

A Corpus Analysis of Intra- and Extralinguistic Factors triggering *ne*-Deletion in Phonic French

Charlotte Meisner

Universität Zürich
 cmeisner@rom.uzh.ch

1 Introduction

Sentential negation – the inversion of the truth value of a proposition – is expressed in standard French by the particles *ne* and *pas* (or emphatically *point*) which embrace the inflected verb (cf. (1) a.). However, in informal and/or phonic¹ French (in combination often referred to as ‘spoken French’), negation is prevalently expressed by *pas* alone or by other *forclusifs* like *plus*, *rien*, *jamais* or *personne* that negate parts of a proposition (cf. (1) b.). This contribution based on a corpus analysis aims at a better understanding of different factors, which operate simultaneously on the (non-)realization of the proclitic element *ne* in phonic French.

- (1) a. *Elle ne vient pas.*
 b. *Elle vient pas.*

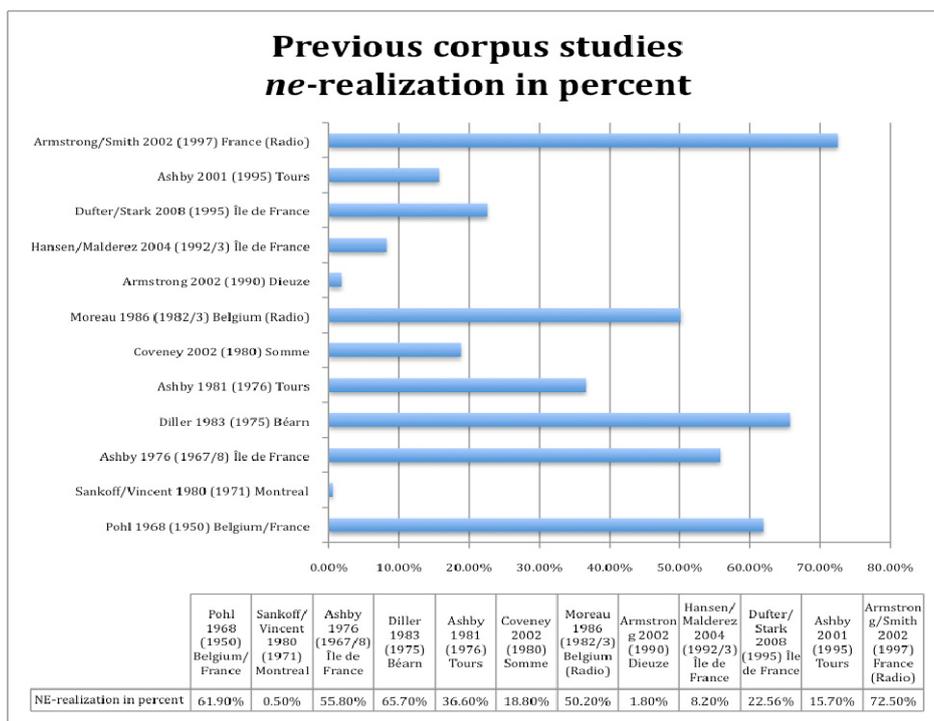


Diagram 1 Previous corpus studies: *ne*-realization in percent

A multitude of intra- and extralinguistic conditions which may influence the realization or delition of *ne* have been proposed in former analyses (cf. Ashby 1976, 1981, 2001, Armstrong 2002, Armstrong/Smith, 2002; Coveney, 2002; Hansen/Malderez, 2004; Dufter/Stark, 2008). Diagram 1 sums up, in top down order starting with the most recent² study, the overall *ne*-realization found in previous corpus analyses during the last five decades. The results presented in diagram 1 show considerable variation; *ne*-realization ranges from 0,5% *ne*-realization in Montreal (cf. Sankoff/Vincent, 1980) to 72,5% in Armstrong/Smiths (2002) study of a French Radio corpus. These two extreme values indicate two important (external) factors on *ne*-realization: diatopic (or geographic) and diaphasic (or stylistic) variation. First, whereas in phonic Quebec French the particle *ne* seems to be almost categorically absent (cf. Sankoff/Vincent 1980), in Europe its realization varies according to the intra- and extralinguistic factors listed under (2) and (3). Second, as diagram 1 shows, the two radio corpora cited here (Moreau, 1986 and Armstrong/Smith 2002) exhibit rather high *ne*-realization rates, which have not been attained in corpora containing *face-to-face* data, since the early eighties (cf. Diller, 1983). Hence, excluding the two radio corpora from the overview in diagram 1, we could also suspect a certain tendency of *ne*-regression during the last half of the 20th century in phonic face-to-face speech. In a diachronic perspective, many researchers argue in favor of an ongoing language change or a shift towards the total absence of *ne* (cf. Posner, 1997). Sometimes normative pressure is found to be the only artificial recovery for the ‘endangered’ particle. The possible regression of *ne*-realization is often described with regard to the so-called *Jespersen Cycle*. The Danish linguist Otto Jespersen (cf. Jespersen, 1917, 1924) characterized the diachronic development (accomplished in other Indo-European languages, for instance in English and German) including the phonetic weakening of a former independent preverbal negation particle, over a stage of pre- and postverbal negation, towards an exclusively postverbal negation, as a cycle. As Völker (2003) reports, in 13th century (formal graphic) French, negation was expressed mainly by the preverbal particles *ne/non* and occasionally supported by postverbal negation markers. Dufter/Stark (2008) note that at the beginning of the 17th century in (informal graphic) French, negation was prevalently expressed by the *forclusif* and *ne*-absence was already available (for an overview cf. also Martineau, 2009). In sum, the assumption of total *ne*-disappearance should be treated with caution for at least two reasons: firstly, the absence of *ne* is not an abrupt recent phenomenon, rather it seems to develop slowly and steadily since the 17th century and is still far from being categorical in modern French, as the radio corpora show. Secondly, overall figures of *ne*-realization (as shown in diagram 1) may mask a more subtle variation, since different intra- and extralinguistic factors operate in complex interaction on *ne*-realization in modern phonic French. Concerning the intra-linguistic factors, the different studies mentioned in diagram 1 suggest the statistic tendencies shown in table 1.

| | Subject type | Presence of other clitics | Frequency of <i>forclusif</i> | Verb tense | Clause type | Phonologic environment |
|--|--------------|---|-------------------------------|----------------------------|-------------------------------|--|
| + favors <i>ne</i>-realization | lexical | no other clitics between <i>ne</i> and the verb | rare (e.g. <i>point</i>) | compound or frequent forms | embedded clause or imperative | intervocalic position (to prevent a hiatus, e.g. <i>Marie n’aime pas</i>), postpausally |
| - hinders <i>ne</i>-realization | clitic | other clitics (e.g. <i>me, te, y, en</i>) between <i>ne</i> and the verb | frequent (e.g. <i>pas</i>) | simple or rare forms | main clause or declarative | |

Table 1 Probabilistic influences of the intralinguistic factors on *ne*-realization

Similarly, in table 2 constellations of external factors have been established which may favor or prevent the realization of *ne* statistically:

| | Age | Social class/ Education | Communication Situation |
|--------------------------------------|-------|----------------------------|----------------------------|
| + favors <i>ne</i> - realization | old | privileged | formal |
| - hinders <i>ne</i> - realization | young | underprivileged | informal |

Table 2 Probabilistic influences of the extralinguistic factors on *ne*-realization

Generally, in a synchronic perspective, the absence of *ne* is seen as a sociolinguistic indicator for lower social class and young speakers (cf. Ashby, 1976, 1981, 2001; Coveney, ²2002; Hansen/Malderez, 2004) and also as a marker of informal speech style (cf. Ashby, 1976, 1981; Armstrong, 2002; Coveney, ²2002) (cf. section 5.2).

However, the descriptions of *ne*-variation often only indicate general and sometimes contradictory tendencies (cf. section 2 for the influence of the communication situation). Furthermore, especially for the external factors, the definitions for categories like (age, social class, formality degree etc.) differ considerably between analyses, which hinders comparisons and allow only fuzzy and unsatisfactory labels like old vs. young and privileged vs. underprivileged. Due to this fact, a corpus-based reconsideration of these tendencies seems inevitable. Nevertheless, a simultaneous examination of all listed factors would be methodologically delicate and is beyond the scope of this paper. Instead, in order to gain a better understanding of the mechanisms determining the presence or absence of *ne*, we propose a detailed examination of two selected factors influencing *ne*-realization in the C-ORAL-ROM reference corpus for spoken French in France (cf. Cresti/Moneglia, 2005) as a showcase. The analyzed factors are the subject type (intra-linguistic) and the formality degree of the communication situation (extralinguistic) both seem to play a major role for *ne*-realization (cf. Ashby 1976, Armstrong/Smith, 2002; Coveney, ²2002; Moreau, 1986; Hansen/Malderez, 2004; Dufter/Stark, 2008). The results presented in this paper will demonstrate exemplarily the complexity and the importance of the co-variation of intra- and extralinguistic factors for linguistic variation in French.

2 Research design

While the general assumptions and methods underlying the present research design are grounded on earlier analyses, the explicit crossing of intra- and extralinguistic factors is a somewhat new approach and requires the development of appropriate analysis tools. Previous analyses have pointed out the particular influence of two factors shown under (2) and (3) on *ne*-realization.

- (2) The intra-linguistic factor: subject type (a small typology)

Type A) *je, tu, il, ils, ce, (on)*³

Clitic personal subject pronouns: i.e. elements, which depend syntactically and phonologically on a (in French verbal) host, which means that clitics can neither be stressed nor conjoined, nor appear alone e.g. as answer to a constituent question. French subject clitics are always attached to the left of the inflected verb (proclisis) and cannot be separated from it except by another clitic (cf. Jones ³2007).

Type B) *elle, nous, vous, elles*

Personal subject pronouns traditionally described as belonging to the clitic pronoun paradigm (type A), but which cannot be identified as clitics via the typical tests (i.e. they can be stressed, separated from the inflected verb and appear alone e.g. as answer to a constituent question, cf. Zwicky 1977).

Type C) *qui, ça/cela*

Non-clitic subject pronouns are syntactically and phonologically independent, that is, they may occur in isolation and in the same position as normal NPs (i.e. noun phrases).

Type D) Lexical subject-NPs, i.e. syntactically and phonologically independent constituents.

The selection of this factor is motivated by the fact that the subject type shows a strong and stable influence on the realization of *ne* throughout earlier analyses (Ashby 1976, Armstrong/Smith 2002, Coveney, ²2002; Moreau, 1986; Hansen/Malderez, 2004; Dufter/Stark, 2008). The division into four subject types (cf. (2)) - instead of only two, namely pronominal vs. lexical - might seem somewhat unusual, but this internal subdivision of the subject pronouns is motivated as well on empiric as on linguistic grounds. Firstly, in previous analyses, only the subject type A seems to constantly hinder the realization of *ne*, while B, C and D seem to favor it. Secondly, a number of tests (cf. Zwicky, 1977) prove clearly, that French clitic subject (and object) pronouns behave differently from subject NPs, proper names and disjunctive pronouns. They can neither be stressed nor conjoined (cf. *Moi/*Je et Jean arriverons bientôt.*), they are not available as short answers (cf. *Qui est arrivé? –Moi/*Je*) and they appear in different positions than NPs (cf. *Jean/*Il, mon voisin, a vu un chien/*le*). For this reason they are said to have a special syntax (cf. the notion of *special clitics*, Zwicky 1977, Anderson 2005). Anderson (2005) identifies two different approaches to define a *clitic*, the first is mainly a phonological one and goes back to the Indo-Europeanists, it defines a *clitic* as a stressless ‘little word’ that depends phonologically (more precisely prosodically) on the adjacent word. The second one is a (morpho)syntactic approach, which consists mainly in the description of the special syntax of certain (pronominal) elements, which behave differently from NPs and disjunctive pronouns. Anderson (2005) adds that the syntactic dimension often, but not necessarily, coincides with the phonological one. This implicates, that syntactic clitics are often but not always stressless (cf. Anderson, 2005 : 1-3, 76). Hence, our type B pronouns are not phonological clitics, but what about their morphosyntactic status as clitics? The application of the typical tests suggests that they can be separated from the verb like normal NPs (cf. *Nous et Jean arriverons bientôt; Qui est arrivé? – Elle*). But do these elements really appear in the same paradigm as the clearly clitic personal pronouns (type A)? In fact we find, in traditional descriptions, on the one hand, the series of syntactically and phonologically independent disjunctive (= non-clitic) pronouns: *moi, toi, lui, elle, nous, vous, eux, elles*, and on the other hand the series of clitic pronouns *je, tu, il, elle, nous, vous, ils, elles*. Both series contain the underlined type B pronouns. It would be desirable to clarify whether they are to be seen as cases of syncretism (i.e. one form fulfilling simultaneously multiple grammatical functions, cf. Bußmann, ³2008 : 708, e.g. the form *nous* for NOM_{non-clitic}/NOM_{clitic}/ACC/DAT), or, as it is often assumed, we deal with cases of homonymy (i.e. accidental formal analogy with distinct meaning, cf. Bußmann, ³2008 : 268) within the two series of French subject pronouns. With regard to *ne*-realization, the remaining key questions (which cannot be addressed in this paper) are whether these pronouns exist in both paradigms in a clitic and a non-clitic version or whether they exist only once, and, if that’s the case, whether they are clitic or not. Since the theoretical questions concerning the clitic or non-clitic status of type B pronouns, cannot be clarified here, we opted in favor of a detailed subdivision of subject types into four groups, so as to detect eventual differences between the types with regard to *ne*-realization.

- (3) The extralinguistic factor: the formality degree of the communication situation (formal vs. informal)

This second factor shown in (3) has been chosen because the achieved results regarding its influence on *ne*-realization are rare, contradictory and methodologically rather superficial. Ashby (1981) simply divides his interviews into a first and a second half, comparing the *ne*-realization rates of the two parts, while he assumes the first to be formal and the latter to be informal because the informants might be more

relaxed the longer they talk to him. Additionally, he reports that three speakers “dropped *ne* roughly twice as often”, when he was able to record them out of their office and in company of friends and family “after food and wine were amply consumed” (cf. Ashby, 1981 : 681-682). Similarly, an observation of Coveney (²2002) leads us to presume that the communication situation has indeed considerable influence on the realization of *ne*. One of Coveney’s (²2002) male informants carries out a massive *style shift* from a formal situation (recorded in the office and in the presence of his boss), in which he realizes *ne* in 50% of the cases, as compared to an informal situation (recorded outdoors, in the presence of only the researcher and small children) in which his *ne*-realization rate declines to 11,4% (cf. Coveney, ²2002 : 88). However, Armstrong’s (2002) analysis suggests the opposite tendency. His data indicate a rather limited influence of the factor communication situation: in informal situations the tested speakers (adolescents) only realize 1,1% of all possible *ne*, which does not differ considerably from the *ne*-realization rate of 2,9% in formal situations (cf. Armstrong, 2002 : 158). Methodologically, Armstrong (2002) makes a difference between *interviews* and *peer conferences* to distinguish between formal and informal situations. The former ones are conducted and recorded by the researcher himself, whereas the latter ones are structured and recorded by the speakers without the researcher. Probably, Armstrong’s (2002) results are only decisive for very young speakers under similar conditions (the only difference between the interview settings is the absence/presence of the researcher). Facing these contradictory results, the influence of the communication situation requires a subtle reconsideration.

The two factors selected for the present paper will enable the corpus based verification of the following three hypotheses:

The hypotheses concerning the intra-linguistic factor i.e. the subject type are:

- I. a. The subject type’s influence on *ne*-realization is crucial.
b. Clitic subjects (type A) favor *ne*-omission, while non-clitic subjects (Type B-D) favor *ne*-realization.

The hypotheses concerning the extralinguistic factor i.e. the formality degree of the communication situation are:

- II. a. Formal discourses show a relatively high presence of *ne*.
b. Informal discourses show a relatively low presence of *ne*.

The hypothesis concerning both factors is:

- III. The subject-type’s frequency differs according to the formality degree of the communication situation: this is the underlying factor, triggering different *ne*-realization rates in formal and informal discourses.

In order to test these hypotheses, a new corpus analysis technique has been adopted: the French part of the C-ORAL-ROM spoken language reference corpus (about 440.000 words) is divided into 8 sub-corpora, which form a continuum between extremely formal and very informal data, cf. diagram 2. The corpus division is based on Koch/Oesterreicher’s (1990) ‘immediate and distant communication’ model (cf. *kommunikative Nähe und Distanz*, Koch/Oesterreicher, 1990). The model foresees ten communicational parameters to determine the formal (i.e distant) or informal (i.e. immediate) character of a communication situation. The information given in the C-ORAL-ROM meta-data covers five of Koch/Oesterreicher’s (1990) ten parameters. Hence, the following five parameters provide the basis of the corpus division: known vs. unknown interlocutor; private vs. public communication; spontaneous vs. planned communication dialogue vs. monologue and face-to-face vs. telephone communication.

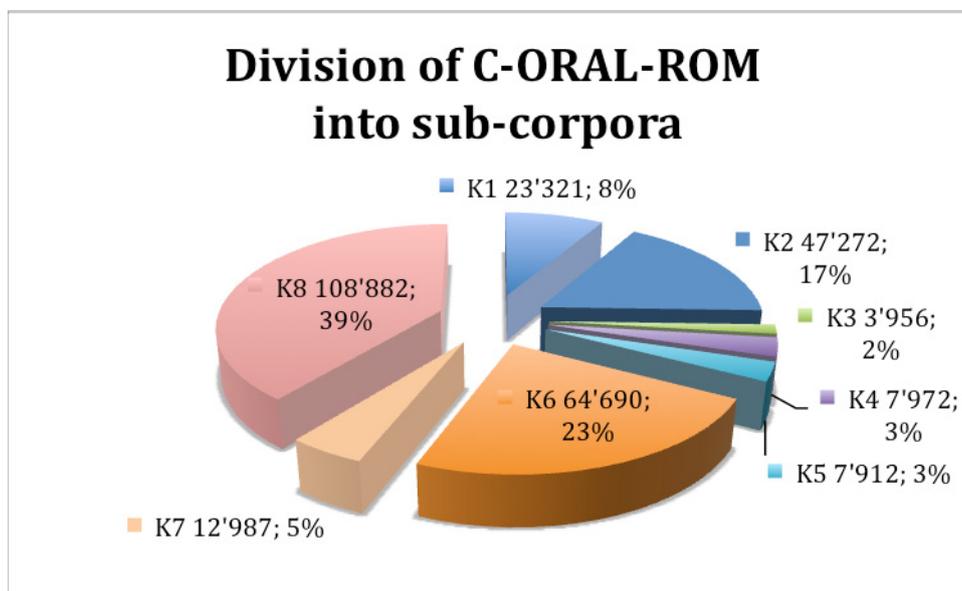


Diagram 2 The division of the C-ORAL-ROM corpus into rather formal (extreme K1) and rather informal (extreme K8) sub-corpora

The quantitative breakdown given in diagram 2 shows the number of words of every sub-corpus and its percentage of the C-ORAL-ROM corpus as a whole. The range of colors in diagram 2 indicates different qualitative characteristics of the sub-corpora: K1 covers extremely formal discourses, that is, data of situations which correspond to all distant criteria (i.e. planned public monologues, directed to unknown interlocutors) like political speeches, sermons or guided tours in a museum. Subsequently the sub-corpora of K2, K3, K4 etc. match decreasingly the criteria of formal communication and increasingly the ones of informal communication. For instance, K2 contains data of political debates and university lectures, in K3 we find professional advices, in K4 and K5 professional dialogic communication via telephone and face-to-face, K6 contains private monologues (e.g. travel report) and K7 stores private telephone conversations. Finally, the sub-corpus K8 contains the most informal discourses, like data from face-to-face conversations between friends and family members. In every sub-corpus and in the corpus as a whole the *ne*-realization rate is tested and calculated via concordance and statistic programs (*Antconc*, *Excel*, *SPSS*). Following the best practice of previous analyses (cf. Ashby 1976, 1981, 2001, Armstrong 2002, Armstrong/Smith, 2002; Coveney, 2002; Hansen/Malderez, 2004; Dufter/Stark, 2008) and the additional considerations of Krötsch (2007), have been excluded all occurrences of negative elements that are not unambiguously expressions of sentential negation and offer a position for *ne*. The following expressions are or contain homonyms of *forclusifs* and have hence been excluded: *pas mal* ('beaucoup'), *pas* ('step'), *pas vrai* ('not true'), *pas grand chose* ('no big thing'), *pourquoi pas* ('why not'), *de rien* ('you are welcome'), *si jamais* ('if ever'), *plus* [plys] ('more'), *plus grand* ('bigger'), *non plus* ('neither') and all non restrictive uses of *que*, like *alors que*, *bien que*, *parce que*, *pendant que* as well as the relative pronoun, the conjunction and the interrogative pronoun *que*. Furthermore, all cases in which the phonological context, for example threw *liaison*, might inhibit the identification of *ne* being realized or not, have been excluded. Hence, the clitic subject pronoun *on* had to be categorically omitted from the study, since with *on* there is usually *liaison* if the following verb starts in a vowel, which is extremely frequent (cf. *on a pas envie* vs. *on n'a pas envie* both are realized as [ɔnapãvi]).

3 The results of the corpus analysis

2432 negation token (i.e. occurrences of sentential negation), which form the database for the following analysis, have been detected in the corpus as a whole. The overall rate of *ne*-realization in the corpus is

696 *+ne*. This corresponds to 29% of all negation token. Hence, in the vast majority of cases (1736 *-ne*, 71%) in the corpus the negation is expressed only by the *forclusif* while *ne* is not realized. These overall results fully meet the expectations established on grounds of previous analyses.

3.1 Verification of hypothesis I

For the purpose of testing hypothesis I. a. “The subject type’s influence on *ne*-realization is crucial”, we need to evaluate the overall *ne*-realization in the corpus (29%, 696 *+ne*) in correlation with the different subject types. The results are shown in table 3 in top-down-order from frequent to rare *ne*-realization.

| | Subject type | | <i>+ne</i> | % | <i>-ne</i> | % | total |
|------|--------------|-------------------------------------|------------------|---------|------------|--------|-------|
| | frequent | D | Lexical subjects | 167 | 87% | 26 | 13% |
| | | Subjectless (Infinitivals, gerunds) | 74 | 80% | 19 | 20% | 93 |
| | | Pronominal subjects | | | | | |
| B | | <i>nous</i> | 15 | 71% | 6 | 29% | 21 |
| C | | <i>qui</i> | 86 | 66% | 44 | 34% | 130 |
| B | | <i>elles</i> | 4 | 57% | 3 | 43% | 7 |
| B | | <i>vous</i> | 27 | 42% | 37 | 58% | 64 |
| A | | <i>il</i> | 100 | 30% | 232 | 70% | 332 |
| C | | <i>ça/ cela</i> | 35 | 26% | 101 | 74% | 136 |
| | | <i>ils</i> | 20 | 21% | 77 | 79% | 97 |
| rare | B | <i>elle</i> | 20 | 20% | 79 | 80% | 99 |
| | A | <i>je</i> | 116 | 15% | 651 | 85% | 767 |
| | A | <i>tu</i> | 11 | 9% | 113 | 91% | 124 |
| | A | <i>ce</i> | 21 | 6% | 348 | 94% | 369 |
| | | total | 696 | 28.62 % | 1736 | 71.38% | 2432 |

Table 3 Overall *ne*-realization according to subject type

For completeness sake table 3 not only contains negative clauses with finite verbs but also negative infinitival constructions and gerunds. However, infinitival and gerund constructions are not directly concerned by our three hypotheses, they require a different grammatical analysis than finite clauses and will rather be considered as additional information than directly included in the analysis. The results shown in table 3 are illustrated in diagram 3.

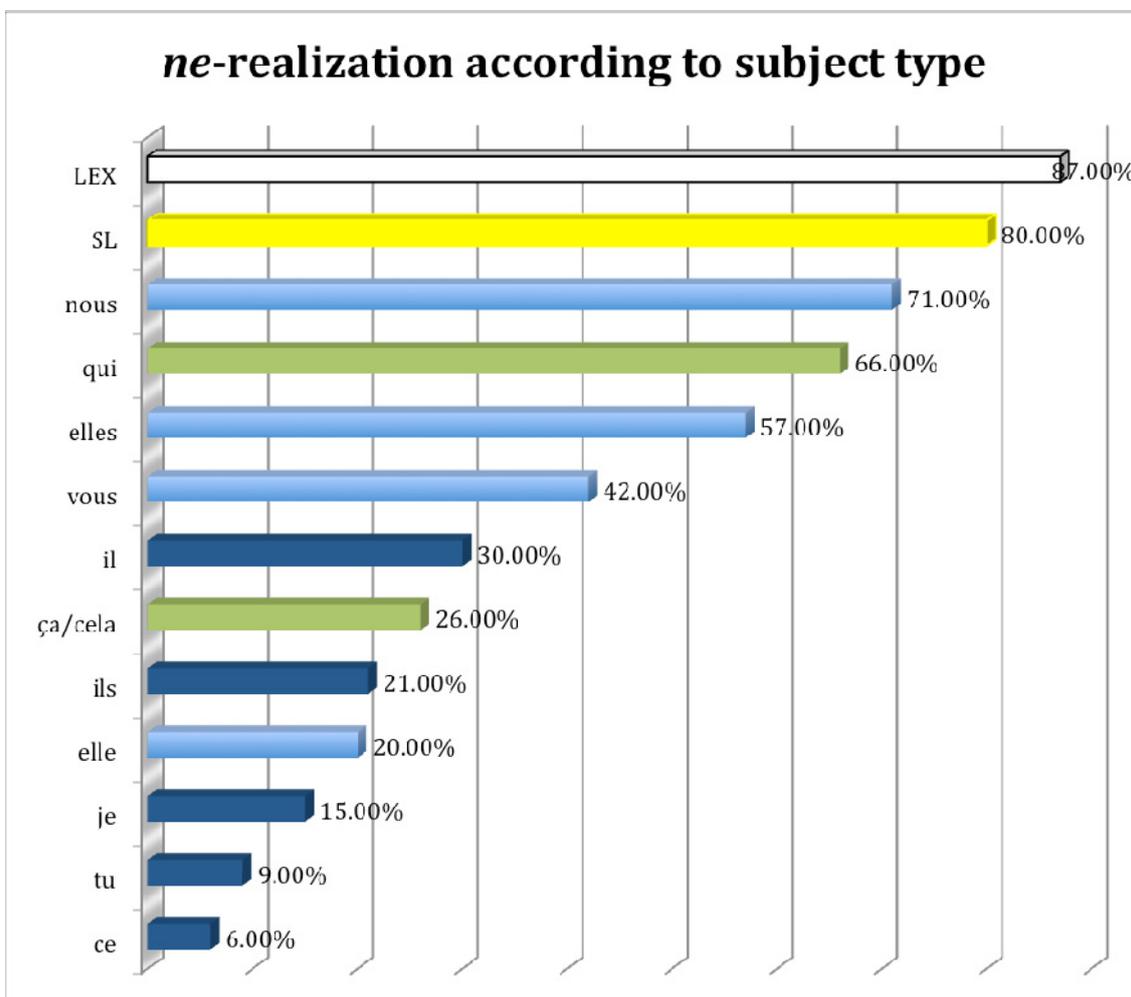


Diagram 3 Overall ne-realization according to subject type

Clitic subjects (type A) are presented (in the online publication of this paper) in dark blue, stressable but syntactically clitic subject pronouns (type B) are shown in light blue and non-clitic pronouns (type C) in green, while lexical subjects (type D) are marked in white and infinitival constructions/gerunds (SL for subjectless or at least without overt subject) appear in yellow. The extreme values range from a fairly low *ne*-realization rate of 6% *+ne*, shown by the clitic subject *ce*, on the left, to the lexical subjects on the right, which attain a very high *ne*-realization rate of 87% *+ne*. This large margin leads us to conclude that the influence of the subject type on *ne*-realization is actually crucial. This view is supported by the statistical chi-square test ($\chi^2 746.5$, $p < 0.000$)⁴, which shows that *ne*-realization significantly depends on the subject type (or on the type of morphosyntactic construction in general). The hypothesis I. a. is therefore confirmed by the corpus data.

The verification of hypothesis I. b. “Clitic subjects (type A) favor *ne*-omission, while non-clitic subjects (Type B-D) favor *ne*-realization” is more delicate and requires a more detailed analysis of the data shown in table 3 and diagram 3. On the one hand, in the majority of the negative propositions with a clitic subject (type A, dark blue) the particle *ne* is not realized: apart from *il* (30% *+ne*), their rates do not surpass the overall average of 29% *ne*-realization. It is notable that *il* elicits despite its clitic nature (and its high discourse frequency; cf. section 6.3) comparatively high *ne*-realization rates. This fact may be

explained by the presence of impersonal constructions (cf. *il (n')y a pas*) that differ in their referential properties from the personal pronoun *il*⁵ and may eventually trigger higher *ne*-realization rates than personal *il*. On the other hand, lexical subjects (type D, white, 87% +*ne*) and some of the stressable (type B, light blue, cf. *nous* 71% +*ne*, *elles* 57% +*ne* and *vous* 42% +*ne*), as well as the non-clitic pronouns (type C, green, cf. *qui* 66% +*ne* and *ça/cela* 26% +*ne*) show much higher *ne*-realization rates. We can insofar conclude that the hypothesis I. b. is also confirmed: generally speaking, clitic subjects hinder *ne*-realization while lexical and to a lesser extent non-clitic and stressable subjects pronouns seem to favor it.

The results have justified, so far, the somehow unusual sub-division of pronominal subjects into A, B and C types, which has been undertaken in section 2. Especially the status of the stressable (type B) pronouns *elle*, *elles*, *nous* and *vous* within the pronominal system of phonic French requires further discussion. Also with regard to actual *ne*-realization, the influence of type B pronouns remains rather vague. Since they dominate the center of diagram 3 we cannot deduce any conclusions about their clitic or non-clitic nature from the data.

3.2 Verification of hypothesis II

In order to verify the hypotheses II a. “Formal discourses show a relatively high presence of *ne*” and b. “Informal discourses show a relatively low presence of *ne*”, which concern the external factor under examination here, i.e. the formality degree of the communication situation, we need to compare the *ne*-realization rates within the 8 sub-corpora ranging from very formal to very informal data. The results are shown in table 4 and in diagram 4.

| | Communication situation: formal vs. informal sub-corpora | +ne | % | -ne | % | Total |
|----------|--|-----|--------|------|--------|-------|
| formal | Corpus K1 unknown interlocutor, public, planned, monologue | 95 | 55.56% | 76 | 44.44% | 171 |
| | Corpus K2 unknown interlocutor, public, spontaneous, dialogue | 248 | 63.75% | 141 | 36.25% | 389 |
| | Corpus K3 unknown interlocutor, private, planned, monologue | 4 | 28.57% | 10 | 71.43% | 14 |
| | Corpus K4 unknown interlocutor, private, spontaneous, dialogue, telephone | 10 | 16.95% | 49 | 83.05% | 59 |
| | Corpus K5 unknown interlocutor, private, spontaneous, dialogue, face-to-face | 24 | 33.80% | 47 | 66.20% | 71 |
| | Corpus K6 known interlocutor, private, spontaneous, monologue | 185 | 40.31% | 274 | 59.69% | 459 |
| informal | Corpus K7 known interlocutor, private, spontaneous, dialogue, telephone | 6 | 4.17% | 138 | 95.83% | 144 |
| | Corpus K8 known interlocutor, private, spontaneous, dialogue, face-to-face | 122 | 10,86% | 1001 | 89,14% | 1123 |
| | Total | 696 | 28.62% | 1736 | 71.38% | 2432 |

Table 4 *Ne-realization according to sub-corpora of rather formal and rather informal discourses*

The chi-square test ($\chi^2 = 547,4$, $p > 0.000$) shows that the realization of *ne* significantly depends on the sub-corpus. In table 4 the results for each sub-corpus are shown in top-down order from the most formal sub-corpus K1 to the most informal one K8.

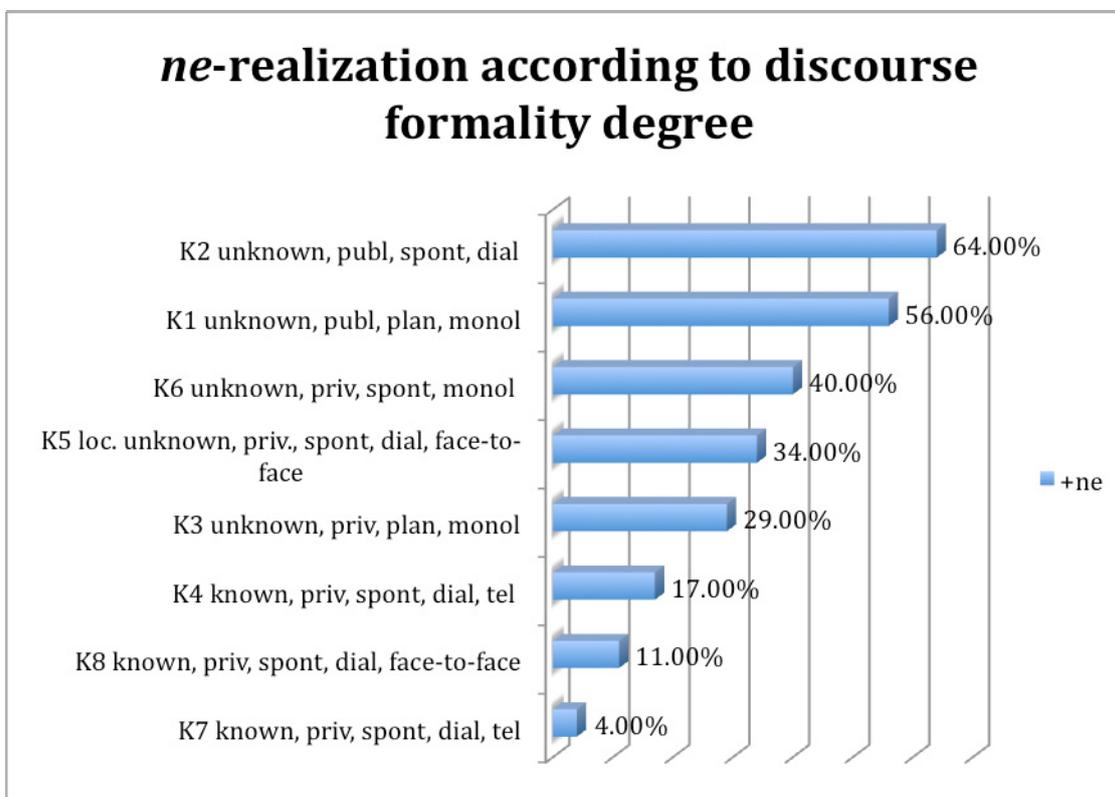


Diagram 4 *Ne-realization according to sub-corpora of rather formal and rather informal discourses*

Diagram 4 shows the *ne*-realization rates of the sub-corpora in top-down order. First of all, the results confirm the anticipations of the hypotheses II a. and b., insofar that the more informal sub-corpora K7 and K8, on the left of diagram 4, exhibit much lower *ne*-realization rates than the more formal ones, K1 and K2, on the right. The parameters of *known vs. unknown interlocutor* and *public vs. private conversation* turn out to be particularly crucial for *ne*-realization (cf. K7, K8, K1, K2), while *spontaneous vs. planned communication* and *monologue vs. dialogue* do not seem to have the same importance, since the rather formal corpus K2, which shows the highest *ne*-realization rate, contains data of spontaneous and dialogic communication. Similarly, telephone communication does not seem to imply a higher degree of formality than face-to-face communication, since the lowest *ne*-rates are measured in K7, a corpus with telephone data. The order of the average corpora K3, K4, K5 and K6 appears quite random and does not allow any conclusion about the influence of isolated parameters. In addition, after considering the limited size of K3, K4 et K5 (cf. diagram 2) and the number of negation token which does not surpass hundred cases in these three corpora (cf. table 4), their results ought to be treated carefully. Generally speaking, we can conclude, that the tendencies predicted by the hypotheses II a. and b. are verified by the data, that the parameters of *known vs. unknown interlocutor* and *private vs. public communication* seem to be the most important ones, but that the influence of isolated parameters remains rather fuzzy.

3.3 Verification of hypothesis III

We have seen that the *ne*-realization in C-ORAL-ROM significantly depends on both factors: the subject type and the discourse type. To deduce continuative information out of these results it would be interesting to know whether there is a dependency relation between the two former variables, too; namely between the subject type and the discourse type. Following the predictions of hypothesis III “The subject-

type's frequency differs according to the formality degree of the communication situation: this is the underlying factor, triggering different *ne*-realization rates in formal and informal discourses”, such relation exists. In order to test this hypothesis, we may currently disregard *ne*-realization rates and instead focus on subject type frequency in every sub-corpus. According to hypothesis III, the distribution of different subject types within the sub-corpora might be quite irregular. Indeed this presumption is confirmed by the results shown in diagram 5.

Distribution of subject types according to sub-corpora

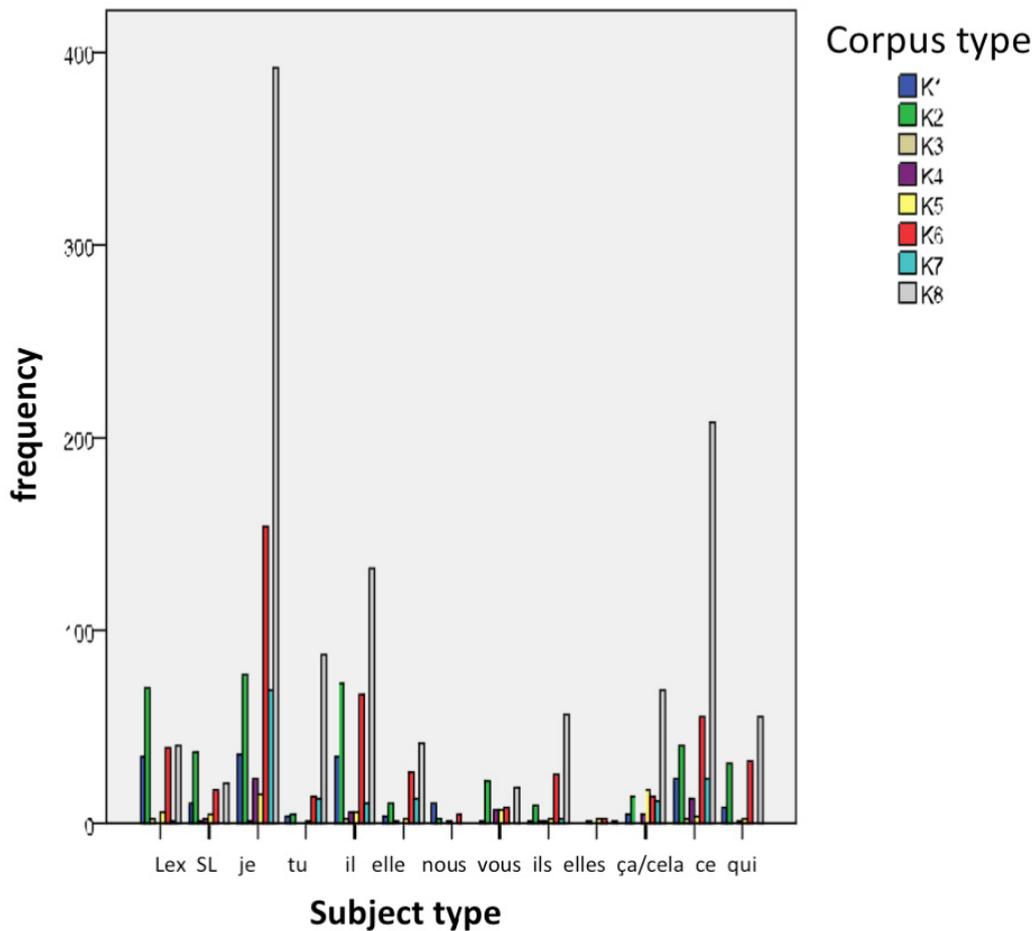


Diagram 5 Distribution of subject types according to sub-corpora

The frequency of the type B subject pronouns *nous* and *elles* seems to be particularly weak compared to the other subject types. To a lesser extent, this is also true for *elle*, *vous* and *ils* (a part from the informal corpus K8). In the informal sub-corpus K8, some of the clitic (type A) pronouns, like *je*, *tu*, *il* and *ce*, attain strikingly high frequencies. By taking into account the huge situational differences between the sub-corpora (cf. section 2) the irregular distribution of subject types within the sub-corpora does not come as a surprise. According to Du Bois (1987, 2003) the frequency of lexical subjects in discourse is generally weak (cf. the concept of *Preferred Argument Structure* from Du Bois 1987, 2003), which is confirmed by the data in diagram 5, especially for the informal corpora K7 and K8. Also the chi-square test

($\chi^2 = 620.6$, $p < 0.000$) shows that the distribution of subject types in the different sub-corpora is not regular, but that it depends significantly on the corpus type. Summing up the results in diagram 5, the implication of hypothesis III of an irregular distribution of subject types across the sub-corpora is confirmed. However, we still do not know whether this really is the underlying factor determining *ne*-distribution in discourse. In order to find it out, we will propose a more detailed analysis of the data in diagram 5.

Since diagram 5 does not allow a direct comparison of the established subject types A to D, the data was rearranged in diagram 6, with respect to the established subject types: lexical subjects (type D) appear in white, and stressable (type B) as well as non-clitic (type C) pronominal subjects appear in light blue. On grounds of statistic reliability - this time - we only compare two of the larger sub-corpora: the formal corpus K2, showing the highest, and the informal corpus, K8 showing the lowest *ne*-realization rate.

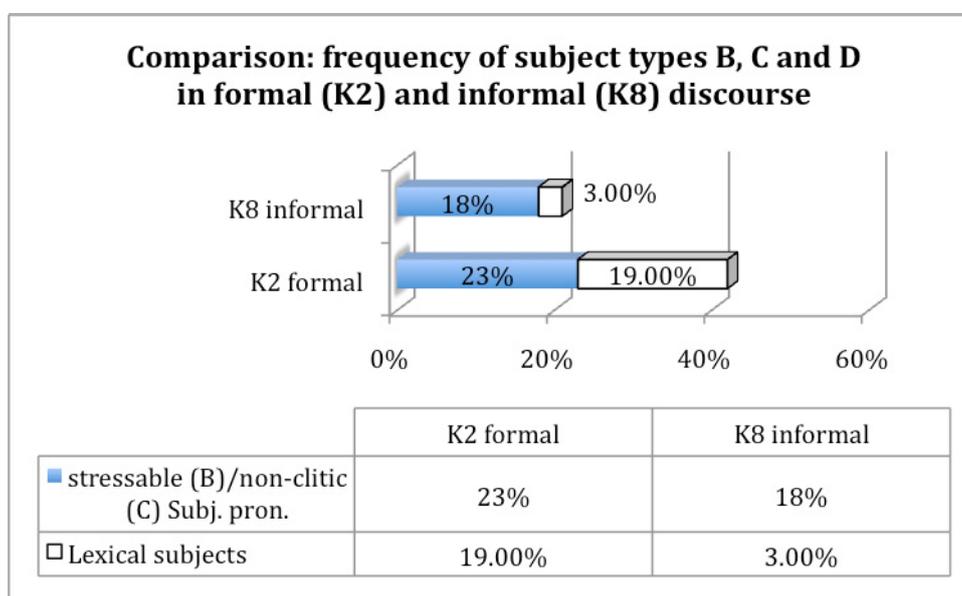


Diagram 6 Comparison: frequency of subject types B, C, D in formal (K2) and informal (K8) discourse

On the one hand, in diagram 6, the formal corpus K2 shows in total 42% of type B, C and D subjects. This is quite a high frequency for these subject types which are in general rather weakly represented. In the break down, we observe in K2 19% of lexical (type D) subjects and 23% of stressable and non-clitic pronouns (type B and C). On the other hand, in the informal corpus K8, only 3% of the subjects are lexical NPs (type D) and 18% are stressable and non-clitic (type B and C) subject pronouns. While the pronoun value does not differ too much from the one of the formal corpus, it is the lexical subject rate, which makes the enormous difference between the two corpora. Consequently, the B, C, and D type subjects in the informal corpus only constitute a total of 21% of all subjects. However, considering the rates of clitic (type A) subjects shown in diagram 7 below, we note that, even in the formal corpus – where the lexical subjects (type D) are particularly frequent – the clitic subjects (type A) still constitute a majority of 58%. In the informal corpus they even attain 79%. Summing up the results regarding the sub-corpora K2 and K8, we can say that the frequency of lexical subjects differs extremely from formal (19%) to informal (3%) discourses, that the B and C type pronouns are distributed almost equally, as well as that the frequency of clitic subjects (type A) is constantly high in formal discourses and rises even higher in informal ones. The implication of hypothesis III, predicting that the frequency of subject types differs with respect to the formality degree, is therefore particularly true for lexical (type D) subjects.

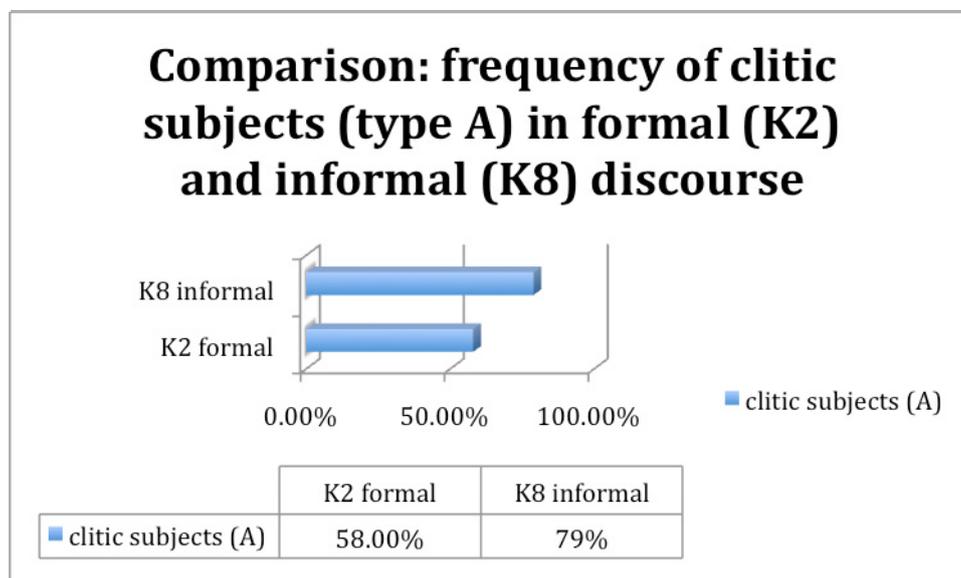


Diagram 7 Comparison: frequency of clitic subjects (type A) in formal (K2) and informal (K8) discourse

With regard to hypothesis III, we may conclude by resuming the three statistically significant relations detected throughout this analysis in the following way: *ne*-realization depends firstly on the subject type and secondly on the formality degree. Thirdly, there is an interconnection between the formality degree and the frequency of different subject types. Hence we can suggest the following interdependency between the three variables: formality degree determines the frequency of different subject types and the different subject types determine the (non-)realization of *ne*. The proposed hierarchy backs up the idea of an underlying key-factor, put forward in hypothesis III. The key-factor is a preference for (clitic) subject pronouns in informal discourses, which unites different, seemingly unrelated factors that operate on *ne*-realization. This pronoun preference, termed (alongside with the preference for known entities) *Preferred Argument Structure* (PAS) by Du Bois (1987, 2003), was discovered in *Sacapultec Maya* and is hence not a specific trait of (phonic) French; on the contrary it seems to be a universal discourse characteristic.

4 Intermediate conclusion

The detailed analysis presented in this paper has shown that both, the subject type and the formality degree are crucial factors for the realization of the negation particle *ne* in phonic French. In order to serialize the various influences of intra- and extralinguistic factors, I have proposed the following hierarchy:

1. The formality degree (extralinguistic factor) determines the selection of subject types (intralinguistic factor). In formal situations the types B, C and D are preferred, and in informal situations the type A.
2. The subject type (intralinguistic factor) determines the realization of *ne*: clitic subjects (type A) somehow seem to hinder *ne*-realization, while lexical subjects encourage *ne*-realization with high probability. Stressable (type B) and non-clitic (type C) subject pronouns are to be situated in an intermediate zone with a slight tendency towards *ne*-realization. Thus, subject type seems to govern – by its respective frequency (cf. PAS) – the variation in *ne*-realization rates between formal and informal discourses.

5 Discussion

The present analysis has shown evidence from corpus (*parole*) data for at least one intralinguistic factor (subject-type) triggering absence/presence of the French negation particle *ne*, which beyond doubt

touches systemic linguistic (*langue*) theory. Thus, the quite clear-cut empiric results of the present corpus analysis may shed light on some current (socio)linguistic questions or lead towards new issues, which we are going to discuss in the following debate.

1. Has *ne* really disappeared from phonic French?
2. Is *ne*-absence/presence really a marker of (in)formal speech style?
3. What role does frequency play?
4. Is there a co-occurrence-restriction between subject clitics and *ne*?

5.1 Has *ne* really disappeared from phonic French?

Considering the claim of an ongoing language change (cf. section 1), our corpus data clearly suggests that *ne*-realization is quite consistent with lexical and some pronominal subjects. While many researchers conjure the total disappearance of *ne* from ‘spoken French’ grammar (cf. Auger, 1994), we observe in modern phonic French surprisingly high *ne*-rates, elicited by certain morphosyntactic contexts (e.g. lexical/non-clitic subjects, possibly also infinitivals and inverted questions). The widespread impression of *ne*-regression in modern French might be caused by the high overall frequency of clitic subjects in speech production. In earlier research on phonic French, lacking the technical finesses of recorders, transcription systems, electronic data processing and concordance programs, neither the percentage of *ne*-realization nor the relative frequencies of subject types could be established reliably. Data from older periods concern graphic realization of French (cf. Völker, 2003) and are thus only partially or not at all comparable to recent ones (cf. Dufter/Stark, 2008, for a methodological discussion of this point). Older or less detailed analyses may furthermore obscure the fact that *ne*-regression systematically evades certain syntactic areas. Despite the low overall rates of *ne*-realization in discourse, we can define quite specific morphosyntactic circumstances that lead to a robust maintenance of *ne*. If *ne*-realization were gradually and uniformly decreasing, this would be rather unexpected. Hence, we may explain *ne*-realization not only by normative pressure, which would effect all negated clauses, regardless of the subject type, but rather consider the morphosyntactic subject type (or the syntactic construction as a whole) (cf. section 5.4)

5.2 Is *ne*-absence/presence really a marker of (in)formal speech style?

With regard to the ongoing debate about whether the absence of *ne* might be an unconscious sociolinguistic *indicator* of less privileged classes, a stylistic *marker* of both less privileged classes and informal speech styles, or a *stereotype*, i.e. a marker that has attracted conscious attention (cf. Trudgill, 2003), the results presented in this paper suggest that none of these three hypotheses holds. Since *ne*-realization turns out to be highly dependent on the intra-linguistic factor of subject type, it seems implausible to argue that social classes may differ in their use of subject types. However, the same is not true for different speech styles. As we have seen in section 3.3, formal and informal discourses differ considerably with respect to subject type frequency. Nevertheless, the absence of *ne* can neither be regarded as a stylistic marker for informal speech styles nor as a stereotype, since – with clitic subjects – it occurs frequently and perfectly unmarked in formal discourses, too. At best, we could label it something like a ‘hearer-marker’: since hearers constantly perceive the lower *ne*-rates in informal discourses (that are provoked by the subject type ratio), they may mistake this epiphenomenon for an independent characteristic feature of informal discourse. Vice versa, it seems more convincing to consider the occasional presence of *ne* with clitic subjects as a stylistic marker of formal discourse. Speakers might mark formal speech style by a kind of hypercorrection, i.e. by realizing *ne* actively even in statistically improbable and morphosyntactically at least marked positions, e.g. with clitic subjects (cf. section 5.4).

5.3 What role does frequency play?

Moreau (1986) has been the first to argue convincingly in favor of certain lexicalization tendencies that can occur when negation is expressed by the most frequent negative element, combined with the most frequent verbs, in the most frequent forms. Very often, these negative sequences seem to be lexicalized without the element *ne* (cf. the notion of *sequence préformée* from Moreau, 1986). Indeed, within the C-ORAL-ROM corpus we find three structures that are likely to be ‘performed sequences’ without *ne*. They occur far more often than the corresponding structures containing *ne*: *c’est pas* (234 token, 95%) vs. *ce n’est pas* (12 token, 5%); *je sais pas* (182 token, 95%) vs. *je n(e) sais pas* (10 token, 5%) and *(il) y a pas* (45 token, 75%) vs. *il n’y a pas* (15 token, 25%). In the special case of *je sais pas* being realized as [ʃəpa] the introduction of *ne* becomes even ungrammatical (*[ʃnəpa]), which clearly suggests that [ʃəpa] is a lexicalized or performed sequence⁶. Grounded on the quantitative results for these three examples, we might predict a general correlation between high discourse frequency of a given linguistic item and low *ne*-realization rates in its environment. For certain pronouns such as *je* (15%+*ne*) and *ce* (6%+*ne*) this seems to be the case. They have rather low *ne*-realization rates and occur quite frequently in subject positions of negative constructions: *je* 33%+*ne* and *ce* 16%+*ne*. However, there are also clitic subjects with equally low *ne*-realization rates, which provide counter evidence to our frequency-inhibits-*ne*-realization-hypothesis, since they occur rarely in discourse. The pronouns *tu* and *ils* manifest 9% and 20% of *ne*-realization while their discourse frequency does not pass 5% of all subject positions in negated clauses. Taken together, these results suggest that there is no bi-directional correlation between discourse frequency and *ne*-realization. High discourse frequency may imply low *ne*-realization rates (and even lexicalization of structures without *ne*) but not vice versa, since low *ne*-realization rates may also occur with comparatively rare (clitic) items like *tu* or *ils*.

5.4 Is there a co-occurrence-restriction between subject clitics and *ne*?

Since, on the one hand, *ne* is almost categorically realized (cf. section 3.1, 87%+*ne*) with lexical subjects and, on the other hand, very often absent with clitic subjects (cf. section 3.1, *ne*-realization ranges from *ce* 6%+*ne* to *il* 30%+*ne*), we might be able describe these tendencies in terms of clitic-constraints or co-occurrence-restrictions. The fact that the order of clitic pronouns (clitic-cluster) in declarative sentences differs from the order of lexical constituents expressing the same syntactic function is widely recognized (cf. the notion of *special clitics*, Zwicky, 1977, Anderson, 2005). Furthermore the syntactic constraints with regard to the order and co-occurrence of elements that occur in clitic-clusters have been described in detail (cf. Perlmutter, 1971). In French, for instance, we find two basic rules that determine the order and co-occurrence of the elements in table 5 below.

Basic rules:

1. Not more than one element per column
2. Elements of the third and the fifth column never co-occur

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|-----------|--|--------------------------------------|---------------------------|----------|-----------|
| Nom | Neg | 1st/2ndAcc/Dat/Refl DO/IO | 3rdAcc DO | 3rdDat IO | Loc | Gen |
| <i>je</i> <i>tu</i> <i>il/elle</i> <i>nous</i> <i>vous</i> <i>ils/elles</i> <i>ce</i> | <i>ne</i> | <i>me</i> <i>te</i> <i>se</i> <i>nous</i> | <i>le</i> <i>la</i> <i>les</i> | <i>lui</i> <i>leur</i> | <i>y</i> | <i>en</i> |

Table 5 *Clitic-cluster in French* (cf. Jones³2007 : 252)

Apparently a result of person-case-constraints, the elements of one column are mutually exclusive (cf. **il tu va*). Furthermore, elements of the columns 3 and 5 are mutually exclusive (cf. (4) a.-d.). Whereas the

first rule sums up the paradigmatic incompatibility of elements expressing the same case and different persons, the second rule describes complex syntactic constraints with regard to case and person. We observe an asymmetry between the 1st and the 2nd person on the one hand and the 3rd person on the other hand. That is, elements of the 1st and 2nd person expressing accusative (DO) or dative (IO) (as well as the reflexive paradigm) situated in column 3 behave differently from elements expressing the third person in accusative (DO) or dative (IO) in the columns 4 and 5. In isolation, the forms of column 3 can express as well accusative (DO), cf. *il me voit* as dative (IO), cf. *il me parle*. However, in combination with column-3-forms they can only be dative (IO), cf. *il me les rendra*; conversely with forms from column 6 and 7 they always express accusative (DO), cf. *il m'y accompagnera*. The concrete effects of basic rule 2. are shown in (4) and (5). Whereas the examples (4) a. to d. are ungrammatical and have to be reformulated by use of the preposition *à* in combination with the corresponding non-clitic pronouns (cf. 4 a.'-b.'), the ones in (5) a. to c. are fine.

- (4) a. **Il me lui présente.*
a.' *Il me présente à lui.*
b. **Il te leur présente.*
b.' *Il te présente à eux.*
c. **Il se lui présente.*
c.' *Il se présente à lui.*
d. **Il nous lui présente.*
d.' *Il nous présente à lui.*

- (5) a. *Il le lui présente*
b. *Il la leur présente*
c. *Il les leur présente.*

It is beyond the scope of this paper to explain why these constraints operate within the French clitic-cluster (for further discussion please see Perlmutter, 1971; Morin, 1981; Miller/Sag, 1997; Miller/Monachesi, 2003; Pomino, 2009). With respect to our analysis, the important question that arises from the existing constraints sketched above is, whether we can formulate the tendencies discovered in the corpus analysis, namely the apparent incompatibility of type A subject clitics with the negation clitic *ne*, as an additional 'rule' or co-occurrence-restriction. Since in table 5 the first column contains as well the 'real' clitic (type A) pronouns as also the stressable (type B) ones, it unites elements that are from our perspective heterogeneous (cf. subject-typology in (2), section 2). Consequently, the first step in this direction is a sub-division of the first column into 1a and 1b as shown in table 6.

Basic rules:

1. Not more than one element per column
2. Elements of the third and the fifth column never co-occur

Additional tendency:

3. Elements of column 1a hardly co-occur with the element of column 2

| 1b | 1a | 2 | 3 | 4 | 5 | 6 | 7 |
|---------------------------|-------------------|-----------|--|---------------------|---------------------|----------|-----------|
| type B | type A | | | | | | |
| Nom _{stressable} | Nom _{cl} | Neg | 1 st /2 nd /Ref 1 | 3 rd Acc | 3 rd Dat | Loc | Gen |
| <i>elle(s)</i> | <i>je</i> | <i>ne</i> | <i>me</i> | <i>le</i> | <i>lui</i> | <i>y</i> | <i>en</i> |
| <i>nous</i> | <i>tu</i> | | <i>te</i> | <i>la</i> | <i>leur</i> | | |
| <i>vous</i> | <i>il(s)</i> | | <i>se</i> | <i>les</i> | | | |
| | <i>ce</i> | | <i>nous</i> | | | | |

Table 6 Co-occurrence restriction tendency for (pro-)nominal elements in French?

In table 6 we reproduce the syntactic order of elements as it is recognized for French in general (shown in table 5) by dividing the nominative column into two sub-columns which correspond to our pre-established subject types A clitic and B stressable. In the first column 1b we find stressable personal pronouns that are traditionally described as clitics but do not behave as such in the typical tests (cf. Zwicky 1977) like *elle(s)*, *nous* and *vous* (cf. type B of our analysis), while in 1a are listed the ‘real’ clitics (type A) *je*, *tu*, *il(s)* and *ce* that have a tendency to exclude the particle *ne* in the following column 2. Our reconsideration of the French clitic-cluster does not touch the formerly established basic rules. Since the distinction between the columns 1a and 1b is grounded rather on morpho-phonological than on syntactic criteria, they still form a single paradigm (hence the dotted line in table 6). In other words, the basic rule 1., describing the incompatibility of elements in one column, treats column 1a and b as single syntactic column⁷. But there is another important difference between the tables 5 and 6: we added to the two basic rules a third one, namely the additional tendency 3. “Elements of column 1a hardly co-occur with the element of column 2”, which sums up our corpus results. We prefer calling it a tendency rather than a rule, since counter evidence is available. The combinations of elements in the columns 1a and 2 (cf. *je ne veux pas*) are a part of French grammar (also of phonic French). They are just statistically unlikely, stylistically marked (cf. section 5.2) and could be the result of spelling pronunciation rather than of spontaneous speech. For the combination of elements in column 1b with *ne* (column 2) speakers show mixed preferences. The realization rates of these pronouns range from *nous* 71%+*ne* to *vous* 42%+*ne*. Compared to the ‘real’ clitics in column 1a they seem to be less dependent on the verb. While type A clitics seem to compete with *ne* for the preverbal position and often exclude the negation particle from realization, type B clitics are not phonologically dependent on the verb, they leave the unaccented preverbal position to *ne* and favor hence its realization. Additionally, the assumption that clitics phonologically compete for realization in the preverbal position is coherent with the findings of earlier analyses suggesting that the presence of the clitics in column 3 to 7 might hinder *ne*-realization as well (cf. section 1 and Ashby, 1981; Hansen/Malderez, 2004). The lack of competition encouraging *ne*-realization with type B pronouns is even more distinctive with non-clitic (type C) and lexical (type D) subjects (which do not appear in the clitic-clusters shown in table 5 and 6 because they are not clitic). Since these constituents are syntactically and phonologically independent, they don’t compete with *ne* for the preverbal position and hence lead to its realization with high probability. Taken together, the subject types A, B, C and D form a continuum from very unlikely *ne*-realization to likely *ne*-realization. In sum, we developed a co-occurrence restriction tendency based on our corpus results and formulated it in terms of the additional tendency 3. We would like to underline that, unlike the basic rules 1. and 2., the tendency in 3. is not a grammatical rule or a person-case restriction whose disregard generates ungrammatical structures (like for instance the disregard of rule 1. would generate structures like **je tu veux pas*). Hence, our additional tendency 3. predicts the preference of *ne*-absence with ‘real’ clitic subjects in a synchronic perspective on modern French (which could eventually lead to a grammatical rule in the future).

6 Conclusion

Throughout the present analysis have been tested the respective influences and possible interrelations of two intra- and extralinguistic factors, namely the subject type and the formality degree of the communication situation, on the realization of the clitic negation particle *ne*. Both factors turned out to have considerable influence, while the subject type seems to be even more decisive. Furthermore, the results clearly suggest that the two tested factors are united by an underlying key-factor, that is the frequency of different subject types, which varies according to the communication situation and hence determines the overall frequency of *ne* in discourse. Since clitic subjects, which tend to hinder *ne*-realization, are generally frequent and occur especially in informal communication situations (cf. section 3.3), the frequency of lexical subjects makes the crucial difference between formal and informal discourses. Lexical subjects are typical for formal discourses and seem to provide morphosyntactic conditions in which *ne* is stably realized. However, not every *ne*-particle found in the data co-occurs with a lexical subject. Also non-clitic and stressable subject pronouns (whose clitic/non-clitic status remains to be clarified) seem to have considerable influence on *ne*-realization, and can often lead to its realization, while the realization with clitic subjects is very improbable (but can occur as a marker of formal discourse or as a hyper-correction phenomenon). Hence, the overall type/token-ratio of subjects sets out the scene for *ne*-realization in discourse. In the ensuing discussion we have addressed current topics in French socio-stylistics and corpus linguistics, such as the thesis of a total disappearance of *ne* from phonic French, its status as a sociolinguistic indicator or stylistic marker, and the significance of frequency effects e.g. the formation of preformed sequences. The evaluation of these questions in the light of our results has shown that *ne* has not disappeared from phonic French, on the contrary, its realization is quite robust with lexical subjects, some kinds of non-clitic and stressable subjects as well as with infinitival constructions and gerunds. For this reason *ne*-absence/presence can neither be termed an indicator for certain social classes, nor marks its absence informal communication situations (even if this impression might come up, due to perceived low *ne*-frequency in informal discourse). The formation of preformed sequences, caused by high discourse frequency of some forms, may catalyze *ne*-absence and make it irreversible in certain cases, such as [ʃepa] vs. *[ʃnɛpa], but this effect cannot be generalized. It appears, therefore, that a morphosyntactic and/or phonologic explanation for *ne*-absence/presence is the fundamental one. Ultimately, we tried to explain this mechanism in terms of a co-occurrence-restriction tendency that inhibits *ne*-realization with clitic subjects. The tendency evaluates the likelihood of certain clitic-clusters and predicts that clitic subjects (type A) are likely to exclude *ne*-realization because of competition inside the clitic-cluster. The co-occurrence restriction tendency sketched in section 5.4 allows a unified approach to the two intra- and extralinguistic factors tested in the corpus analysis through the key-factor PAS (i.e. respective frequency of subject types). Furthermore, since the subject types A to D seem to form a continuum according to *ne*-realization, the co-occurrence restriction tendency could be included in a constraint-based approach (like e.g. Optimality Theory), which might be able to capture the whole range of variation in *ne*-realization with different subject types. Despite a number of questions that could not be treated profoundly in the present paper (for instance the morphosyntactic status of stressable pronouns, the influence of infinitival and gerund constructions, the interaction with other intra- and extralinguistic factors etc.), we hope to have pointed out some possible directions for further research and to have shown that corpus linguistics and linguistic theory are not contradictory, but complementary approaches to language.

References

- Armstrong, N. (2002). Variable deletion of French *ne*. A cross-stylistic perspective, in: *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 French³ language Studies*, 12, 23–41.
- Anderson, S. R. (2005). *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.
- Ashby, W. J. (1976). The Loss of the Negative Morpheme NE in Parisian French. *Language*, 57, 647–687.

- Ashby, W. J. (1981). The Loss of the Negative Particle *ne* in French: a Syntactic Change in Progress. *Language*, 75, 674–687.
- Ashby, W. J. (2001). Un nouveau regard sur la chute du *ne* en français parlé tourangeau. S'agit-il d'un changement en cours ?. *Journal of French language Studies*, 11, 1-22.
- Auger, J. (1994). Pronominal Clitics in Québec Colloquial French: A Morphological Analysis. Doctoral Dissertation: University of Pennsylvania.
- Bußmann, H. (2008). *Lexikon der Sprachwissenschaft*, Stuttgart: Kröner.
- Chomsky, N. (1991). Some Notes on Economy of Derivation and Representation, Freidin, R. (ed), *Principles and Parameters in Comparative Grammar*. Cambridge, MA/London: MIT Press, 417-54.
- Coveney, A. (2002). Variability in spoken French. A sociolinguistic study of Interrogation and Negation. Bristol UK, Portland USA: Elm Bank.
- Cresti, E./Moneglia, M. (eds.) (2005). C-oral-Rom. Integrated reference corpora for spoken Romance languages. Amsterdam: Benjamins.
- De Cat, C. (2007). French dislocation. Interpretation. Syntax. Acquisition. Oxford: Oxford University Press.
- Diller, A.-M. (1983). Subject NP and variable constraints: The case of *ne* deletion. Fasold, R. W. (ed.), *Variation in the Form and Use of Language. A Sociolinguistic Reader*. Washington DC: Georgetown University Press, 167-174.
- Du Bois J. W. (1987). The Discourse Basis of Ergativity. *Language*, 63-4, 805-855.
- Du Bois, J. W. (2003). Argument structure. Grammar in use. Du Bois, J. W. / Kupf, L. E./Ashby, W. J. (eds.), *Preferred Argument Structure. Grammar as Architecture for function*. Amsterdam: Benjamins, 11-60.
- Dufter, A./Stark, E. (2008). La linguistique variationnelle et les changements linguistiques 'mal compris': Le cas de la 'disparition' du *ne* de négation. Combettes, B. / Marchello-Nizia, C. (eds.), *Etudes sur le changement linguistique en français*. Nancy: Presses Universitaires de Nancy, 115-128.
- Gries, Stefan Th. (2009), *Statistics for Linguistics with R. A Practical Introduction*, Berlin: De Gruyter.
- Haegeman, Liliane (1995). *The Syntax of Negation*. New York: Cambridge University Press.
- Hansen, A. B./Malderez, I. (2004). Le *ne* de négation en région parisienne: une étude en temps réel. *Language et Société*, 107, 5-30.
- Jespersen O. (1917). *Negation in English and Other Languages*. København: Hoest.
- Jespersen O. (1924). *The Philosophy of Grammar*. London: Allen & Unwin.
- Jones, M. A. (2007). *Foundations of French Syntax*. New York: Cambridge University Press.
- Kaiser, G. A. (1992). Die klitischen Personalpronomina im Französischen und Portugiesischen. Eine synchronische und diachronische Analyse. Frankfurt a. M. : Veruert.
- Kaiser, G. A. (2008). Zur Grammatikalisierung der französischen Personalpronomina. Stark, E. / Schmidt-Riese R. / Stoll, E. (eds.), *Romanische Syntax im Wandel*, Tübingen: Narr.
- Koch, P./Oesterreicher, W. (1990). *Gesprochene Sprache in der Romania. Französisch, Italienisch, Spanisch*. Tübingen: Niemeyer.
- Krötsch, M. (2007). Le statut de (*ne*) *pas* dans les textes parlés et écrits en français contemporain. Kablitz, A. / König, B. /Küpper, J. et al. (eds.), *Romanistisches Jahrbuch*, 58, Berlin/New York: De Gruyter, 92-131.
- Martineau, F. (2009). Modeling Change : a historical sociolinguistics perspective on French negation, Kawaguchi, Y. / Durand, J. / Minegishi, M. (eds.), *Corpus and Variation in Linguistic Description and Language Education* : Benjamins, 159-178.
- Meisner, C. (2008). Die Realisierung und Auslassung der Negationspartikel *ne* im gegenwartssprachlichen Französisch. Eine Analyse sprachinterner und externer Faktoren anhand eines Korpus und einer Sprecherbefragung, Magisterarbeit, Freie Universität Berlin: unpublished.
- Miller, Philip & Paola Monachesi (2003). Les pronoms clitiques dans les langues romanes. *Les langues romanes*:

- Problèmes de la phrase simple*, Godard, D. (ed.). Paris: Editions du CNRS.
- Miller, P. H. / Sag, I. A. (1997). French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15-3, 573-639.
- Moreau, M.-L. (1986). Les séquences préformées: entre les combinaisons libres et les idiomatismes. Le cas de la négation avec ou sans *ne*. *Le français moderne*, 54, 137-160.
- Morin, Y.-C. (1981). Some Myths about Pronominal Clitics in French. *Linguistic Analysis* 8-2, 95-109.
- Perlmutter, D. M. (1971). *Deep and Surface Structure Constraints in Syntax*. New York et al.: Holt, Rinehart and Winston.
- Pomino, N. (2009). Grundlegende Aspekte französischer Objektclitika im Rahmen der *Distributed Morphology*. Talk held at the Freie Universität Berlin: unpublished manuscript.
- Posner, R. (1997). *Linguistic Change in French*. Oxford: Clarendon Press.
- Sankoff, G./Vincent, D. (1980). The productive use of *ne* in spoken Montreal French. Sankoff, G. (ed), *The Social Life of Language*. Philadelphia: University of Pennsylvania, 311-345.
- Söll, L. (1974). *Gesprochenes und geschriebenes Französisch*. Berlin: Schmidt.
- Sportiche, D. (1996). Clitic constructions. Rooryck, J./ Zaring, L. (ed), *Phrase Structure and the lexicon*, Dordrecht/Boston/London, Kluwer.
- Trudgill, P. (2003). *A Glossary of Sociolinguistics*. Edinburgh: Edinburgh Univ. Press.
- Völker (2003). *Skripta und Variation: Untersuchungen zur Negation und zur Substantivflexion in altfranzösischen Urkunden der Grafschaft Luxemburg (1237-1281)*. Tübingen: Niemeyer.
- Zwicky, A. (1977). On Clitics. *Indiana University Linguistics Club*, 1-40.

¹ Following Söll (1974) we use the term *phonic*, in order to distinguish the canal of realization from *graphic*. This allows us to avoid less clear-cut expressions like ‘spoken’ or ‘oral’ French, which are not only associated with informal phonic speech but also with some kinds of informal graphic speech production like chat or sms communication.

² We indicate two dates for each study presented in diagram 1. The first one is the year of publication while the second one indicates the year(s) of corpus constitution.

³ The clitic *on* had to be excluded from the corpus analysis for methodological reasons (cf. section 2).

⁴ The Pearson chi-square tests in this paper have been calculated via the statistics program SPSS. A chi-square test can show whether the observed frequencies differ significantly from a uniform distribution. The null hypothesis (H_0) is the default hypothesis assuming a uniform or random distribution of variation (in this case of *ne*-presence/absence), which is tested against an alternative Hypothesis H_1 , stating that the variation is not due to chance. The p-value indicates the probability of H_0 being true: the lower the p-value, the less like H_0 is true (cf. Gries 2009 for a comprehensive overview of statistics in linguistics).

⁵ Since the distinction between personal and impersonal *il* has not been made in the present analysis. The actual number of impersonal constructions remains unclear.

⁶ Unfortunately the notational conventions of C-ORAL-ROM do not allow making a difference between the two realizations [ʒepa] and [zəsepɑ], since both are transcribed as *je sais pas*.

⁷ The clitic pronoun *on* has been excluded from the templates in table 3 and 4 since it is not part of the corpus analysis. However, if it were included, it would appear in column 1/1a and its presence would handicap the rigid application of basic rule 1 since it frequently co-occurs with *nous*. For a discussion of the question whether examples such as *nous on va au cinéma* are to be seen as topicalized dislocations or as unmarked constructions in which the subject clitic is reanalyzed as preverbal agreement affix cf. De Cat 2007 and Kaiser 1992, 2008.