Language Variation in Online Personal Ads from Quebec: The Case of ne

Rémi A. van Compernolle
University of North Texas

Abstract

This article analyzes the distribution and variable deletion of French ne, the first marker of verbal negation, in a corpus of online personal advertisements from Quebec. While ne deletion rates are nearly categorical in informal spoken Canadian French, ne use remains strong in online personals. At the same time, online dating site members productively demonstrate variation in their personals, in that they delete ne in over 20% of variable contexts. VARBRUL analyses reveal that the variable presence of ne is conditioned by the age of the advertiser and the presence vs. absence of second-person address. The results are discussed within the broader context of the progressive loss of ne in French as well as self-presentation and audience design in online communication contexts.

Introduction

Verbal negation in Modern French is characterized by its two "embracing" particles. The first particle (ne, n') precedes a conjugated verb, and one or more negative complements (e.g., pas 'not', rien 'nothing', jamais 'never') follows, as in examples (1) and (2).1 However, pro-clitic (i.e., pre-verbal) ne is often omitted in informal or everyday conversational French, leaving the post-verbal negative complement as the sole negator, as in examples (3) and (4).

1. Paul ne veut pas venir.
   'Paul does not want to come.'
2. tu n'aimes pas le chocolat?
   'Do you not like chocolate?'
3. Paul Ø veut pas venir.
   'Paul doesn't want to come.'
4. tu Ø aimes pas le chocolat?
   'Don't you like chocolate?'

While prescriptive grammars and learner textbooks present and portray two-particle negation, such as in (1) and (2), as the "right way" to negate a verb, several recent corpus-based studies have shown that ne is very frequently deleted throughout much of France (Armstrong, 2002; Ashby, 1981, 2001; Coveney, 1996; Hansen & Malderez, 2004; Moreau, 1986). In addition, nearly categorical ne deletion has been observed in large corpora of spoken Montreal French (Sankoff & Vincent, 1977 [1980]), Ottawa-Hull French (Poplack & St-Amand, 2007), and Swiss French (Fonseca-Greber, 2000, 2007). Armstrong and Smith (2002) have even reported that ne deletion is spreading to highly monitored speech contexts, such as radio interviews with

Language@Internet, 5 (2008), article 1. (www.languageatinternet.de, urn:nbn:de: 0009-7-13486, ISSN 1860-2029)
politicians, although *ne* retention rates in formal settings remain extremely high when compared to everyday conversational French.

The ultimate fate of *ne*, however, is unknown at this point in time. Some researchers have suggested that the declining rates of *ne* retention reported in the literature imply a change in progress, and that, upon completion of that change, *ne* will disappear from the French language altogether (Ashby, 1981, 2001). Other scholars have argued that the variation is stable, and that *ne* continues to be an important sociolinguistic resource for speakers (Hansen & Malderez, 2004). While literature on *ne* deletion in spoken French is abundant, there exist relatively few comparisons of variable *ne* presence vs. absence as it occurs in writing and electronic communication. Such comparisons have the potential to provide insight into the spread of *ne* deletion in these and other types of under-studied discourse.

The present study provides a brief discussion of variable *ne* realization in a corpus of online personal advertisements from Quebec, focusing specifically on the effect of selected social and linguistic factors: age of advertiser, gender, and choice of address pronoun. Variable rule (VARBRUL) analyses reveal that age and choice of address pronoun have significant main effects on *ne* presence versus absence, while gender is a nonsignificant factor. Partitioned analyses considering each age group (i.e., 18-25 years, 26-35 years, and 36-45 years) as an independent data set provide a more nuanced interpretation of the results. In concluding, the somewhat surprising finding that dating site members productively demonstrate *ne* variation is discussed within the broader context of self-presentation, identity, and language style online.

**Literature Review**

*Previous Research on Internet Dating*

Online dating sites have received increasing attention from social scientists and communication researchers over the past few years (Ben-Ze'ev, 2004; Ellison, Heino, & Gibbons, 2006; Fiore & Donath, 2004; Gudelunas, 2005; Groom & Pennebaker, 2005; Hancock, Toma, & Ellison, 2007; Hardey, 2004; Hollander, 2004; Sahib, Koning, & van Witteloostuijn, 2006; Smith & Stillman, 2002; Whitty & Carr, 2006; Yurchisin, Watchravesringkan, & McCabe, 2005). The Internet allows communication to occur relatively quickly and efficiently between people in spite of various geographical and time constraints; this has broad implications for the development and maintenance of online relationships. Ben-Ze'ev (2004, p. 7) notes that increased levels of synchronicity help emotions develop more quickly, since interested parties can send and respond to messages rapidly through electronic mail or real-time chat.

Much of the literature on online dating has focused on self-presentation and potential deception on dating sites (e.g., Ellison et al., 2006; Hancock et al., 2007). This is in line with research concerning relationship maintenance and self-presentation in computer-mediated environments more generally (e.g., Bargh, McKenna, & Fitzsimons, 2002; McLaughlin, Osbourne, & Ellison, 1997; Parks & Floyd, 1996; Roberts & Parks, 1999; Utz, 2000). While similarities exist between
dating sites and other forms of CMC (e.g., discussion fora, chat, blogs), Ellison et al. (2006) note that "the online dating forum is qualitatively different from many other online settings due to the anticipation of face-to-face interaction inherent in this context (Gibbs, Ellison, & Heino, 2006) and the fact that social practices are still nascent" (p. 416).

Ellison et al.'s (2006) study found that online dating participants attempt to circumvent the limitations of the technology (e.g., the lack of nonverbal information), while simultaneously taking advantage of the capacities of the online medium in order to present their "ideal" self. At the same time, Ellison et al.'s informants reported that the anticipation of face-to-face interaction leads them "to balance their desire for self-promotion with their need for accurate self-presentation" (p. 430). By the same token, participants in Ellison et al.'s study were weary of the potential for deception in the profiles of other members, leading them to develop strategies for determining the credibility of self-presentations. These online dating site members reported adopting "various strategies to demonstrate the credibility of their [own] identity claims, recursively applying the same techniques they employed to uncover representational ruses in others" (pp. 430-431).

Only a limited number of researchers have undertaken analyses of the discourse of online personal ads (e.g., Groom & Pennebaker, 2005), and still fewer have done so from a sociolinguistic perspective (e.g., van Compernolle, forthcoming a, forthcoming c). Groom and Pennebaker (2005) explored the language used by online dating site members, comparing men's and women's personals along rather stereotypical lines. For example, they found that men focused more on physical traits, money and possessions, and personal success, while women were more concerned with personalities and building interpersonal relationships. These results are in line with what has been shown previously in analyses of print personal ads (e.g., Koestner & Wheeler, 1988; Smith & Stillman, 2002).

The vast majority of research on online dating has dealt with speakers of English. Thus, little is known about practices used by online dating participants' who are speakers of other languages. van Compernolle (forthcoming a) analyzed selected discursive and pragmatic norms in a corpus of French-language online personals from men and women from Quebec. He found five types of recurring information/utterances: 1) greeting; 2) purpose; 3) self-promotion; 4) description of partner; and 5) invitation to contact. Self-promotion was by far the most often occurring type of information, followed by invitations to contact and greetings. Descriptions of the ideal partner were most frequent in young men's personals, and a purpose and/or justification for joining an online dating site was most frequently provided by older members (especially women). It was also found that (younger) men were more likely to include a description of their (ideal) partner than were women, leading van Compernolle to note that (young) men may be seen not only as advertisers of a product (i.e., themselves) but also as discerning consumers of a product (i.e., potential partners). van Compernolle also briefly explored address strategies, finding that singular second-person address was preferred in general. However, young men tended to use plural address at high frequencies. van Compernolle concluded that this may be indicative of young men's desire to "play the field" and to correspond with several (or many) potential partners.
simultaneously. This result is in line with Groom and Pennebaker's (2005) finding that young men tended to contact many women indiscriminately, while women preferred to correspond with one man at a time.

van Compernolle (forthcoming c) conducted a qualitative analysis of the pragmatics of second-person address in online personals from Quebec, from which he presents four conclusions. First, the presence of second-person address is indicative of online dating participants' perception of the personal ad as a form of interpersonal communication, as they are addressing their (potential) reader(s) directly. Second, singular address (i.e., tu) is used in the anticipation of one-to-one private communication, where the use of vous or other plural address form could lead to linguistic ambiguity and/or be seen as sociopragmatically inappropriate. Third, the use of plural address indicates that at least some members exploit the one-to-many participation framework in order to contact the greatest number of potential partners. Fourth, second-person pronouns (familiar tu, formal vous-singular or vous-plural) are used strategically as a means of either creating familiarity with one's reader (i.e., by using tu) or avoiding an overly informal style of discourse (i.e., by using vous-singular or vous-plural).^4

**Variable 'ne' Deletion in Spoken French**

Variable ne deletion has been extensively studied in spoken French. Although most research has drawn from corpora of metropolitan French (e.g., Armstrong, 2002; Ashby, 1981, 2001; Coveney, 1996; Hansen & Malderez, 2004), other scholars have reported on variable ne in Montreal French (Sankoff & Vincent, 1977 [1980]), Ottawa-Hull French (Poplack & St-Amand, 2007), and Swiss French (Fonseca-Greber, 2000, pp. 297-305, 2007). The present study is principally concerned with variable ne realization in Quebec French, a variety in which ne deletion has become almost categorical, at least in everyday conversation.

In the first large-scale study of this type of sociolinguistic variation in French, Sankoff and Vincent (1977 [1980]) investigated the productive use of ne in a corpus of Montreal French. The data were collected from sociolinguistic interviews with 60 informants—evenly distributed across demographic categories (e.g., gender, age, social class)—recorded in 1971 (Sankoff-Cedergren corpus). Sankoff and Vincent reported that ne was retained in only 0.5% of verbal negations (N ~ 10,000), and that 45 of the 60 speakers categorically deleted ne during their interviews (1980). No clear pattern of variation was found across demographic groups, suggesting that near categorical ne deletion has become more or less generalized in Montreal. The highest ne retention rate reported was in the speech of a 61-year-old woman from a wealthy business family, who had spent several years in France as a student. Nonetheless, she retained ne in only 8% of verbal negations. Occurrences of two-particle negation seemed to crop up during the discussion of serious topics, such as religion, the punishment of children, the telling of proverbs, and language, leading Sankoff and Vincent to conclude that ne presence functions as a stylistic feature of discourse.
Poplack and St-Amand's (2007) study offers insight into the origins of ne deletion in Canadian French, drawing from the Récits du français québécois d'autrefois (RFQ)—folktales told by rural speakers of French from Quebec born between 1846 and 1895—and another corpus of Ottawa-Hull French constructed in the second half of the 20th century. In the RFQ, ne was retained in only 0.1% of verbal negations (N = 9,438), which suggests that, at least among rural speakers, ne was nearly categorically deleted in 19th century Quebec. In the Ottawa-Hull corpus, ne was retained in only 0.2% of verbal negations (N = 61,316), demonstrating that "no change has taken place with regard to rate" (p. 723).

Following Sankoff and Vincent (1977 [1980]), Poplack and St-Amand (2007) analyzed the rare occurrences of ne in both corpora. While in the RFQ ne tended to be used randomly, tokens of it were concentrated in parts of interviews discussing serious topics and during proverbial-style discourse ("soapbox" style) in the Ottawa-Hull corpus. The researchers concluded that although no change has occurred with respect to frequency of ne retention, a significant change has occurred in the productive use of ne; that is, the negative morpheme has been pragmaticized in 20th century French. Poplack and St-Amand refer to this as "specialized ne" use.

In spoken European French, by way of comparison, ne variation is still considered "productive." Studies of metropolitan French indicate that ne use is more common than in Quebec, although its frequency of occurrence appears to be diminishing. Coveney (1996) conducted sociolinguistic interviews with 30 adult speakers working in several summer camps (colonies de vacances) in Northern France in the mid-1980s. He reported an overall frequency of ne retention of 18.8% (N = 2,932) and presented several important findings with regard to differences among demographic groups. First, younger speakers aged 17 to 22 years retained ne with the lowest frequency (8.4%), while those aged 24 to 37 years retained ne in 23.9% of verbal negations. The oldest informants—all middle-aged women who worked as cooks—retained ne at the highest frequency (28.8%). Coveney does, however, report that one 19-year-old male informant retained ne at an extremely high frequency; he was excluded from the results. In this case, Coveney argues that since this young man was already an assistant camp director, he may have felt the need to represent the camp favorably and therefore retained ne at a high frequency (pp. 83-84). This is in line with previous research on variation as a function of the "linguistic market," which takes into account "specifically how speakers' economic activity, taken in its widest sense, requires or is necessarily associated with, competence in the legitimized language" (Sankoff & Laberge, 1978, p. 239; emphasis in original). In addition, at the beginning of the interview, the 19-year-old assistant director asked whether he should speak more slowly and more distinctly so that English speakers would be able to understand better, and Coveney notes that he did in fact speak in a rather careful style.

Speaker gender was not found to influence variation in Coveney's (1996) study, which is inconsistent with research conducted by Ashby (1981), who found that women were more likely to delete ne than their male counterparts. However, social class (i.e., "working" vs. "intermediate") was reported to be an important differentiating factor, as working class speakers
retained *ne* in only 9.2% of negations, while intermediate class speakers retained it 16.4% of the time. In his conclusion, Coveney suggests that:

> certain aspects of the social differentiation found . . . do not point straightforwardly to a change in progress. What seems more probable is that there is now a pattern of age-grading, whereby each generation of speakers has virtually a zero rate of *ne* retention as children and adolescents, but then as they become older modify their speech, under pressure from, and in the direction of, the written language. (p. 90)

Ashby (2001) conducted a real-time study of *ne* retention in Tours (central France), drawing on his corpus of spoken French from 1976 (see Ashby, 1981) together with a new corpus of recorded interviews of comparable speakers from 1995. Ashby found some evidence of a change in progress, implying the progressive loss of *ne* in metropolitan French, as percentages of *ne* retention fell from 37% in the 1976 corpus to 18% in the 1995 corpus. VARBRUL analyses (using GoldVarb 2.0) were performed in order to test the significance of a number of social factors. In both corpora, social class was found to be a determining factor: Speakers in the upper and intermediate classes favored *ne* retention, while working class speakers disfavored its presence. GoldVarb 2.0 selected the speaker's gender as significant in the 1976 corpus (*ne* retention was favored in the speech of men, but disfavored for women) but not in the 1995 corpus. The results of the 1995 corpus corroborate those reported by Coveney (1996) for speaker gender. Age was selected as significant in both corpora: While *ne* retention rates have fallen across all groups, older speakers' rates remain significantly higher than those of younger speakers. Ashby posits that *ne* is not necessarily in danger of being lost; rather, he writes of a "shift" in progress, whereby lower frequencies of *ne* retention represent an express decision across social classes and ages to use everyday language in the same circumstances that, in 1976, required a higher level of formality (2001; see also Posner, 1997).

Hansen and Malderez (2004) also conducted a real-time study of *ne* deletion in France, drawing from two corpora of recorded interviews with speakers from the Paris region. The first (the Péretz-Juillard corpus) was established in the 1970s, and the Hansen-Malderez corpus included interviews conducted in the late 1980s and early 1990s. Hansen and Malderez found a significant decline in *ne* retention rates, from 15.8% in the 1970s to 8.2% in the 1990s. In line with previous research (e.g., Ashby, 2001, Coveney, 1996), age was the most influential social factor, as the youngest informants—aged 4 to 16 years—categorically omitted *ne*, while the oldest informants—51 to 64 years—retained *ne* 22.3% of the time. A rather straightforward linear progression according to age was found; that is, the older the speaker was, the more frequently he or she retained *ne*.

Level of education, although less influential than age, was also found to be significant, as university educated speakers retained *ne* more frequently than those who had only completed secondary-level coursework. However, while significant drops in *ne* retention rates were found between the 1970s and 1990s, Hansen and Malderez are reluctant to interpret their findings as a positive sign of a change in progress. They note that in the written school assignments of young
children, *ne* was retained with frequencies as high as 75%. Citing Blanche-Benveniste (1997), who reported that young girls were perfectly capable of using *ne* appropriately when imitating the speech of elegant ladies, Hansen and Malderez (2004) argue that *ne* is not disappearing from French; rather, its association with formal discourse, the speech of members of upper levels of society, and the (formal) written language make it an important sociolinguistic resource for all speakers.

*Variable 'ne' Deletion in Written French*

While *ne* deletion in spoken Quebec French is nearly categorical, *ne* deletion has not made its way into the formal written language, such as print publications, in either Canadian or European varieties of French (Martineau & Mougeon, 2003). Moreover, only one previous study (Hansen & Malderez, 2004) has considered *ne* deletion in unpublished writing, in this case, young French children's school assignments. Hansen and Malderez report *ne* retention as high as 75% in some cases, even though the children nearly categorically deleted *ne* in informal speech. However, Coveney (1996) notes that as speakers of French are exposed to more formal education, *ne* retention becomes categorical in most written contexts.

*Variable 'ne' Deletion in Electronic French*

Recently, some research has been published concerning variable *ne* deletion in French CMC environments, which involve writing, but vary in their degree of synchronicity and planning. van Compernolle (2007) conducted a descriptive distributional analysis of *ne* retention versus deletion in a 78,125-word corpus of unmoderated Internet Relay Chat (IRC) collected from two different channels over four non-consecutive days in late 2005. He found that rates of *ne* retention in IRC align with those reported in everyday conversational (European) French (~16%). A qualitative analysis revealed that most tokens of *ne* occurred during ludic or emphatic contexts. Ludic contexts included jokes, role-playing, imitations, and quoting; emphatic contexts were associated with oppositions and contrasts or found during arguments and disputes. When emphatic and ludic tokens of *ne* were excluded from the statistics, *ne* retention fell to only 4% in the corpus. Following Bell's (1984, 2001) model of stylistic variation as audience/referee design, van Compernolle (2007) concluded that *ne* deletion has become the norm in French-language chat, but that IRC participants rely on *ne* to "infuse the flavor" (Bell, 2001) of more formal communicative contexts into IRC discourse in response to or in initiation of a style shift.

Using the same corpus of IRC, van Compernolle (forthcoming b) investigated variable *ne* deletion in relation to a number of internal linguistic factors, namely the nature of the subject (i.e., noun phrase, pronoun, or [-overt subject] environments), negative complement type (e.g., *pas, rien, jamais*), preceding and following phonological environment (i.e., vowel vs. consonant), and sentence type (i.e., declarative, interrogative, imperative). Subject type and preceding phonological environment had significant main effects; however, these two factors interact. This interaction led van Compernolle (forthcoming b) to consider each subject type as an independent data set in a partitioned VARBRUL analysis. This revealed that *ne* retention was most likely to
occur with noun phase subjects, followed by [-overt subject] environments (e.g., imperatives), and then pronoun subject environments. Within the [-overt subject] category, negative complement type was significant: pas disfavored ne retention, and all others tended to favor it. The surrounding phonological environment was a significant factor only within the pronoun subject category, which van Compernolle speculates is caused by the grammaticalization of a number of clitic-verb sequences. This is especially true in the case of the sequences c'est 'it is', tu es 'you are', and tu as 'you have', which are often reduced to c, t, and ta in text-based chat. Thus, orthography (i.e., absence or presence of orthographic variation) may also be a contributing factor in the variation (see also van Compernolle & Williams, 2007).

A number of other studies of French-language CMC using different corpora have also remarked on the use of ne. van Compernolle and Williams' (2007) cross-CMC study of orthographic variation in IRC, hobby-oriented discussion fora, and moderated chat discussions (i.e., synchronous chat discussions with an invited guest in which a moderator fields questions from participants) illustrates the dramatic stylistic differences among three CMC types. Although they focused primarily on variation in the spelling of selected subject-verb sequences, the authors note that ne deletion is widespread in IRC but that ne is nearly categorically retained in moderated chat discussions, because of the formality of the context (interview-style discourse) and the fact that a designated moderator was responsible for which questions appeared and were answered by the guests. The asynchronous discussion fora fell along the middle of the stylistic continuum, with ne retention rates ranging from 50 to 75%. Lawrence Williams (personal communication) has also indicated that while ne deletion occurs relatively frequently in asynchronous discussion fora, the rates do not approach those found in synchronous chat. Williams notes, however, that variation across discussion fora on different sites is considerable, as ne retention in some fora is at approximately 50%, while in others it reaches 80% or higher. This variability is greater than that observed across different unmoderated chat channels (e.g., IRC) or different moderated chat discussions (nearly categorical ne retention), which, Williams explains, may be due to variations in group or community identity across different websites, which have different focuses and themes around which communities are formed.

Research Questions and Hypotheses

The results reported in previous studies of ne deletion serve as the basis for the research questions and hypotheses explored in the present study. The first research question to be explored is:

RQ1: How often, if at all, is ne deleted in online personal ads? If advertisers productively demonstrate variation, how do deletion rates compare to those reported in studies of spoken Quebec French and forms of electronic and non-electronic written or typed communication?
Gender is a classic sociolinguistic variable, and much variationist research over the past 40 years has shown that, in the case of stable variation, women tend to use standard forms more frequently than do men (Labov, 1990, p. 210, Principle I). In addition, gender tends to be a significant factor in self-presentation and the discourse produced in French-language online personals (van Compernolle, forthcoming a) and a determining factor in choice of address pronoun or address strategy (van Compernolle, forthcoming c). However, the effect of gender on *ne* use may be different. Although Ashby’s (1981) Tours study found that men were more likely to retain *ne* than were women, more recent research (e.g., Ashby, 2001; Coveney, 1996; Hansen & Malderez, 2004) has shown that men and women delete *ne* with approximately equivalent frequencies, and it is now generally accepted that gender is a nonsignificant factor. Therefore, consistent with the literature on *ne* variation, the first hypothesis is:

**H1:** Gender will be statistically nonsignificant in this study.

The most influential factor in many of the studies summarized above has been the speaker's age: Younger speakers tend to delete *ne* more often than older ones. However, age may not be a factor in Quebec, since all speakers appear to delete *ne* at very high frequencies. The second research question to be investigated is thus:

**RQ2:** How, if at all, does age favor or disfavor *ne* presence in online personals?

In line with the research summarized above, it is hypothesized that:

**H2:** Younger online dating participants will delete *ne* at higher rates than older ones.

In light of recent variationist research that has considered co-occurrence rules (e.g., Bell, 2001; Coveney, 2003), the present study will also explore the use of second-person pronouns (i.e., *tu,* *vous,* or [no second-person address]) as an independent variable. Previous research on French-language Internet dating networks has suggested that the use of a second-person pronoun (or related form, e.g., imperative, object pronoun, disjunctive pronoun) personalizes the ad, in that the advertiser is directly addressing the reader (van Compernolle, forthcoming c). In contrast, those ads in which only 1st- and/or 3rd-person references are made tend to be associated with more formal written styles. Thus, the third research question is:

**RQ3:** How, if at all, does second-person address use influence *ne* retention?

Based on van Compernolle's (forthcoming c) analysis of online personals, it is hypothesized that:

**H3:** The presence of a second-person form of address will disfavor *ne* retention, while the absence of second-person address will favor it.
Data and Method

The Social and Linguistic Environment of Netclub.com

Netclub.com is a multinational Internet dating site owned by Trader Classified Media. As of February 2008, the site boasted a total membership of over 4,600,000 people across North America and Europe (Netclub, 2008a). The Quebec site, based in Montreal, is accessed by over 50,000 unique visitors per month (Netclub, 2008b). A basic membership is freely available with the registration of a valid email address; a VIP membership is available for a nominal fee that depends on the length of subscription (month, year, etc.). The only difference between a basic membership and a VIP membership is the ability to contact other members through the site's numerous communication tools (real-time chat, in-network electronic mail, etc.). VIP members can contact any member they wish; basic members, while able to browse anyone's profile, can only contact VIP members. In other words, two non-paying "basic" members cannot contact each other without a fee.

On Netclub.com, members create a profile that includes a variety of personal information (e.g., age, gender, location, zodiac sign, hair color, hobbies), as well as a personal narrative. The socio-biographic information and physical descriptors are chosen from a series of drop-down menus and/or by checking boxes. This information is stored by the site, and it is searchable by other members. In other words, a young woman who seeks a man between 25 and 30 years of age, who is single with no children, and likes to go camping, can enter these criteria into a search tool. The profiles of men fitting her description will appear, and she can begin browsing for one she would like to contact.

Members are also given space to write up to 2,000 characters (~300 words) describing themselves, their interests, what they are looking for, and so forth; however, the average personal narrative is shorter than 120 words (van Compernolle, forthcoming a). Information provided in the personal narrative (called son annonce 'his/her ad' on Netclub.com) often expands on the biographical information included in the searchable data (e.g., age, location, interests, physical characteristics). For the purposes of the present study, biographical information about members was collected from the profiles, and the variable deletion of ne in the personal narrative section was analyzed.

Data Collection

The Quebec data were collected during late Fall 2006 over the course of one week. In order to access the personals, the author created two basic membership accounts: one as a man seeking a women and one as a woman seeking a man (this was done in order to view profiles of both men and women). In order to diminish the (potential) influence of regional variation, only those ads posted by men and women from Montreal were considered. Ads were selected based on the date on which they were last updated, in order from most recent to least recent, as they appeared
Language Variation in Online Personal Ads

during a basic profile search. Six profile searches were conducted, three each as a man seeking a woman and as a woman seeking a man. Each search targeted a specific age group (i.e., 18-25 years, 26-35 years, and 36-45 years). Fifty ads from each search were retained for analysis, for a total corpus of 300 personals evenly distributed across demographic categories.

While a basic membership was required in order to collect data, the author never published a profile, nor did he contact any members. Thus, none of the data was collected through deceptive means (e.g., by contacting a member and pretending to be interested in meeting). In order to protect members' privacy, pseudonyms, photographs, and other personally identifiable data were not recorded or copied. Socio-biographic information (i.e., gender, age, location) was simply noted for each of the personal narratives, which were copied and pasted into a text file for further review. Following data collection, the text files were reviewed, and each occurrence of verbal negation was identified and coded as described in the following section.

Coding and Statistical Procedures

Each instance of negation was identified and coded as either two-particle or single-particle (i.e., ne present or ne absent). Tokens of non-verbal single-particle negation, such as in the case of elliptical negations (e.g., pas vrai! 'not true!' or 'no way!'), were excluded from the analysis, since they do not represent loci of variation in the ne present vs. absent variable. In addition, instances of pas mal 'not bad' were excluded, since this expression has been lexicalized without ne in modern French (Coveney, 1996; Hansen & Malderez, 2004; van Compernolle, forthcoming b). Also excluded were instances of two-particle negation in fixed expressions (e.g., n'est-ce pas?). All remaining instances of verbal negation representing variable contexts were then coded for variable rule (VARBRUL) analysis using GoldVarb X (Sankoff, Tagliamonte, & Smith, 2005).

GoldVarb X (henceforth GV) is a multivariate regression analysis program developed for linguistic variation research; it is capable of determining the extent to which any number of factors condition variable rules. In the present study, the ne present vs. absent variable is explored according to the advertiser's gender, age group, and use of second-person pronouns. In the following analyses, GV factor weights (FWs) indicate the probability of ne being present with respect to these independent variables. GV functions on a scale of 0.0 to 1.0, with 0.50 being the point at which the rule—in this case, ne presence—is likely to occur in a given context. Thus an FW less than or equal to 0.49 indicates that ne presence is disfavored, while an FW greater than or equal to 0.50 means that ne presence is favored.

Results

Variable 'ne' Deletion in Online Personal Ads

The first level of analysis aims to compare rates of ne retention in the online personals with those reported in previous analyses of spoken Quebec French and electronic and non-electronic written or typed communication. Table 1 displays overall ne retention rates in spoken Quebec French,
French-language IRC and discussion fora, French children's written school assignments, and online personals from Quebec. It is important to note that formal writing for publication (e.g., in newspapers) requires that *ne* be retained; thus, it is generally accepted that there is no productive variation in this type of written discourse, in that the absence of *ne* would be seen as a grammatical error. The only exception to this rule might be in reported discourse, for example, where an interviewee or witness to an event happened to delete *ne* when speaking to a journalist.

<table>
<thead>
<tr>
<th>Study</th>
<th>Data</th>
<th>% <em>ne</em> retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sankoff &amp; Vincent (1977 [1980])</td>
<td>Montreal French</td>
<td>0.5</td>
</tr>
<tr>
<td>Poplack &amp; St-Amand (2007)</td>
<td>Ottawa-Hull French</td>
<td>0.2</td>
</tr>
<tr>
<td>van Compernolle (forthcoming b)</td>
<td>French-language IRC</td>
<td>16.1</td>
</tr>
<tr>
<td>Williams (p.c.)</td>
<td>French-language discussion fora</td>
<td>~50-80</td>
</tr>
<tr>
<td>Hansen &amp; Malderez (2004)</td>
<td>Children's written school assignments</td>
<td>~60-75</td>
</tr>
<tr>
<td></td>
<td>(France)</td>
<td></td>
</tr>
<tr>
<td>Present study</td>
<td>Online personals from Quebec</td>
<td>77.6</td>
</tr>
</tbody>
</table>

Table 1. Retention of *ne* across discourse types

The data indicate that *ne* retention rates in the Quebec online personals align closely with those reported in analyses of asynchronous discussion fora and French children's written school assignments. In comparison to French-language IRC and spoken Quebec French, *ne* retention is more heavily favored in the online personals. This difference is most likely due to level of synchronicity (Herring, 2001, 2002, 2007). Because of their asynchronous nature, personals can be planned in advance and after being typed, they can be revised by the author. The result may be a more monitored style of discourse relative to spontaneous speech or IRC. Therefore, in response to RQ1, *ne* retention rates in online personals align more closely with those reported for non-electronic writing and asynchronous electronic communication than with those reported in studies of synchronous CMC and spontaneous speech. This is perhaps an unsurprising finding, given that online personals are more-or-less static texts, which are naturally associated with the written French language. What is perhaps more unexpected is that online dating site members productively demonstrated variation, in that *ne* was absent from 22.4% of variable contexts. In the formal written language, one would expect *ne* retention to be categorical. Thus, it is clear that at least some degree of orality, or at least informality, is present in the discourse of online personals. This point will be discussed in more detail in the conclusion, following the VARBRUL analyses presented below.

**VARBRUL Analysis of 'ne' Deletion**

The first coding of the VARBRUL analysis of *ne* deletion in the Quebec personals considers gender, age, and use of second-person pronouns as independent variables. The results are provided in Table 2.
The data indicate, first, that gender is a nonsignificant factor in the variation, which corroborates results reported in previous studies of *ne*. Thus, H1 is supported. Nonetheless, there is a nonsignificant tendency for women to use *ne* more than men, which is consistent with the classic notion in variationist sociolinguistics that women use standard variants more often than men in situations of stable sociolinguistic variation (Labov, 1990). It may be that this difference is found only in this selection of personals; a larger corpus may be needed in order to confirm or reject the conclusion that gender is nonsignificant.

Second, age was selected as significant. The youngest age group disfavors the use of *ne*, while the middle and older groups favor *ne* retention. This difference may be explained by the fact that many of the men and women in the 18-25 year age group were students, and they may have been more inclined than their older counterparts to use informal or everyday language even in a static, "written" environment. Thus, H2 is supported. In addition, those advertisers in the 26-35 year age group may tend to retain *ne* at higher frequencies because many of them have begun their careers and may therefore feel the need to use a more careful style of discourse across situations. It is also worth noting that age has been considered the most influential social predictor of *ne* variation in almost every previous study of negation. A comparison of ranges (i.e., the difference between the highest and lowest FW across independent variables) confirms that age is the most significant factor (range = .27 vs. range for second-person pronoun = .20; gender was nonsignificant).

Finally, second-person pronoun use was selected as significant. Those members who used either *tu* or *vous* disfavor *ne* retention, while those who did not address their reader(s) with a second-person form of address favored *ne* retention. In response to RQ3, then, we can say that the use of a second-person pronoun increases *ne* deletion, which confirms H3. Second-person address personalizes the text, in that the advertiser is directly addressing his or her interlocutor, which may lead to a more conversational-style personal ad (van Compernolle, forthcoming a,

Table 2. VARBRUL analysis: First coding

<table>
<thead>
<tr>
<th>Factor</th>
<th>n</th>
<th>%</th>
<th>FW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>197/265</td>
<td>74.3</td>
<td>[ ]</td>
</tr>
<tr>
<td>Women</td>
<td>206/254</td>
<td>81.1</td>
<td>[ ]</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-25 yrs.</td>
<td>112/172</td>
<td>65.1</td>
<td>.33</td>
</tr>
<tr>
<td>26-35 yrs.</td>
<td>155/185</td>
<td>83.8</td>
<td>.58</td>
</tr>
<tr>
<td>36-45 yrs.</td>
<td>136/162</td>
<td>84.0</td>
<td>.60</td>
</tr>
<tr>
<td>2nd-pers. pr.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tu</em></td>
<td>157/206</td>
<td>76.2</td>
<td>.44</td>
</tr>
<tr>
<td><em>vous</em></td>
<td>143/192</td>
<td>74.5</td>
<td>.47</td>
</tr>
<tr>
<td>No 2nd pers. pr.</td>
<td>103/121</td>
<td>85.1</td>
<td>.64</td>
</tr>
<tr>
<td>Total</td>
<td>403/519</td>
<td>77.6</td>
<td>---</td>
</tr>
</tbody>
</table>
forthcoming c). Incidentally, little difference was observed between *tu* and *vous* (i.e., both disfavor *ne* retention), which suggests that the choice of address form is nonsignificant.

A second coding was done in light of an unbalanced distribution in the data found in a cross-tab analysis. The second coding considered each age group as independent data sets in a partition analysis (see Paolillo, 2002, pp. 89-93) in order to provide a more nuanced interpretation of the variation. These results are given in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>18-25 years</th>
<th></th>
<th>26-34 years</th>
<th></th>
<th>36-45 years</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>n</em></td>
<td>%</td>
<td><em>FW</em></td>
<td><em>n</em></td>
<td>%</td>
<td><em>FW</em></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>57/95</td>
<td>60.0</td>
<td>[ ]</td>
<td>76/91</td>
<td>83.5</td>
<td>[ ]</td>
</tr>
<tr>
<td>Women</td>
<td>55/77</td>
<td>71.4</td>
<td>[ ]</td>
<td>79/94</td>
<td>84.0</td>
<td>[ ]</td>
</tr>
<tr>
<td><strong>2nd pers. pr.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tu</em></td>
<td>23/47</td>
<td>48.9</td>
<td>.32</td>
<td>62/74</td>
<td>85.1</td>
<td>[ ]</td>
</tr>
<tr>
<td><em>Vous</em></td>
<td>49/79</td>
<td>62.0</td>
<td>.44</td>
<td>51/60</td>
<td>85.0</td>
<td>[ ]</td>
</tr>
<tr>
<td>Ø</td>
<td>40/46</td>
<td>87.0</td>
<td>.76</td>
<td>41/51</td>
<td>80.4</td>
<td>[ ]</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>112/172</td>
<td>65.1</td>
<td>.33</td>
<td>155/185</td>
<td>83.8</td>
<td>.58</td>
</tr>
</tbody>
</table>

Table 3. VARBRUL analysis: Second coding

What emerges from the results shown in Table 3 is that second-person pronoun use is a contributing factor only among *Netclub* members aged 18 to 25 years. Those who use *tu* and *vous* disfavor *ne*, while the absence of second-person address favors *ne*. Second-person pronoun use was nonsignificant in the two other age groups. In light of this finding, it seems reasonable to presume that age is the most significant factor in the variation, while second-person pronoun use may be a secondary contributing factor, at least among those aged 18 to 25 years. Gender was not selected as a significant factor for any of the age groups. This finding corroborates the interpretation of the results of the first coding discussed above.

**Conclusion**

The results reported in this study provide insight into variable *ne* use in a written (typed) genre of Quebec French, a phenomenon that until now has mainly been studied in informal speech. First, the data indicate that online dating site members from Quebec do indeed productively demonstrate variation in the use of *ne* in this type of asynchronous CMC. At the same time, rates of *ne* retention remain relatively high (77.6%), similar to those reported for asynchronous discussion fora (van Compernolle & Williams, 2007; Williams, p.c.) and French children's written school assignments (Hansen & Malderez, 2004). Second, despite the tendency for women to use *ne* more often relative to men, gender was found to be a nonsignificant factor in the variation.
This result corroborates previous studies of *ne* in speech that have found no link between gender and *ne* retention rates in both France and Quebec (Armstrong, 2002; Ashby, 2001; Coveney, 1996; Hansen & Malderez, 2004; Poplack & St-Amand, 2007; Sankoff & Vincent, 1977). Third, age group was found to be significant. Those dating site members in the 18-25 year age group were disfavored to retain *ne* compared to advertisers aged 26 years or older. This is consistent with previous research on *ne* variation in France (e.g., Ashby, 2001; Coveney, 1996; Hansen & Malderez, 2004) that reported a tendency for younger speakers to delete *ne* more frequently than their older counterparts. Finally, the use of second-person address was selected as a significant factor in the first coding, in that those members who used *tu* or *vous* (or a related form) deleted *ne* more often than those who avoided second-person address. However, a partition analysis considering each age group as an independent data set revealed that this factor operated only among members in the 18-25 year age group. Thus, the advertiser's age emerges as the most influential factor considered in this study.

A surprising result of this study is that online dating site members demonstrated variation in the use of *ne* in a static, written (typed) communication context. Even more unexpected is the relatively high percentage of negations produced without *ne*: 22.4%. This finding may be indicative of two phenomena: (i) that the tendency to delete *ne* may be spreading to other types of discourse, such as writing and asynchronous CMC; and (ii) that online dating site members may use informal or non-standard linguistic forms in their personals as a means of presenting part of their offline (linguistic) identities, and that the style chosen is designed for their (perceived) audience (Bell, 2001).

Although nearly categorical *ne* deletion has been the norm in colloquial Quebec French since at least the latter half of the 19th century (Poplack & St-Amand, 2007; Sankoff & Vincent, 1977), *ne* remains part of the grammar of standard writing (Martineau & Mougeon, 2003). However, the results of the present study indicate that *ne* deletion has spread to written (typed) asynchronous electronic communication, and that this trend is most prevalent among the youngest online dating site members. Given that rates of *ne* retention increase in older online advertisers' personals, it may seem reasonable to presume that these results are indicative of a change in progress, within an apparent time construct (Labov, 1972, 1994). However, rates of *ne* retention are nearly identical in the middle and older age groups. If these results were indeed indicative of a change in progress, one might expect to find noticeable differences across all groups.

What seems more probable, therefore, is that the data indicate a pattern of age-grading among online dating site members, whereby the cut-off point for frequent *ne* deletion is somewhere in the mid-twenties. As discussed in the analysis above, most of the youngest advertisers were students, while those in the older two age groups had already begun or were well into their careers. Consistent with the notions of age-grading (Labov, 1972, 1994) and the linguistic market (Sankoff & Laberge, 1978), it may be that once individuals enter the workforce their linguistic behaviors move in the direction of the standard written language, at least in some contexts. This conclusion is also in line with Coveney's (1996) findings with regard to variable *ne* in the informal spoken French of France.
Explanations for the motivations to delete *ne* can only be speculative, but one promising hypothesis is that informal linguistic forms are used as a means of creating intimacy between the advertiser and his or her interlocutor (i.e., the reader, who is a potential partner). van Compernolle (forthcoming c) argues that second-person address, especially the use of informal *tu*, is motivated by online advertisers' desire to establish familiarity with (potential) readers. In contrast, those ads in which no second-person address was used tended to be associated more closely with standard, impersonal writing. In the present study, the presence of second-person address in the online personal was found to disfavor *ne* retention among members in the 18-25 age group. Thus, if the online personal is viewed as the opening turn in a conversation with one's interlocutor (as indicated by the presence of second-person address), it seems reasonable that other informal or conversational-style linguistic forms should co-occur. The (categorical) use of standard written forms could actually seem sociopragmatically inappropriate, since it is unlikely that (hyper)formal linguistic variants (e.g., *ne* presence) would be used in face-to-face or live voice communication, at least within certain demographic groups.

As for the higher rates of *ne* use among the middle and older age groups, it may be that these dating site members wish to convey to their audience a certain level of literacy or the ability to conform to standard written norms. This hypothesis would explain why second-person address was a nonsignificant factor in *ne* retention rates among the two older age groups. This last finding is consistent with Bell's (1984, 2001) model of stylistic variation as audience design, which holds that "[s]peakers design their style primarily for and in response to their audience" (2001, p. 143).

Despite the insights provided by this study for French speakers in Quebec, there is still much that is not known about the sociolinguistic, pragmatic, and discursive norms found in online personal ads. One particularly interesting avenue to pursue in future research involves the investigation of co-occurrence rules (Bell, 2001; Coveney, 1996; Ervin-Tripp, 1972). As shown in this study, *ne* deletion co-occurred at significant rates with second-person address among younger advertisers; still other variable features of discourse could be analyzed within the same framework. For example, orthographic variation, which has been shown to be conditioned by expected or perceived levels of formality in French-language CMC (van Compernolle & Williams, 2007), could be investigated in order to determine to what extent grammatical variation, address strategies, and writing systems are interrelated. In addition, informal observations made during the analysis of this corpus indicated that discourse markers were widely used among advertisers aged 18 to 25 years, while they were markedly less frequent in the personals of older advertisers. Such analyses could have interesting implications for researchers interested in sociolinguistic variation, including co-occurrence rules, and the spoken/written divide (Crystal, 2001) in different forms of CMC. Finally, more research is needed that addresses individual motivations for using selected linguistic forms, as these analyses would almost certainly prove insightful in explaining the patterns of sociolinguistic and pragmatic variation observed in CMC environments.
Notes

1. Pro-clitic *ne* descends from Latin *non*. In many Romance languages, *non* is the default negator, while other negative complements sporadically reinforce verbal negation. In French, however, this embracing structure became grammaticalized as early as the 17th century (see Ashby, 1981 and van Compernolle, forthcoming b for overviews).

2. Some researchers, however, have explored *ne* deletion in older texts, such as plays and correspondence from Classical French (ca. 17th and 18th centuries). See Martineau and Mougeon (2003) for an overview and analysis.

3. Separation and divorce were among the most common reasons given for joining the online dating site; these were concentrated among advertisers 36-45 years old (van Compernolle, forthcoming a).

4. For an analysis of second-person address strategies in synchronous CMC (i.e., chat) that differs considerably from the analysis presented in van Compernolle (forthcoming c), see Williams and van Compernolle (2007).

5. For discussions of the ethics of data collection from Internet sources, see the articles in Herring (1996).

References


Coveney, A. (2003). 'Anything you can do, *tu* can do better': *Tu* and *vous* as substitutes for indefinite *on* in French. *Journal of Sociolinguistics*, 7(2), 164-191.


**Biographical Note**

Rémi A. van Compernolle [compernolle@gmail.com] is currently an Adjunct Faculty member in French in the Department of Foreign Languages and Literatures at the University of North Texas. His research interests include sociolinguistic and pragmatic variation, second language acquisition, and computer-mediated communication and discourse.