FROM "Y AS PLUS PERSONNE QUI PARLE" TO "PLUS PERSONNE NE DIT RIEN": THE VARIABLE USE OF THE NEGATIVE PARTICLE *NE* IN SYNCHRONOUS FRENCH CHAT

Rémi A. van Compernolle, B.A.

Thesis Prepared for the Degree of

MASTER OF ARTS

UNIVERSITY OF NORTH TEXAS

May 2007

APPROVED:

Lawrence F. Williams, Major Professor
Marijn S. Kaplan, Committee Member
Marie-Christine W. Koop, Committee Member and Chair of the Department of Foreign Languages and Literatures
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

van Compernolle, Rémi A., From "y as plus personne qui parle" to "plus personne ne dit rien": The variable use of the negative particle *ne* in synchronous French chat. Master of Arts (French), May 2007, 114 pp., 30 tables, 3 figures, references, 62 titles.

This study analyzes negative particle variation (i.e., the variable presence or absence of the negative particle *ne*) in synchronous French chat discourse within a labovian-inspired framework. Selected morphosyntactic, lexical, and phonological constraints are considered. Multivariate analyses performed by GoldVarb 2001 revealed that subject type (i.e., NP, [- overt] subject environment, pronoun) and the phonological environment preceding the position of *ne*—regardless of its presence or absence—are determining factors in the variation. In addition, discursive-pragmatic effect was explored in a sub-sample of data. The results indicate that *ne* is seldom present in verbal negation during explanatory discourse style, yet it is very likely to be retained in ludic, emphatic, and proverbial styles.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1 Aim and Scope

The emergence of various new technologies has enabled communication to occur in a variety of new social contexts "through the medium of written language" (Werry, 1996, p. 47). However, communication in electronic environments is often different from communication in more traditional contexts since interlocutors do not see or hear one another and therefore "do not have access to non-verbal information about how others are responding" (Herring, 1999). Nonetheless, much of the literature on Computer-Mediated Communication[1] (CMC) attempts to compare discursive, communicative, and linguistic features of electronic discourse with those found in either written or spoken language.

CMC can be divided into two broad categories: asynchronous CMC (e.g., electronic mail, discussion forums,[2] etc.) and synchronous CMC[3] (e.g., public chat rooms and instant messaging services). The language used in asynchronous CMC is often likened to that of more traditional forms of written language (e.g., written letters, newspapers, literature, etc.), while the language used in synchronous CMC often appears to resemble everyday conversational speech,[4] at least to some extent.

---

[1] The more recent—and more inclusive—label *Computer-Mediated Discourse* (CMD) might be more appropriate, but Computer-Mediated Communication is used in this thesis. See Herring (2001) for an overview of CMD.

[2] I have used "forums" instead of "fora" since my informal observations of the lexicon used on the Internet suggest that "discussion forums" is the preferred plural form.

[3] Garcia & Jacobs (1999) proposed the term "quasi-synchronous" since the recipient of the message must wait for the sender to complete his or her message and press "send" or "enter" before having access to the content of the message.

[4] One exception to this comparison is moderated chat (Williams, 2006; van Compernolle & Williams, in press).

Although communication researchers and linguists have studied many aspects of CMC, there exists relatively little empirical data concerning specific linguistic traits and variables that have been explored extensively in more traditional communicative environments (e.g., written and spoken discourse). Since CMC environments—especially synchronous text-based chat (see Collot & Belmore, 1996; Werry, 1996; Anis, 1999; Dejond, 2002; Pierozak, 2003a)—are, to say the least, becoming increasingly popular in parts of the world where access to networked technologies is widespread, it follows that the observation and documentation of specific linguistic features of discourse used in this context could prove rather insightful and have the potential to inform research in linguistics and communication studies, among other fields. The present study aims to explore one of the most well known grammatical variables in the modern French language: the use of the negative particle *ne*.

Although the *ne* paradigm (i.e., *ne* present vs. *ne* absent) has been studied in spoken French by many linguists over the course of the past thirty-five years, it has not yet been explored in synchronous, French-language CMC. [5] The data in this study suggest that *ne* use in synchronous, French-language CMC discourse is very similar to *ne* use in informal spoken discourse; that is, *ne* is often omitted from verbal negation, yet it remains an important sociolinguistic resource in communication since its presence usually correlates with a number of stylistic and pragmatic features of discourse.

In addition to comparing the extent to which French language chat participants use *ne* in a variety of syntactic, stylistic, and pragmatic environments with results reported in similar studies on spoken French, this study highlights a number of discursive features particular to French-language synchronous CMC that co-occur and co-vary with *ne*. In this way, the effect of these features of CMC discourse on the *ne* paradigm will be explored.

[5] I have previously reported various findings of this research at the Colloque International: La Langue de la Communication Médiatisée par les Technologies de l'Information et de la Communication in Bordeaux, France in May 2006 and at the 2006 Conference of the Association for French Language Studies in Bristol, England in September 2006 (with Lawrence Williams).

1.2 Formal Explanation and History of Verbal Negation in French

In prescriptive Modern French, verbal negation is achieved by combining the proclitic (i.e., pre-verbal particle) *ne* and another word or adverb that has a negative meaning. When a conjugated verb form is used, *ne* precedes the verb, as well as its pronominal complements, and the second-negative (Neg2) follows (Grevisse, 1993). This type of negation will be referred to as two-particle negation (2Neg).

2Neg has not, however, always been required in the French language. Indeed, verbal negation in French has undergone a number of developments and changes throughout the history of the language. In order to understand the current state of negation in French, the following paragraphs provide a summary of the development of 2Neg.[6]

1.2.1 Old French and Middle French

French—like a number of other European languages—inherited much of its grammar and syntax from Latin, including the negation *non* (Brunot, 1966; Dauzat, 1953, 1964; Pope, 1961; Ewert, 1969; Rohlfs, 1970). In Vulgar Latin, which was spoken throughout much of France into the Middle Ages, *non* could be used alone with a conjugated verb form to express negation (Sancier-Chateau, 1993, p. 93). However, as the French language distinguished itself from Vulgar Latin, *non* weakened to *nen* in pre-verbal position and eventually to *ne* (or *n'* immediately preceding a vowel). Accented *non* was restricted to elliptic use in negative responses to questions and a certain number of archaisms (Rohlfs, 1970; Sancier-Chateau, 1993).[7]

In Old French, "*ne* constituted sufficient negation in itself" (Rickard, 1989, p. 54) and a variety of words (e.g., *pas, point, mie,* etc.) could be added for emphasis. Toward the end of the 15th century, the words *pas* and *point* emerged as the default second-negatives (Dauzat, 1967; Ewert, 1969), but their use remained emphatic, and they were often omitted in favor of

[6] For a general treatment of negation in French, see Pohl (1968) and Kayne (1983).
[7] For a general treatment of phonology and morphology in Old French, see Pope (1961) and Rohlfs (1970).

single-particle negation with *ne*. According to Cohen (1967), two-particle negation had become widespread by the 16th century; however, this usage had not yet become regular and single-particle negation was still preferred (Brunot, 1966, v. 2, p. 472).

As second-negatives were used more frequently in verbal negation, they began to lose their emphatic quality and "gradually . . . came to be invested with a negative meaning" (Ewert, 1969, p. 260). Ashby (1981) has referred to the grammaticalization of second-negatives as "an innovation in French" since, in other Romance languages, "*non* and its descendents are only sporadically reinforced, especially by nouns denoting smallness or insignificance" (p. 674). Dauzat (1967) offered the following explanation for such a development.

> Toutes les langues romanes ont éprouvé le besoin de renforcer la négation qui accompagne le verbe; mais c'est en français que cette tendance s'est développée au maximum, surtout parce que *non*, atone, s'était affaibli en *nen*, puis *ne, n',* particule phonétiquement trop faible pour exprimer une négation énergique. (p. 196)[8]

By the 16th century, second-negatives had become so common and understood as negative that *ne* was often omitted in direct interrogatives involving subject-verb inversion (e.g., *Vient-il pas*; see Brunot, 1966, v. 2; Sancier-Chateau, 1993). However, by the 17th century "les théoriciens cessent de considérer *pas* comme suffisant dans les interrogations directes" (Brunot, 1966, v. 4, p. 1,039),[9] and two-particle negation was recommended.

In addition, grammarians began to consider two-particle negation as required in all instances of verbal negation. For example, François de Malherbe (official poet of the court under Henri IV and later Louis XIII) was of the opinion that "the negative consist[ed] of two

[8] Translation: "All Romance languages have experienced the need to reinforce the negation that accompanies the verb; but it is in French that this tendency was developed the most, especially because *non*, an atonic [unstressed or unaccented] syllable, had weakened to *nen*, then *ne, n'*, a particle that was too phonetically weak to express an energetic negation."
[9] Translation: "theorists cease to consider *pas* as sufficient in direct interrogatives."

parts, not *ne* alone, but with *pas* or *point*" (Rickard, 1989, p. 102), and it is clear that this was the preferred structure of the time. However, the second-negative was still omitted fairly often well into the 17th century (Cohen, 1967). Brunot (1966, v. 4) commented, however, that "la suppression de *pas* est en général dans le style marotique, ce qui achève de donner à cet archaïsme son caractère" (p. 1,034).[10]

Another question arose concerning the placement of the second-negative with infinitives and, by extension, object pronouns. Until the 17th century, the second-negative had always followed the verb, even with infinitival forms. However, "[a]u fur et à mesure que *pas, point* devenaient les compléments indispensables de *ne*, ils devaient, suivant une loi invariable, s'en rapprocher" (Brunot, 1966, v. 4, p. 1,039),[11] and *pas* was placed in front of the infinitival form. Pronouns, on the other hand, were often placed between the two negative particles.

1.2.2 Classical and Modern French

According to Rickard (1989), it was not until the 18th century that the rules of negation became fixed. The *Académie française* decided that two-particle negation was required in direct interrogatives, and two-particle negation (especially *ne...point*) became the standard. *Ne* was seldom used alone; this usage was limited to a few select verbs (e.g., *pouvoir, savoir, empêcher*) and in a limited number of syntactic environments (Grevisse, 1993).

Negation in Modern French follows Classical French usage, although *pas* has replaced *point* as the most common second-negative (Dauzat, 1967). The two negative particles "embrace" (i.e., surround) the conjugated verb or, in the case of an infinitival form, they precede the verb and its pronouns. Simple negation with *ne* is, however, still in usage, albeit

[10] Translation: "the suppression of *pas* is in general Marotian in style, which gives this archaism its character." (Marotian refers to the writing style of the 16th century French poet Clément Marot.)
[11] Translation: "as *pas, point* became the indispensable complements of *ne*, they had to get closer, following an invariable law."

in a very limited number of contexts, one of which is proverbs (see Grevisse, 1993, p. 1,448, § 974).

According to Dauzat (1954) each part of negation has a specific function: "le premier élément [*ne*] énonce une *discordance*, complétée par le second [*pas, point, rien,* etc.] qui exprime la *forclusion*. La porte s'ouvre par une particule qui prépare la négation, le verbe passe, et le vantail se referme sur une précision de la valeur négative" (p. 218).[12] Nonetheless, it appears that negation is expressed principally by the second particle, rendering *ne* somewhat redundant (Ashby, 1981).

The weakening of *ne* and the strengthening of second-negatives has led to a new development in modern (spoken, primarily informal or everyday) French: single-particle negation with the second-negative alone. Ewert (1969) noted that "*ne*, being a mere proclitic and incapable of bearing a stress, is . . . weakened and tends to be omitted as unessential" (p. 260). This usage has not, however, been well received by purists of the French language. Dauzat (1954) argued, for example, that "en concentrant la négation sur un seul mot, on enlève au français une de ses élégances, une finesse d'expression propre à notre langue et que nous devons avoir à cœur de conserver" (pp. 218-219).[13] Nonetheless, it is clear that "[d]ans la **langue parlée,** surtout familière, le *ne* disparaît avec des fréquences variables" (Grevisse, 1993: p. 1462, § 982b, bold in original).[14]

1.2.3 Pleonastic *ne*

Dauzat (1954) has argued that pleonastic *ne* ("*ne* explétif"; e.g., *ne...Ø*) is still an important part of the modern French language. Although single-particle negation with *ne* is not usually considered to be a complete negation in modern French as it was in Old and

[12] Translation: "the first element announces discordance, completed by the second which expresses the debarment. The door is opened by a particle that prepares the negation, the verb passes through and the door is closed on a precision of the negative value."
[13] Translation: "by concentrating the negation on a single word, one removes from French one of its elegances, a finesse of expression particular to our language and that we must be committed to conserve."
[14] Translation: "in the spoken language, especially informal, *ne* disappears at variable frequencies."

Middle French, it can be used to help express a number of nuances.

> Dans les propositions dites complétives, *ne* exprime la crainte . . . , la précaution . . . , l'empêchement. . . . Loin d'être explétive, comme le croyait à tort l'ancienne grammaire, cette particule a une valeur affective très nette et permet, là où un flottement est possible, d'exprimer des nuances, ainsi entre "avant qu'il vienne", plus positif, et "avant qu'il ne vienne", qui entrebâille la porte au doute. (Dauzat, 1954, p. 218)[15]

This usage does not appear to be widespread, and it is usually associated with formal or literary discourse. Indeed, Grevisse (1993) considers this type of *ne* use to be optional.

1.2.4 Summary

The development of verbal negation in French can be divided into four distinct stages, as shown in Table 1.

Table 1. Development of verbal negation in French[16]

Period	**Negation**
Old French	*ne* + verb
Middle French	*ne* + verb (+ second-negative)
Classical French	*ne* + verb + second-negative
Modern French	(*ne*) + verb + second-negative

In Old French, *ne* is used alone in verbal negation. Later, *ne* is sometimes reinforced with another word or adverb in Middle French. In Classical French, both *ne* and a second-negative are required, but as Modern French develops, *ne* is sometimes omitted in speech. Although it is not possible to hypothesize at this point whether *ne* will disappear from the French language in the future (Ashby, 1981; Hansen & Malderez, 2004), there is evidence to support that its use is becoming increasingly dependant on a number of social, linguistic, and stylistic factors. In other words, the presence or absence of *ne* is variable, and the variation

[15] Translation: "In completive clauses, *ne* expresses fear, precaution, impedance. Far from being expletive, as the traditional grammar wrongly believed, this particle has a very clear value and allows, where a hesitation is possible, the expression of nuances, such as between 'avant qu'il vienne', more positive, and 'avant qu'il ne vienne', which holds the door open to doubt."

[16] Adapted from Ashby (1981).

can be linked to various aspects of the speaker's social identity, the formality of the communicative context, the syntactic environment surrounding the *ne* position, and a number of pragmatic features of discourse, among other factors.

1.3 Theoretical Framework

Before continuing to a review of the relevant literature concerning *ne* use in modern French, an explanation of the underlying theoretical notions and principles used for the analysis of the corpus is sketched in the following sections.

1.3.1 The Variationist Tradition

No language on earth is invariable, and there is no monolithic speaker of any language (Sax, 2003). Indeed, variations in language occur constantly. Labov (1972) noted that:

> Most such variations occur only once, and are extinguished as quickly as they arise. However, a few recur, and, in a second stage, they may be imitated more or less widely, and may spread to the point where the new forms are in contrast with the older forms along a wide front. (p. 2)

In the present study, I am operating under the hypothesis that such variations (i.e., those that recur and spread) in any given language are the direct result of social interactions.

The notion that language variation results from social interaction was first advanced in the 1960s and 1970s by Labov, according to whom, "one cannot understand the development of a language change apart from the social life of the community in which it occurs. Or to put it another way, social pressures are continually operating upon language, not from some remote point in the past, but as an immanent social force acting in the living present" (1972, p. 3). Labov (1972) provided three properties of a linguistic variable:

> First, we want an item that is frequent, which occurs so often in the course of undirected natural conversation that its behavior can be charted from unstructured contexts and brief interviews. Secondly, it should be structural:

> the more the item is integrated into a larger system of functioning units, the greater will be the intrinsic linguistic interest of our study. Third, the distribution of the feature should be highly stratified: that is, our preliminary explorations should suggest an asymmetric distribution over a wide range of age levels or other ordered strata of society. (p. 8)

This type of variation "presuppose[s] the option of saying 'the same thing' in several different ways: that is, the variants are identical in referential or truth value, but opposed in their social and/or stylistic significance" (Labov, 1972, p. 271).

The speaker's age, gender, and social class are, among many others, some of the social factors that appear to influence linguistic variables. In addition, a certain number of internal and external linguistic factors—such as clause and sentence type, phonological environment and attention paid to speech—have been shown to be influential. These notions have been defended, critiqued, and reformulated by numerous sociolinguists (Lavandera, 1978; Sankoff, 1980; Bell, 1984; Wolfram, 1991; Coveney, 1996; Eckert & Rickford, 2001; among others) over the past 30 years.

Although Labov's work was principally concerned with phonological variation, his theory and method have since been applied to the study of syntactic, morphological, and lexical variation. The application of variationist theory to non-phonological variation has not, however, gone without criticism. Lavandera (1978) warned that "it is inadequate at the current state of sociolinguistic research to extend to other levels of analysis of variation the notion of sociolinguistic variable originally developed on the basis of phonological data" (p. 171). However, she does not dismiss the idea of applying such analysis to non-phonological data; rather, she assigns "a different status to such data because they need further interpretation" (p. 3).

Sankoff (1972; discussed in Lavandera, 1978) was one of the first to suggest that the variationist framework could be applied to non-phonological data. She posited that "the extension of probabilistic considerations from phonology to syntax is not a conceptually difficult jump" (p. 58). Her analysis was supported by three examples of non-phonological variation: the placement of the future marker in New Guinea Tak Pisin *bai* and, in Montreal French, the use of the complementizer *que* and indefinite *on*.[17] Sankoff's study supported the call for the application of the variationist framework to "cases in which the variation seem[s] not to be the carrier of social and stylistic significance" (Lavandera, 1978, p. 173). In her conclusion, however, Lavandera (1978) reiterates her reticence to draw a parallel between *syntactic alternation* and *sociolinguistic variation*, unless the following conditions hold:

> (1) that [the variables] can be proven to be carriers of non-referential information, to have social and stylistic or other significance . . . and (2) that they prove to be a kind of device of the language similar to phonological variables, that is, elements whose defining property is a quantifiable covariation and for which the frequency relationships are the very signals of those differences. (p. 181)

Milroy (1987) added that the study of syntactic variation can be problematic since "the difficulty [is not] in obtaining *tokens* of a variable, but in obtaining the *full range* of realizations associated with it" (p. 144, italics in original). It is, however, possible to obtain the "full range of realizations" of the variable *ne*; that is, in a negative sentence, *ne* is either absent or present. *Ne* is clearly an example of a syntactic variable that can be analyzed within a framework inspired by the variationist tradition.

[17] *Que*, used as a complementizer (i.e., the subordinating conjunction *que*), is sometimes omitted in informal speech (e.g., je pense [*que*] c'est une bonne idée 'I think [that] it's a good idea'). *On* is the default indefinite pronoun in French, yet it is being replaced by the second-person pronouns *tu/vous*. *On* is also used as the first-person plural pronoun at the expense of *nous*. For a general treatment of the pronouns *on, tu,* and *vous*, see Peeters (2006).

1.3.2 Stylistic Variation and Register

Although this study does not specifically address stylistic variation and register (i.e., language use in two or more different communication environments is not being compared), I have provided a brief overview of these notions.

In general, every speaker of every language is thought to have at his or her disposal a variety of speech styles, each of which is judged appropriate or not for a given communicative context (Labov, 1970 & 1972; Trudgill, 1974; Hymes, 1984; Sax, 2003). This is not only true for phonological variables, but for syntactic, morphological, and lexical variations as well.

Halliday (1976) argued that grammar is composed of a system of choices; in other words, "[t]he speaker of a language, like a person engaging in any kind of culturally determined behaviour, can be regarded as carrying out, simultaneously and successively, a number of distinct choices" (p. 3). The key notion to remember is that language use is "culturally determined." In addition, language choices (i.e., which forms are used when) are not free; rather they are dependent upon a number of factors, including communicative environment, perception of formality, and attention paid to speech (Labov, 1972). Presumably, as speakers move from one context to another, they shift or modify their style of speech so that it may be appropriate for their interlocutors (Bell, 1984, 2001).

Speech style, as defined by Labov (1972) and Bell (1984, 2001), is often used synonymously with the term register (Sax, 2003). For the purposes of the present study, I wish to distinguish *register* from *style*. I have adopted Biber & Finegan's (1994) definition of register. Broadly defined, a register "is a language variety viewed with respect to its context of use" (Biber & Finegan, 1994, p. 4). According to Fischer (1958) levels of formality—both perceived and real—are often associated with registers of language. Register, then, can be regarded as a language variety that is dependent upon the level of formality perceived to be

appropriate in a given communicative context. I will therefore save the term *style* to refer to different ways of speaking within the same register. It will be shown later that several styles can exist at the same level of formality and the variable use of *ne* depends in large part on which style is used.

1.3.3 Speech Communities

Much of the sociolinguistics literature (Labov, 1972; Milroy, 1987; Biber & Finegan, 1994; Chambers, 2003) makes use of the term *speech community* to describe a group of people that share a common set of linguistic behaviors. According to Labov (1972):

> The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms; these norms may be observed in overt types of evaluative behavior, and by uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage. (pp. 120-121)

Recent research (Pierozak, 2003a, 2003b & 2003c, van Compernolle, 2006; Williams, 2006; Williams & van Compernolle, 2007; van Compernolle & Williams, in press) has suggested that CMC users participate in a number of shared norms, some of which are closely related to norms found in non-electronic environments, while other norms have developed specifically in and for electronic environments. It appears that "the lack of geographical constraints and the ability to self-select one's on-line community and communication environment have played very important roles in defining many of the norms for behaviors and practices of Internet users" (Williams & van Compernolle, 2007, p. 815).

The definition of community has changed since the advent of the Internet and other networked technologies. According to Mosco (2004), "existing communities are strengthened and whole new 'virtual' communities arise from the creation of networks of people who share interests, commitments, and values" (p. 31). In addition, new virtual communities in which

people *chose* to participate might be better than traditional ones that are constructed by geographical proximity rather than shared interests, commitments, and values since "a community that hasn't been chosen is a community of lesser quality" (Whittle, 1997, p. 240). Some have even argued that networked technologies allow people to form "smaller, more caring communities" (Hearn et al., 1998, pp. 62-63).

The freedom to choose one's acquaintances and terms of participation is one of the most important aspects of on-line communities (Whittle, 1997). "[E]ach participant has the choice to stay or to leave; therefore, those who stay have chosen to abide by whatever norms have been established by the community as a whole" (Williams & van Compernolle, 2007, p. 38). Since synchronous CMC occurs through the medium of written language, these communities are essentially linguistic (Pierozak, 2003a); therefore, the norms that have been constructed by participants in these on-line communities are, in general, linguistic in nature (Whittle, 1997). It follows that one might consider on-line communities to be speech communities that exist in electronic environments, and whose social norms consist of principally written linguistic behaviors.

1.3.4 Applying a Variationist-Inspired Framework to Synchronous CMC

One of the major limitations of variationist studies in the past has been the "observer paradox" (Labov, 1972). Since the informant is conscious of the presence of the observer (regardless of how informal or familiar the context may be), it is uncertain that the interviewee's style reflects that which he or she uses when the observer is absent. In the present study, this limitation appears to be minimized.

Synchronous CMC offers a rather anonymous communication environment that allows data collection to take place without the informants being aware of the observer. Moreover, informants are speaking to one another, which is similar to the group session method of observation. This type of observation usually reveals more about the vernacular of

the speakers than one-to-one interviews since informants are not directly addressing the observer (Labov, 1972). It follows that the language observed in the present study is representative of the vernacular of the synchronous CMC environment.

One limitation must, however, be noted. The anonymity provided by this form of communication prevents observers from gathering (credible) demographic data about their informants. It is therefore impossible to explore variation according to informant's gender, age, or social class. However, research has shown that, at least in France, the average synchronous CMC participant is young (under 25) and belongs to the middle or upper-middle class (Pierozak, 2003c).

In light of this limitation, the analysis will focus on a variety of linguistic and stylistic factors that may co-occur and co-vary with the use of *ne* in synchronous CMC. In addition, a certain number of pragmatic features particular to synchronous CMC discourse will be examined.

1.4 Research Questions

The present study is motivated by two objectives. First, this research aims to contribute to existing scholarship concerning the variable *ne* in modern French. To this end, the distribution of two-particle negation (2Neg) and single-particle negation (1Neg) will be explored in a variety of grammatical, phonological, and pragmatic environments. Second, this study aims to distinguish chat discourse as one type of modern French that requires more attention from linguists; therefore, I will discuss briefly a number of discursive features particular to synchronous electronic environments.

The present study will focus on four general questions:

(1) How do overall rates of *ne* retention in chat discourse compare to those reported in studies of spoken French?

(2) How do selected linguistic factors (e.g., Neg2 type, subject type, phonological environment, etc.) favor or disfavor *ne* retention?

(3) To what extent is the use of *ne* influenced by sociopragmatic features of discourse (e.g., emphasis, humor, etc.)?

(4) Which features of discourse particular to chat—which may differ from those observed in spoken, primarily informal, discourse—seem to influence *ne* use in this type of communication?

1.5 Organization of the Thesis

In Chapter 1, I have presented the aim and scope of the present study and defined the notions and analytical framework to be used. I have also provided a formal explanation of the history of *ne* and Neg2 usage in the French language. In Chapter 2, I review relevant literature, and in chapter 3, I provide a description of the data collection method, as well as the counting and coding procedures that I have used. In chapter 4, I present my results in comparison with results reported in previous studies of negation in spoken French. Finally, chapter 5 includes a discussion of my results, and the research questions are directly addressed. In addition, I have included a discussion of a certain number of sociopragmatic features of chat discourse that influence the use of *ne* and other linguistic variables that require further investigation.

CHAPTER 2

REVIEW OF THE RELEVANT LITERATURE

The *ne* paradigm is, according to Coveney (1996), "possibly the best known sociolinguistic variable in contemporary French" (p. 55). Much has been written on negation in European and Canadian French, and it seems that *ne* use is under the influence of a certain number of social factors, including the speaker's age, gender, level of education, and social class. In addition to these social factors, *ne* use is very much dependent on a variety of phonological, syntactic, and pragmatic features of discourse.

In the following sections, I review results reported in a number of studies of *ne* in modern French. The review is organized by influence (e.g., social factors, phonology, syntactic constraints, etc.) instead of chronology of when studies were done. This type of review provides a clearer picture of the complex of factors that affect the *ne* paradigm.

2.1 Overall Retention Rates in Previous Studies

A number of studies (Ashby, 1981, 2001; Coveney, 1996; Armstrong & Smith, 2002; Hansen & Malderez, 2004) have suggested that *ne* use is in decline throughout much of France. Although it is unclear whether *ne* will disappear altogether from the French language in the future, it is clear that there has been a continuous tendency to omit *ne* in spoken French at higher frequencies over the course of the past half century. Table 2 illustrates this tendency according to the results reported in a number of recent studies on the variable use of *ne*.

Table 2. Recent decline of French *ne* in France.

Corpus	Year of Data Collection	Location	Overall Retention Rate
Ashby (1981)	1976	Tours, France	37.0%
Coveney (1996)	1989	Somme, France	18.8%
Hansen & Malderez (2004)	1989-1993	Ile-de-France/Oise, France	8.2%
Ashby (2001)	1995	Tours, France	18.0%
Armstrong & Smith (2002)	1997	France	72.5%

Although Armstrong & Smith's (2002) study seems to contradict the hypothesis that *ne* use is declining in modern French, it is important to note that they explored *ne* retention rates in radio shows recorded in 1997 (interviews with politicians, celebrities, etc.), with a similar corpus from 1960-61 and did find a noticeable decline in *ne* retention rates. The relatively high rate of *ne* retention is most likely attributed to the serious nature of the radio programs; as a result, the interviewees and hosts may have felt the need to show a higher level of formality, which often appears to favor *ne* retention. Nonetheless, Armstrong & Smith (2002) concluded that the results of their study "suggest strongly that *ne* deletion is spreading to highly monitored speech styles" (p. 39). The other studies constitute recorded conversations between the informants and researchers that took place in more informal, conversational contexts (i.e., less highly monitored), which often seems to disfavor *ne* retention.

The difference between *ne* retention rates in Ashby (1981) and (2001) clearly shows that French speakers in the city of Tours have begun to omit *ne* at higher frequencies. Hansen & Malderez (2004) demonstrate this trend as well, comparing a corpus of spoken French recorded by Péretz-Juillard between 1972 and 1974 (see Péretez-Juillard, 1977) in the Paris area. Rates dropped from 15.8% in the Péretz-Juillard corpus to only 8.2% in the Hansen & Malderez corpus. Finally, as mentioned above, Armstrong & Smith (2002) also found that *ne* retention rates had dropped even in formal, highly monitored speech contexts where one might expect *ne* to be retained more frequently.

Although results reported in the three diachronic studies (Ashby, 2001; Armstrong & Smith, 2002; Hansen & Malderez, 2004) demonstrate that *ne* deletion is becoming more acceptable at many levels of society and in a variety of social contexts, *ne* does not seem to be disappearing from the French language; rather, its use is "dependent on internal linguistic, stylistic, and social factors" (Ashby, 1981, p. 686; see also Hansen & Malderez, 2004).

Although Coveney (1996) did not conduct a diachronic study of *ne* retention, his results demonstrate that *ne* is used at a rather low frequency in his corpus, which focused on many social and linguistic factors that had previously been shown to influence *ne* use (e.g., Ashby, 1981). These factors will be discussed in the following sections as they have been observed and documented in previous studies.

2.2 Social Factors

Previous research (Ashby, 1981, 2001; Coveney, 1996; Armstrong & Smith, 2002; Hansen & Malderez, 2004) has shown that the presence or absence of *ne* is—at least to some extent—determined by the speaker's social identity.

> Along the social-group or interspeaker dimensions of linguistic variation, *ne* is a grammatical variable of the type that responds to a variationist analysis; that is, the variable occurs frequently enough for speakers to be able to employ in a probabilistic way . . . the (non)standard variant to signal various aspects of their social identity. (Armstrong & Smither, 2002, p. 28)

Specifically, age "is indeed an important variant with *ne* retention or deletion" (Ashby, 1981, p. 683). In Ashby (1981) and (2001), the age group 51-64 retained *ne* at relatively high rates (52% in the 1976 data and 25% in 1995), while younger speakers, ages 14 to 21, retained *ne* at rates of only 19% and 15% in the respective corpora. Although overall *ne* retention rates had dropped in the period between the two studies, the data suggest strongly that older speakers tend to use *ne* more frequently than do younger ones. Another

possible explanation is that the observer's paradox may be more pronounced during interviews with older informants, and it is possible that non-retention would be more manifest in the observer's absence. Ashby (2001) does not, unfortunately, address this issue.

Coveney (1996) reported similar results, as informants between the ages of 50 and 70 years (all women) used *ne* at a rate of 28.8%. Not surprisingly, the youngest informants, aged between 17 and 24 years, had the lowest retention rates—only 8.4%. Once again, "age of speaker emerges as the most important differentiating factor" (Coveney, 1996, pp. 86-87). Results reported in Hansen & Malderez (2004) confirm that age is the most important social factor concerning *ne* use. The oldest speakers in their corpus—aged 51 to 64 years—retained *ne* at a rate of 22.3%, while informants aged between 15 and 23 years, retained *ne* in only 4.6% of occurrences of negation. By comparing the four studies mentioned above, we see that *ne* retention rates appear to be declining, yet are consistently higher among older speakers.[1]

Table 3. Age of speaker and *ne* retention.

Corpus	Older Speakers	Younger Speakers	Overall Retention Rate
Ashby (1981)	52%	19%	37%
Coveney (1996)	28.8%	8.4%	18.8%
Ashby (2001)	25%	15%	18%
Hansen & Malderez (2004)	22.3%	4.6%	8.2%

In addition to the speaker's age, other social factors—including education level and social class—appear to influence the use of *ne*. Ashby (1981) and (2001), as well as Coveney (1996) and Hansen & Malderez (2004), identify social class as an important predictor of *ne* retention rates. Upper-middle class speakers tend to retain *ne* at higher rates than do their middle-class and working-class counterparts. However, in Ashby (2001), social class was found to be less influential among younger speakers than among older speakers where "la

[1] Coveney (1996) and Hansen & Malderez (2004) also provide data for different age groups. These results were not included in the table in order make the comparison between Ashby's (1981) and (2001) two generation-based age groups.

variable (*ne*) reste un indicateur d'appartenance sociale" (p. 13).[2] This suggests that age is, in fact, a more important co-variant of *ne* than social class.

Following Ashby (2001), Hansen & Malderez (2004) also support the claim that, all things being equal, age is the one social factor that is of most importance. Although education level was explored and found to be somewhat influential, the authors conclude that "l'âge des locuteurs semble structurer la variation de manière beaucoup plus convaincante que le niveau d'études" (p. 19).[3] Geographic origin was also examined; however, its influence on the variable *ne* remains uncertain since those informants coming from different parts of France normally arrived in the Paris area after the age of 25 for work-related reasons (no children were classified as coming from outside the Paris area). Thus, age seems to be, once again, the most important social factor.

A final social factor appears to have some influence on the variable use of *ne*: the speaker's gender. Ashby (1981) reported that women were more likely to omit *ne* than men, and he provides overall retention rates of 30% and 42% for women and men, respectively. Ashby (1981) noted that "[i]f one accepts the theory that *ne* is indeed now being lost in French, it appears that women are in the vanguard of this change" (p. 685); however, Ashby (2001) found that the speaker's gender was not as important as it had been in the corpus from 1976, as men and women retained *ne* in 20% and 17% of all negations, respectively, in the 1995 corpus. Coveney (1996) also reported that *ne* retention rates were approximately equivalent between men and women (16% and 15% respectively). Similar results were found by Hansen & Malderez (2004), who found that the speaker's gender "ne semble pas jouer de rôle décisif" (p. 18).[4]

[2] Translation: "the variable (*ne*) remains an indicator of social class."
[3] Translation: "the speaker's age seems to structure the variation much more convincingly than does his or her level of education."
[4] Translation: "does not seem to play a decisive role."

Although these results might suggest that the speaker's gender has become less influential over the past two and a half decades, they also support Ashby's (1981) claim that women are leading change in progress if, of course, one accepts that men have simply "caught up" linguistically to women. However, it is unfortunate that Ashby (2001) does not offer any other explanations concerning the effect of the speaker's gender on *ne* retention in either of his corpora. It is entirely possible, for example, that Ashby's female informants were—for reasons unknown—more at ease with him during the interviews than their male counterparts in the 1976 data, and, therefore, they produced fewer tokens of *ne*.

2.3 Phonological Factors

Ashby (1981) provided evidence that the variable use of *ne* is also influenced by a certain number of phonological factors. Ashby noted that *ne* was "likely to be retained postpausally . . . and in intervocalic position, provided one of the vowels is nasal" (p. 677). However, Ashby found no evidence of regressive nasal assimilation when a consonant precedes the *ne* position, which differs from what appears to occur in Montreal French (Sankoff & Vincent, 1977).[5]

It is also surprising, according to Ashby (1981), that *ne* retention was not favored between two nasal vowels; however, he argues that had there been more than six such tokens in the corpus, the probability for *ne* retention in such a phonological environment would have been higher. Coveney (1996), following Pohl's (1968) observation, suggested that this phonological environment did indeed merit investigation.

Another phonological factor that affects the variable use of *ne* is coalescent assimilation. For example, Coveney (1996) provides the following cases: *je sais pas* [ʃɛ pɑ] and *je suis pas* [ʃɥi pɑ]. In these cases, not only is the schwa deleted, but [ʒ] and [s] are

[5] Assimilation is the process by which a phoneme is changed to match an adjacent one, usually to facilitate pronunciation in rapid speech. In regressive nasal assimilation, a following nasal vowel influences the preceding phoneme (e.g., *tu en as* → *t'en as*).

combined, producing [ʃ]. According to Coveney (1996), "[i]n such instances, when the phonological environment is so radically different according to whether the *ne* is present or absent, it seems more reasonable to say that the grammar is constraining the phonology, rather than vice versa" (p. 78). Coveney even proposed that one possible solution to this problem would be to exclude such cases from quantitative studies of *ne*; however, this would eliminate a large number of tokens of negation. He decided, therefore, not to "attempt any general quantification of phonological environment" (p. 78), but to bear it in mind, as it may be a contributing factor in a number of syntactic and grammatical environments.

Tokens of negation found in one phonological environment have, however, been considered impossible to study: the presence of [n] preceding a verb that begins with a vowel (or vowel sound). Since *ne* [nə] becomes *n'* [n] in the prevocalic position, it is very difficult—if not impossible—to distinguish prevocalic [n] of the negative particle *ne* from [n] resulting from elision. For example, the sentence *on n'est pas* "one is not" or "we are not"[6] is, phonetically, the same as *on est pas* (i.e., both sentences are pronounced [ɔ̃nɛpa]). This particular environment does not lend itself to the study of *ne* use in spoken French; thus, tokens of negation in this phonological environment have been excluded from previous studies (see Ashby, 1981 and 2001; Coveney, 1996; Armstrong & Smith, 2002; Hansen & Malderez, 2004).

2.4 Linguistic Factors

A certain number of internal linguistic factors have also been demonstrated to influence the variable use of *ne*: specifically, Neg2 type, lexicalization, grammatical subject, and sentence type have been found to be important differentiating factors (see Ashby, 1981; Coveney, 1996; Armstrong & Smith, 2002; Hansen & Malderez, 2004).

[6] The pronoun *on* can be used for indefinite reference to mean "one" and for definite reference to mean "we." Peeters (2006) provides an overview of the various possible referents for the pronoun *on*.

According to Armstong & Smith (2002), more frequently occurring Neg2s—most notably *pas*—co-occur less frequently with tokens of *ne*, confirming what Ashby (1981) and Coveney (1996) had previously found in their respective corpora. Results reported by Hansen & Malderez (2004), who found that "plus les éléments entrant dans l'expression négative sont fréquents, plus il y a chute du *ne*" (p. 25),[7] suggest that this occurs in Parisian French as well. Table 4 shows *ne* retention rates according to Neg2 type and frequency as reported by Ashby (1981), Coveney (1996), Armstrong & Smith (2002) and Hansen & Malderez (2004).

Table 4. Retention rates of *ne* reported in previous studies according to Neg2 type.

Neg2	**Ashby (1981)**		**Coveney (1996)**		**Armstrong & Smith (2002)**		**Hansen & Malderez (2004)**	
	Tokens	*Retention*	*Tokens*	*Retention*	*Tokens*	*Retention*	*Tokens*	*Retention*
pas	2,330	33%	2,317	16.4%	1,748	70.5%	941	8.2%
plus	127	51%	209	25.8%	85	77.6%	96	9.4%
rien	104	34%	146	21.2%	57	80.7%	61	6.6%
jamais	73	36%	84	26.2%	57	80.7%	35	11.4%
que	115	59%	109	34.9%	60	95%	23	30.4%
personne	20	75%	24	33.3%	3	33.3%	10	0.0%
aucun	N/A	N/A	33	21.2%	50	82%	N/A	N/A
Multi.[8]	24	41%	N/A	N/A	N/A	N/A	N/A	N/A

Table 4 shows clearly that there is a general tendency to omit *ne* more often when the Neg2 occurs more frequently. However, *que* and *plus* seem to contradict this pattern. Ashby (1981) explained that this may occur since "[t]hese second negatives may not be used elliptically, as *jamais*, *rien*, and *personne* may be" (p. 679).[9] It also seems that semantic factors may play a role since "*que* and *plus* may be less categorically negative than the other second negatives" (Ashby, 1981, p. 679). Armstrong & Smith (2002) commented on this phenomenon as well, suggesting that *ne* may be used in order to avoid confusion or ambiguity; especially in the case of *plus* since *plus* can be used as a Neg2 or in a positive sense. However, the authors admitted that the affirmative *plus* is most usually pronounced

[7] Translation: "the more frequent the negative elements are, the more *ne* is dropped."
[8] "Multi." refers to instances of multiple negation; that is, two or more Neg2s are used (e.g., *je n'ai plus rien*).
[9] For example, the question "Qui est venu ce soir?" 'Who came this evening?' can be answered elliptically by the Neg2 *Personne* ('Nobody').

with the final [s] (i.e., [plys]), whereas the Neg2 is not (i.e., [ply] or even [py]) (p. 37).

Hansen & Malderez (2004) added another Neg2 type: *pas* followed by an adverb (see also Coveney, 1996). For example, *pas tellement, pas vraiment, pas du tout* were included in the [*pas* + adverb] category, and *ne* retention was only 3.9%—a rather striking difference when compared to the 8.2% retention rate reported in the *pas* category. Additionally, Hansen & Malderez (2004) treated the expression *pas mal* as a separate category; and, of the 35 occurrences of this expression in the corpus, not one included *ne*. Coveney (1996) had previously reported similar results, which suggests that a certain number of [*pas* + adverb] combinations (especially *pas mal* and *pas cher*) "might almost be thought of as single lexical items" (p. 80). Following Coveney (1996), Hansen & Malderez (2004) went even further, suggesting that "[i]l s'agit peut-être d'une espèce de lexicalisation récente des expressions de ce type qui feraient perdre le statut proprement négatif de *pas*" (p. 23).[10] Although far from conclusive, the data in the Hansen & Malderez corpus suggest that a number of [*pas* + adverb] combinations are becoming lexicalized in modern French since little or no difference was observed in the Péretz-Juillard corpus between [*pas* + adverb] retention rates and *pas*-alone rates. If the lexicalization of other *pas* + adverb combinations is indeed occurring, it is most likely following the model of *pas mal* (Coveney, 1996; Hansen & Malderez, 2004).

In addition to the expression *pas mal* and other [*pas* + adverb] combinations, it appears that certain expressions are "pre-formed sequences" as opposed to novel sentence structures. These sequences disfavor *ne* retention since they occur at such high frequencies that they seem to have undergone the process of lexicalization, albeit not entirely in all cases (Coveney, 1996). In particular, the sequences *je (ne) sais (pas), je (ne) suis (pas), ce (n') est (pas), il (n') y a (pas),* and *il (ne) faut (pas)* have been classified as pre-formed sequences (see Ashby, 1981; Moreau, 1986; Coveney, 1996; Armstrong & Smith, 2002; Hansen &

[10] Translation: "It may be a sort of recent lexicalization of this type of expression, which would cause *pas* to lose its negative status."

Malderez, 2004). In addition, a certain number of other subject-verb sequences have also been selected as candidates for pre-formed sequences (see Coveney, 1996; Hansen & Malderez, 2004). *Ne* is frequently omitted in these cases because these sequences are used at such high frequencies in affirmative clauses.

> Since affirmative clauses are far more numerous than negative ones, it could be argued that it is essentially the subject + verb sequence which speakers operate with, regardless of whether a negative item follows or not. (Coveney, 1996, p. 79)

Hansen & Malderez (2004) concurred with Coveney (1996), noting that the "hyperfréquence des verbes *être, avoir, savoir et pouvoir*" (p. 25, italics in original),[11] which co-occur frequently with the subject clitics *je, tu, il,* and *ce,* and the Neg2 *pas*, provides an environment that strongly disfavors *ne* retention. Hansen & Malderez (2004) reported that *ne* retention with frequently occurring verbs (i.e., verbs that occurred more than 20 times in the corpus) was at only 7.2%, whereas less frequently occurring verbs collocated with *ne* at a rate of 13.3%. Ashby (1981) and Coveney (1996) reported on *ne* retention in several specific sequences, as shown in Table 5.

Table 5. *Ne* retention rates in preformed sequences (Ashby, 1981; Coveney, 1996).

Sequence	**Ashby (1981)** **Overall Retention = 37%**	**Coveney (1996)** **Overall Retention = 18.8%**
je sais	14%	8.6%
je suis	N/A	14.3%
c'est	7%	3.6%[12]
il y a	24%	2.4%
il faut	16%	10.5%

Both Ashby (1981) and Coveney (1996) demonstrated clearly that *ne* retention in these sequences was well below overall retention rates in their respective corpora. Incidentally, the sequences *il y a* and *il faut* are somewhat problematic since the subject clitic subject *il* is

[11] Translation: "The extremely high frequency of the verbs *être* (to be), *avoir* (to have), *savoir* (to know), and *pouvoir* (to be able)."
[12] Coveney (1996) includes *ça est* in this category.

often dropped (either completely or partially)[13] in spoken French (Coveney, 1996; Sax, 2003). Ashby (1981) did not mention this, and it is unclear whether or not his data included instances of both *il y a* [ilja] or [ija] and *y a* [ja]. Coveney (1996) considered both forms in his data; however, he noted that "once the impersonal pronoun *il* has been omitted (as it is extremely frequently), there is a very strong tendency to also omit *ne*, in order to avoid the rare initial cluster [nj]" (p. 81). Hansen & Malderez (2004) took a different position and excluded from their data examples in which the subject clitic had been dropped (e.g., *y a pas, faut pas*) since, according to them, *ne* omission was, in this case, obligatory.

In addition, other types of subjects appear to influence *ne* use. For example, clauses in which the subject is a noun phrase (NP) tend to favor *ne* retention, while those in which a subject pronoun is used tend to favor *ne* omission. Table 6 shows this tendency as reported in previous studies.

Table 6. *Ne* retention according to subject type as reported in previous studies.

Subject Type	Ashby (1981)	Coveney (1996)	Armstrong & Smith (2002)	Hansen & Malderez (2004)
NP	78%	67.2%	N/A	56.4%
non-clitic pronoun[14]	57%	N/A	89.2%	N/A[15]
clitic pronoun	28%	14.6%	61.7%	5.8%

Table 6 clearly demonstrates that clauses whose subject is a NP favor *ne* retention, as these rates are well above the overall rate reported in all three studies that examined this factor. It is also clear that *ne* retention is not favored in clauses containing a clitic subject pronoun, as these rates are consistently well below the overall rate. Ashby (1981)—among others (specifically Fonseca-Greber & Waugh, 2003a, 2003b)—have suggested that clitics "are now in the process of becoming bound to the verb at the morphological level" (p. 680), which may

[13] Clitic *il* can either be dropped completely, or the [l] may be deleted, leaving only the [i] audible (e.g, *i faut* [ifo] or *i y a* [ija]).
[14] This category includes instances of *cela* and *quelqu'un* (Ashby, 1981).
[15] Hansen & Malderez (2004, pp. 21-22) count non-clitic and clitic pronouns together; their results are, therefore, somewhat difficult to compare to the other studies.

be contributing to the loss of *ne*.

> As the subject clitic and verb grow more and more closely bound, *ne*, which can only occur between them, may be progressively squeezed out. And since it has become only a redundant mark of negation in modern French, *ne* can easily be dispensed with. (p. 681)

Coveney's (1996) data support the hypothesis that the use of a subject clitic disfavors *ne* retention, yet he offered a different line of explanation.

> Clitic subjects collocate with negated verbs far more frequently than do NPs, not only because they are about ten times more frequent in discourse, but also because they form a closed class, whereas NPs are, of course, an open, indeed, infinite set. (p. 73)

Although *ne* retention appears to be disfavored when a subject clitic is used, *ne* retention rates appear to vary between the various subject clitics. More specifically, *je, tu, on,* and *ce* have been found to collocate with *ne* at a very low frequency, regardless of the Neg2 and verb used in the negation. On the other hand, *il, elle,* and *ils* collocate with *ne* at a relatively higher rate, (see, for example, Hansen & Malderez, 2004). Extremely low rates of *ne* retention appear to be, according to Ashby (1981), a result of another continuing change in modern French: "the fusion of the clitic pronoun and the verb" (p. 868).

Another linguistic factor that appears to influence the variable *ne* is clause type; and, by extension, verb tense and form (e.g., simple or compound, personal or impersonal, etc.). Ashby (1981) reported that *ne* was retained 31% of the time in declarative clauses, while *ne* was retained categorically in imperative clauses. Armstrong & Smith (2002) reported a high rate of *ne* retention in imperative clauses as well (95%).

Along the same lines, it has been demonstrated that *ne* retention remains relatively high in co-occurrence with negated impersonal verb forms (i.e., those that are not conjugated

with a subject, such as present participles and infinitives). Ashby (1981) reported a *ne* retention rate of 68% in this environment and Hansen & Malderez (2004) reported a rate of 26.3%. Coveney (1996) and Armstrong & Smith (2002) separated infinitives and present participles and reported the following retention rates.

Table 7. *Ne* retention in [- overt] subject environments as reported by Coveney (1996) and Armstrong & Smith (2002).

Corpus	Retention with pres. part.	Retention with infinitives
Coveney (1996)	50%	62.5%
Armstrong & Smith (2002)	66.7%	93.6%

Armstrong & Smith (2002) offered one possible explanation for higher rates of *ne* retention in this particular environment.

> It seems likely that the relative rarity of negated verbs lacking overt subjects leads the speakers to focus on the negative markers in the utterance more than they would do in expressions with subjects, and this greater self-monitoring would naturally entail a higher rate of *ne* retention. (p. 36)

Other verb forms have also been taken into consideration; in particular, retention rates in compound and simple tenses have been compared. Ashby (1981) provided evidence that *ne* retention was more likely in compound tenses, as retention rates with the auxiliary verbs *être* and *avoir* were 55% and 50%, respectively, whereas *ne* was retained 35% of the time with lexical verbs. Hansen & Malderez (2004) reported similar findings, as retention rates in compound tenses reached 13.2%, while *ne* was retained only 7.8% of the time in simple tenses. These results suggest that compound tenses favor *ne* retention. Ashby (1981) also found evidence that *ne* was less likely to be retained with the aspectual auxiliary *aller* (29% of all instances).

2.5 Stylistic Factors

It has been frequently suggested in the literature that *ne* is dependent on a number of stylistic factors; primarily formality and discourse topic. The influence of formality and

discourse topic is particularly evident when one compares results reported in Armstrong & Smith (2002) with those reported in studies concerning more informal speech (Ashby, 1981; Coveney, 1996; Hansen & Malderez, 2004).

Ashby (1981) attempted to quantify this factor by comparing *ne* retention in the first half of each interview with informants with that of the second half. He asserted that interviewees became more at ease with him as the conversation progressed, and, as a result, their speech became less guarded in the later half of the conversation. *Ne* retention should, therefore, be lower in the second half of the interview than in the first half. This was shown to be true—at least to some extent—as retention rates dropped on average by 2% (from 37% to 35%). Although these results suggested that speakers retain *ne* at higher frequencies when they believe that the context requires a more formal or more highly monitored speech style, they are less than conclusive. However, Ashby (1981) was able to follow three informants as they "moved between two widely different social settings" (p. 681). *Ne* retention rates for these three speakers dropped from an average of 37% in the formal setting to only 16% in the informal setting, thereby confirming Ashby's presumption.

It has also been suggested that discourse topic may have an effect on the variable use of *ne*. Indeed, Ashby (1981) explored this aspect by selecting five speakers who had the lowest rates of *ne* retention. He identified 24 tokens of *ne* and attempted to determine why these tokens were produced by these speakers. Four *ne* tokens occurred when the speaker was responding to or repeating a question in which *ne* had been used. Another two tokens occurred after a false start, which suggested self-monitoring on the part of the speaker. One token was used in a direct quotation and five were produced while talking about disciplining children. Another five tokens were classified as proverbial, following the example of Sankoff & Vincent (1977),[16] and two more were produced during the discussion of serious topics.

[16] Proverbial refers to truisms or general statements or rules (see Sankoff & Vincent, 1977).

Only four tokens, then, seemed to escape explanation, but Ashby did not comment on any other linguistic or phonological factors that might have led to their presence. Results reported in Armstrong & Smith (2002) provide the most conclusive evidence that *ne* retention is strongly favored in highly monitored speech (i.e., formal contexts).

Coveney (1996) found that one informant in particular had *ne* retention rates that differed greatly from the scores of the others. In his opinion, this was due—at least in part—to the fact that this young male was an assistant camp director and may have felt the need to represent the camp favorably by speaking more formally. Further, the informant asked during the interview if he should speak slowly in order to make his speech more easily understood by native speakers of English, which presumably resulted in a more formal register of speech (i.e., less similar to his vernacular since more attention was paid to his speech). The speaker's belief that a more formal register of speech was required led Coveney (1996) to consider his scores separately; however, this supported the hypothesis that speakers retain *ne* at higher frequencies in situations that, in their view, require more a formal register of speech; for example, sociolinguistic interviews with a researcher (Labov, 1972).

Ashby (2001) acknowledged the effect that his presence—as well as the presence of the tape recorder—might have had on informants, yet insists that speakers in his study were at ease during the interviews.

> Ces conditions n'ont certes pas favorisé un registre tout à fait familier, où la chute du *ne* serait probablement encore plus manifeste, mais dans l'ensemble, au bout de quelques minutes les locuteurs avaient l'air assez à l'aise, comme s'ils oubliaient la présence du magnétophone, et ne manifestaient aucune réticence à parler. (p. 8)[17]

[17] Translation: "These conditions certainly did not favor a completely familiar register of speech, where *ne* omission would probably be more manifest, but as a whole, after a few minutes the speakers seemed at ease, as if they had forgotten about the tape recorder, and showed no reluctance to speak."

As Coveney (1996) and Ashby (2001) pointed out, sociolinguistic interviews may be perceived as a more formal speech environment by some speakers; much like the radio interviews analyzed by Armstrong & Smith (2002), albeit to a lesser extent.

Hansen & Malderez (2004) brought another interesting stylistic factor into perspective: *ne* production in children's written school assignments. Since children between the ages of 5 and 14 years had omitted *ne* categorically in their corpus of informal speech, the authors decided to explore *ne* use in written school assignments of young children. Hansen & Malderez (2004) reported that *ne* retention in these written assignments was as high as 75% with noun subjects and 50% with pronominal subjects—a striking difference between rates reported in the speech of similarly aged informants. Likewise, Blanche-Benveniste (1997) noted that children in her corpus were perfectly capable of using *ne* appropriately when imitating the speech of elegant women. These results suggest strongly that *ne* is not being completely lost from the French language. Hansen & Malderez (2004) posited that "on est encore très loin d'un tel aboutissement pour la simple raison que les locuteurs se trouvent exposés à l'usage de ce *ne* dans une diversité de contextes qui le retiennent certainement dans leur système linguistique; d'abord dans le français parlé formel" (p. 26)[18]

2.6 Summary of the Relevant Literature

Previous studies clearly indicate that *ne* use is indeed dependent on a number of social, linguistic, and stylistic factors. Age, in particular, stands out as an important demographic factor, as does the social class to which the speaker belongs. It is also clear that *ne* use is a sociolinguistic variable that might be considered a marker of social identity since retention rates tend to be higher among older speakers of the higher social class (Ashby, 2001).

[18] Translation: "we are still far from such an outcome for the simple reason that speakers are exposed to the use of this *ne* in diverse contexts, which retains it in their linguistic system; firstly, in formal spoken French."

Among the phonological and linguistic factors, subject type might be considered a determining factor, especially when a subject clitic is used and assimilation is possible (see Ashby, 1981; Coveney, 1996). Other important factors include verb type and form, Neg2 type, and frequency of [subject + verb] sequences.

Style remains a determining factor since highly monitored speech styles appear to favor *ne* retention (Armstrong & Smith, 2002). It is also important to remember that the speaker's perception of the required level of formality of a given communicative context plays a role (Coveney, 1996). Finally, as Hansen & Malderez (2004) noted, even young children who omitted *ne* categorically in informal speech were able to use it appropriately in written assignments. In addition, children appear to be capable of using 2Neg in imitation of certain stylistically marked contexts, such as the speech of elegant women (Blanche-Beneveniste, 1997).

The present study attempts to address many of the same factors explored in the aforementioned research. Although interpersonal variation and demographic factors have always been taken into consideration in previous variationist studies, the present study does not attempt any such quantification since informants in chat environments are able to remain more or less anonymous, and it is impossible to gather (credible) personal data about such informants. However, as mentioned in 1.3.4, the anonymity granted by the chat environment provides an ideal context in which to observe informants' vernacular (at least as far as synchronous CMC is concerned). It is reasonable to assume that the informants' language is less guarded than that of informants in sociolinguistic interviews since it is less directly observable (i.e., informants in chat environments do not participate in interviews with a researcher and are more or less unaware that observation is taking place).

CHAPTER 3

METHODOLOGY

3.1 The IRC Environment

Internet Relay Chat (IRC) is a protocol that enables users to chat in real time with one or more interlocutors by sending text-based messages when logged on to a specific chat room (i.e., chat channel). Chat rooms may be public or private and are generally hosted by a network, which is itself connected to a server. IRC servers are freely accessible, provided that the user has access to an Internet connection and has downloaded one of several IRC clients available on the Internet.[1] The IRC environment is rather anonymous, and participants see only the pseudonyms (i.e., screen names) of their interlocutors.

Once logged on to a channel, the user simply types a message and presses the "enter" button to send it. Any person connected to the chat channel will see the message. Since communication in the IRC environment occurs in real-time and several (sometimes dozens) of participants are often sending and responding to messages simultaneously, chat discussions can appear to be random and are potentially difficult to follow for the non-initiated. Excerpt 1 provides an example of data taken during an IRC discussion.

Excerpt 1.
<Ayame> jme suis acheté une épée XD
<Okko> Qu'est-ce que j'ai 'core fait ?
<KaM> bisouuuuuuuuuuuus ma tite Ayame au fait :p
<Okko> Ayame oO
<Okko> pourquoi faire ? couper les carottes ? :o
<romanticboy> lu Ayame
<romanticboy> tu va bien?
<romanticboy> lu Okko

[1] An IRC client is a program which connects a user to his or her choice of IRC servers. Pierozak (2003c) provides a detailed description of the French-language IRC environment.

<Okko> salut romanticboy
* taku`[Oforum] is now known as taku
<romanticboy> j t ai deja di hier Ayame ca te monte a la tete ces jeux!
<romanticboy> lollllllllllll
<romanticboy> ;)
<Ayame> ça va et toi romanticboy ?
<luka> moi aussi g une épée :)
<Ayame> Okko pour décorer ^^
<Ayame> mecmsn20 dégage
<Okko> Tu sais Ayame, pour le self-defense, y a mieux que les épées, c'est pas très discret pour sortir dans la rue, surtout en cette saison, j'te vois mal avec un pardessus pour cacher le bazar :o
['<Ayame> i bought myself a sword XD
<Okko> What have i done now?
<KaM> mwaaaaaaa my little Ayame actually :p
<Okko> Ayame oO
<Okko> to do what? cut carrots? :o
<romanticboy> hey Ayame
<romanticboy> how're you?
<romanticboy> hey Okko
<Okko> hey romanticboy
* taku`[Oforum] is now known as taku
<romanticboy> i already told you yesterday Ayame these games are going to you head!
<romanticboy> lollllllllllll
<romanticboy> ;)
<Ayame> fine and you romanticboy ?
<luka> i also have a sword :)
<Ayame> Okko for decoration ^^
<Ayame> mecmsn20 get out of here
<Okko> You know Ayam, there's better than swords for self-defense, it's not very discreet for going out, especially at this time of year, I have trouble picturing you in a trench coat to hide it :o']

As excerpt 1 illustrates, several discussions are taking place at the same time, and in the space of approximately two minutes, six different users send messages. As Herring (1999) noted, communication in the chat environment can be somewhat incoherent. Despite the apparent lack of interactional coherency, however, IRC and other forms of synchronous CMC have become popular communicative environments around the globe.

3.2 Data Collection

The data used in this study were collected from two different age-based chat channels (*#18-25ans* and *#25-35ans*) found on the public IRC server EpikNet.[2] Data was collected over the course of four days—selected at random—for approximately four hours each time during the fall of 2005. The chat discussions were recorded using the transcript recording function provided in the mIRC software[3] and saved as text files for analysis. The corpus constitutes a wide range of conversation topics and discussions held during different times of day (afternoon and evening), as well as during both weekday and weekend times.

During data collection times, I did not actively engage in the discussion, nor did I reveal my identity as a researcher. The chat participants were—as far as can be known—unaware that observation was taking place.[4] However, it must be noted that since this type of communication occurs in a public space, the participants are certainly aware of the possibility that any number of people could be following the chat session or reading the log.

Following data collection, the transcripts of data were reviewed and analyzed, and all server- and human-generated turns were counted.[5] In order to determine the size of the corpus, every human-generated word was counted using a concordance program,[6] which will be discussed in the following section.

Table 8 shows the number of turns (both server- and human-generated) and the number of human-generated words found in the corpus.

---

[2] EpikNet is a free-access text-based IRC server with a large selection of French language chat channels. More information about EpikNet can be found at the following URL: http://www.epiknet.org/

[3] mIRC software is one of several programs that allows IRC users to access different IRC severs. Information about mIRC software can be found at the following URL: http://www.mirc.com/

[4] This project—including data collection method—was reviewed and approved by the Institutional Review Board for the Protection of Human Subjects of the University of North Texas. Project application #05-357.

[5] I have opted to define a turn as a message sent by a participant when he or she hits the Enter key (human-generated) or when the server sends a message (server-generated). This definition is not, however, unproblematic. For a discussion of turn-taking in chat, see Thorne (1999, ch. 5) or Williams (2003, ch. 3). For a general treatment of turns, see van Lier (1988).

[6] Information about Concordance© software can be found at the following URL: http://www.concordancesoftware.co.uk. A detailed explanation is provided in section 3.3.

Table 8. Turns and Human-generated Words

Chat Channel	Total No. of Turns	Human Gen. Turns	Server Gen. Turns	Human Gen. Words
#18-25ans	8,774 (47.78%)	6,864 (78.23%)	1,910 (21.77%)	37,752 (48.32%)
#25-35ans	9,591 (52.22%)	7,690 (80.18%)	1,901 (19.82%)	40,373 (51.68%)
Total	**18,365 (100%)**	**14,554 (79.25%)**	**3,811 (20.75%)**	**78,125 (100%)**

Server-generated turns were not analyzed further since the objective of this study is to investigate the variable use of *ne* by human participants engaged in chat discussions. The human-generated turns were reviewed and analyzed, and every token (i.e., occurrence) of negation was identified and coded as described in the following sections.

3.3 The Concordance© Software and Counting Methodology

Concordance© is a program that allows the user to compile a complete concordance of all words found in a text document, the frequency of each word, and where it can be found in the text. Figure 1 is a screen capture of the Concordance© software.

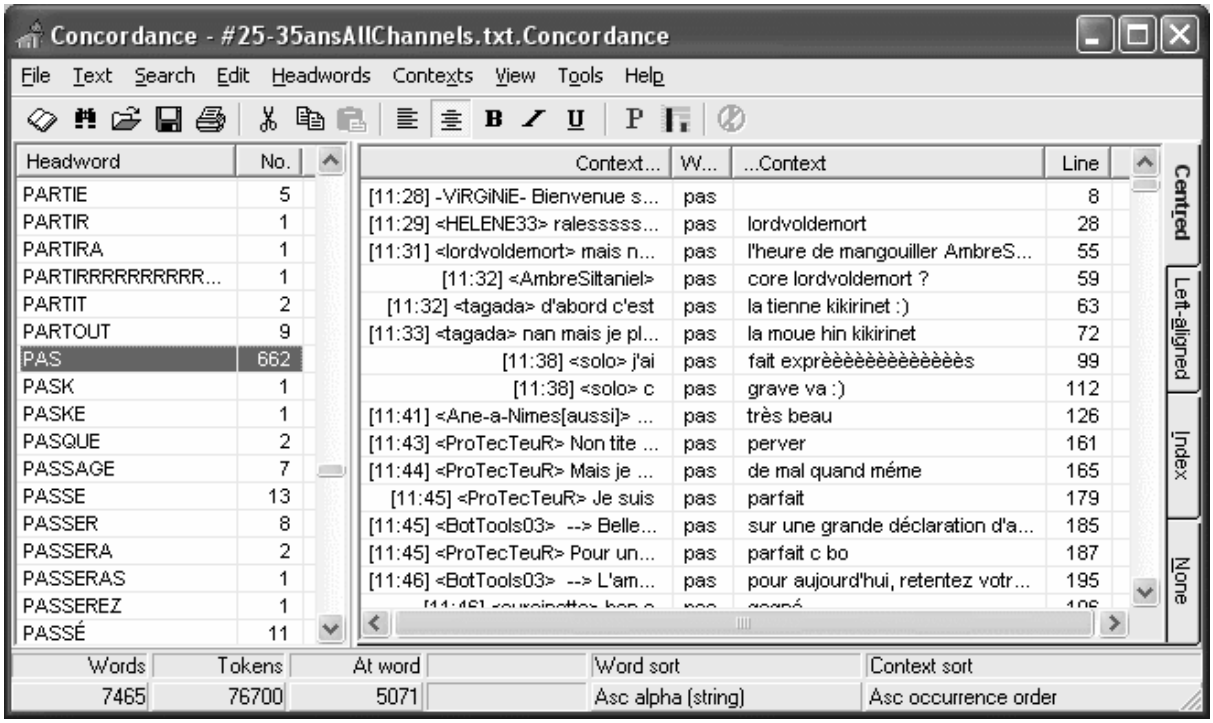

Figure 1. Concordance© software: headword list and occurrences.

A list of headwords found in the file is shown on the left side of the screen and the number of occurrences is provided. By selecting one of the headwords from the list, the user is able to view all occurrences of that word to the right. In Figure 1, the word *pas* has been

selected and all occurrences of *pas* found in the text are shown to the right, as well as where it appears in the text.

In addition to providing a list of headwords and their location in the text file, Concordance© allows the user to view each occurrence in its context by selecting one of the examples found in the right window. A separate window appears in which the selected occurrence is highlighted, as shown in Figure 2.

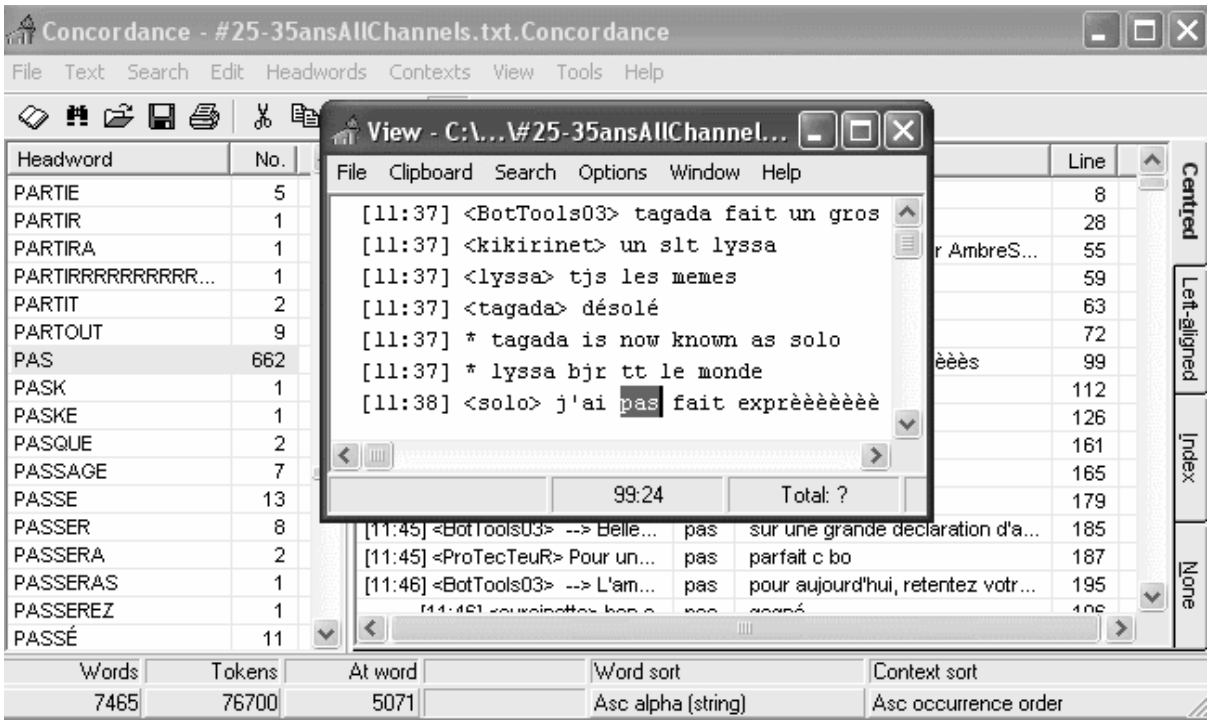

Figure 2. Concordance© software: text viewer.

The text viewing function is especially useful when reviewing excerpts of data since the user is able to view the token in its context (i.e., in order to read turns preceding and following the excerpt).

Although Concordance© is very efficient, certain problems did arise; specifically, typographical errors and abbreviations commonly found in chat discourse, as well as conjunctions with punctuation, were counted inconsistently by the program. This inconsistency often resulted in two or more words being counted as one (e.g., *pask'elle* or *parce qu'elle* was counted as one word), among other anomalies. In addition, *ne* was often omitted from the head-word list because of punctuation or spacing errors (e.g., *jene sais pas*

would have been counted as three words: *jene, sais, pas*).

In light of these problems, the head word list and many examples were carefully reviewed in order to ensure an accurate word count. All occurrences of Neg2s were also carefully reviewed and every example of negation was counted and coded as described in section 3.4.

3.4 Examples of Negation Found in the Corpus

Every token of negation in the corpus was identified and classified as either two-particle negation (2Neg) or single-particle negation (1Neg). Tokens of 2Neg were reviewed and instances of fixed expressions (e.g., *n'est-ce pas, ne serait-ce pas*, etc.) were eliminated.[7] Instances of 1Neg were then classified as one of six types shown in Table 9.

Table 9. Types of 1Neg found in the corpus.

Type	**Description**
A	Non-verbal sentence
B	Non-verbal sentence preceded by a verbal clause
C	Fixed expressions involving a Neg2
D	Verbal clause (non-fixed expression)
E	Verbal clause in which a clitic pronoun has been deleted
F	Non-traditional form

The following excerpts of data illustrate the different 1Neg types found in the corpus. The negation of interest has been underlined. The term *sic* has not been used to indicate that a grammatical, orthographic, or typographic error has merely been reproduced. A literal English translation is provided in brackets immediately below each excerpt. A second translation in more "natural" English is also provided (in single quotation marks).

1Neg Type A is usually a short response to a question or statement sent by another participant, as illustrated by Excerpt 2. This type of 1Neg can also be a question or statement by itself, especially when followed by an adjective, adverb or past participle (i.e., the subject and verb are not present). Excerpt 3 provides an example of this use.

[7] Hansen & Malderez (2004) provide this methodology since these expressions require *ne* to be present.

Excerpt 2.
<Eliot> pas vraiment non
[<Eliot> [Neg2] really no]
'<Eliot> not really no'

Excerpt 3.
<Eliot> pas trop oqp ange_away?
[<Eliot> [Neg2] too busy ange_away?]
'<Eliot> not too busy ange_away?'

1Neg Type B is a non-verbal clause, which is similar to Type A. However, Type B is immediately preceded by a verbal clause to which the Type B negation is normally linked, as shown in Excerpt 4.

Excerpt 4.
<Prue> n'empêche que c'est pas forcément vrai, mais pas faux non plus Hugo :)
[<Prue> [Neg1]-stop[8] that it's-[Neg2] necessarily true, but [Neg2] false either Hugo :)]
'<Prue> albeit it's not necessarily true, but not false either Hugo :)'

1Neg Type C includes instances of the expression *pas mal*. This expression is generally considered to have been lexicalized without *ne* in modern French (see Coveney, 1996), although there is some debate (see Hansen & Malderez, 2004). Excerpt 5 provides an example of *pas mal* use in a verbal clause. In addition, this expression can also be used as a quantifying adverb in lieu of *beaucoup* as shown in Excerpt 6. For these reasons, instances of *pas mal* have been counted separately from other instances of 1Neg. A number of examples of the expression *pas cher* have also been included in this category.[9]

Excerpt 5.
<romanticboy> c pa mal
[<romanticboy> it's [Neg2]-bad]
'<romanticboy> it's not bad'

Excerpt 6.
<ToUfOu> j'ai pas mal bossé
[<ToUfOu> i've [Neg2]-bad worked]
'<ToUfOu> i worked alot'

[8] Neg1 refers to the preverbal *ne*. The expression *n'empêche que* always includes *ne* (Grevisse, 1993). Such examples have not been counted as tokens of the negative particle *ne*.
[9] Most instances of *pas cher* have been counted as 1Neg Type D. However, it has been argued that *pas cher* is becoming lexicalized in French (see Coveney, 1996) since it can be used in an expression such as *on peut trouver des voitures à pas cher*. Such examples of *pas cher* were counted as Type C.

Type D of 1Neg is the most often occurring type, which includes a negated verb form as well as a subject (e.g., a noun phrase or pronoun), except in the case of imperatives and impersonal moods. Examples of 1Neg Type D constitute instances of *ne* deletion since the presence of *ne* was possible (i.e., between the subject and the verb), but *ne* was not used (i.e., omitted), as shown in Excerpt 7. Thus, I have opted to define 1Neg Type D as the variant of 2Neg. In other words, the paradigm [*ne* present vs. *ne* absent] could be re-written [2Neg vs. 1Neg Type D].

Excerpt 7.
<KaM> les gens vulgaire j'aime pas moi
[<KaM> the people vulgar i like-[Neg2] me]
'<KaM> I don't like vulgar people'

1Neg Type E occurs when chat participants attempt to imitate the spoken form of a certain number of expressions; specifically, *il faut* (Excerpt 8) and *il y a* (Excerpt 9). As in spoken French, the indefinite clitic pronoun *il* is very often deleted in chat. Other examples of subject clitic deletion were also counted as Type E (Excerpt 10). This is another example of how participants in this type of communication environment attempt to imitate certain traits common in spoken discourse.

Excerpt 8.
<Prue> y a pas de fontaine
[<Prue> there have-[Neg2] any fountain]
'<Prue> there are no fountains'

Excerpt 9.
<Salizar> Ben faut pas Petite-Peste xD
[<Salizar> [DM][10] better-[Neg2] Petite-Peste xD
'<Salizar> Well better not Petite-Peste xD']

Excerpt 10.
<Devotion> ah non suis pas gentil moi :(
[<Devotion> ah no am-[Neg2] nice me :(
'<Devotion> ah no i'm not nice :(']

[10] Discourse Marker

Although it seems that *ne* cannot be present in 1Neg Type E, these instances have not simply been discarded. I have opted to analyze separately 1Neg Type E and to consider such examples as indicative of the relationship between IRC discourse and everyday conversational speech. In addition, I have included 1Neg Type E in Chapter 5 as part of the discussion of features of discourse particular to chat.

1Neg Type F includes instances of non-traditional verb negation (i.e., grammatically incorrect usage); specifically, imperative clauses in which the Neg2 precedes the verb, which imitates certain spoken forms (e.g., a command for a family pet or other very informal contexts).

Excerpt 11.
<KaM> puis pas touche a ma ange_away
[<KaM> [DM] [Neg2]-touch to my ange_away
'<KaM> [DM] no touching my ange_away']

3.5 Coding of Tokens

After identifying every occurrence of 2Neg and 1Neg Type D, each token was coded for analysis. GoldVarb 2001 was used in order to analyze a number of linguistic (syntactic) factors: Neg2 type, subject type, phonological environment, and sentence type. GoldVarb 2001 is a multivariate analysis program that allows the user to test the influence of a number of independent variables on the dependent variable (in this case *ne*-present vs. *ne*-absent) and to determine the significance of the influence of each independent variable. Thus, it was possible to analyze not only the variable use of *ne*, but also other variables that co-occur and/or co-vary with *ne*. In Chapter 4, I describe GoldVarb coding in more detail during the discussion of each factor. I have also provided a detailed coding chart in Appendix A.

Following the analysis of syntactic factors, it became clear that *ne* use was determined to a large extent by discourse style (e.g., ludic, emphatic, proverbial, and explanatory). In order to test this hypothesis, a second GoldVarb analysis was conducted for a sub-sample of the data. I have provided more detailed information about this sample in 4.3 and the coding in

Appendix D.

Each token of negation was analyzed in its context. Turns preceding and following the example in question were reviewed in order to determine which pragmatic function (e.g., humor, emphasis, etc.) the token fulfilled. The text-viewing function of Concordance© software proved useful for the analysis of these tokens since it was often necessary to read several (sometimes many) turns above and/or below the example in order to understand the context of the conversation, which was not always immediately obvious given the interactionally incoherent nature of synchronous CMC (Herring, 1999). It was, in fact, often the presence of punctuation, smileys,[11] capitalization, a shift in the use of pronouns, or the reaction of another participant in a subsequent turn that made classification possible.

[11] "Smileys" are a type of emoticon. The Oxford English Dictionary defines an emoticon as "a representation of a facial expression formed by a short sequence of keyboard characters (usually to be viewed sideways) and used in electronic mail, etc., to convey the sender's feelings or intended tone." For a general treatment of smileys in French-language chat, see Pierozak 2003c.

CHAPTER 4

ANALYSIS OF THE CORPUS

In the following sections, I examine a limited number of linguistic factors in order to draw a comparison of the use of *ne* in speech and IRC discourse. In addition to exploring a variety of syntactic and phonological environments, I present and analyze a number of pragmatic features of discourse that appear to influence strongly the use of *ne* in this communicative context.

I am operating under the assumption that IRC constitutes a more or less informal communication environment (i.e., similar to everyday conversational speech). Moreover, the language observed in this environment should be relatively unguarded since the participants were unaware that observation was taking place.[1] It can therefore be assumed that the data collected and analyzed in this chapter represent the "vernacular" of the IRC environment.

4.1 Distribution of 2Neg and 1Neg

Before proceeding to the analysis of linguistic and pragmatic factors that influence *ne* use, let us first consider the overall distribution of 2Neg and 1Neg in the corpus. Table 10 demonstrates clearly that most second-negatives did not co-occur with tokens of *ne*.

Table 10. Distribution of 2Neg and 1Neg

Total Negation	**2Neg**	**1Neg**
1,648 (100%)	195 (11.83 %)	1,468 (88.17%)

Instances of 1Neg were divided into several types, depending on the nature of the negation (see 3.4). This was done in order to determine which examples could be compared

[1] It must be noted that the language used in this communicative environment is somewhat guarded since there are a number of operators, also known as "ops", who enforce the rules and regulations described in the netiquette. Profanity and the use of capital letters, for example, are prohibited and ops have the power to kick-out and/or ban any participant who does not respect these rules. This does not appear, however, to elicit a less-than-informal speech style where non-traditional forms and structures might be perceived as inappropriate.

with instances of 2Neg (i.e., examples of the variant of 2Neg). Table 11 divides tokens of 1Neg into six possible types of 1Neg.[2]

Table 11. 1Neg distribution.

1Neg Type A	1Neg Type B	1Neg Type C	1Neg Type D	1Neg Type E	1Neg Type F	Total
247 (17.00%)	50 (3.44%)	22 (1.51%)	1,017 (69.99%)	107 (7.36%)	10 (0.69%)	1,453 (100%)

Table 11 shows clearly that 1Neg Type D is the most often occurring type of 1Neg, representing approximately 70% of all 1Neg tokens. In the remainder of the present study, the distribution of 2Neg and 1Neg Type D (i.e., *ne* present vs. *ne* absent) will be explored. Table 12 illustrates the overall distribution of 1Neg Type D and 2Neg.

Table 12. Overall distribution of 1Neg Type D and 2Neg.

2Neg	1Neg Type D	Total
195 (16.09%)	1,017 (83.91%)	1,212 (100%)

It is clear that *ne* retention rates are low in this corpus (16.09% overall), which corroborates the findings of Coveney (1996), Ashby (2001), and Hansen & Malderez (2004). This in itself seems to support the assumption that, at least as far as the variable *ne* is concerned, the discourse of French-language IRC is similar to that of everyday conversational speech.[3] It is therefore reasonable to assume that IRC provides a communication environment in which informal language and non-traditional structures are perceived as appropriate (or at least not inappropriate).[4]

[2] Type A = non-verbal sentence; Type B = non-verbal sentence preceded by a verbal clause; Type C = fixed expressions; Type D = verbal sentence (non-fixed expression); Type E = verbal sentence, subject omitted; Type F = non-traditional structure. See 3.4 for a full description and examples of each 1Neg type.

[3] Armstrong & Smith (2002) provided evidence that, although *ne* deletion seems to be spreading to highly-monitored speech styles, high rates of *ne* retention are often associated with formal discourse, such as radio programs. Hansen & Malderez (2004) also reminded us that children, who often have the lowest retention rates, are capable of using *ne* correctly in written school assignments.

[4] Williams & van Compernolle (2007) demonstrated, for example, that the informal second-person pronoun *tu* is overwhelmingly preferred to the more formal *vous*-singular in IRC environments, and that *vous*-singular use is often perceived as strange or socially inappropriate.

4.2 GoldVarb 2001 and Statistical Procedures

A certain number of internal linguistic factors appear to structure the variable use of the negative particle *ne* in everyday conversational speech (Ashby, 1981; Coveney, 1996; Hansen & Malderez, 2004). In section 4.2, I explore a limited number of these factors (e.g., second-negative type (Neg2 type), subject type, and preceding and following phonological environment) in a corpus of synchronous French-language CMC. The data were analyzed with GoldVarb 2001 in order to demonstrate test for the statistical significance of each factor group.

Analyzing data with GoldVarb 2001 requires three basic steps. First, all tokens of the dependent variable must be coded according to the factor groups established by the researcher. For the purposes of this study, each occurrence of 1Neg Type D and 2Neg was coded by Neg2 type, subject type, preceding and following phonological environment, and sentence type. The GoldVarb 2001 coding-key for the analysis of the first coding is provided in Appendix A.

In the second step, a one-level binomial analysis is performed by the program. GoldVarb 2001 calculates the weight of each factor, or, in other words, the "probability of the dependent variable occurring in the context" (Tagliamonte, 2006, p. 220). A GoldVarb 2001 probability score equal to or greater than .500 indicates that the dependent variable is favored, while a score equal to or less than .499 indicates that the dependent variable is disfavored. GoldVarb 2001 reports these scores based on the order of the dependent variables when doing the analysis. In the present study, the analysis was performed on the bias of the presence of *ne*. Therefore, GoldVarb 2001 scores reported in the following analysis state the probability that *ne* will be present (i.e., $\geq .500$ = 2Neg favored, $\leq .499$ = 2Neg disfavored) according to each factor group. (The one-level binomial analysis report for the first coding is provided in Appendix B.)

The third and final step of a GoldVarb 2001 analysis involves what is called a "step-up/step-down analysis," which determines the "best fit of the model to the data" (Tagliamonte, 2006, p. 228). It is at this point that GoldVarb 2001 calculates the significance of each factor group considered and provides additional evidence of overlapping and interaction of factor groups. (The step-up/step-down analysis report for the first coding is provided in Appendix C.)

In addition to GoldVarb 2001 probability scores, chi-square was used where I had not coded for the multivariate analysis since "the chi-square test enables us to compare the frequencies we actually observe with those we should expect to observe on the basis of some theoretical model" (Butler, 1985, p. 112). For example, GoldVarb 2001 had been used to analyze the distribution of one- and two-particle negation according to subject type (e.g., noun phrase, pronoun, [- overt] subject); however, I wished to examine the variation within the pronoun category since much could be learned by exploring the distribution of negation according to pronoun type (e.g., clitic vs. non-clitic). Therefore, using chi-square, I was able to determine the statistical significance of a certain number of sub-factor groups that were not necessarily represented in the GoldVarb 2001 analysis.[5] Although chi-square has been used in previous studies of negation in French (e.g., Hansen & Malderez, 2004), it is used sparingly in this thesis only in order to understand linguistic variation within the system as a whole since chi-square cannot account for intraspeaker variation or weigh the frequency of each speaker's contribution of different types of tokens.[6]

[5] In the remainder of this paper, any mention of chi-square should be understood as a traditional or Pearson's chi-square test.
[6] For an overview of the assumptions underlying chi-square, see Hatch & Lazaraton (1991, ch. 14), Butler (1985, ch. 9), and Wilcox (1996, pp. 82-85).

4.3 Linguistic Factors

4.3.1 Analysis of First Coding

All tokens of 2Neg and Type D of 1Neg were coded for a variety of internal linguistic factors, including Neg2 type, subject type, preceding phonological environment, following phonological environment, and sentence type. After reviewing the data, it was obvious that too few tokens of Neg2s other than *pas* were available for independent statistical analysis. Therefore, all Neg2s other than *pas* were collapsed into one category, thereby eliminating a number of empty cells and extremely low cell counts. In addition, subject type was divided into three categories: noun phrase (NP), pronoun, and [- overt] subject (i.e., imperatives). Negated infinitives were not considered in the final instantiation of the analysis for this study due to a number of methodological problems, namely a lack of examples.

Most infinitives that were preceded by a negative particle followed a modal auxiliary, which is an environment where the presence of *ne* can alter the semantic interpretation of the negation, yet its absence leaves the phrase rather ambiguous. For example, *on ne peut pas aimer ce film* ('one $_{[Neg1]}$-can-$_{[Neg2]}$ like this film') conveys a different message than *on peut ne pas aimer ce film* ('one can $_{[Neg1]}$-$_{[Neg2]}$-like this film'). While the former expresses an *impossibility* (i.e., it is impossible to like this film) by concentrating the negation on the modal auxiliary *pouvoir*, the latter expresses a *possibility* (i.e., it is possible to not like this film) by focusing the negation on the following infinitive *aimer*. Yet when *ne* is absent (i.e., *on peut pas aimer ce film*), the possible nuance illustrated above disappears. In light of this limitation, it was assumed that the modal auxiliary was negated, not the following infinitive.

After recoding the data accordingly (i.e., excluding negated infinitives), a one-step binomial analysis was run in order to test each of the factor groups independently of one another (see Appendix B). The input probability score of 0.143 for this analysis indicated that the chances are 14.3% that "any given token" (Young & Bayley, 1996, p. 270) will retain *ne*.

"The input probability is the likelihood that the [2Neg] rule will operate in any circumstance, irrespective of conditioning factors" (p. 270). The data shown in Table 13 provide factor weights for all factor groups from the one-step binomial analysis of the first coding.

Table 13. One-level binomial: first coding.

Factor Group	**Weight**
Neg2 Type	
pas	.492
Neg2 other than *pas*	.530
Subject Type	
NP	.892
Pronoun	.454
[- overt] subject	.511
Preceding Phonological Environment	
Vowel	.447
Consonant	.486
Nasal Vowel	.544
No Preceding Phonological Environment	.863
Following Phonological Environment	
Vowel	.468
Consonant	.524
Sentence Type	
Declarative	.543
Interrogative	.446
Imperative	.172

The one-step binomial report revealed a number of problems in the distribution of the data. First, no convergence was found after 20 iterations, which indicates that the data do not fit within the expected theoretical model. Second, relatively high error scores were present in a number of cells. According to Preston (1996, p. 11), an error score of equal to or less than 2.0 indicates a good fit. In this first coding of the data, no less than eight cells had error scores greater than 2.0. Third, the total chi-square calculated was over 48.19, and in order for the results to pass the goodness-of-fit test (Young & Bayley, 1996, pp. 272-273), the total chi-square produced by Goldvarb 2001, the total chi-square (p = .05 with 9 degrees of freedom) should have been less than 16.919. These problems suggest that two or more factor groups may be interacting. A cross-tab analysis of factors groups 2 and 3 (i.e., subject type and preceding phonological environment) showed very unbalanced distribution in some

coding strings, which also suggests interaction between factor groups.

The step-up/step-down analysis of the first coding (see Appendix C) was reviewed in order to determine whether factor weight (i.e., constraint) rankings change as factor groups are run against each other. The step-up/step-down report revealed that whenever factor groups 2 and 3 (subject type and preceding phonological environment, respectively) were co-present during the regression analysis, no convergence was reached, which strongly suggests that the two factor groups are inextricably linked. Further, factor weight rankings for subject type actually changed in a number of runs. As Table 13 indicates, the constraint ranking should be, in order from *ne* most likely to least likely: NP, [- overt] subject, pronoun. However, in four runs where both subject type and preceding phonological environment are present (#8, #16, #20, and #23; Appendix C), pronouns and [- overt] subject environments are reversed. Subject type and preceding phonological environment are clearly interacting in this study. It is therefore necessary to separate the two interacting factor groups and run two separate analyses, a remedy suggested by Tagliamonte (2006, p. 234).

4.3.2 Analysis of Second Coding: Phonological Environment

Preceding and following phonological environment were first examined independently of all other factor groups in the second coding. The one-step binomial analysis for phonological environment (see Appendix D) indicated as 15.3% the probability that the rule (i.e., the presence of *ne*) would be applied to any given token, regardless of the environment. In addition, the total chi-square was below the value required ($p = .05$, 4 degrees of freedom), indicating that the data are a good fit to the model. In addition, only one cell had an error score greater than 2.0; however, Tagliamonte (2006, p. 221) argues that this is not uncommon in a distributional analysis such as this. Table 14 provides the factors weights reported by GoldVarb 2001 in the one level analysis.

Table 14. One-level analysis of phonological environment factor weights: second coding.

Factor Group	**Weight**
Preceding Phon. Environ.	
Vowel	.445
Consonant	.712
Nasal	.604
No preceding phon. environ.	.594
Following Phon. Environ.	
Vowel	.484
Consonant	.512

Given the apparent good fit of the distribution of 2Neg and Type D of 1Neg, the step-up/step-down analysis was performed for phonological environment (see Appendix E). GoldVarb 2001 found preceding phonological environment to be significant. Table 15 provides the distribution of 2Neg and Type D of 1Neg according to this factor.

Table 15. 2Neg and 1Neg distribution according to preceding phonological environment.

Environment	**2Neg**	**1Neg Type D**	**Total**	**Varbrul Score**
Vowel	114 (12.67%)	786 (87.33%)	900 (100%)	.444
Consonant	45 (30.82%)	101 (69.18%)	146 (100%)	.710
Nasal vowel	14 (21.88%)	50 (78.12%)	64 (100%)	.607
No preceding phon. environ.	22 (21.57%)	80 (78.43%)	102 (100%)	.602
Total	**195 (16.09%)**	**1,017 (83.91%)**	**1,212 (100%)**	**—**

The data in Table 15 indicate that the retention of *ne* is disfavored when the phonological environment immediately preceding the *ne* position is a vowel (or vowel sound). However, when a consonant or nasal vowel precedes *ne*, it is likely that 2Neg will be used. In addition, *ne* retention is favored when there is no preceding phonological environment, such as in the case of imperatives.

GoldVarb 2001 did not find following phonological environment to be significant, which is attributed to a lack of significant difference in *ne* retention rates within this factor group. Table 16 gives the overall distribution of 2Neg and Type D of 1Neg according to following phonological environment.

Table 16. 2Neg and 1Neg distribution according to following phonological environment.

Environment	**2Neg**	**1Neg Type D**	**Total**
Consonant	117 (16.93%)	574 (83.07%)	691 (100%)
Vowel	78 (14.97%)	443 (85.03%)	521 (100%)
Total	**195 (16.09%)**	**1,017 (83.91%)**	**1,212 (100%)**

A combination of the results for preceding and following phonological environment suggests that the surrounding phonological environment might determine—at least to some extent—whether *ne* is present or absent. These results are shown in Table 17. An example of each phonological environment is also provided.

Table 17. *Ne* retention according to surrounding phonological environment.

Environment	**2Neg**	**1Neg Type D**	**Total**
Consonant_Consonant (C_C) [elle (ne) veut pas] 'she does not want to'	18 (25.00%)	54 (75.00%)	72 (100%)
Consonant_Vowel (C_V) [il (n') est pas là] 'he is not there'	27 (36.49%)	47 (63.51%)	74 (100%)
Vowel_Consonant (V_C) [tu (ne) sais pas] 'you do not know'	72 (14.91%)	411 (85.09%)	483 (100%)
Vowel_Vowel (V_V) [tu (n') es pas là] 'you are not there'	42 (10.07%)	375 (89.93%)	417 (100%)
Nasal_Consonant (N_C) [on (ne) fait pas ça] 'one does not do that'	9 (20.45%)	35 (79.55%)	44 (100%)
Nasal_Vowel (N_V) [on (n') est pas là] 'one is not there'	5 (25.00%)	15 (75.00%)	20 (100%)
Ø_Consonant (Ø_C) [(ne) parlez pas] 'do not speak'	18 (19.57%)	74 (80.43%)	92 (100%)
Ø_Vowel (Ø_V) [(n') achetez pas ça] 'do not buy that'	4 (40%)	6 (60%)	100 (100%)
Total	**195 (16.09%)**	**1,017 (83.91%)**	**1,212 (100%)**

In order to determine whether selected phonological environments were influential on *ne* retention rates, a series of chi-square tests was performed. Table 18 shows clearly that, in many cases, the leading and following phonological environments interact and influence, together, whether *ne* will be present or absent. The abbreviations shown in Table 17 (e.g.,

C_C, C_V, V_C, and so forth) have been used in the remainder of the discussion of surrounding phonological environment.

Table 18. Interaction of selected phonological environments.

Environments	**Chi-square results**
C_C vs. V_C	χ^2 (1, N = 555) = 4.69, p = .05
C_C vs. V_V	χ^2 (1, N = 489) = 12.71, p = .05
C_V vs. V_C	χ^2 (1, N = 557) = 20.45, p = .05
V_C vs. V_V	χ^2 (1, N = 900) = 4.73, p = .05

Let us consider, for example, C_C and V_C environments, both of which were included in the "following consonant" category. Table 17 suggests that *ne* retention is dependent upon whether a vowel or consonant precedes the *ne* position (25% for C_C and 14.91% for V_C) and the results given in Table 18 show the significance of the distribution.

A comparison of C_V and V_V environments—which had drastically different *ne* retention rates (36.49% for C_V and 10.07% for V_V; see Table 17)—also reveals that the presence or absence of pre-vocalic *ne* is determined by the sound (a consonant or vowel) that precedes it. In addition, the data suggest that *ne* retention rates with preceding vowels are, at least to some extent, dependent upon the phonological environment that follows the *ne* position. The higher rate of *ne* retention for V_C environments (14.91%; see Table 17) compared with V_V environments (10.07%) was also found to be significant (see Table 18).

What emerges from the analysis of selected surrounding phonological environments is that the variable presence or absence of *ne* is determined to some extent by both the environment immediately preceding the *ne* position and the environment that immediately follows it. Specifically, *ne* retention rates are significantly lower in intervocalic position than when *ne* follows a vowel and precedes a consonant.

4.3.3 Analysis of the Second Coding: Morphosyntactic and Lexical Environment

The second part of the analysis of the second coding considers Neg2 type (i.e., *pas* vs. Neg2 other than *pas*), subject type (e.g., NP, pronoun, [- overt] subject), and sentence type (e.g., declarative, interrogative, imperative). After reviewing the one-level binomial analysis

(see Table 19 below for factor weights and Appendix F for full results), it was obvious that by separating the two phonological factor groups from the other structural factor groups considered in the second coding, the results fit well within the expected model, according to the criteria explained by Young and Bailey (1996) and Tagliamonte (2006).

Table 19. Factor weights from one-level binomial analysis for Neg2 type, subject type, and sentence type: second coding.

Factor Group	**Weight**
Neg2 Type	
pas	.491
Neg2 other than *pas*	.535
Subject Type	
NP	.895
Pronoun	.446
[- overt] subject	.580
Sentence Type	
Declarative	.505
Interrogative	.433
Imperative	.511

A step-up/step-down analysis was run for the three factor groups considered in this part of the second coding (see Table 25 below for significant Goldvarb scores and Appendix G for full results). After reviewing all runs reported in this part of the analysis, subject type emerged as a significant factor. GoldVarb 2001 did not, however, find Neg2 type or sentence type to be significant. Incidentally, when run independently of phonological environment, factor weight constraints within the subject type factor group did not change (as they had in the step-up/step-down procedure in section 4.3.1), which further supports the hypothesis that subject type and preceding phonological environment are inextricably linked.

Neg2 Type

The data shown in Table 19 suggest that *ne* retention does not depend on the type of Neg2 present. Table 20 shows clearly that *ne* retention rates are similar among most of the different Neg2 types. Nonetheless, the slight difference in *ne* retention rates between *pas* and Neg2s other than *pas* suggests that *ne* occurs at relatively higher frequencies with items that occur less frequently in discourse (Coveney, 1996; Hansen & Malderez, 2004).

Table 20. Distribution of 1Neg Type D and 2Neg according to Neg2 type: second coding.

Neg2	**2Neg**	**1Neg Type D**	**Total**
pas	143 (14.91%)	816 (85.09%)	959 (100%)
Neg2 other than *pas*	52 (20.55%)	201 (79.45%)	253 (100%)
Total	**195 (16.09%)**	**1,017 (83.23%)**	**1,212 (100%)**

Although GoldVarb 2001 did not find Neg2 type to be significant, the distribution of 2Neg and Type D of 1Neg was found to be significant by a chi-square procedure when all single Neg2s (e.g., *pas, rien, jamais,* etc.) were collapsed and tested against the negations in which more than one Neg2 had been used: χ^2 (1, N = 1,210) = 5.04, p = .05. This finding strongly suggests that *ne* retention is favored when more than one Neg2 is used in a negation (e.g., *plus rien, plus personne,* etc.). This distribution is shown in Table 21.

Table 21. Single vs. Multiple Neg2s: second coding.

	2Neg	**1Neg Type D**	**Total**
Single Neg2s	**186 (15.72%)**	**997 (84.28%)**	**1,183 (100%)**
Multiple Neg2s	**9 (31.03%)**	**20 (68.97%)**	**29 (100%)**
plus rien	2	7	9
plus personne	2	3	5
personne...rien	2	0	2
pas que	0	3	3
pas rien	0	3	3
jamais que	2	0	2
plus jamais	1	0	1
jamais rien	0	2	2
jamais personne	0	2	2
Total	**195 (16.09%)**	**1,017 (83.91%)**	**1,212 (100%)**

Although retention rates are higher when more than one Neg2 is used, the results of this analysis reveal that the variable use of *ne* does not normally depend on the type of Neg2.[7] This finding is in itself rather important since it suggests that *ne* non-retention has become more or less generalized regardless of which second negative is present. Nonetheless, it is important to note that a corpus with more Multiple Neg2 tokens would provide a more solid basis for making this conclusion.

[7] This finding does not corroborate results reported in Hansen & Malderez (2004), who found that more frequently occurring Neg2s (principally *pas*) were less likely to co-occur with *ne* than less frequently occurring Neg2s.

Sentence Type

The results of the step-up/step-down procedure for the second coding also reveal that the variable use of *ne* does not depend on sentence type. The distribution does, however, reveal, once again, to what extent *ne* non-retention has been generalized. Table 22 divides tokens of negation by sentence type: declarative, interrogative, and imperative.

Table 22. *Ne* retention according to sentence type: second coding.

Sentence type	**2Neg**	**1Neg Type D**	**Total**
Declarative	166 (16.27%)	854 (83.73%)	1,020 (100%)
Interrogative	10 (10.75%)	83 (89.25%)	93 (100%)
Imperative	19 (19.19%)	80 (80.81%)	99 (100%)
Total	**195 (16.09%)**	**1,017 (83.91%)**	**1,212 (100%)**

The data shown in Table 22 indicate that *ne* retention rates are similar in all three environments. It is therefore reasonable to assume that the presence or absence of *ne* does not depend on the type of sentence; however, the analysis of a larger corpus with more tokens of negative interrogative and imperative sentences might prove insightful.

Subject Type

Tokens of negation with [+ overt] subjects were also considered. Table 23 divides [+ overt] subject type into two main categories: noun phrases (NP) and pronouns. In addition to substantives, pronominalized negative particles that function as subjects (e.g., *personne, rien,* and so forth) were also included in the NP category since they are, historically, substantives (Ewert, 1969; Ashby, 1981; Rickard, 1989). Although pronominalized negative particles often function as objects (e.g., *Je n'entends personne* 'I hear no one'), as subjects, they appear to be more closely related to nouns and, therefore, to their historical meaning (Ashby, 1981; Grevisse, 1993, p. 1,076). The pronoun category includes all instances of clitic subject pronouns (e.g., *je, tu, on,* etc.), as well as relative pronouns (e.g., *qui* and *ce qui*) and demonstrative pronouns (e.g., *ça* and *cela*).

There exists a noticeable difference in *ne* retention rates between NPs and pronouns. Table 23 shows clearly that there is a preference to retain *ne* when an NP is present (59.55%),

while *ne* retention does not appear to be preferred when a subject pronoun is used (12.01%). A chi-square was calculated, which found the difference in *ne* retention rates between NPs and pronouns to be significant: χ^2 (1, N = 1,119) = 140.09, p = .05. These results corroborate the findings of Ashby (1981), Coveney (1996), and Hansen & Malderez (2004).

Table 23. Distribution of 2Neg and 1Neg Type D according to [+ overt] subject type: second coding.

[+ Overt] Subject	**2Neg**	**1Neg Type D**	**Total**
NP	**53 (59.55%)**	**36 (40.45%)**	**89 (100%)**
Noun[8]	45 (60.81%)	29 (39.19%)	74 (100%)
Negative particle	8 (53.33%)	7 (46.67%)	15 (100%)
Pronoun	**123 (12.01%)**	**901 (87.99%)**	**1,024 (100%)**
je	60 (12.79%)	409 (87.21%)	469 (100%)
tu	10 (6.71%)	139 (93.29%)	149 (100%)
il	8 (16.00%)	42 (84.00%)	50 (100%)
elle	4 (14.29%)	24 (85.71%)	28 (100%)
on	9 (15.79%)	48 (84.21%)	57 (100%)
ce	4 (2.76%)	141 (97.24%)	145 (100%)
nous	2 (100%)	0 (0.00%)	2 (100%)
vous	2 (11.11%)	16 (88.89%)	18 (100%)
ils	1 (10.00%)	9 (90.00%)	10 (100%)
elles	1 (33.33%)	2 (66.67%)	3 (100%)
ça[9]	9 (14.06%)	55 (85.94%)	64 (100%)
qui[10]	11 (40.74%)	16 (59.26%)	27 (100%)
[+ Overt] Subject Total	**176 (15.81%)**	**937 (84.19%)**	**1,113 (100%)**

Fifteen examples of Neg2s were included in the NP subject category, eight of which (53.33%) co-occurred with *ne*. A closer analysis of the examples of Neg2 subjects revealed that [*personne*$_{(\text{subject})}$ + Neg2] was used in four of the eight tokens of 2Neg in this category; in fact, not one example of [*personne*$_{(\text{subject})}$ + Neg2] was counted as 1Neg. The remaining four tokens of 2Neg were examples of *personne* and *rien* (two examples each). The data suggest that when more than one Neg2 is used in negation, *ne* retention is favored, provided that one of the Neg2s present is the subject of the clause.

[8] One instance of *les miennes* (2Neg) was included as a NP.
[9] One instance of *cela* (2Neg) was included with tokens of *ça*.
[10] Four instances of *ce qui* (three 2Neg and one 1Neg) have also been included in this category.

Ne retention according to the various types of pronouns identified in the corpus was also considered. Although *ne* retention is not favored when a pronoun is present in general (see Table 23), *ne* retention rates vary depending upon which pronoun is used. Table 24 divides subject pronouns into clitic and non-clitic.[11]

Table 24. Distribution of 2Neg and 1Neg Type D with clitic and non-clitic subject pronouns: second coding

Pronoun type	**2Neg**	**1Neg Type D**	**Total**
Clitic	103 (11.04%)	830 (88.96%)	933 (100%)
Non-clitic	20 (21.98%)	71 (77.02%)	91 (100%)
Total	**123 (12.01%)**	**901 (87.99%)**	**1,024 (100%)**

The data shown in Table 24 indicate that *ne* co-occurs more frequently with non-clitic subject pronouns than with clitic subject pronouns, which corroborates results reported in previous studies (Coveney, 1996; Armstrong, 2002; Armstrong & Smith, 2002; Hansen & Malderez; 2004). A chi-square test revealed that the distribution of 2Neg and 1Neg Type D in Table 24 was significant: χ^2 (1, N = 1,024) = 9.39, p = .05.

One possible explanation for this difference is that relative pronouns—which account for 27 of the 91 tokens of non-clitic pronouns—occur less frequently than other types of pronouns. If one accepts the hypothesis that frequency of occurrence determines, at least to some extent, the likelihood that *ne* will be absent or present (see Coveney, 1996; Hansen & Malderez, 2004), it seems reasonable to assume that the low frequency of *qui* and *ce qui* in this corpus contributes to the high rate of co-occurrence with *ne*.

In addition, there is a noticeable difference in *ne* retention rates between *qui* and *ça* (40.74% for *qui* and 14.06% for *ça*). This may be explained by the functional difference between relative and demonstrative pronouns. The demonstrative pronoun *ça* is often used in pre-formed sequences, in both affirmative and negative phrases (e.g., *ça va* or *ça [ne] va pas*), which may account for the low rate of *ne* retention. Relative pronouns, however, are usually present in novel sentence structures and refer to a NP or an idea introduced in the

[11] Non-clitic pronouns include instances of demonstrative and relative pronouns.

preceding clause, which appears to favor *ne* retention (see Coveney, 1996).

In light of Armstrong & Smith's (2002) suggestion that [- overt] subject environments focus the speaker's attention on the negation and result in higher *ne* retention rates, environments in which no overt subject is present were explored. In this analysis, only imperatives were considered (see 4.3.1). *Ne* retention rates are relatively high in the [- overt] subject category (19.19%).

In [+ overt] subject environments, *ne* retention appears to be disfavored (see Table 23), yet it is favored—at least to some extent—in [- overt] subject environments (see Table 24). A chi-square was calculated for these two environments and found that this factor was indeed significant: $\chi^2(1, N = 1{,}240) = 7.19$, $p = .05$. In addition, GoldVarb 2001 found that the factor group "subject type" (i.e., NP, pronoun, [- overt] subject) to be significant. These results are shown in Table 25.

Table 25. Distribution of 2Neg and 1Neg Type D according to subject type: second coding.

Subject environment	**2Neg**	**1Neg Type D**	**Total**	**GoldVarb Score**
NP	53 (59.55%)	36 (40.45%)	89 (100%)	.899
Pronoun	123 (12.01%)	901 (87.99%)	1,024 (100%)	.455
[- overt] subject	19 (19.19%)	80 (80.81%)	99 (100%)	.583
Total	**195 (16.09%)**	**1,017 (83.91%)**	**1,212 (100%)**	**—**

The data shown in Table 25 reveal that *ne* retention is strongly favored when the subject of the verb is an NP. Incidentally, a GoldVarb 2001 score of .899 is extremely high, which indicates that the rule (i.e., the use of 2Neg) will be applied very often in this particular environment; yet, since the GoldVarb 2001 score for pronoun is less than .500, *ne* retention is disfavored in this environment. In addition, the GoldVarb 2001 score for [- overt] subject environment indicates that the rule will be applied since the score is over .500. Although these results are not surprising when compared to those reported by Ashby (1981), Coveney (1996), and Hansen & Malderez (2004), they are indicative of the relationship between the discourse of IRC and everyday conversational speech. This in itself suggests that participants

are writing in this communicative environment as they would speak in informal settings, at least as far as the variable *ne* is concerned. However, writing as it is traditionally understood (e.g., letters, literature, print publications, and so forth) must be distinguished from the spontaneous communication that occurs in this synchronous text-based environment.

> Il faut sans doute se méfier de la vue étroite et idéalisée que l'on a en général de l'écrit, identifié au texte publié, élaboré dans la durée et corrigé par des professionnels. Il est sans doute difficile de trancher entre une influence de l'oral sur l'écrit et un rapprochement déterminé par l'élargissement du domaine d'usage de l'écrit. De plus, le partage entre la contamination spontanée et le recours intentionnel à des effets d'oralité est pratiquement impossible à effectuer. (Anis, 1999, p. 75)[12]

4.3.4 Interaction of Subjects and Phonological Environment

Although the analysis presented in 4.2 highlights that subject type and phonological environment interact to such an extent that it is necessary to consider the two separately when performing a multivariate analysis such as that performed by GoldVarb 2001, upon review of the results of the analysis it became clear that this interaction could not simply be left unexplored. A cross-tabulation of subject type and preceding phonological environment revealed that no less than 874 pronouns, which appear to disfavor *ne* retention (see Table 25), were counted in the preceding vowel category. Table 23 provides the distribution of 2Neg and Type D of 1Neg with pronouns ending in a vowel (or vowel sound) according to following phonological environment (e.g., consonant or vowel).

[12] Translation: "We must be weary of the strict and idealized view of writing, which is associated with published texts that are elaborated over time and corrected by professionals. It is without a doubt difficult to distinguish between an influence of spoken language on written language and a rapprochement that is determined by the expansion of the use of writing. In addition, it is practically impossible to make a distinction between spontaneous contamination and the intentional use of oral characteristics."

Table 26. Pronoun with final vowel and following phonological environment.

Following environment	**2Neg**	**1Neg Type D**	**Total**
Consonant	67 (14.19%)	405 (85.81%)	472 (100%)
Vowel	29 (7.21%)	373 (92.79%)	402 (100%)
Total	**96 (10.98%)**	**778 (89.02%)**	**874 (100%)**

It is clear that the phonological environment following a pronoun ending in a vowel is influencing *ne* retention. Table 23 shows that *ne* occurs at a higher rate when a consonant follows than when a vowel follows. A chi-square procedure found the distribution to be significant: χ^2 (1, N = 874) = 10.82, p = .05. One possible explanation for this difference is the fact that contractions are very often made between certain pronouns and a following vowel (or vowel sound). Indeed, contractions are required with *je* and *ce* when a vowel follows (e.g., *j'ai* or *c'est*), while other more informal (optional) contractions are made, for example, with *tu* (e.g., *t'es*). Since such contractions occur rather frequently in affirmative phrases, and it can be argued that speakers most usually operate with the [pronoun + verb] sequence (Ashby, 1981; Coveney, 1996), it is not surprising that in negative phrases, *ne* retention rates are extremely low in the intervocalic position (approximately one-half the overall retention rate).

Ne retention with pronouns ending in a consonant (e.g., *il, elle, ils, elles*) was also considered. Table 27 gives the distribution of 1Neg Type D and 2Neg for these pronouns according to following phonological environment.

Table 27. Pronoun with final consonant and following phonological environment.

Following environment	**2Neg**	**1Neg Type D**	**Total**
Consonant	4 (9.09%)	40 (90.91%)	44 (100%)
Vowel	11 (23.40%)	36 (76.60%)	47 (100%)
Total	**15 (16.48%)**	**76 (73.52%)**	**91 (100%)**

Although a chi-square test did not find the distribution shown in Table 27 to be significant, the distribution suggests that *ne* retention is not common in the context of a pronoun with a final consonant followed by another consonant; yet, the results shown in

Tables 17 and 18 (section 4.3.2) indicate that *ne* retention is favored in C_C environments. This discrepancy is most likely explained by the fact that examples of C_C environments were predominately tokens of NPs, which strongly favor *ne* retention.

Ne retention also appears to occur more often when a vowel follows a pronoun ending in a consonant. This may suggest that *ne* is used in order to avoid liaison is certain cases. In addition, an analysis of environments in which liaison would result in a [z] (e.g., *vous [n']avez pas, ils [n']ont pas*) was preformed and revealed that *ne* was retained in 33.33% of such instances.

One final observation was made regarding an environment that has traditionally been omitted from studies of *ne*: the clitic *on* followed by a vowel (or vowel sound). This environment has been considered difficult (if not impossible) to study in spoken French since it is unclear whether [n] is the result of liaison or pre-vocalic *ne* (e.g., *on n'a pas de pommes* [õnɑpɑdəpɔm] 'we don't have any apples'). This obstacle does not exist in the chat environment since the written transcription of the discussion is available for analysis.

Fifty-seven tokens of *on* were found in the corpus, fourteen of which were followed by a vowel sound. Of these fourteen tokens, only one (7.14%) co-occurred with *ne*. Although the limited number of tokens does not allow for any definite conclusions to be drawn from this analysis, the results do provide some evidence that the [n] observed in spoken discourse might be the result of liaison rather than pre-vocalic *n'*. It is also possible that speakers themselves do not know whether the [n] is pre-vocalic *n'* or the result of liaison. If this is true, it appears that they have generalized non-retention to this environment in synchronous French-language chat.

4.3.5 Summary of Results: Second Coding

The results from the two separate analyses of phonological environment and Neg2 type, subject type, and sentence type suggest that both the phonological environment

immediately preceding the *ne* position and subject type are significant factors. The data reported in 4.3.2 indicate the *ne* retention is disfavored when a vowel preceded the *ne* position, yet it is favored when a consonant or a nasal vowel is present. In addition, it appears that 2Neg is preferred when there is no preceding phonological environment. The results shown in 4.3.3 suggest that the absence of *ne* is rather generalized and does not depend upon the type of Neg2 used or sentence type. The one factor that emerges as significant in this part of the analysis is subject type; specifically, *ne* retention is very strongly favored when an NP is present, while 1Neg Type D appears to be preferred with subject pronouns. The retention of *ne* also seems to be favored in [- overt] subject environment, which corroborates the findings reported in 4.3.2 concerning the factor "no preceding phonological environment."

Last, a review of the data revealed that the vast majority of tokens counted as preceding vowels in 4.3.2 were instances of pronouns (see 4.3.4), which disfavor the use of 2Neg. It was found that pronouns ending in consonants were more likely to co-occur with *ne* than those that end in vowels. Moreover, a large percentage of tokens that had been counted in the preceding consonant category were found to be NPs, which strongly favor *ne* retention. Although these results are far from conclusive, type of subject emerges as the most influential internal linguistic factor, while phonology is most likely a contributing or underlying factor in the variation.

4.4 Discursive-Pragmatic Effect

4.4.1 Overview of Style-Shifting

The results shown in 4.3 indicate that a number of structural and phonological factors determine—at least to some extent—whether *ne* is present or absent; yet, it is possible that other, non-structural, factors influence the variation as well. Ashby (1981) posited that discourse topic (i.e., subject of discussion) was an influential factor in *ne* retention. He found that *ne* retention co-occurred frequently with serious topics, such as the punishment of

children and religion, as well as with the telling of general truths (e.g., proverbs, common sayings, etc.). In addition, Armstrong & Smith (2002) advanced the hypothesis that *ne* could be used in order to reinforce or add emphasis to a negation, and Coveney (1996, pp. 88-89) reported an extreme style shift by one 35-year-old male informant, presumably because the informant felt at one point that he should speak more clearly so that non-native speakers of French could understand him more easily. However, discussions and analyses of the intraspeaker dimension and discursive-pragmatic effect remain anecdotal in previous studies. Indeed, Armstrong (2002) is one of the few researchers to have undertaken qualitative analyses of the variable *ne* along the intraspeaker, stylistic dimension.

Armstrong's (2002) quantitative results do not indicate an observable pattern or system since some of the speakers retained *ne* at higher rates in interview style, yet others used *ne* at higher rates in conversation style. What emerges from Armstrong's analysis is that style-shifting was occurring on the micro-level; in other words, individual speakers appeared to produce tokens of *ne* during brief episodes of formal-style discourse in both conversation and interview styles, which suggests a diminished importance of the traditional binary paradigm of formal vs. informal.

> Several of the stylistic effects produced by the Dieuze informants in conversation style through their use of *ne* . . . contradict a simplex formal-informal analysis of style variation. Micro-style variation . . . is reflected in the use of *ne* through a reduction in the degrees of style shift, since the 'formal' episodes in conversation style call for the use of *ne* quite frequently relative to interview style. (Armstrong, 2002, p. 171)

In light of Bell's (1984, 2001) theory of variation as audience design and Armstrong's (2002) analysis of negative particle use along the intraspeaker dimension, I have chosen to explore the *ne* retention rates of six different chat participants in order to compare the

variation found in individuals with that of the corpus as a whole. Bell (2001) defines audience design as "a strategy by which speakers draw on a range of linguistic resources available in their speech community to respond to different kinds of audiences" (p. 145) and argues that style shift is a responsive, yet *active*, behavior (pp. 143-144). Figure 3 provides an illustration of how variation as audience design functions.

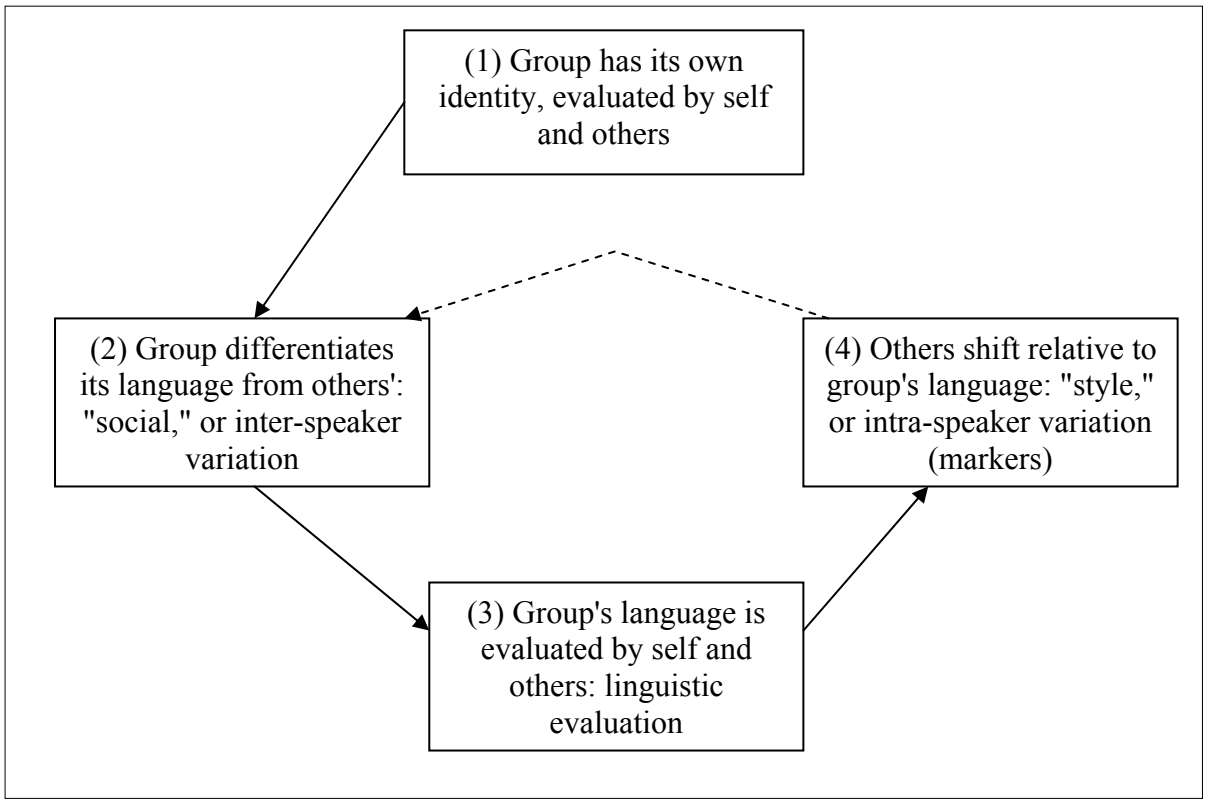

Figure 3. Bell's model of audience design (adapted from Bell, 2001, p. 142).

Bell (1984, 2001) assumes that each group (i.e., speech community) is seen as an independent entity (1) that has established a set of linguistic norms that are particular to that group (2). The language used within the group (i.e., among group members) is constantly evaluated (3) by both the group as a whole and individual members. Those who participate in the group shift their use of language relative to the expected norms of the group (4). Thus, the underlying principal of variation as audience design presupposes that "style is oriented to people rather than to mechanisms or functions" (Bell, 2001, p. 141).

Of primary interest to the analysis of discursive-pragmatic effect in the present study is the notion that style-shifting can occur according to the topic of discussion, a shift that

"derives its meaning and direction of shift from the underlying association of topics or setting with typical audience members" (Bell, 2001, p. 146). In addition, style-shifts may occur when "the individual speaker . . . uses language resources often from beyond the immediate speech community" (Bell, 2001, p. 147). As the analysis presented earlier has indicated, there exists an overwhelming preference for 1Neg Type D, which suggests that the discourse of synchronous CMC is closely associated with that of everyday conversational speech. Therefore, those who engage in chat discussions shift their style of discourse in the direction of everyday conversational speech, indicated here by the overwhelming preference for single-particle negation. Yet instances of 2Neg were observed in this analysis, and although patterns were found concerning a number of syntactic factors, an analysis of tokens of 2Neg might reveal other underlying factors in the variation; specifically shifts in the topic or tone of the discussion that led to the use of 2Neg.

4.4.2 Overview of Selected Participants

The six participants chosen were the top six contributors in the corpus (i.e., they produced the highest number of turns). Additionally, each of the six participants was present on at least three occasions during data collection. These six participants alone produced 11,438 words (over 14% of the data) and approximately 16.21% of all tokens of negation (both 1Neg Type D and 2Neg). Table 28 shows clearly that the overall rate of 2Neg for these six participants (16.92%) is approximately equivalent to the overall frequency of *ne* retention (16.77%) in this corpus. It therefore seems reasonable to assume that *ne* use among these six participants is representative of the corpus.

Table 28. Distribution of 2Neg and 1Neg Type D for top six contributors.

Participant	**2Neg**	**1Neg D**	**Total Neg**
<angelina>	1 (1.67%)	59 (98.33%)	60 (100%)
<salizar>	12 (44.44%)	15 (55.56%)	27 (100%)
<prue>	2 (7.41%)	25 (92.59%)	27 (100%)
<ange>	3 (9.38%)	29 (90.63%)	32 (100%)
<bruluin>	1 (4.76%)	20 (95.24%)	21 (100%)
<nomade>	15 (44.12%)	19 (55.88%)	34 (100%)
Total	**34 (16.92%)**	**167 (83.08%)**	**201 (100%)**

Along the intraspeaker dimension, rates of *ne* retention are unevenly distributed. While no single participant categorically omitted *ne*, <angelina> and <bruluin> produce only one token of 2Neg. These findings are not, however, uncommon (see Armstrong, 2002). In addition, two participants produce an extremely high number of 2Neg tokens (<salizar> and <nomade>). How can the differences in *ne* retention rates shown in Table 26 be explained? What factors or influences can we identify? In which contexts is *ne* being used? In order to address these questions, it is necessary to examine the tokens of negation produced by these six participants.

4.4.3 Definitions and Examples of Discourse Styles

A preliminary analysis of the sub-sample of data considered in this section revealed that tokens of 2Neg often correlated with a certain number of discourse topics, including arguments, jokes, imitations/role-plays, general truths, and policing. In order to test whether the topic of discourse is in fact influential, a statistical analysis was performed using GoldVarb 2001. The tokens of negation (N = 201) produced by <angelina>, <salizar>, <prue>, <ange>, <bruluin>, and <nomade> were first re-coded according to the internal linguistic factors discussed in 4.2, including subject type (NP, pronoun, [- overt] subject), preceding and following phonological environment, and sentence type (declarative, interrogative, imperative). In addition, the tokens were coded according to four general categories of "discourse style": ludic, emphatic, proverbial, and explanatory.

Ludic discourse style includes examples of jokes, imitations, and role-playing, as illustrated by Excerpts 11 and 12. Although it was not always immediately apparent that a message had been intended to convey humor, it was often the reaction of other participants that made classification possible. For example, smileys and other graphemes are often included in messages in order to convey emotion (Pierozak, 2003c & 2006). (In the following excerpts of data, *sic* has not been used to indicate that a grammatical, orthographic, or other type of mistake has been merely reproduced. Tokens of negations have been underlined and English translations in italics have been provided below data excerpts.

Excerpt 11.
<ange_away> Eliot de koi tu te plains krib a dit ke ct pa toi le boulet dont il parlait
[...]
<ange_away> :p
<ange_away> Eliot what are you complaining about krib said that you weren't the idiot he was talking about
[...]
<ange_away> :p

Excerpt 12.
<Anonyme952046> yen a pa ki chatte un pe la?
<Salizar> Je ne pense pas.
<Salizar> (Donc je ne suis pas)
<Anonyme952046> aren't there any who are chatting a little?
<Salizar> I do not think [so].
<Salizar> (Therefore I am not)

The second line of Excerpt 11 includes the smiley :p, which represents a stuck-out-tongue smiley face, which is most usually associated with humor, much as sticking one's tongue out while smiling (jokingly) would be in face-to-face communication. Although Excerpt 12 does not include any examples of such graphemes, it is nonetheless a clear-cut example of ludic discourse. <Salizar> responds to <Anonyme952046> by quoting Descartes' statement "je pense, donc je suis" (from his *Discours de la méthode*, 1637), but turns the phrase in adding the negation.

Emphatic discourse style normally occurs during arguments and disputes. This style is also used during "policing" (i.e., when a participant is warned or punished for his or her

behavior, which is usually related to the use of profanity or unwelcome statements). Excerpt 13 provides another example of Emphatic discourse style since <angelina> does not appear to appreciate <zidzid>'s comment about HELENE33|repo's breathing troubles.

Excerpt 13.
<angelina> est alors ta respiration ca va ma HELENE33|repo ?
<zidzid> elle etouf
<angelina> zidzid chute un peu de respect
[...]
<angelina> parle mieu ou tu gercle te moque pas des gents
<angelina> so how's your breathing my HELENE33|repo ?
<zidzid> she's suffocating
<angelina> zidzid shhhh a little respect
[...]
<angelina> speak better or you'll get it don't make fun of people

Proverbial style is similar to Laberge & Sankoff's (1977) "morals and truisms" category. Examples of general truths, rules, and proverbs have been included in this category. Although a certain number of the examples counted as proverbial have a somewhat ludic function, I believe that it is important to distinguish between ludic style, which involves jokes, and proverbial style, which exposes truths and generalities. Excerpts 14 and 15 provide examples of proverbial discourse style.

Excerpt 14.
<Salizar> Comme je le dis toujours : Un couple qui ne se dispute plus c'est un couple qui s'en fout
<Salizar> As I always say: A couple that no longer fights, is a couple that does not care

Excerpt 15.
<nomade> il n y a aucune obligation a distibuer des pelles et des kiss
<nomade> there is no obligation to give out kisses

Although Excerpt 14 is much more closely related to a proverb than 15, both examples have the same basic function: they provide general information and value judgments of a situation. Such examples were categorized as proverbial style.

Explanatory discourse style is what one might be tempted to label "normal conversation style." General explanations, observations, descriptions, and questions have been included in the Explanatory discourse style category. Although this may appear to be a

vast (even vague) definition, a preliminary analysis of the data did not find that this type of discourse was any more common than ludic, emphatic, or proverbial style. Excerpts 16, 17, and 18, provide examples of explanatory discourse style.

Excerpt 16.
<Prue> n'empêche que c'est pas forcément vrai, mais pas faux non plus Hugo :)
<Prue> albeit it's not necessarily true, but not false either Hugo :)

Excerpt 17.
<ange_away> je l'ai trouvée
<ange_away> la robe
<ange_away> :p
<ange_away> ah nan
<ange_away> g ^pa trouvé
<ange_away> i found it
<ange_away> the dress
<ange_away> :p
<ange_away> ah no
<ange_away> I didn't find it

Excerpt 18.
<angelina> er je suis desoler je pouvais rien faire rien ne marcher pardon les gents
<angelina> er i'm sorry i couldn't do anything nothing was working sorry guys

The three examples above demonstrate clearly that explanatory discourse style is a rather general, indeed *explanatory*, type of discourse. It differs from the other three discourse styles identified in the corpus in that only linguistic information is communicated. In other words, no extra- or meta-linguistic information—such as humor, anger, or judgment—is conveyed in the message.

4.4.3 Results

The results of the GoldVarb 2001 analysis for the 201 tokens of negation considered in this part of the analysis indicate that, once again, subject type emerges as a significant factor group.

Table 29. Distribution of 2Neg and 1Neg Type D according to subject type.

Subject type	**2Neg**	**1Neg Type D**	**Total**	**GoldVarb score**
NP	10 (76.92%)	3 (23.08%)	13 (100%)	.962
Pronoun	20 (12.12%)	145 (87.88%)	165 (100%)	.469
[- overt]	4 (17.39%)	19 (82.61%)	23 (100%)	.282
Total	**34 (16.92%)**	**167 (83.08%)**	**201 (100%)**	**—**

It is clear that *ne* retention is favored with NP subjects, while it is disfavored with subject pronouns. However, in this sub-sample of data, [- overt] subjects appear to disfavor *ne* retention. Phonological environment (both preceding and following) and sentence type were not found to be statistically significant, which corroborates the results reported earlier in 4.2.3 and 4.2.4. Low cell counts for second negatives other than *pas* made a quantification of this factor impossible. Overall, the results from the six participants selected for analysis appear to reflect what was found for the whole corpus, and it is therefore reasonable to assume that the results concerning discourse style will be rather insightful.

The four different types of discourse style co-occurred with *ne* at various frequencies, yet *ne* was neither retained nor omitted categorically in any one of them. GoldVarb 2001 found discourse style to be statistically significant. Table 31 provides Varbrul scores for the independent variables in this factor group and *ne* retention rates according to discourse style for the six participants chosen for analysis.

Table 30. 2Neg and 1Neg Type D distribution according to discourse style.

Discourse style	**2Neg**	**1Neg Type D**	**Total**	**Varbrul Score**
Explanatory	2 (1.79%)	110 (98.21%)	112 (100%)	.130
Proverbial	4 (50.00%)	4 (50.00%)	8 (100%)	.865
Emphatic	5 (22.73%)	17 (77.27%)	22 (100%)	.886
Ludic	23 (38.98%)	36 (61.02%)	59 (100%)	.930
Total	**34 (16.92%)**	**167 (83.08%)**	**201 (100%)**	**—**

It is clear that discourse style determines—at least in part—whether *ne* is retained. *Ne* retention is strongly disfavored in the explanatory discourse style, while it is favored in proverbial, emphatic, and ludic styles. In addition, 23 of the 34 examples of 2Neg (67.65%) were counted in the ludic category. This should not be surprising, however, given the

"increased opportunities for language play" (Herring, 1999) provided by the synchronous chat environment (see also Pierozak, 2003a). The playful aspect of this type of communication may even be a reason for its popularity. According to Herring (1999):

> On-line surveys reveal that humorous messages are the most highly appreciated types of messages in computer-mediated environments, even in those one would not necessarily characterize as 'recreational'. . . . The potential for humorous play inherent in the . . . computer-mediated environment constitutes one of the biggest attractions of CMC for many users.

Table 31 also shows that emphatic discourse style favors *ne* retention, which supports the claim that *ne* can be used to add secondary emphasis to a negation (Armstrong & Smith, 2002), especially during disputes and arguments. Further, it should not be surprising that *ne* retention is favored in proverbial style since it is reasonable to assume that a speaker might be more careful in his or her speech when citing a proverb or telling a general truth. This provides further evidence for the hypothesis that more "standard" forms and structures are used when more attention is paid to speech (Labov, 1972).

In addition, following Bell's (2001) model of audience design, it can be argued that "the individual speaker . . . uses language resources often from beyond the immediate speech community" (p. 147). Although *ne* retention is generally disfavored in chat environments, the analysis presented in the preceding paragraphs strongly suggest that group members (i.e., chat participants) draw upon their linguistic resources (i.e., 2Neg) "from beyond the immediate speech community" in order to express humor and anger, and to imitate the discourse of proverbs and truisms.

CHAPTER 5

CONCLUSION

5.1 Summary of Results

Comparisons of the discourse of synchronous French-language CMC and what one would expect to observe in everyday conversational speech reveal a number of linguistic, stylistic, and pragmatic similarities. In the case of *ne* use, synchronous CMC participants omit *ne* from verbal negation at frequencies similar to those reported in previous research concerning everyday conversational French (e.g., Coveney, 1996; Hansen & Malderez, 2004). In other words, *ne* is retained at variable rates depending on the linguistic and/or stylistic environment (see ch. 4).

The results of this research suggest that subject type and phonological environment are inextricably linked and that it may be necessary to consider these two factors separately when doing a multivariate analysis of the use of *ne*. In this study, subject type and preceding phonological environment emerged as influential factors in the variation of *ne* use when separated. Although following phonological environment was not found to be significant, a closer analysis of the phonological environment surrounding the *ne* position revealed that the presence of *ne* does indeed depend on—at least to some extent—the sound following its position. In addition, neither second-negative type nor sentence type was found to be significant, which suggests that the variable use of *ne* is not influenced by these factors. Overall, it appears that the type of subject present (e.g., NP or pronoun) or the lack of an overt subject (e.g., imperatives) determines whether *ne* will be present or absent, while the phonological environment preceding the *ne* position is an underlying and contributing factor,

specifically when a subject pronoun is used (see 4.3.4).

A qualitative analysis of tokens of negation produced by the six top contributors in the corpus revealed that discourse style was a determining factor in the variation. While ludic, emphatic, and proverbial styles appear to favor strongly the retention of *ne*, the presence of *ne* is disfavored in explanatory style. These findings support Ashby's (1981) hypothesis that discussion topic and discourse-pragmatic effect are influential factors as well as Armstrong's (2002) suggestion that style-shifting occurs on a micro-level.

Results concerning selected syntactic factors reported in this study are not surprising; overall, they corroborate results reported in previous research of *ne* in everyday conversational speech (Ashby, 1981; Coveney, 1996; Hansen & Malderez, 2004). The similarities between the discourse of synchronous French-language CMC and informal spoken French are indicative of the relationship between these two types of discourse. In other words, the language used in synchronous CMC resembles in many ways that of informal spoken French; in particular the overwhelming preference for the omission of *ne*. It therefore seems reasonable to say that synchronous CMC provides an environment in which informal and non-traditional[1] forms and structures are accepted as the norm.

Although some research (Ashby, 1981, 2001) has suggested that *ne* is disappearing from the French language, another point of view maintains that we are far from seeing the completion of such a change (see Hansen & Malderez, 2004). The present study supports the latter of these two hypotheses since synchronous CMC participants use the variable *ne* in a probabilistic way in certain syntactic and pragmatic environments. Indeed, the evidence supporting the hypothesis that discourse style is a determining factor in the variation suggests that *ne* has become an important pragmatic resource for synchronous CMC participants. How is such a development possible? More specifically, which factors and influences can be

[1] The term "non-standard" has been avoided since it could be argued that the absence of *ne*, although grammatically incorrect according to prescribed grammar, is, in fact, the standard in this type of communication.

identified to explain the evolution of *ne* from the default marker of negation to a sociopragmatic tool capable of communicating humor and anger, among other extra-linguistic information?

The literature on the history of the French language—specifically the development of verbal negation—suggests that the phonetic weakening of Latin *non* rendered *ne* incapable of bearing an energetic negation, which led to the grammaticalization of second-negatives and the development of two-particle negation. Ashby (1981) highlights the effect of the grammaticalization of second-negatives, which has contributed to the disposability of *ne* since it is a redundant marker of negation. Further, "[i]t appears that this change . . . is given impetus by another continuing change in French: the fusion of the clitic pronoun and the verb" (Ashby, 1981, p. 686; see also Coveney, 1996; Fonseca-Greber & Waugh, 2003).

This final conclusion made by Ashby (1981) appears to be true for synchronous French-language CMC. Recent research (Pierozak, 2003c; Williams, 2006; van Compernolle & Williams, in press) has identified a number of fused [subject clitic + verb] sequences; specifically *c'est, tu es,* and *j'ai* ('it is', 'you are', and 'I have', respectively). Electronic environments (both synchronous and asynchronous) allow the user to abbreviate these sequences; *c'est* [sɛ] becomes *c* [sɛ] (pronounced as the letter "c"), *tu es* (often contracted to *t'es* [tɛ] in spoken French) becomes *t* [tɛ] (pronounced as the letter "t"), and *j'ai* [ʒɛ] becomes *g* (pronounced as the letter "g").

Abbreviations such as these, which are often referred to as syllabograms,[2] occur at such high frequencies in affirmative sentences—presumably in order to save time when engaging in synchronous CMC (see Pierozak, 2003b; van Compernolle & Williams, in press)—that synchronous CMC users appear to operate principally with the [clitic + verb] syllabogram in negative sentences as well. By reducing the [clitic + verb] sequence to a

[2] Pierozak (2003c) proposed the French term *syllabogramme*.

single-syllable grapheme, the presence of *ne* is made impossible (see van Compernolle & Williams, in press). If it is true that syllabograms are used in order to save time even in negative sentences, than it is reasonable to assume that the use of more traditional forms (e.g., *tu n'es pas*) is deliberate. If such use is indeed deliberate, synchronous CMC users must be aware of the pragmatic functions of *ne*—at least in this type of communication—and the various discourse styles (e.g., ludic, emphatic, proverbial, and explanatory) that favor or disfavor its presence.

5.2 Directions for Future Research

Although this study has undertaken a formal investigation of *ne* use in only one form of synchronous CMC, informal observations suggest that *ne* retention rates are similar in other synchronous CMC environments. In addition, recent research (van Compernolle & Williams, in press) has suggested that *ne* use in IRC is very different from what can be observed in discussion forums and moderated chat. These differences appear to be due—at least in part—to the synchronicity of communication and the perceived level of formality of the communication context. Future studies of the use of *ne* in CMC are needed in order to determine to what extent the discourses of different types of synchronous and asynchronous CMC are similar. In addition, more research on the effect of abbreviated orthographic forms on the use of *ne*—among other linguistic variables—might prove insightful in the study of sociolinguistic variation in CMC. Last, discourse-pragmatic effect on *ne* retention must be explored further in both CMC and spoken French in order to understand more completely the complex of linguistic and pragmatic factors that co-occur and co-vary with this variable.

APPENDIX A

GOLDVARB 200 CODING KEY (ALL DATA)

Factor Group 1: Neg2 Type
1 = *pas*
2 = Neg2 other than *pas*

Factor Group 2: Subject Type
q = NP
e = Pronoun
k = [- overt] subject

Factor Group 3: Preceding Phonological Environment
f = Vowel
d = Consonant
s = Nasal Vowel
N = No preceding phonological environment

Factor Group 4: Following Phonological Environment
V = Vowel
C = Consonant

Factor Group 7: Sentence Type
z = Declarative
x = Interrogative
c = Imperative

APPENDIX B

ONE-STEP BINOMIAL: FIRST CODING

```
Binomial Varbrul, 1 step
========================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.

 - One-level analysis only:One-level binomial analysis:

Run # 1, 40 cells:
No Convergence at Iteration 20
Input 0.143

Group Factor Weight App/Total Input&Weight

   1:       2    0.530    0.21       0.16
            1    0.492    0.15       0.14

   2:       e    0.454    0.12       0.12
            q    0.892    0.60       0.58
            k    0.511    0.19       0.15

   3:       f    0.447    0.13       0.12
            s    0.544    0.22       0.17
            d    0.486    0.31       0.14
            N    0.863    0.22       0.51

   4:       C    0.524    0.17       0.16
            V    0.468    0.15       0.13

   5:       z    0.543    0.16       0.17
            c    0.172    0.19       0.03
            x    0.446    0.11       0.12

  Cell  Total      App'ns    Expected     Error
 2qsVz     3          2          1.982     0.000
 2qsCz     1          1          0.708     0.412
 2qfVz     6          6          3.411     4.553
 2qfCz     5          3          3.110     0.010
 2qdVz     7          5          4.245     0.341
 2qdVx     1          0          0.510     1.042
 2qdCz    11          5          7.237     2.021
 2qdCx     1          1          0.565     0.769
 2kNCc     5          3          1.103     4.184
 2esVz     4          0          0.657     0.786
 2esCz    10          1          1.970     0.595
 2efVz    64          7          7.519     0.041
 2efVx     1          0          0.083     0.090
 2efCz   104         12         14.819     0.625
 2efCx     7          1          0.707     0.135
```

2edVz	10	3	1.347	2.346
2edVx	1	0	0.095	0.105
2edCz	11	2	1.789	0.030
2edCx	1	0	0.116	0.131
1qsVz	3	2	1.876	0.022
1qfVz	8	6	4.245	1.546
1qfCz	6	2	3.511	1.568
1qdVz	18	10	10.246	0.014
1qdVx	1	1	0.472	1.119
1qdCz	16	8	9.960	1.022
1qNVz	1	1	0.898	0.114
1kNVc	9	3	1.466	1.919
1kNCc	85	13	16.607	0.974
1esVz	10	1	1.443	0.159
1esCz	27	7	4.694	1.371
1esCx	6	0	0.748	0.854
1efVz	306	21	31.357	3.811
1efVx	32	2	2.294	0.041
1efCz	326	52	40.662	3.612
1efCx	35	2	3.077	0.414
1edVz	34	7	4.003	2.542
1edVx	2	1	0.166	4.583
1edCz	29	2	4.141	1.291
1edCx	3	0	0.304	0.338
1eNCx	2	2	0.857	2.666

Total Chi-square = 48.1946
Chi-square/cell = 1.2049
Log likelihood = -477.718

APPENDIX C

STEP-UP/STEP-DOWN ANALYSIS: FIRST CODING

```
Binomial Varbrul
================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.
Threshold, step-up/down: 0.050001

 # Stepping up:
 # Stepping up:

---------- Level # 0 ----------

Run # 1, 1 cells:
Convergence at Iteration 2
Input 0.161
Log likelihood = -534.667

---------- Level # 1 ----------

Run # 2, 2 cells:
Convergence at Iteration 4
Input 0.160
Group # 1 -- 2: 0.576, 1: 0.480
Log likelihood = -532.419 Significance = 0.037

Run # 3, 3 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.445, q: 0.899, k: 0.583
Log likelihood = -483.652 Significance = 0.000

Run # 4, 4 cells:
Convergence at Iteration 5
Input 0.154
Group # 3 -- f: 0.444, s: 0.607, d: 0.710, N: 0.602
Log likelihood = -518.983 Significance = 0.000

Run # 5, 2 cells:
Convergence at Iteration 3
Input 0.161
Group # 4 -- C: 0.515, V: 0.480
Log likelihood = -534.242 Significance = 0.372

Run # 6, 3 cells:
Convergence at Iteration 4
Input 0.160
Group # 5 -- z: 0.505, c: 0.555, x: 0.388
Log likelihood = -533.230 Significance = 0.243
```

```
Add Group # 2 with factors eqk

---------- Level # 2 ----------

Run # 7, 6 cells:
Convergence at Iteration 5
Input 0.145
Group # 1 -- 2: 0.536, 1: 0.490
Group # 2 -- e: 0.445, q: 0.896, k: 0.590
Log likelihood = -483.250 Significance = 0.387

Run # 8, 9 cells:
No Convergence at Iteration 20
Input 0.144
Group # 2 -- e: 0.481, q: 0.903, k: 0.235
Group # 3 -- f: 0.453, s: 0.557, d: 0.491, N: 0.830
Log likelihood = -479.949 Significance = 0.063

Run # 9, 6 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.446, q: 0.902, k: 0.563
Group # 4 -- C: 0.527, V: 0.464
Log likelihood = -482.605 Significance = 0.158

Run # 10, 5 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.446, q: 0.898, k: 0.574
Group # 5 -- z: 0.505, c: 0.510, x: 0.430
Log likelihood = -483.267 Significance = 0.684

No remaining groups significant

Groups selected while stepping up:  2
Best stepping up run:   #3
-------------------------------------------

 # Stepping down:
 # Stepping down:

---------- Level # 5 ----------

Run # 11, 40 cells:
No Convergence at Iteration 20
Input 0.143
Group # 1 -- 2: 0.530, 1: 0.492
Group # 2 -- e: 0.454, q: 0.892, k: 0.511
Group # 3 -- f: 0.447, s: 0.544, d: 0.486, N: 0.863
Group # 4 -- C: 0.524, V: 0.468
Group # 5 -- z: 0.543, c: 0.172, x: 0.446
Log likelihood = -477.718

---------- Level # 4 ----------

Run # 12, 23 cells:
No Convergence at Iteration 20
Input 0.144
Group # 2 -- e: 0.454, q: 0.894, k: 0.505
Group # 3 -- f: 0.447, s: 0.547, d: 0.486, N: 0.861
Group # 4 -- C: 0.525, V: 0.467
```

```
Group # 5 -- z: 0.543, c: 0.173, x: 0.443
Log likelihood = -478.023 Significance = 0.450

Run # 13, 26 cells:
No Convergence at Iteration 20
Input 0.152
Group # 1 -- 2: 0.566, 1: 0.483
Group # 3 -- f: 0.409, s: 0.563, d: 0.675, N: 0.887
Group # 4 -- C: 0.512, V: 0.484
Group # 5 -- z: 0.550, c: 0.148, x: 0.418
Log likelihood = -512.328 Significance = 0.000

Run # 14, 18 cells:
Convergence at Iteration 6
Input 0.144
Group # 1 -- 2: 0.530, 1: 0.492
Group # 2 -- e: 0.446, q: 0.898, k: 0.577
Group # 4 -- C: 0.526, V: 0.465
Group # 5 -- z: 0.507, c: 0.494, x: 0.432
Log likelihood = -481.920 Significance = 0.041

Run # 15, 23 cells:
No Convergence at Iteration 20
Input 0.144
Group # 1 -- 2: 0.534, 1: 0.491
Group # 2 -- e: 0.454, q: 0.889, k: 0.513
Group # 3 -- f: 0.446, s: 0.550, d: 0.486, N: 0.865
Group # 5 -- z: 0.542, c: 0.179, x: 0.448
Log likelihood = -478.572 Significance = 0.194

Run # 16, 28 cells:
No Convergence at Iteration 20
Input 0.144
Group # 1 -- 2: 0.534, 1: 0.491
Group # 2 -- e: 0.481, q: 0.903, k: 0.228
Group # 3 -- f: 0.453, s: 0.549, d: 0.490, N: 0.830
Group # 4 -- C: 0.523, V: 0.469
Log likelihood = -478.705 Significance = 0.385

Cut Group # 1 with factors 21

---------- Level # 3 ----------

Run # 17, 15 cells:
No Convergence at Iteration 20
Input 0.153
Group # 3 -- f: 0.409, s: 0.569, d: 0.680, N: 0.881
Group # 4 -- C: 0.514, V: 0.481
Group # 5 -- z: 0.551, c: 0.147, x: 0.412
Log likelihood = -513.956 Significance = 0.000

Run # 18, 10 cells:
Convergence at Iteration 5
Input 0.144
Group # 2 -- e: 0.446, q: 0.901, k: 0.571
Group # 4 -- C: 0.527, V: 0.464
Group # 5 -- z: 0.507, c: 0.492, x: 0.429
Log likelihood = -482.192 Significance = 0.042

Run # 19, 13 cells:
No Convergence at Iteration 20
```

```
Input 0.144
Group # 2 -- e: 0.454, q: 0.892, k: 0.507
Group # 3 -- f: 0.447, s: 0.553, d: 0.486, N: 0.863
Group # 5 -- z: 0.542, c: 0.181, x: 0.445
Log likelihood = -478.963 Significance = 0.178

Run # 20, 16 cells:
No Convergence at Iteration 20
Input 0.144
Group # 2 -- e: 0.481, q: 0.906, k: 0.226
Group # 3 -- f: 0.454, s: 0.551, d: 0.491, N: 0.827
Group # 4 -- C: 0.524, V: 0.468
Log likelihood = -479.055 Significance = 0.367

Cut Group # 5 with factors zcx

---------- Level # 2 ----------

Run # 21, 8 cells:
Convergence at Iteration 5
Input 0.153
Group # 3 -- f: 0.445, s: 0.604, d: 0.712, N: 0.594
Group # 4 -- C: 0.512, V: 0.484
Log likelihood = -518.745 Significance = 0.000

Run # 22, 6 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.446, q: 0.902, k: 0.563
Group # 4 -- C: 0.527, V: 0.464
Log likelihood = -482.605 Significance = 0.073

Run # 23, 9 cells:
No Convergence at Iteration 20
Input 0.144
Group # 2 -- e: 0.481, q: 0.903, k: 0.235
Group # 3 -- f: 0.453, s: 0.557, d: 0.491, N: 0.830
Log likelihood = -479.949 Significance = 0.186

Cut Group # 4 with factors CV

---------- Level # 1 ----------

Run # 24, 4 cells:
Convergence at Iteration 5
Input 0.154
Group # 3 -- f: 0.444, s: 0.607, d: 0.710, N: 0.602
Log likelihood = -518.983 Significance = 0.000

Run # 25, 3 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.445, q: 0.899, k: 0.583
Log likelihood = -483.652 Significance = 0.063

Cut Group # 3 with factors fsdN

---------- Level # 0 ----------

Run # 26, 1 cells:
Convergence at Iteration 2
```

```
Input 0.161
Log likelihood = -534.667 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down:  1  5  4  3
Best stepping  up  run: #3
Best stepping down run: #25
```

APPENDIX D

ONE-STEP BINOMIAL: SECOND CODING

(PHONOLOGICAL ENVIRONMENT)

```
Binomial Varbrul, 1 step
========================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.

 - One-level analysis only:One-level binomial analysis:

Run # 1, 8 cells:
Convergence at Iteration 5
Input 0.153

Group Factor Weight App/Total Input&Weight

   1:       f    0.445    0.13       0.13
            s    0.604    0.22       0.22
            d    0.712    0.31       0.31
            N    0.594    0.22       0.21

   2:       C    0.512    0.17       0.16
            V    0.484    0.15       0.15

  Cell  Total       App'ns    Expected     Error
    sV    20          5          4.112     0.242
    sC    44          9          9.886     0.102
    fV   417         42         50.044     1.469
    fC   483         72         63.991     1.155
    dV    74         27         21.907     1.682
    dC    72         18         23.051     1.628
    NV    10          4          1.989     2.536
    NC    92         18         20.019     0.260
Total Chi-square = 9.0759
 Chi-square/cell = 1.1345
Log likelihood = -518.745
```

APPENDIX E

STEP-UP/STEP-DOWN ANALYSIS: SECOND CODING

(PHONOLOGICAL ENVIRONMENT)

```
Binomial Varbrul
================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.
Threshold, step-up/down: 0.050001

 # Stepping up:
 # Stepping up:

---------- Level # 0 ----------

Run # 1, 1 cells:
Convergence at Iteration 2
Input 0.161
Log likelihood = -534.667

---------- Level # 1 ----------

Run # 2, 4 cells:
Convergence at Iteration 5
Input 0.154
Group # 1 -- f: 0.444, s: 0.607, d: 0.710, N: 0.602
Log likelihood = -518.983 Significance = 0.000

Run # 3, 2 cells:
Convergence at Iteration 3
Input 0.161
Group # 2 -- C: 0.515, V: 0.480
Log likelihood = -534.242 Significance = 0.372

Add Group # 1 with factors fsdN

---------- Level # 2 ----------

Run # 4, 8 cells:
Convergence at Iteration 5
Input 0.153
Group # 1 -- f: 0.445, s: 0.604, d: 0.712, N: 0.594
Group # 2 -- C: 0.512, V: 0.484
Log likelihood = -518.745 Significance = 0.493

No remaining groups significant

Groups selected while stepping up:  1
Best stepping up run:   #2
-------------------------------------------

 # Stepping down:
```

```
 # Stepping down:

---------- Level # 2 ----------

Run # 5, 8 cells:
Convergence at Iteration 5
Input 0.153
Group # 1 -- f: 0.445, s: 0.604, d: 0.712, N: 0.594
Group # 2 -- C: 0.512, V: 0.484
Log likelihood = -518.745

---------- Level # 1 ----------

Run # 6, 2 cells:
Convergence at Iteration 3
Input 0.161
Group # 2 -- C: 0.515, V: 0.480
Log likelihood = -534.242 Significance = 0.000

Run # 7, 4 cells:
Convergence at Iteration 5
Input 0.154
Group # 1 -- f: 0.444, s: 0.607, d: 0.710, N: 0.602
Log likelihood = -518.983 Significance = 0.493

Cut Group # 2 with factors CV
```

APPENDIX F

ONE-STEP BINOMIAL ANALYSIS: SECOND CODING

(MORPHOSYNTACTIC ENVIRONMENT)

```
Binomial Varbrul, 1 step
========================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.

 - One-level analysis only:One-level binomial analysis:

Run # 1, 10 cells:
Convergence at Iteration 5
Input 0.145

Group Factor Weight App/Total Input&Weight

   1:       2    0.535    0.21       0.16
            1    0.491    0.15       0.14

   2:       e    0.446    0.12       0.12
            q    0.895    0.60       0.59
            k    0.580    0.19       0.19

   3:       z    0.505    0.16       0.15
            c    0.511    0.19       0.15
            x    0.433    0.11       0.11

  Cell  Total       App'ns    Expected     Error
   2qz    33         22          20.776     0.195
   2qx     2          1           1.120     0.029
   2kc     5          3           1.096     4.239
   2ez   203         25          28.012     0.376
   2ex    10          1           1.070     0.005
   1qz    52         29          30.561     0.193
   1qx     1          1           0.516     0.938
   1kc    94         16          17.908     0.251
   1ez   732         90          86.639     0.148
   1ex    80          7           7.303     0.014
Total Chi-square = 6.3867
 Chi-square/cell = 0.6387
Log likelihood = -482.901
```

APPENDIX G

STEP-UP/STEP-DOWN ANALYSIS: SECOND CODING

(MORPHOSYNTACTIC ENVIRONMENT)

```
Binomial Varbrul
================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.
Threshold, step-up/down: 0.050001

 # Stepping up:
 # Stepping up:

---------- Level # 0 ----------

Run # 1, 1 cells:
Convergence at Iteration 2
Input 0.161
Log likelihood = -534.667

---------- Level # 1 ----------

Run # 2, 2 cells:
Convergence at Iteration 4
Input 0.160
Group # 1 -- 2: 0.576, 1: 0.480
Log likelihood = -532.419 Significance = 0.037

Run # 3, 3 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.445, q: 0.899, k: 0.583
Log likelihood = -483.652 Significance = 0.000

Run # 4, 3 cells:
Convergence at Iteration 4
Input 0.160
Group # 3 -- z: 0.505, c: 0.555, x: 0.388
Log likelihood = -533.230 Significance = 0.243

Add Group # 2 with factors eqk

---------- Level # 2 ----------

Run # 5, 6 cells:
Convergence at Iteration 5
Input 0.145
Group # 1 -- 2: 0.536, 1: 0.490
Group # 2 -- e: 0.445, q: 0.896, k: 0.590
Log likelihood = -483.250 Significance = 0.387

Run # 6, 5 cells:
```

```
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.446, q: 0.898, k: 0.574
Group # 3 -- z: 0.505, c: 0.510, x: 0.430
Log likelihood = -483.267 Significance = 0.684

No remaining groups significant

Groups selected while stepping up:  2
Best stepping up run:   #3
-------------------------------------------

 # Stepping down:
 # Stepping down:

---------- Level # 3 ----------

Run # 7, 10 cells:
Convergence at Iteration 5
Input 0.145
Group # 1 -- 2: 0.535, 1: 0.491
Group # 2 -- e: 0.446, q: 0.895, k: 0.580
Group # 3 -- z: 0.505, c: 0.511, x: 0.433
Log likelihood = -482.901

---------- Level # 2 ----------

Run # 8, 5 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.446, q: 0.898, k: 0.574
Group # 3 -- z: 0.505, c: 0.510, x: 0.430
Log likelihood = -483.267 Significance = 0.410

Run # 9, 6 cells:
Convergence at Iteration 4
Input 0.159
Group # 1 -- 2: 0.578, 1: 0.479
Group # 3 -- z: 0.503, c: 0.572, x: 0.396
Log likelihood = -530.912 Significance = 0.000

Run # 10, 6 cells:
Convergence at Iteration 5
Input 0.145
Group # 1 -- 2: 0.536, 1: 0.490
Group # 2 -- e: 0.445, q: 0.896, k: 0.590
Log likelihood = -483.250 Significance = 0.706

Cut Group # 3 with factors zcx

---------- Level # 1 ----------

Run # 11, 3 cells:
Convergence at Iteration 5
Input 0.145
Group # 2 -- e: 0.445, q: 0.899, k: 0.583
Log likelihood = -483.652 Significance = 0.387

Run # 12, 2 cells:
Convergence at Iteration 4
Input 0.160
```

```
Group # 1 -- 2: 0.576, 1: 0.480
Log likelihood = -532.419 Significance = 0.000

Cut Group # 1 with factors 21

---------- Level # 0 ----------

Run # 13, 1 cells:
Convergence at Iteration 2
Input 0.161
Log likelihood = -534.667 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down:  3  1
Best stepping  up  run: #3
Best stepping down run: #11
```

APPENDIX H

GOLDVARB 2001 CODING KEY

(SELECTED DATA)

Factor group 1: Subject type
- q = Noun phrase
- w = Pronoun
- e = [- overt] subject

Factor group 2: Preceding phonological environment
- a = Vowel (or vowel sound)
- s = Consonant
- d = Nasal vowel
- f = No preceding phonological environment

Factor group 3: Following phonological environment
- g = Vowel (or vowel sound)
- h = Consonant

Factor group 4: Sentence type
- z = Declarative
- x = Interrogative
- c = Imperative

Factor group 5: Discourse style
- v = Ludic
- m =Explanatory
- b = Emphatic
- n = Proverbial

APPENDIX I

GOLDVARB 2001 RESULTS FOR SELECTED DATA:

PHOLOGICAL ENVIRONMENT

```
Binomial Varbrul, 1 step
========================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.

 - One-level analysis only:One-level binomial analysis:

Run # 1, 8 cells:
Convergence at Iteration 5
Input 0.142

Group Factor Weight App/Total Input&Weight

   1:       a    0.404    0.10       0.10
            f    0.577    0.21       0.18
            s    0.883    0.56       0.56
            d    0.804    0.40       0.40

   2:       g    0.436    0.13       0.11
            h    0.552    0.20       0.17

  Cell  Total       App'ns    Expected     Error
    sh    10          5           6.059     0.470
    sg     8          5           3.934     0.568
    fh    21          4           4.563     0.089
    fg     3          1           0.446     0.808
    dh     5          2           2.275     0.061
    dg     5          2           1.722     0.068
    ah    75         11           9.090     0.457
    ag    74          4           5.911     0.671
Total Chi-square = 3.1910
 Chi-square/cell = 0.3989
Log likelihood = -79.451


Binomial Varbrul
================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.
Threshold, step-up/down: 0.050001

 # Stepping up:
 # Stepping up:

---------- Level # 0 ----------
```

```
Run # 1, 1 cells:
Convergence at Iteration 2
Input 0.169
Log likelihood = -91.363

---------- Level # 1 ----------

Run # 2, 4 cells:
Convergence at Iteration 5
Input 0.144
Group # 1 -- a: 0.400, f: 0.610, s: 0.881, d: 0.798
Log likelihood = -80.034 Significance = 0.000

Run # 3, 2 cells:
Convergence at Iteration 4
Input 0.167
Group # 2 -- g: 0.435, h: 0.553
Log likelihood = -90.607 Significance = 0.223

Add Group # 1 with factors afsd

---------- Level # 2 ----------

Run # 4, 8 cells:
Convergence at Iteration 5
Input 0.142
Group # 1 -- a: 0.404, f: 0.577, s: 0.883, d: 0.804
Group # 2 -- g: 0.436, h: 0.552
Log likelihood = -79.451 Significance = 0.284

No remaining groups significant

Groups selected while stepping up:  1
Best stepping up run:   #2
-------------------------------------------

 # Stepping down:
 # Stepping down:

---------- Level # 2 ----------

Run # 5, 8 cells:
Convergence at Iteration 5
Input 0.142
Group # 1 -- a: 0.404, f: 0.577, s: 0.883, d: 0.804
Group # 2 -- g: 0.436, h: 0.552
Log likelihood = -79.451

---------- Level # 1 ----------

Run # 6, 2 cells:
Convergence at Iteration 4
Input 0.167
Group # 2 -- g: 0.435, h: 0.553
Log likelihood = -90.607 Significance = 0.000

Run # 7, 4 cells:
Convergence at Iteration 5
Input 0.144
Group # 1 -- a: 0.400, f: 0.610, s: 0.881, d: 0.798
Log likelihood = -80.034 Significance = 0.284
```

```
Cut Group # 2 with factors gh

---------- Level # 0 ----------

Run # 8, 1 cells:
Convergence at Iteration 2
Input 0.169
Log likelihood = -91.363 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down:  2
Best stepping  up  run: #2
Best stepping down run: #7
```

APPENDIX J

GOLDVARB 2001 RESULTS FOR SELECTED DATA:

MORPHOSYNTACTIC ENVIRONMENT AND DISCOURSE STYLE

```
Binomial Varbrul, 1 step
========================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.

 - One-level analysis only:One-level binomial analysis:

Run # 1, 15 cells:
Convergence at Iteration 12
Input 0.051

Group Factor Weight App/Total Input&Weight

   1:       w    0.434    0.12       0.04
            e    0.421    0.17       0.04
            q    0.981    0.77       0.74

   2:       z    0.505    0.18       0.05
            x    0.439    0.07       0.04
            c    0.500    0.18       0.05

   3:       v    0.917    0.39       0.37
            m    0.140    0.02       0.01
            b    0.853    0.23       0.24
            n    0.947    0.50       0.49

  Cell  Total       App'ns    Expected     Error
   wzv    39         14          12.400     0.303
   wzn     7          3           3.018     0.000
   wzm    90          0           0.616     0.621
   wzb    15          2           2.965     0.392
   wxv     3          1           0.789     0.077
   wxm    10          0           0.052     0.053
   wxb     1          0           0.159     0.188
   qzv     7          6           6.788     3.012
   qzn     1          1           0.981     0.019
   qzm     4          2           1.284     0.588
   qzb     1          1           0.944     0.059
   ezm     1          0           0.006     0.007
   ecv    10          2           3.022     0.495
   ecm     7          0           0.045     0.045
   ecb     5          2           0.931     1.507
Total Chi-square = 7.3658
 Chi-square/cell = 0.4911
Log likelihood = -55.390
```

```
Binomial Varbrul
================
Name of cell file: Untitled.cel

Using fast, less accurate method.
Averaging by weighting factors.
Threshold, step-up/down: 0.050001

 # Stepping up:
 # Stepping up:

---------- Level # 0 ----------

Run # 1, 1 cells:
Convergence at Iteration 2
Input 0.169
Log likelihood = -91.363

---------- Level # 1 ----------

Run # 2, 3 cells:
Convergence at Iteration 5
Input 0.151
Group # 1 -- w: 0.437, e: 0.542, q: 0.949
Log likelihood = -78.589 Significance = 0.000

Run # 3, 3 cells:
Convergence at Iteration 5
Input 0.166
Group # 2 -- z: 0.517, x: 0.279, c: 0.527
Log likelihood = -90.742 Significance = 0.543

Run # 4, 4 cells:
Convergence at Iteration 6
Input 0.076
Group # 3 -- v: 0.886, m: 0.181, b: 0.781, n: 0.924
Log likelihood = -66.821 Significance = 0.000

Add Group # 3 with factors vmbn

---------- Level # 2 ----------

Run # 5, 11 cells:
Convergence at Iteration 12
Input 0.052
Group # 1 -- w: 0.433, e: 0.419, q: 0.982
Group # 3 -- v: 0.917, m: 0.140, b: 0.854, n: 0.948
Log likelihood = -55.417 Significance = 0.000

Run # 6, 10 cells:
Convergence at Iteration 6
Input 0.075
Group # 2 -- z: 0.526, x: 0.348, c: 0.407
Group # 3 -- v: 0.888, m: 0.179, b: 0.788, n: 0.918
Log likelihood = -66.345 Significance = 0.629

Add Group # 1 with factors weq

---------- Level # 3 ----------
```

```
Run # 7, 15 cells:
Convergence at Iteration 12
Input 0.051
Group # 1 -- w: 0.434, e: 0.421, q: 0.981
Group # 2 -- z: 0.505, x: 0.439, c: 0.500
Group # 3 -- v: 0.917, m: 0.140, b: 0.853, n: 0.947
Log likelihood = -55.390 Significance = 0.974

No remaining groups significant

Groups selected while stepping up:  3  1
Best stepping up run:   #5
-------------------------------------------

 # Stepping down:
 # Stepping down:

---------- Level # 3 ----------

Run # 8, 15 cells:
Convergence at Iteration 12
Input 0.051
Group # 1 -- w: 0.434, e: 0.421, q: 0.981
Group # 2 -- z: 0.505, x: 0.439, c: 0.500
Group # 3 -- v: 0.917, m: 0.140, b: 0.853, n: 0.947
Log likelihood = -55.390

---------- Level # 2 ----------

Run # 9, 10 cells:
Convergence at Iteration 6
Input 0.075
Group # 2 -- z: 0.526, x: 0.348, c: 0.407
Group # 3 -- v: 0.888, m: 0.179, b: 0.788, n: 0.918
Log likelihood = -66.345 Significance = 0.000

Run # 10, 11 cells:
Convergence at Iteration 12
Input 0.052
Group # 1 -- w: 0.433, e: 0.419, q: 0.982
Group # 3 -- v: 0.917, m: 0.140, b: 0.854, n: 0.948
Log likelihood = -55.417 Significance = 0.974

Run # 11, 5 cells:
No Convergence at Iteration 20
Input 0.150
Group # 1 -- w: 0.456, e: 0.399, q: 0.951
Group # 2 -- z: 0.493, x: 0.343, c: 0.654
Log likelihood = -78.299 Significance = 0.000

Cut Group # 2 with factors zxc

---------- Level # 1 ----------

Run # 12, 4 cells:
Convergence at Iteration 6
Input 0.076
Group # 3 -- v: 0.886, m: 0.181, b: 0.781, n: 0.924
Log likelihood = -66.821 Significance = 0.000

Run # 13, 3 cells:
```

```
Convergence at Iteration 5
Input 0.151
Group # 1 -- w: 0.437, e: 0.542, q: 0.949
Log likelihood = -78.589 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down:  2
Best stepping  up  run: #5
Best stepping down run: #10
```

REFERENCES

Anis, J. (Ed.). (1999). *Internet communication et langue française* [Internet communication and the French language]. Paris: Hermès Science.

Armstrong, N. (2002). Variable deletion of French *ne*: A cross-stylistic perspective. *Language Sciences, 24,* 153-173.

Armstrong, N., & Smith, A. (2002). The influence of linguistic and social factors on the recent decline of French *ne*. *Journal of French Language Studies, 12,* 23-41.

Ashby, W. (1981). The loss of the negative particle *ne* in French: A syntactic change in progress. *Language, 57,* 674-687.

Ashby, W. (2001). Un nouveau regard sur la chute du *ne* en français parlé tourangeau: S'agit-il d'un changement en cours? [A new look at the loss of *ne* in spoken Tours French: A change in progess?] *Journal of French Language Studies, 11,* 1-22.

Bell, A. (1984). Language style as audience design. *Language in Society, 13,* 145-204.

Bell, A. (2001). Back in style: Reworking audience design. In P. Eckert & J. R. Rickford (Eds.) *Style and sociolinguistic variation* (pp. 139-169). Cambridge, England: Cambridge University Press.

Biber, D., & Finegan, E. (1994). *Sociolinguistic perspectives on register.* Oxford, England: Oxford University Press.

Blanche-Benveniste, C. (1997). La notion de variation syntaxique dans la langue parlée [The notion of syntactic variation in speech]. *Langue Française, 115,* 19-29.

Brunot, F. (1966). *Histoire de la langue française: Des origines à nos jours* [History of the French language: From the origins to our days]. Paris: Armand Colin.

Butler, C. (1985). *Statistics in linguistics.* Oxford, England: Basil Blackwell Ltd.

Chambers, J. K. (2003). *Sociolinguistic theory: Linguistic variation and its significance.* Oxford, England: Blackwell Publishing.

Cohen, M. (1967). *Histoire d'une langue: Le français (des lointaines origines à nos jours)* [History of a language: French (from the origins to our days)] (3rd ed.). Paris: Editions Sociales.

Collot, M., & Belmore, N. (1996). Electronic language: A new variety of English. In S. Herring (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives* (pp. 13-28). Philidelphia: John Benjamins.

Coveney, A. (1996). *Variability in spoken French: A sociolinguistic study of interrogation and negation.* Exeter, England: Elm Bank.

Dauzat, A. (1954). *Le génie de la langue française* [The genius of the French language]. Paris: Payot.

Dauzat, A. (1967). *Tableau de la langue française* [A painting of the French language]. Paris: Payot.

Dejond, A. (2002). *La cyberlangue française* [The French cyberlanguage]. Tournai, France: La Renaissance du Livre.

Eckert, P., & Rickford, J. (Eds.). (2001). *Style and sociolinguistic variation.* Cambridge, England: Cambridge University Press.

Ewert, A. (1969). *The French language.* London: Faber & Faber Limited. (Original work published in 1933).

Fischer, J. (1958). Social influence of a linguistic variant. *Word, 14,* 47-56.

Garcia, A. C., & Jacobs, J. B. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction, 32,* 337-367.

Grevisse, M. (1993). *Le bon usage: Grammaire française* [The right usage: French grammar]. 13th ed. Paris: Duculot.

Halliday, M. A. K. (1976). A brief sketch of systemic grammar. In G. R. Kress (Ed.), *Halliday: System and function in language* (pp. 3-6). Oxford, England: Oxford University Press.

Hansen, A., & Malderez, I. (2004). Le *ne* de négation en région parisienne: Une étude en temps réel [The *ne* of negation in the Paris area: A real-time study]. *Langage & Société, 107,* 5-30.

Hatch, Evelyn, & Lazaraton, Anne. (1991). *The research manual: Design and statistics for applied linguistics.* New York: Newbury House.

Hearne, G., Mandeville, T., & Anthony, D. (1998). *The communication superhighway: Social and economic change in the Digital Age.* St. Leonards: Allen & Unwin.

Herring, S. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication, 4*. Retrieved November 25, 2006, from http://jcmc.indiana.edu/vol4/issue4/herring.html/

Herring, S. C. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 612-634). Oxford, England: Blackwell

Hymes, D. (1974). *Foundations of sociolinguistics.* Philadelphia: University of Pennsylvania Press.

Kayne, R. (1983). Connectedness. *Linguistic Inquiry, 14,* 223-249.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Lavandera, B. (1978). Where does the sociolinguistic variable stop? *Language in Society, 7,* 171-182.

Lewis, D. M. (2005). Arguing in English and French asynchronous online discussion. *Journal of Pragmatics, 37,* 1,801-1,818.

Milroy, L. (1987). *Observing and analysing natural language: A critical account of sociolinguistic method.* Oxford, England: Basil Blackwell.

Mosco, V. (2004). *The digital sublime: Myth, power, and cyberspace.* Cambridge: MIT.

Peeters, B. (2006). "Nous on vous tu(e)": La guerre (pacifique) des pronoms personnels ["Nous on vous tu(e)": The (peaceful) war of personal pronouns]. *Zeitschrift für romanische Philologie, 122,* 201-220.

Peretz-Juillard, C. (1977). *Les voyelles orales à Paris dans les dynamiques des âges de la société* [Oral vowels in Paris in the dynamics of ages of society]. Unpublished doctoral dissertation, Université de Paris V.

Pierozak, I. (2003a). Le "français tchaté": Un objet à géométrie variable? ["Chatted French": An object of variable geometry?] *Langage et Société,104,* 123-144.

Pierozak, I. (2003b). La variation à la marge sur Internet: Pseudos et (re)présentations [Variation on the edge on the Internet: Pseudos and (re)presentations]. *Cahiers du Français Comptemporain, 8,* 195-222.

Pierozak, I. (2003c). *Le français tchaté. Une étude en trois dimensions—sociolinguistique, syntaxique et graphique—d'usages IRC* [Chatted French. A study in three dimensions—sociolinguistic, syntactic, and graphic—of IRC usages]. Unpublished doctoral dissertation, Université d'Aix-Marseille I.

Pohl, J. (1975). L'ommission de *ne* dans le français parlé contemporain [The omission of *ne* in contemporary spoken French]. *Le Français dans le Monde, 111,* 17-23.

Pope, M. K. (1961). *From Latin to Modern French with especial consideration of Anglo-norman: Phonology and morphology.* Manchester, England: Manchester University Press. (Original work published in 1934).

Rickard, P. (1989). *A history of the French language.* 2nd ed. London: Unwin Hyman.

Rohlfs, G. (1970). *From Vulgar Latin to Old French: An introduction to the study of the Old French language* (V. Almazan & L. McCarthy, Trans.). Detroit: Wayne State University Press. (Original work published in 1960).

Sancier-Chateau, A. (1993). *Introduction à la langue du XVII^{e} siècle:* S*yntaxe* [Introduction to the language of the XVII^{e} century: Syntax]. (C. Thomasset, Dir.). Paris: Nathan.

Sankoff, G. (1972). Above and beyond phonology in variable rules. In C.-J. Baily & R. W. Shuy (Eds.), *New ways of analyzing variation in English.* Washington, D. C.: Georgetown University Press.

Sankoff, G., & Vincent, D. (1977). L'emploi productif de *ne* dans le français parlé à Montréal [The productive use of *ne* in Montreal spoken French]. *Le Français Moderne, 45,* 243-256.

Sankoff, G. (1980). *The social life of language.* Philadelphia: University of Pennsylvania Press.

Sax, K. (2003). Acquisition of stylistic variation by American learners of French. *Dissertation Abstracts International, 65* (02), 491A. (UMI No. 200505889)

Tagliamonte, S. A. (2006). *Analysing sociolinguistic variation.* Cambridge, England: Cambridge University Press.

Thorne, S. (1999*).* An activity theoretical analysis of foreign language electronic discourse. *Dissertation Abstracts International, 61* (03), 920A. (UMI No. 200107598)

Trudgill, P. (Ed.). (1984). *Applied sociolinguistics.* London: Academic Press Inc.

van Compernolle, R. (2006, May). *The variable use of* ne *in public French-language chat.* Paper presented at CMT 2006, Bordeaux, France.

van Compernolle, R., & Williams, L. (in press). De l'oral à l'électronique: la variation orthographique comme ressource sociostylistique et pragmatique dans le discours électronique. *Glottopol.*

van Lier, L. (1988). *The classroom and the language learner: Ethnography and second-language classroom research.* London: Longman Publishing Group.

Werry, C. (1996). Linguistic and interactional features of Internet Relay Chat. In S. Herring (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives* (pp. 47-64). Philadelphia: John Benjamins.

Whittle, D. B. (1997). *Cyberspace: The human dimension.* New York: Freeman.

Williams, L. (2003). The nature and complexities of chat discourse: A qualitative case study of multi-level learners of French in an electronic environment. *Dissertation Abstracts International, 64* (09), 3,275A. (UMI No. 200500146)

Williams, L. (2006, May). *Orthographic and sociopragmatic norms and variation in French-language web forums, moderated chat, and non-moderated chat.* Paper presented at CMT 2006, Bordeaux, France.

Williams, L., & van Compernolle, R. (2007). Second-person pronoun use in on-line French-language chat environments. *The French Review, 80,* 804-820.

Wolfram, W. (1991). The linguistic variable: Fact and fantasy, *American Speech, 66,* 22-32.