIWC, especially those extracted from the texts of long-term relationship seekers.

First of all, LIWC showed that long-term relationship seekers use more positive emotion words than casual relationship seekers. Presumably, such positive emotion words are more important to mention when looking for a long-term relationship because they emphasize long-term internal qualities. In addition, words such as “trustworthy,” “honest,” and “serious,” all words distinctive for the texts of long-term relationship seekers, highlight personality traits particularly valued in a long-term relationship partner. For those people looking for a long-term relationship, trust is considered an important factor when determining dating desirability and likelihood of contacting (Wotipka and High, 2016). Mentioning these words gives an

indication of what a profile owner finds important, either as an own personality trait or as a trait sought for in a potential partner (Van Berlo and Ranzini, 2018). Some of these distinctive content-specific features of long-term texts provided by the word-based classifier matched with the words considered distinctive for long-term relationship seekers within the Positive Emotion LIWC category. The log-likelihood ratio analysis presented positive emotion words such as “sweet” and “careful” to be distinctive for long-term relationship texts. This analysis for the words in the Status category also revealed that long-term relationship seekers tend to use the words “work,” “job,” and “company” more often than casual relationship seekers, which can give a tentative indication of long-term relationship seekers valuing resources more in their relationships than casual relationship seekers. Taken together, this suggests that online daters looking for a high-involved, long-term relationship act strategically by emphasizing personal attributes that are (considered to be) important when being involved in a long-term relationship (e.g., Gibbs et al., 2006; Eastwick and Finkel, 2008).

At the same time, profile owner’s language use seems to be affected by the intended relationship goal online daters have in mind, without them being aware of this. LIWC found long-term relationship seekers to use more I-references and also the word-based classifier provided “my (profile)” and “I” as distinctive features for texts of long-term relationship seekers. These findings are not line in with our hypothesis, which stated that casual relationship seekers use more self-references because of higher self-focus and self-involvement. One of the explanations for this contrasting finding may be that using I-references can also be interpreted as a sign of increased levels of self-disclosure, where self-disclosure promotes intimacy and closeness (Laurenceau et al., 1998; Slatcher et al., 2008). Serious relational objectives, such as relationships with higher chances of leading to offline encounters, can affect the type and amount people self-disclose in their self-presentation (Gibbs et al., 2006; Ranzini and Lutz, 2017). Since long-term relationship seekers are more tempted to involve in intimate and close relationships, it may perhaps not be surprising that they self-disclose more by using more I-references. The fact that long-term relationship seekers wrote on average longer profile texts supports the assumption that long-term relationship seekers engage in higher levels of self-disclosure, by being more honest, providing more personal information and making more conscious and intentional disclosures to others online (Gibbs et al., 2006; Toma and Hancock, 2011). Casual relationship seekers are in general considered to be more impersonal and have a lower need to invest in others, which may lead them to disclose less information about themselves and asking for less information about others (Gibbs et al., 2006; Toma and Hancock, 2011). Higher levels of self-disclosure may help as a mean to reduce uncertainty, with people who believe there will be future (offline) contact relying more heavily on more information and reduced uncertainty (Gibbs et al., 2006).

Taken together, results of both LIWC and the word-based classifier seem to imply that during the profile writing process, long-term relationship seekers’ linguistic behavior is affected both consciously and unconsciously. On the one hand, long-term

relationship seekers consciously mention topics and internal characteristics and qualities that become more important over the long-term. On the other hand, they are distinctive in their use of I-references, pronouns that are automatically produced. Conscious and unconscious processes affecting language use in parallel corresponds with earlier indications in Toma and Hancock's (2012) study on deceptive linguistic behavior. They argued that lying profile owners make strategic decisions about what topics to mention, by using more work- and achievement-related words to emphasize more truthful aspects of themselves, but also use fewer self-references and more negations because of an unconscious process of psychological distancing from the deceptive behavior.

Linguistic Behavior of Casual Relationship Seekers

All findings discussed so far concerned profile texts of long-term relationship seekers. It is, however, much more challenging to get a grip on the content of profile texts written by casual relationship seekers. Identifying patterns among the distinctive features provided by the word-based classifier for texts of casual relationship seekers was fairly difficult. This raises the question what casual relationship seekers do write about in their profiles, as they seem to contain less self-disclosing and personal information. Presumably, profile texts of casual relationship seekers have less in common and are more diffuse. Their profiles may be less in-depth and remain more on the surface. Perhaps, the fact that casual relationship seekers do not necessarily have long-term perspectives, makes them less motivated to self-disclose.

The word-based classifier has shown that it was harder to classify the texts of casual relationship seekers than those of long-term relationship seekers. This corresponds with the earlier indication that profile texts of casual relationship seekers share less commonalities and are more diffuse. In addition to that, when looking at the texts of profile owners with casual relationship goals, it appears that fewer of these distinctive content-specific features are important for explaining the variance between the two groups of texts, compared to texts written by those looking for a long-term relationship partner. To illustrate, "to send," the tenth most important distinctive feature for the texts of casual relationship seekers had an importance score of 0.005, compared to 0.022 for the tenth word "to enjoy" for the long-term relationship texts. This indicates that relatively few words were meaningful to classify casual texts as such, with the exception of the word "date," with an importance score of 0.087.

LIWC and the Word-Based Classifier

The second goal of this study was to determine what a data-driven word-based classifier would add to LIWC when analyzing dating profile texts. Both classification methods deal differently with linguistic input and use other vocabulary approaches; the word-based classifier uses an open-vocabulary and LIWC a closed-vocabulary approach. In our study, we see two main advantages of adding a word-based classifier to the LIWC analysis. First, the classifier finds words relevant

within the domain of online dating that do not occur in the predefined LIWC word list. In our case, four from the twenty content-specific features provided by the word-based classifier in **Table 3** were not part of the LIWC dictionary (i.e., "date," "feel like," "my profile," "serious" (adjective)). This shows, that although the word list of LIWC seems to capture most of the content-specific features that distinguishes texts of casual and long-term relationship seekers, some important distinctive content-specific features were uncovered by the classifier. Second, while LIWC offered a broader perspective, reflective of underlying psychological processes that may have influenced profile owners' language use, the word-based classifier zoomed in more specifically on the content-specific features within a category that may be important to indicate differences between profile texts of the two relationship seeking groups. For example, as LIWC showed long-term relationship seekers to use more positive emotion words, a category that contains a vast amount of words all related to positive emotions, the word-based classifier identified the content-specific features that may play an important part within this broad category of positive emotions (i.e., "trustworthy" and "honest").

The size of the word lists in some LIWC categories makes it difficult to interpret differences between two groups of texts. Therefore, an additional log-likelihood ratio analysis was conducted, which examined which words within the LIWC categories were used significantly more by long-term or casual relationship seekers. This analysis showed comparable results with the distinctive content-specific features provided by the word-based classifier, particularly for the long-term relationship seeking group. Some words that were distinctive for long-term relationship seekers within the Positive Emotion category were the same as the distinctive content-specific features of long-term relationship seekers provided by the word-based classifier ("to enjoy," "honest," "calm," and "trustworthy"). This indicates that similar words may have driven linguistic differences between the two relationship groups within this LIWC category as for the word-based classifier. Positive emotion words that were distinctive for casual relationship texts within this category (e.g., "adventurous," "festivals," and "adventure") can be an indication of their more casual attitude toward dating and relationships, and can be associated with distinctive content-specific features like "date," "feel like," and "to eat" provided by the word-based classifier. While there is a considerable degree of overlap between the distinctive LIWC category words and the content-specific features of the word-based classifier for the long-term relationship seekers, this is not the case for the other group of texts. This can imply that positive emotion words are not the words through which casual relationship seekers distinguish themselves from long-term relationship seekers. This corresponds with the relatively low frequency scores of these distinctive positive emotion words (e.g., "adventurous" and "festival" occurred respectively 50 and 34 times in casual relationship texts). Moreover, that long-term and casual relationship seekers showed minimal differences in use of words within the LIWC categories Body & Sexuality and Status corresponds with the fact that no body, sexuality, and status words were part of the content-specific feature list of the

word-based classifier. Overall, results of the two methods reveal similar patterns, with these patterns being more pronounced in texts of long-term relationship seekers than in those of people looking for a more casual relationship.

Suggestions for Future Work

Although most profile owners follow dating sites' guidelines and hold on to emerged conventions among profile owners, there are still many different ways of constructing a profile text. It may therefore not be surprising that the effect sizes are rather small and that the classifier predicts accurately only around 10% above chance. Our results are observed on the basis of a large collection of dating profiles and although there seem to be indications of linguistic differences between the two groups of texts, we cannot make strong claims about the extent to which an online dater's relationship goal can be immediately derived from a linguistic analysis of this person's profile text. The use of a specific textual element is not restricted to only one relationship type category: texts of casual relationship seekers do contain features distinctive for texts of long-term relationship seekers and vice versa. To examine whether individuals are able to identify profile owners' relationship goals based on the profile texts that are written, a more experimental, controlled experiment should be conducted with human judges.

Moreover, it would be interesting to investigate what combination of linguistic and other (visual) cues in dating profiles affect online daters' behaviors and what combination of cues are effective to achieve the intended relationship goal. As the current data set did not include any information about the number of contact messages in response to each dating profile, it was impossible to investigate the profiles' success and effectiveness. However, it would be interesting to include this, so as to examine how the way profile owners textually present themselves plays a role in establishing desired romantic relationships. In addition, follow-up research could further examine the collocations with which particular words occur to provide how and in what context these words are more likely to be used (e.g., "I" from a self-centered or self-disclosing perspective or "honest" as an own personality trait or a desired trait in a partner). A next step could then be to

investigate whether linguistic differences between relationship seeking groups can also be observed at propositional, sentence, or text levels, for example, in verb argument constructions, question and declarative sentence constructions, and topics that are mentioned. With a better understanding of preferences of online daters who seek either a long-term or casual relationship, it could be possible for profile owners to increase chances of finding the intended relationship partner by adjusting the profile (texts) based on the relationship goal.

CONCLUSION

Results of LIWC and the word-based classifier in this study suggest that relationship goals can leave linguistic traces in the textual component of the dating profile, including both strategic linguistic cues as well as unconsciously leaked cues. As far as we are aware, this is the first study which shows that online dating profile owners differ in how they linguistically present themselves dependent on the type of relationship sought. Although the present study shows that intentions and relationship goals of online daters can affect language use in dating profiles, additional empirical work is necessary to shed better light on the effects of writers' intentions and goals on linguistic behavior.

AUTHOR CONTRIBUTIONS

TvdZ, AS, MM, and EK conceived the original idea and developed the theoretical framework. TvdZ took the lead in conducting the analyses and writing the manuscript and was supported by AS and MM, who supervised during the project. CvdL worked out most of the machine learning details. Both EK and CvdL contributed to the final version of the manuscript. All authors discussed the results, provided critical feedback and helped shape the research, analysis and manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2019.00022/full#supplementary-material>

REFERENCES

- Aggarwal, C. C., and Zhai, C. (2012). "A survey of text classification algorithms," in *Mining Text Data*, eds C. C. Aggarwal and C. Zhai (Boston, MA: Springer), 163–222. doi: 10.1007/978-1-4614-3223-4_6
- Agnew, C. R., Van Lange, P. A., Rusbult, C. E., and Langston, C. A. (1998). Cognitive interdependence: commitment and the mental representation of close relationship. *J. Pers. Soc. Psych.* 74, 939–954. doi: 10.1037/0022-3514.74.4.939
- Akehurst, J., Koprinska, I., Yacef, K., Pizzato, L., Kay, J., and Rej, T. (2011). "Explicit and implicit user preferences in online dating," in *Proceedings of 15th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining*, Berlin: Springer, 15–27. doi: 10.1007/978-3-642-28320-8_2
- Boot, P., Zijlstra, H., and Geenen, R. (2017). The dutch translation of the Linguistic Inquiry and Word Count (LIWC) 2007 dictionary. *Dutch J. Appl. Ling.* 6, 65–76. doi: 10.1075/djtal.6.1.04boo
- Boyd, R. L. (2017). "Psychological Text Analysis in the Digital Humanities," in *Data Analytics in Digital Humanities, Multimedia Systems and Applications*, ed S. Hai-Jew (New York, NY: Springer), 161–189. doi: 10.1007/978-3-319-54499-1_7
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*. Monterey, CA: Wadsworth Publishing.
- Buunk, B. P., Dijkstra, P., Fetchenhauer, D., and Kenrick, D. T. (2002). Age and gender differences in mate selection criteria for various involvement levels. *Pers. Relatsh.* 9, 271–278. doi: 10.1111/1475-6811.00018
- Chan, L. S. (2017). Who uses dating apps? Exploring the relationships among trust, sensation-seeking, smartphone use, and the intent to use dating apps based on the integrative model. *Comput. Hum. Behav.* 72, 246–258. doi: 10.1016/j.chb.2017.02.053
- Clark, H. (1996). *Using Language*. Cambridge, NY: Cambridge University Press. doi: 10.1017/CBO9780511620539

- Clemens, C., Atkin, D., and Krishnan, A. (2015). The influence of biological and personality traits on gratifications obtained through online dating websites. *Comput. Hum. Behav.* 49, 120–129. doi: 10.1016/j.chb.2014.12.058
- Davis, E. M., and Fingerman, K. L. (2016). Digital dating: online profile content of older and younger adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 71, 959–967. doi: 10.1093/geronb/gbv042
- Diaz, F., Metzler, D., and Amer-Yahia, S. (2010). “Relevance and ranking in online dating systems, in *Proceedings of the 33rd international ACM SIGIR conference on Research and Development in Information Retrieval*, (New York, NY: ACM), 66–73. doi: 10.1145/1835449.1835463
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Comput. Linguist.* 19, 61–74.
- Eastwick, P. W., and Finkel, E. J. (2008). Sex differences in mate preferences revisited: do people know what they initially desire in a romantic partner? *J. Pers. Soc. Psych.* 94, 245–264. doi: 10.1037/0022-3514.94.2.245
- Eastwick, P. W., Keneski, E., Morgan, T. A., McDonald, M. A., and Huang, S. A. (2018). What do short-term and long-term relationships look like? Building the relationship coordination and strategic timing (ReCAST) model. *J. Exp. Psychol. Gen.* 147, 747–781. doi: 10.1037/xge0000428
- Ellison, N. B., Hancock, J. T., and Toma, C. L. (2012). Profile as promise: a framework for conceptualizing veracity in online dating self-presentations. *N. Media Soc.* 14, 45–62. doi: 10.1177/1461444811410395
- Gallant, S., Williams, L., Fisher, M., and Cox, A. (2011). Mating strategies and self-presentation in online personal advertisement photographs. *J. Soc. Evol. Cult. Psychol.* 5, 106–121. doi: 10.1037/h0099272
- Gallois, C., Ogay, T., and Giles, H. (2005). “Communication accommodation theory: a look back and a look ahead,” in *Theorizing About Intercultural Communication*, ed W. B. Gudykunst (Thousand Oaks, CA: Sage), 121–148.
- Gibbs, J. L., Ellison, N. B., and Heino, R. D. (2006). Self-presentation in online personals: the role of anticipated future interaction, self-disclosure, and perceived success in Internet dating. *Communic. Res.* 33, 152–177. doi: 10.1177/0093650205285368
- Giles, H., Coupland, N., and Coupland, J. (1991). “Accommodation theory: Communication, context, and consequence,” in *Contexts of Accommodation*, eds H. Giles, J. Coupland, N. Coupland (New York, NY: Cambridge University Press), 1–68. doi: 10.1017/CBO9780511663673.001
- Gill, A. J., French, R. M., Gergle, D., and Oberlander, J. (2008). “The language of emotion in short blog texts,” in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. (New York, NY: ACM), 299–302. doi: 10.1145/1460563.1460612
- Groom, C. J., and Pennebaker, J. W. (2005). The language of love: sex, sexual orientation, and language use in online personal advertisements. *Sex Roles* 52, 447–461. doi: 10.1007/s11199-005-3711-0
- Hancock, J. T., Landrigan, C., and Silver, C. (2007). “Expressing emotion in text-based communication,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (New York, NY: ACM), 929–932. doi: 10.1145/1240624.1240764
- Laurenceau, J. P., Barrett, L. F., and Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *J. Pers. and Soc. Psych.* 74, 1238–1251. doi: 10.1037/0022-3514.74.5.1238
- Li, N. P., and Kenrick, D. T. (2006). Sex similarities and differences in preferences for short-term mates: what, whether, and why. *J. Pers. and Soc. Psych.* 90, 468–489. doi: 10.1037/0022-3514.90.3.468
- Lo, S. K., Hsieh, A. Y., and Chiu, Y. P. (2013). Contradictory deceptive behavior in online dating. *Comput. Hum. Behav.* 29, 1755–1762. doi: 10.1016/j.chb.2013.02.010
- Luo, S., and Tuney, S. (2015). Can texting be used to improve romantic relationships?—The effects of sending positive text messages on relationship satisfaction. *Comput. Hum. Behav.* 49, 670–678. doi: 10.1016/j.chb.2014.11.035
- McAdams, D. P., Healy, S., and Krause, S. (1984). Social motives and patterns of friendship. *J. Pers. Soc. Psych.* 47, 828–838. doi: 10.1037/0022-3514.47.4.828
- Nguyen, D., Gravel, R., Trieschnigg, D., and Meder, T. (2013). “How old do you think I am?” A study of language and age in Twitter,” in *Proceedings of the Seventh International Conference on Weblogs and Social Media*. (Palo Alto, CA: AAAI Press), 857–862.
- Paltoglou, G., and Thelwall, M. (2012). Twitter, MySpace, Digg: Unsupervised sentiment analysis in social media. *ACM Trans. Intell. Syst. Technol.* 3:66. doi: 10.1145/2337542.2337551
- Pennebaker, J. W. (2011). *The Secret Life of Pronouns: What Our Words Say About Us*. New York, NY: Bloomsbury Press.
- Pennebaker, J. W., Booth, R. J., Boyd, R. L., and Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC 2015 [Computer software]*. Austin, TX: Pennebaker Conglomerates (www.LIWC.net).
- Pennebaker, J. W., and King, L. A. (1999). Linguistic styles: language use as an individual difference. *J. Pers. and Soc. Psych.* 77, 1296–1312. doi: 10.1037/0022-3514.77.6.1296
- Pennebaker, J. W., and Stone, L. D. (2003). Words of wisdom: language use over the life span. *J. Pers. Soc. Psych.* 85, 291–301. doi: 10.1037/0022-3514.85.2.291
- Peter, J., and Valkenburg, P. M. (2007). Who looks for casual dates on the internet? A test of the compensation and the recreation hypotheses. *N. Media Soc.* 9, 455–474. doi: 10.1177/1461444807076975
- Ranzini, G., and Lutz, C. (2017). Love at first swipe? Explaining Tinder self-presentation and motives. *Mob. Media Commun.* 5, 80–101. doi: 10.1177/2050157916664559
- Regan, P. C., Levin, L., Sprecher, S., Christopher, F. S., and Gate, R. (2000). Partner preferences: what characteristics do men and women desire in their short-term sexual and long-term romantic partners? *J. Psychol. Human Sex.* 12, 1–21. doi: 10.1300/J056v12n03_01
- Rosen, L. D., Cheever, N. A., Cummings, C., and Felt, J. (2008). The impact of emotionality and self-disclosure on online dating versus traditional dating. *Comput. Human Behav.* 24, 2124–2157. doi: 10.1016/j.chb.2007.10.003
- Ross, S. (2017). *This is Officially the Most Popular Way People are Meeting Their Spouse*. Retrieved from: <https://www.theknot.com/content/online-dating-most-popular-way-to-meet-spouse> (Retrieved May 13, 2019).
- Russell, A. W., and Schober, M. F. (1999). How beliefs about a partner’s goals affect referring in goal-discrepant conversations. *Discour. Proces.* 27, 1–33. doi: 10.1080/01638539909545048
- Sanderson, C. A., Keiter, E. J., Miles, M. G., and Yopyk, D. J. (2007). The association between intimacy goals and plans for initiating dating relationships. *Pers. Relatsh.* 14, 225–243. doi: 10.1111/j.1475-6811.2007.00152.x
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., et al. (2013). Personality, gender, and age in the language of social media: the open-vocabulary approach. *PLoS ONE* 8:e73791. doi: 10.1371/journal.pone.0073791
- Sharabi, L. L., and Dykstra-DeVette, T. A. (2019). From first email to first date: strategies for initiating relationships in online dating. *J. Soc. Pers. Relat.* 1–19. doi: 10.1177/0265407518822780
- Slatcher, R. B., Vazire, S., and Pennebaker, J. W. (2008). Am “I” more important than “we”? Couples’ word use in instant messages. *Pers. Relatsh.* 15, 407–424. doi: 10.1111/j.1475-6811.2008.00207.x
- Stephure, R. J., Boon, S. D., MacKinnon, S. L., and Deveau, V. L. (2009). Internet initiated relationships: Associations between age and involvement in online dating. *J. Comput. Mediat. Commun.* 14, 658–681. doi: 10.1111/j.1083-6101.2009.01457.x
- Sternberg, R. J. (1986). A triangular theory of love. *Psychol. Rev.* 93, 119–135. doi: 10.1037/0033-295X.93.2.119
- Tausczik, Y. R., and Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* 29, 24–54. doi: 10.1177/0261927X09351676
- Tay, Y., Tuan, L. A., and Hui, S. C. (2018). “CoupleNet: Paying Attention to Couples with Coupled Attention for Relationship Recommendation,” in *Proceedings of the Twelfth International AAAI Conference on Web and Social Media*, (Palo Alto, CA: AAAI Press).
- Toma, C. L. (2014). Towards conceptual convergence: an examination of interpersonal adaptation. *Commun. Q.* 62, 155–178. doi: 10.1080/01463373.2014.890116
- Toma, C. L., and D’Angelo, J. D. (2017). “Connecting profile-to-profile: how people self-present and form impressions of others through online dating profiles,” in *The Impact of Social Media in Modern Romantic Relationships*, eds N. M. Puyanant-Carter and J. S. Wrench (Lanham, MD: Lexington), 147–162.
- Toma, C. L., and Hancock, J. T. (2011). “A New Twist on Love’s Labor: Self-Presentation in Online Dating Profiles,” in *Computer-Mediated*

- Communication in Personal Relationships*, eds K. B. Wright and L. M. Webb (New York, NY: Peter Lang), 41–55.
- Toma, C. L., and Hancock, J. T. (2012). What lies beneath: the linguistic traces of deception in online dating profiles. *J. Commun.* 62, 78–97. doi: 10.1111/j.1460-2466.2011.01619.x
- Van Berlo, Z., and Ranzini, G. (2018). “Big Dating: a Computational Approach to Examine Gendered Self-Presentation on Tinder,” in *Proceedings of the 9th International Conference on Social Media and Society* (New York, NY: ACM), 390–394. doi: 10.1145/3217804.3217951
- Van den Bosch, A., Busser, G. J., Daelemans, W., and Canisius, S. (2007). “An efficient memory-based morphosyntactic tagger and parser for Dutch,” in *Selected Papers of the 17th Computational Linguistics in the Netherlands Meeting*, eds F. van Eynde, P. Dirix, I. Schuurman, and V. Vandeghinste (Leuven, BE: CLIN), 99–114.
- Van der Lee, C., and Van den Bosch, A. (2017). “Exploring lexical and syntactic features for language variety identification,” in *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects* (Stroudsburg, PA: ACL Anthology), 190–199. doi: 10.18653/v1/W17-1224
- Van der Zanden, T., Mos, M., and Schouten, A. (2018). Taalaccommodatie in online datingprofielen: effecten van opleidingsniveau en type datingsite op taalgebruik [Language accommodation in online dating profiles: effects of education level and type of dating site on language use]. *Tijdschr. Taalbeheers.* 40, 83–106. doi: 10.5117/TVT2018.1.zand
- Van Wissen, L., and Boot, P. (2017). “An Electronic Translation of the LIWC Dictionary into Dutch,” in *Electronic Lexicography in the 21st Century: Proceedings of eLex 2017 Conference* (Brno: Lexical Computing CZ), 703–715.
- Walther, J. B. (2007). Selective self-presentation in computer-mediated communication: hyperpersonal dimensions of technology, language, and cognition. *Comput. Hum. Behav.* 23, 2538–2557. doi: 10.1016/j.chb.2006.05.002
- Weidman, A. C., Cheng, J. T., Chisholm, C., and Tracy, J. L. (2015). Is she the one? Personality judgments from online personal advertisements. *Pers. Relatsh.* 22, 591–603. doi: 10.1111/pere.12097
- Whitty, M. T. (2008). Revealing the ‘real’ me, searching for the ‘actual’ you: presentations of self on an internet dating site. *Comput. Human Behav.* 24, 1707–1723. doi: 10.1016/j.chb.2007.07.002
- Wotipka, C. D., and High, A. C. (2016). An idealized self or the real me? Predicting attraction to online dating profiles using selective self-presentation and warranting. *Commun. Monogr.* 83, 281–302. doi: 10.1080/03637751.2016.1198041
- Yarkoni, T. (2010). Personality in 100,000 words: a large-scale analysis of personality and word use among bloggers. *J. Res. Pers.* 44, 363–373. doi: 10.1016/j.jrp.2010.04.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 van der Zanden, Schouten, Mos, van der Lee and Kraemer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.