7-1-2009

The Spread of Change in French Negation

Angus Grieve-Smith
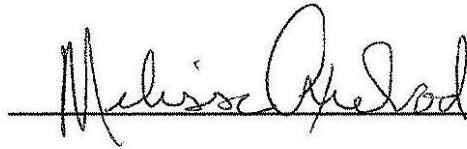
Angus Grieve-Smith

Candidate

Linguistics

Department

This dissertation is approved, and it is acceptable in quality
and form for publication:

Approved by the Dissertation Committee:

Melissa Axelrod, Chair

Catherine Travis

Scott A. Schwenter

William Croft

The Spread of Change in French Negation

Angus B. Grieve-Smith

B.A., Binghamton University, 1993

M.A. University of Chicago, 1994

Dissertation

Submitted in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

Linguistics

The University of New Mexico

Albuquerque, New Mexico

August 2009

Acknowledgements

My chair, Melissa Axelrod, and my other committee members, Bill Croft, Catherine Travis and Scott Schwenter, have given me valuable guidance and suggestions during this project. The dissertation would be worth much less if they had not asked hard questions and pressed me to examine more closely both my data and the theories I used to explain it. This project grew out of a term paper for Joan Bybee's class in Frequency Effects and Emergent Grammar, and I am grateful to Joan and my classmates for guidance and feedback on it. Alan Hudson and Jill Morford have also helped to guide me throughout my doctoral career at the University of New Mexico. Many others in the UNM Linguistics community have provided feedback and moral support, including Karen Naughton, Bonnie Rudy, Cecelia Flood, Jordan Lachler, Sean Burke, Dan Parvaz, Ferdinand de Haan, Natalya Stukova and Jim MacFarlane.

The dissertation also benefited from a Research, Project and Travel Grant from UNM, which helped pay for a trip to Paris for exploratory research and a laptop that helped me to research and write during many late nights in the cafés of Woodside and Sunnyside. Bill Bloom of the UNM Financial Aid Office helped to get that money to me in a reasonable amount of time.

I received several helpful comments when presenting my results at various conferences, notably from Larry Horn, Ambroise Queffelec, Barbara Vance, Paula Chesley and Chad Howe. My friend and neighbor Bill Gratzer, of the Department of Mathematics at Iona College, provided invaluable information and advice about the mathematics of my models and the appropriate use of statistics. Ezra Story, my best friend since childhood, has always been available when I needed programming advice or

The Spread of Change in French Negation

BY

Angus B. Grieve-Smith

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Linguistics

The University of New Mexico

Albuquerque, New Mexico

August 2009

THE SPREAD OF CHANGE IN FRENCH NEGATION

ANGUS B. GRIEVE-SMITH

B.A., Binghamton University, 1993

M.A., University of Chicago, 1994

Ph.D., University of New Mexico, 2009

ABSTRACT

Many varieties of French have changed over the years from expressing predicate negation (Geurts 1998) with *ne* alone, to the embracing construction *ne … pas*, and then to postverbal *pas* alone (Jespersen 1917). When the increase in the frequency of *ne … pas* over time is plotted on a graph, it takes the S shape of the logistic function (Kroch 1989).

Bybee and Thompson (1997) note that "the type frequency of a pattern determines its degree of productivity," but "high frequency forms with alternations resist analogical leveling." These two observations provide an explanation for the logistic progression observed by Kroch (1989). Following Lotka (1925) and Volterra (1926), we can extend this model to take into account the competition between constructions to express the same function.

To test these models, I have compiled a corpus of French theatrical texts from the twelfth to the twentieth century. The logistic function accurately models the use of *ne … pas* in these texts ($R^2 = 0.899$), but the Lotka-Volterra model predicts the post-1600 changes in preverbal *ne* alone and embracing *ne … pas* and *ne … point* with even greater accuracy ($r = 0.948$ and 0.978).

Table of Contents

List of Figures

List of Tables

1 Prologue

Scene I: London, 1160, the gate of old Saint Paul's Cathedral. The priests have long been aware that most of their congregation no longer understands Latin. During a break in the service they are supplementing the Latin mass with a performance of the story of Cain and Abel in the vernacular, which for the elite of England is Norman French. King Henry I and his court are in the audience.

```
1)   Cain: Tu es traïtres tot provez.
     Abel: Certes non sui.
```

Here are their lines translated into present-day English:

```
'Cain: You are a traitor, and it's all proven.
Abel: I most definitely am not.'
```

Scene II: Paris, 1483, the fairgrounds of Saint-Germain. The company of the Clercs de la Bazoche are performing their hit Farce du cuvier, "The Washtub Farce." Jacquinot's wife has fallen into the washtub and can't get out by herself.

```
2)   La Femme: Et sa, la main, mon doulx amy, car de me
     lever ne suis forte.
     Jacquinot: Amy? mais ton grant ennemy; vouldrois
     t'avoir baisée morte.
```

Translated into present-day English:

```
'La Femme: Okay, give me a hand, my sweet friend,
because I'm not strong enough to lift myself up.
Jacquinot: Friend? More like your worst enemy; I wish
I could have killed you with a kiss.'
```

Scene III: Paris, 1542, the newly rebuilt Palais du Louvre. The company of the Confrères de la Passion have arranged a special court performance in honor of the Duc

1

de Vendôme, son-in-law of King François I. They are performing excerpts of their new

play, Le Mistére du Viel testament. In this scene, after killing Abel, Cain is filled with

shame and remorse.

3) Cain: J'ay trop peché villainement; je **ne** suis **pas**
    digne de vivre.

    The English translation of Cain's line:

    'Cain: Wickedly, I have sinned too much; I am not fit
    to live.'

    Scene IV: the Centre Culturel d'Aubange, in the Belgian town of Athus, 2003. The

company "Théâtre des Gaietés d'Athus" is performing Eric Hubert's police drama Le

Parfumeur, for which they will receive the Trophée Royal. FBI Agent White is

interrogating the suspect Marcos.

4) White: Il vous arrive fréquemment de faire vos propres
    tatouages ?
    Marcos: Trop dur. Et puis, je suis **pas** doué en dessin.
    Tant que ce n'est que des lettres.

In English:

    'White: Is that something you do often, make your own
    tattoos?
    Marcos: That's too hard. And I'm not exactly gifted in
    design. But if it's only letters...'

2 Introduction

> J'ay pitié d'un homme qui fait de si grandes différences entre *pas* et
> *point* ...

> *-Jean-louis Guez de Balzac, Socrate chrestien (1662).*

The examples you have just read illustrate an evolution that has taken place in the French language over the past thousand years. The bolded text in each excerpt serves to negate the sentence or clause; in all these cases the result is a phrase meaning "I am not." In the mid-twelfth century, the phrase "non sui" spoken by Abel was already a ᴿᴇʟɪᴄ of Classical Latin *non sum*. By this time, *non* had been reduced to *ne* in most contexts, and people were starting to use *pas* (which is also a noun meaning "step" or "pace") regularly after the verb. In many situations in the mid-sixteenth century, sentences negated with the embracing construction *ne ... pas* were more common than sentences negated with preverbal *ne* alone. We have evidence that people began to stop pronouncing the *ne* in *ne ... pas* in some contexts, and in present-day conversational French, postverbal *pas* alone is now the most common way to negate sentences.

Similar changes happened with nouns like *point*, which meant "point," "stitch" or "point in time," and *mie*, which meant "crumb." They both became part of grammatical constructions *ne ... point* and *ne ... mie*, but in the late Middle Ages *ne ... mie* declined in use and vanished from the language, and in the early modern period *ne ... point* declined and is now itself a relic.

As we will see in Chapter 4, this change in French negation from preverbal *ne* alone to postverbal *pas* alone is one of the most famous examples of syntactic change, and a textbook example of grammaticization (Meillet 1912). It has been discussed by

3

many people, most famously Otto Jespersen. Jespersen (1917) observed that the same pattern where a monomorphemic negation is strengthened by additional material and then reduced to a monomorphemic construction has also occurred in English, Danish and German, and in some languages (such as Latin/French and Greek) has recurred multiple times. It has been found in numerous other languages, including Welsh (Willis 2005) and Arabic (Lucas 2005).

Grammaticization is widely understood to include other processes such as syntactic reanalysis, analogical extension, reduction, semantic bleaching and decategorialization (Heine 2003). There is a significant debate as to whether it is more than the sum of its parts, or can be reduced to reanalysis plus extension, and in Chapter 5 I will summarize the debate and its relevance to the case of French negation. I will briefly discuss reanalysis, synthesizing the work of Traugott (1989), Croft (2000), Eckardt (2007), Detges and Waltereit (2002), Geurts (1998), Israel (2001), Kiparsky and Condoravdi (2006), and Schwenter (2006) into a hypothesized semantic path from literal reference through emphatic denial to predicate negation.

The main focus of this study will be on analogical extension and a cluster of related concepts, such as propagation, selection, lexical diffusion, competition, specification, obligatorification and productivity. Bybee and Thompson (1997) argue that analogical extension is a function of the relative productivity of competing constructions, and as such is motivated by differences in TYPE FREQUENCY, the relative prevalence of the construction among the different lexical items that use it. It affects contexts with low TOKEN FREQUENCY (the number of times that the context has been perceived by the language user) first, and those with high token frequency last.

Kroch (1989) predicts that in syntactic changes, the use of a new variant will often follow an "S-curve" logistic pattern; Labov (1994) attributes this to propagation across language users and not through the lexicon, but as we will see later in Chapter 5, his exclusion of lexical propagation is not justified. The logistic function was developed by Verhulst (1838) to model exponential population increases in limited domains, as described by Malthus (1789). However, its application is limited to a single species, and does not take into account competition for resources. Lotka (1925) and Volterra (1926) extended the model to describe this kind of resource competition.

My aim is to construct a detailed quantitative model of extension and propagation, and to test this model against data from French negation. Kroch's prediction has been tested on data from other changes in the history of French (Kroch 1989) and on English negation by Kallel (2007), but Kallel focused on the last stage (the shift from embracing negations like *ne ... not* to postverbal *not* alone). Bybee and Thompson's (1997) predictions have been applied to changes in verb conjugations (Hooper 1976) and to the productivity of plural markers in English and German (Bybee 1985), but they have not been tested on any area relating to French negation. Martineau and Mougeon (2003) performed a long-term corpus study of the shift from *ne ... pas* to postverbal *pas* alone, but did not measure token or type frequencies. This is thus the first systematic corpus study of the shift from preverbal *ne* alone to the *embracing negation* constructions *ne ... pas* and *ne ... point* in French sentence negation, and the first one to test logistic regression of this change and the role of frequency effects.

In the past, changes in French negation have been examined by looking at usage frequencies in individual texts (Yvon 1959) or by comparing usage in texts selected from

various time periods (Yvon 1962, Kroch 1989) but *ad hoc* text selection makes a study vulnerable to selection bias (Hopkins, Hopkins and Glass 1996). In the context of a synchronic study of English, Lee (1999) recommends taking a representative sampling to avoid selection bias, and his recommendations are valid for all corpus studies. I will extend this principle to the study of change over time by systematically sampling the available texts. The resources needed to create a strictly representative sample are not currently available, but the corpus I have assembled is more consistent and finer-grained than previous corpora used in this area.

One of the major risks in corpus linguistics is genre bias (Lee 1999, Grieve-Smith 2006). To avoid this kind of bias, I have chosen to restrict my corpus to theatrical texts. This limits the generalizability of my findings to theatrical texts as well, but increases their overall reliability. Theatrical texts are also the closest genre to spoken conversation that is available for the period under discussion. I collected 46 texts covering the period 1160 through 1929, totaling over 500,000 words and 6500 negated sentences.

Although the corpus is limited, and thus our power to generalize from the results is also limited, the results are promising for further research. In the Old French data, the use of preverbal *ne* alone and the embracing negations encoded a distinction between predicate negation and presupposition denial, but beginning in the seventeenth century, the playwrights in the corpus stopped thinking of the constructions as different and began using them either interchangeably, or else according to rules set down by grammarians. Syntactically, I found three interacting competitions in this system: the competition among embracing *ne ... pas*, *ne ... point* and *ne ... mie* from Old French through the sixteenth century; competition between *ne ... pas* and *ne ... point* from the time that

6

ne ... mie stopped being used in the corpus through the nineteenth century; and competition between preverbal *ne* alone and embracing *ne ... pas* from the seventeenth through twentieth centuries.

The logistic model is only meaningful if applied to increasing changes, and it fits two such changes closely: the rise of *ne ... pas* and the rise of *ne ... point*. While a bare logistic model of the kind used by Kroch (1989) does not model the interaction between type frequency and productivity, a Lotka-Volterra model that takes into account competition between constructions does fit the data very well. The data are also supportive of the idea that high token frequency constructions resist this kind of change.

I will begin the discussion with an overview of the history of French in general and negation in particular. In Chapter 4 I will discuss the semantic and pragmatic aspects of changes in negation. I will then present the theoretical background on grammaticization, reanalysis and extension. I will describe the methods that I have used for collecting and analyzing the corpus data, and then present the results and discuss their implications.

3 Negation in the History of French

The changes that I will be discussing are a small part of the larger history of the French language, and in later chapters we will see how strongly they were affected by this history. We have over a thousand years of records to draw on, from which I will be examining a selection of theatrical texts. In this chapter I will give a brief overview of the sociopolitical history of the language. I will then discuss the overall morphosyntactic history of negation in French, beginning with its antecedents in Latin and continuing on through the present day; in a later chapter I will discuss the current understanding of the semantic and pragmatic evolution of the term. Where possible I will illustrate the discussion with examples from the corpus I have collected, and if necessary I will supplement these examples with excerpts from more recent texts.

3.1 The Sociopolitical History of French

French is one of the best-known and most widely spoken languages in the world (Nadeau and Barlow 2006). Its documented history spans over a thousand years, from the Strasbourg Oaths in 842 to the present day. As described by Lodge (1993), the language evolved from a contact situation when a population of Latin-speaking Roman citizens moved from Italy into Gaul, and the formerly Celtic-speaking tribes in the area shifted to Latin. During the breakup of the Roman Empire, parts of Gaul were occupied by Germanic-speaking tribes who eventually shifted to speak the language varieties (described as either Vulgar Latin or Proto-Romance) that were spoken by the people who already lived in the area.

Due to the breakdown of central Roman authority and the difficulties of communication during the Middle Ages, the vernacular language developed a significant amount of regional variation, and these varieties were grouped by Dante (1304) into the southern *langue d'oc* and the northern *langue d'oïl*, based on the different words for "yes" in these varieties. It is the *langue d'oïl* variety spoken in the area around Paris that became known as French, but there are several other varieties of *langue d'oïl*, including Norman, Burgundian, Picard and Walloon.

A large body of literature was written in French and Picard in the middle ages, in a number of indigenous genres such as mystery plays and epic and lyric poetry, but Latin was used in many other functions. In the sixteenth century a movement arose to establish French as the language of the royal court, of the law and of literature, replacing Latin. In 1539 King Francis I (François I) signed the Ordinance of Villers-Cotterêts, requiring the use of French in all judicial acts, notarized contracts and official legislation. There was a corresponding literary movement known as the *Pléiade*, who had been involved in the revival of interest in classical literature, and argued that French was not barbaric as some claimed, but in fact capable of the same greatness achieved by Latin and Greek. Eventually French came to be used in worldwide diplomacy and in France and Belgium's former colonies, as well as replacing some of the indigenous varieties within the borders of France, Belgium, Luxembourg, Switzerland and Monaco.

3.2 Prelude to French: Negation in Latin

The usual story of French negation, as found in any number of discussions of the issue (e.g. Detges and Waltereit 2002, Hopper and Traugott 2003, Rostila 2006), goes something like this: In Latin, the negator of choice was *non* alone, but people said *non*

vado passum as a figure of speech to mean "I'm not going one step." They said it often

enough that the word *passum* became polysemous, with a noun retaining the original

meaning of "step" and an adverb that, bleached of the restriction to motion, served to

intensify negations. There were other phrases, such as *non vedo punctum* and *non*

comedo micam, meaning "I don't see one point" and "I'm not eating one crumb." In the

meantime, as we will see in the next section, *non* was reduced to *ne*. The phrase *ne ... pas*

gradually increased in meaning until it replaced *ne* alone and eventually all the other

negators, and then people began to stop saying the *ne* in the vernacular, so now the main

negator is postverbal *pas* alone.

This view of the history French negation leads Rostila (2006) to reject Detges and

Waltereit's (2002) view of grammaticization and argue that high type frequency caused

the grammaticization of *ne ... pas* by prompting its generalization from verbs of

movement to general forms. The essentials of this story are correct, but as with most

stories that are widely told, it is a simplification, and the expanded version contains

several important details.

There is in fact no evidence that anyone has ever said or written *non vado*

passum, except as an example sentence. Schweighäuser wrote (1852: 225), "Observons

toutefois que cette modification apporte au sens du mot *pas* est antérieure aux plus

anciens monuments de la langue."[1] I have been unable to find any more recent citations

or uses of *passum* or its Vulgar Latin variant *passu* with the verb *vadere* or any other

verbs of motion. However, Schweighäuser observes that many languages have forms

---

[1] "Let us note in any case that this modification undergone by the sense of the word *pas* is prior to
the most ancient relics of the language."

using "items of minimal value" to reinforce negation, and gives the following example from Latin:

5) **Non** ego **nunc** emam vitam tuam **vitiosa nuce.**
   'I wouldn't buy your life for a rotten nut.'

Plautus, Miles Gloriosus II, 3, 45

Schweighäuser lists this among eight examples from Plautus, two from Cicero and one from Catullus, using coins (*as, dupoindius, numus, teruncius, triobolus*), food (in addition to the rotten nut, a *ciccus*, the membrane separating the seeds of a pomagranate) and various materials (*floccus*, a bit of fleece; *naucus*, a kind of nut zest; *pilus*, a hair; *pluma*, a feather; *titivillitium*, a piece of lint), and a *libell* or small book.

Although he has no simple examples of *passum* used in negation, Schweighäuser indicates (1852: 224), two examples where Plautus uses the related noun *pes*, "foot" as an item of minimal value.

6) **Ne** iste hercle ab ista **non pedem** discedat
   'Good Lord! The fellow wouldn't move a step from her'

Plautus, Asinaria III, 3, 13, translation by Paul Nixon (1916).

This uses the verb *discedere*, 'to move away from', not *vadere*, 'to go'. More importantly, Schweighäuser also brings up this example, from a document relating to the Albigensian crusade, discussing God's gift (or lack thereof) of land to Abraham:

7) Quod autem dedit **nec passum pedis,** dicit Stephanus.
   'Because he did not give him a single pace, says Stephen'

Collection Doat., vol. XXXVI, folio 120

Schweighäuser does not seem to be aware that the phrase *nec passum pedis* is directly quoted from Stephen's speech as related in the Latin Vulgate (Acts 7:5), where

the Douay-Rheims translation is "not the pace of a foot." This is the only Latin example where *passum* is used in an item of minimal value, so it merits some close attention. It turns out that the Latin phrase is a direct translation of a Greek phrase in the original New Testament, ουδε βημα ποδος. This phrase, in turn, seems to have been used by Stephen to allude to a passage in Deuteronomy 2:5, which was identical in the Septuagint. In the original Hebrew it was גֶל מִדְרַ דְ כַּף־רָ. This is translated in the King James version as "not so much as a foot breadth," and in most translations convey the sense of a space that is just big enough to set foot on. In fact, the Vulgate translation of this passage is a very explicit one:

8) **neque** enim dabo vobis de terra eorum **quantum potest unius pedis calcare vestigium quia**
   'For I will not give you of their land so much as the step of one foot can tread upon'

Deut. 2:5, Vulgate and Douay-Rheims translations

In Latin, *passum* was a Roman unit of distance equivalent to the length of two footsteps (left and right), about five feet. One *passum* was borrowed into English as *pace*, and a thousand paces, *mille passibus*, became our *mile* (Zupko 1985). *Pes* itself was a Roman unit of distance, which fits with Plautus' use of it in the previous example. The phrase in Acts 7:5 of the Vulgate could have referred to a pace or a foot-breadth.

It is not at all clear that the French phrase *ne ... pas* evolved out of this phrase in the Vulgate Bible. It could have arisen independently, and just coincidentally resembled the form in *Acts of the Apostles*. However, medieval France was a very religious society, and there is also a possibility that the phrase could have been borrowed through allusions to that section of Acts. What is clear is that *pas* did not evolve "from step to negation," as claimed by Eckardt (2007), but "from pace (or foot) to negation"; it was already being

used as a measurement of area before it became part of a negative construction. A measurement, especially a small one relative to territory, makes more sense as an item of minimal value, and as a minimal unit of both distance and territory *passu* could be expected to occur in a much wider range of contexts than just verbs of motion.

Diez (1882: 1079), Vaananen (1967: 240), Stein (2002) and Eckardt (2007) give evidence that *punctum, micam* and *gutta*, the antecedents of *point, mie* and *goutte* were already being used as items of minimal value in Latin, particularly in the work of Plautus and Jerome. The expression *ne punctum quidam (temporis)* "not a single point in time" was common in the work of Cicero and other authors; note how, as with *passum*, its sense was a more general, abstract one as opposed to the concrete sense of "stitch" sometimes given. Petronius uses *non micam panis*, "not a crumb of bread"; more interestingly, Diez cites *non micam sanae mentis habere*, "not a crumb of sane mind," but the only uses of that expression I can find are in the writings of the sixteenth-century Dutch philosopher Erasmus. The phrase *neque parata gutta certi consilii* "not a drop of a firm plan" is used by Plautus; interestingly, Lindsay (1900: 30) reports that *gutta* is used in only one of the sources for this play, the one that has been most copied; the other source, the Ambrosian Palimpsest, has *neque paratust quicquam*, "not any kind of ...," a minimizer that is not derived from a noun, but was passed on to French in the form *quelconque*. It is possible that *gutta* could have been introduced by a copyist.

From this incomplete data we see that these minimizing idioms were common in Latin, as in all languages, but no single expression was frequent enough to be a grammatical construction. In the next section we will examine the history of sentence negation in French.

3.3 An Overview of the History of Negation in French

A number of authors have written about the history of negation in French. An extensive overview in Japanese is given by Kawaguchi (2005), and less-complete summaries can be found in papers by Rostila (2006) and Englebert (1984). The Groupe Aixois de Recherche en Syntaxe produced a series of detailed investigations of negation in particular Old French texts (Yvon 1959, 1960a, 1960b; Yvon 1962 compares the medieval examples to the twentieth century poet Paul Valéry). Malcolm Offord (1976) focused on the fourteenth-century prose epic *Berinus*. In general histories of French, Brunot (1901) covers every period up to 1900, and for the fourteenth and fifteenth century, Marchello-Nizia (1992) provides a long section full of examples.

3.3.1 Preverbal Negation from Late Latin to Old French

As Lodge (1996) observes, it is difficult to establish a coherent narrative of the change from Latin to French, because so much of the language during that period is unrecorded. Because of this, the corpus used for this study begins at 1100, but other authors have uncovered some evidence of general trends. Lodge does not hesitate to describe the linguistic situation in the late Roman Empire as diglossic, where the vast majority of written texts were in a conservative Classical Latin, but everyday conversation took place in various dialects of Vulgar Latin or Proto-Romance that diverged more and more from the written form every year. Very little was written in "the Roman language" until the twelfth century. Because of this, we have a lot of evidence of Classical Latin and a significant amount for Old French, but relatively little about the period between these two stages, a period lasting roughly four hundred years. Unfortunately, some of the most interesting changes in negation seem to have taken place

during this gap, such as the evolution of *non/ne ... pas* from a poetic expression to a fixed idiom and then to an emphatic negator, and there is not much we can do to fill the gap other than searching for lost texts.

There are clear descendents of the Latin negator *non* in French, including the homophonous predicate negator. In medieval French texts it appeared in a wider range of contexts. Moignet (1965) describes this evolution in detail, although for the earliest stages (the ninth and tenth centuries) he only has a handful of texts to go on, in all genres combined. In preverbal form, *non* tended to reduce to *no* and then *ne* before a verb beginning in a consonant, and to *nen* and then *n'* before a verb beginning in a vowel; this change was part of a general vowel reduction tendency in Old French that also affected such frequent items as *jo*, 'I', reduced to *je*. In other cases, such as in single-word responses or when negating a noun phrase, *non* remained unreduced. Moignet attributes this to the variable stress system of Old French, which allowed particles like *non* to be stressed, and he suggests that as a regular (i.e. preverbal) predicate negator *non/no/ne* did not ordinarily receive independent stress, and was thus reduced.

Moignet also points out that there was an intermediate stage (confirmed by the data from this study) where a few high-frequency fixed expressions such as *non ai* 'I don't have' and *non ferai* 'I won't do it' continued to take preverbal *non*.

Old French *ne* was a reduced form of *non*, but before a vowel the word was not always reduced as much, resulting in *nen*. In a few instances, it retained stress, and was written as *né* or *në*. In later texts, the word was reduced even further before a vowel, to the simple consonant *n'*. If *ne* was followed by the direct object clitic *l'* (in its prevocalic

reduced form), the two particles could be combined to form *nel*, as in the common refusal *nel ferai* 'I won't do it.'

3.3.2 From Minimizing Idioms to Emphatic Negators

Many of the Latin minimizing idioms that I discussed in the previous section reappeared in Old French, with some new ones added (Schweighäuser 1852 gives a full list). Many of them were still clearly idioms, but some were more frequent, more formulaic and less attached to their original meanings.

These more frequent expressions took the form of EMBRACING NEGATIONS, with a preverbal *non* or *ne* and a postverbal particle. I should point out here that 'postverbal' is a simplification: these particles almost always appeared after the verb, but as I will discuss later, in a few instances they were instead topicalized to a position before the *ne*, and in circumstances where an infinitive was negated but its matrix clause was not, the particles were often put between the *ne* and the finite verb; this is still the rule in Standard French. For sentence negation, this could be one of *pas, point* or *mie*. Some authors (e.g. Ewert 1943, Posner 1985, Gregory 1997) also include *nient*, *giens* and *goutte*, especially since they are used in closely related Romance varieties, but many of them were more commonly used as negative pronominals (see the next section).

Pas, point, mie, etc. derived from nouns meaning 'pace,' 'point,' 'crumb' and so forth, and these nouns normally take articles in French, but the particles in these negative constructions are not used with articles. This suggests that they were grammaticized before articles became obligatory in French, an interpretation which is further supported by the Latin data cited in the previous section.

16

Although by most accounts the embracing negation constructions with *ne* and a postverbal particle formed only a small minority of sentence negations in the earliest Old French texts, they gradually grew in frequency over the years, and now form the overwhelming majority of sentence negations; how this occurred is the focus of this study.

There are no instances of embracing negation in the first two extant examples of Old French, the Strasbourg Oaths of 842 and the ninth-century *Cantilène de Saint Eulalie*, but Strasbourg Oaths are very short, and both texts are not particularly representative of typical language use. The eleventh-century Anglo-Norman epic poem *La Chanson de Roland* contains examples of *ne ... pas*. It does occur with *aller*, 'to go', but also with a number of other verbs including *devoir*, 'to be obligated', *pouvoir*, 'to be able', and *être*, 'to be'. The verb *aller* also occurs with *ne ... mie*, which is the most common negator in the poem; it also occurs with *aimer*, 'to love' and *desotrier*, 'to refuse'. The other sentence negator in the *Chanson de Roland* is *ne ... nient*, which occurs with *aimer* and *falloir*, 'to be necessary'. The eleventh-century Limousin liturgical drama *Sponsus* contains the phrase *n'auret pont* 'he will not have', which is the earliest example of *ne ... point* that I have been able to find.

These are the first attestations of these embracing constructions, and contrary to the simplified history of French negation, most of the time they already occur in contexts that are inconsistent with their original noun meanings. They are partly grammaticized even in Latin, and fully grammaticized emphatic negative particles by the time we get to French. The bulk of the grammaticization takes place off the historical record. This undermines that section of Rostila's (2006) argument, because it was several hundred

17

years (at a minimum) between the time when *ne ... pas* became no longer restricted to verbs of motion and its increase in type frequency in the seventeenth century.

3.3.3 Differences among Sentence Negators in Old French

There are differences in the meaning and usage of the various embracing negation constructions, but they are very subtle, and have frustrated many a grammarian. Price (1997) gives an overview of some of the claimed differences. The strongest difference he finds is that, at least before the seventeenth century, *ne ... point* tended to be used more with the partitive construction, a characteristic feature of French used to express quantity. He found very few examples of *ne ... pas* with the partitive construction in Old French. A superficial investigation of the corpus I use for this study confirms Price's observation, but more detailed counts would be necessary for full confirmation. This is the first unambiguous instance of *ne ... pas* with the partitive construction in my corpus:

```
9)   Je n'y fais pas de grant despesche.
     'I don't make a lot of money here.'
```

Nouveau Pathelin, 1474

Price also found significant regional variation in the use of these negative constructions. While *ne ... pas* predominated in Paris, Normandy and England, *ne ... mie* was more common in Northern and Eastern France, particularly Lorraine and Picardy. In the northern parts of Burgundy that became Belgium, *ne ... nient* was popular. He confirmed this with data from the *Atlas Linguistique de France* collected between 1902 and 1920; in addition, he found that the subjects of the *Atlas* in Picardy used *ne ... point* frequently in non-partitive contexts.

Despite these geographical and usage differences, Price observed that many authors used several sentence negators almost interchangeably in a single text, regardless

of their dialect background. He gives two possible motivations for this. One is for rhyme: since all four negators ended in different sounds, it was easy to substitute one for another to make a better rhyme. Another is to avoid repetition: an author could vary the sound of a text by using different negators. Offord (1967: 332) observes that 'MS [manuscript] variants testify to the interchangeability of the particles,' tallying twelve examples in the *Roman de Berinus* where one manuscript might have *ne ... mie* but another might have *ne ... pas*. He does not mention any similar variation for *ne* alone versus one of the embracing negation constructions.

3.3.4 When the Postverbal Item Doesn't Come After the Verb

So far, I have been describing *pas*, *point* and *mie* as POSTVERBAL particles, but there are some instances where they appear not only before the verb, but before *ne*. This is consistent with a general tendency in Old French to allow syntactic units that normally came after the verb to appear before it, especially objects and these negation particles. This appears to have been done for information structure reasons (Lambrecht 1994), such as to announce a topic or to place the focus on a particular item. It could also be used to make the rhyme and meter fit the author's intention, but it had to be consistent with the information structure constraints as well.

```
10)   Mun défens tu pas ne gardas
      'My prohibition you did not keep'
```

Ordo Representacionis Ade, ca. 1160, line 411

This practice continued through the end of Old (sometimes called Middle) French, but is not found at all after the tightening of word order in Classical French.

3.3.5 Negations that Aren't Really, Part 1: Expletive Negation

Throughout the recorded history of French, there have been constructions that have similar forms to the sentence negators, but have been used for other functions. These are often known as EXPLETIVE NEGATIONS, from the Latin *explere* 'to fill out,' meaning that they were interpreted as empty forms that were necessitated by the syntax of the sentence. The most common uses are in comparatives, as in the following example:

11)   es plus fresche que **n**'est rose
      'you are fresher than a rose'

Ordo Representacionis Ade, ca. 1160

The semantic pathway is fairly clear from 'you are more fresh, and by comparison a rose is not' to 'you are more fresh than a rose is.' This is also found in some varieties of Italian (Stassen 1985). The other common use of expletive negation is in the complement of certain subjunctive constructions:

12)   J'ai paour qu'ele **ne** t'escape
      'I'm afraid it will get away from you'

Jehan Bodel, Jeu de Saint Nicolas, ca. 1200

Croft (2000: 136) interprets this as an example of CRYPTANALYSIS, where the language user interprets the negation as "insufficiently marked" and inserts an overt marking.

In Old French, it was common to use *ne ... pas* and other constructions for expletive negations; in fact, one of the first letters of the *Académie française* described the French language and its destiny with a postverbal construction as follows (Caillet 1857: 467):

13) notre langue, plus parfaite déjà que **pas** une des
    autres vivantes
    'our language, more perfect already than any of the
    other living ones'

<div align="right">*Académie Française, 1637*</div>

However, later in the sixteenth century grammarians established rules requiring only *ne* alone in expletive constructions.

3.3.6 Classical French: a Shift to Embracing Constructions

In the sixteenth and seventeenth centuries, many scholars of French noticed the increase in use of embracing constructions, especially *ne ... pas* and *ne ... point*; Kawaguchi (2005) has compiled a number of citations. One of the earliest was the Englishman John Palsgrave, who published a grammar in 1530. He wrote (Palsgrave 1530: 110):

> For where as they put ne before theyr verbes, so often as they expresse negation, like as we use 'nat' in our tong after our verbes. They put also after theyr verbes pas, poynt or mye, whiche of theym selfe signifye nothyng, but onely be as signes of negation.

The poet and critic François Malherbe was known for 'preferring to consult his ear rather than a grammar' (Malherbe, Chevreau and Ménage 1666, volume 1: 386). In 1606 he published a 225-page critique of the life's work of his older rival Philippe des Portes. Brunot (1891: 466-469) draws attention to eight occasions when des Portes used *ne* alone and Malherbe argued that an embracing construction would be more appropriate. Here is one example:

14)   Apre et sauvage cœur, trop fière volonté
        Dessous une douce, humble, angélique figure.
        Vous **n**'aurez grand honneur de m'avoir surmonté.
        'Harsh and wild heart, and too proud will
        Beneath a sweet, humble, angelic face.
        You will not find great honor in having overcome me.'

Philippe des Portes, Les amours d'Hippolyte, Sonnet XLVI, 1573

Malherbe called the first two lines a "graceless Italianism," and writes this about the last line:

> On ne dit pas: *vous n'aurez grand honneur*, mais: *vous n'aurez pas grand honneur*. Quelle apparence de dire à une fière volonté, qu'elle n'aura pas grand honneur de l'avoir surmonté?[2]

Malherbe (1606 [1862]: 312)

A year later Charles Maupas, who taught French to foreigners in Paris, published a French grammar for non-native learners (Maupas 1607). He tried to describe some of the contexts where *ne* alone was used as opposed to *ne ... pas* or *ne ... point*, such as the following (Maupas 1607: 334):

> Nous les obmettons ordinairement és propos conditionnels, qui expriment la conjonction Latine *Nisi*, avec liaison a un verbe. Exemp. *Si ie ne vous aimais*.[3]

Maupas and other grammarians were generally at a loss to find any single rule that determined whether it was appropriate to use *ne* alone or *ne ... point* instead of

---

[2] People do not say: *vous n'aurez grand honneur*, but *vous n'aurez pas grand honneur*. What a thing for someone to say to a proud will, that she will not find great honor in having overcome them?

[3] We usually omit them [*pas*, *point* and others] in conditional clauses that express the Latin conjunction *nisi*, subordinate to a finite verb. For example, *If I did not love you*.

ne ... pas. Later in the seventeenth century, building on the work of Maupas and

Malherbe, the grammarian Claude Favre de Vaugelas published one of the most

influential grammars of French. He wrote (Vaugelas 1647: 409):

> il est tres-difficile de donner des reigles pour sçauoir quand il faut plustost
> dire *pas*, que *point*, il le faut apprendre de l'Vsage, & se souuenir que *point*
> nie bien plus fortement que *pas*.[4]

It says something about the politics of grammar that many later grammarians and

writers turned the last sentence into a rule, ignoring the rest of what Vaugelas wrote

(Price 1997). Before the sentence above, however, Vaugelas gave a list of the contexts

where he generally found *ne* alone without *pas* or *point*, such as the following:

> On les supprime d'ordinaire auec le verbe *pouuoir*, comme *il ne le peut*
> *faire, il ne pouuoit mieux faire, il ne peut marcher.* Ce n'est pas que l'on ne
> peust dire, *Il ne le peust pas faire, il ne pouuoit pas mieux faire, Il ne peut*
> *pas marcher.* Mais il est incomparablement meilleur & plus elegant sans
> *pas*.[5]

This practice of listing contexts where a non-productive form appears continued

after Vaugelas; some of the contexts listed are quite specific fixed expressions. One of

the most comprehensive lists was compiled by the English grammarian Alfred Ewert

(1943). The following table compares the lists given by Malherbe, Maupas, Vaugelas and

Ewert.

---

[4] It is very difficult to give rules that determine when it is better to
say *pas* than *point*; people need to study the usage of other writers and remember
that *point* negates much more strongly than *pas*.
[5] They are ordinarily omitted with the verb *pouvoir*, as *in he cannot do it, he*
cannot do better, he cannot walk. This does not mean that one cannot say *Il ne le*
peust pas faire, il ne pouuoit pas mieux faire, *Il ne peut pas marcher.* But it is
incomparably better and more elegant without *pas*.

Context	Malherbe (1606)	Maupas (1607)	Vaugelas (1667)	Ewert (1943)
with *craindre* and other verbs of fear, uncertainty and doubt	√	√		
with expressions of prevention and caution		√		
with causative expressions such as *afin que* and *pour que*		√		
presence of other negators such as *ni, jamais* (see next section)	√	√	√	
with *savoir* and *pouvoir*		√	√	√
subordinate to another negative		√		√
exclamatory or final *que*		√		√
conditional *si*		√		√
comparative (see next section)			√	
time expressions such as *ça fait ... que*			√	
with *oser*			√	√
with *cesser*				√
fixed expressions such as *à Dieu ne plaise*, *n'importe*, *n'avoir garde*, *n'avoir cure*				√
exclamatory *qui*				√
condition by inversion (not counted)				√

Table 1. Negative contexts in which preverbal *ne* alone is used, as reported by four grammarians.

3.3.7 Modern French: Widespread Use of *pas* Alone

One of the most striking aspects of modern vernacular French is the use of postverbal *pas* alone to mark sentential negation.

15)   Voyons, c' est **pas** tout ça...
      'Hold on, that's not all...'

Henri Meilhac and Ludovic Halévy, La Cigale, 1877

Ashby (2001) found that it was used 82% of the time in sociolinguistic interviews he conducted in the city of Tours (out of a total of 1891 tokens produced by 29 speakers), with *ne ... pas* used the rest of the time. This has widely been interpreted as a new change in progress, but a number of objections (Valli 1983, Blanche-Benveniste and Jeanjean

1987) have been raised to this idea. One is that there are instances of postverbal *pas, mie, point* and *rien* alone as far back as this example:

16) Te sera-il **pas** bien sauvage
    Garder désormais la maison ?
    'Will it PAS be a real beast for you
    To keep your house in order from now on?'

Farce nouvelle a six personnages, ca. 1425

I have left the *pas* untranslated in this example, because it is not clear *a priori* what its meaning is. A similar sentence with *pas* alone in Modern French would be translated as 'Won't it be a beast ...,' but Eckardt (2007, chapter 5) observes that the earlier 'puzzling' instances of *pas* alone (and other postverbal constructions) were in either questions or comparatives; see the last section in this chapter for further discussion. The only evidence that these might be negations is their superficial resemblance to the *ne ... pas* construction, and they are all consistent with alternative interpretations without negations, such as 'Will it be a bit of a beast ...' for the example. Eckardt makes a strong case that these particles were still being used as 'items of minimal value' to modify the interrogative and comparative constructions, as in English *at all, a bit, a little, something* and *somewhat*.

There is significant historical evidence of postverbal *pas* and *point* being used unambiguously as sentence negators in the childhood speech of Louis XIII as reported by Jean Héroard (Ernst 1985), and even further, in the writings of his great-grandmother Marguerite de Navarre (Brunot 1901) and some of her contemporaries in the early sixteenth century. The earliest dated attestation of postverbal *pas* alone given by Brunot (1901, tome II: 472) comes from Nicolas de Troyes:

25

17) on luy avoit **point** fait faire de service
    'they had not made him perform any service'

Nicolas de Troyes, Le Grand parangon des nouvelles nouvelles, 1536

It should be pointed out that this example is also ambiguous, with an alternative interpretation along the lines of 'they had hardly made him perform any service.'

Jespersen speculates (1917: 19-20) about the course that *ne*-deletion may have taken in French, starting with short responses with no verb (*pas de ça !*, 'none of that!'), and moving through interrogative and impersonal sentences and vowel-initial verbs before diffusing through the entire lexicon. However, Martineau and Mougeon (2003) have conducted a corpus study of the history of postverbal *pas* alone in negations that strongly suggests that frequency plays a greater role than prosody in this area. In their corpus, the phenomenon expanded significantly in the seventeenth century with reductions of high-token-frequency phrases like *ce n'est pas*, "it isn't" to *c'est pas* and *il ne faut pas*, 'you (generic) shouldn't,' to *(il) faut pas*, only then moving to common verbs with the first person subject clitic *je*, and later with the second person subject clitic *tu/t'* and infinitives. The usage-based theories of language change laid out by Bybee and Thompson (1997) and Haiman (1997) and described in the next chapter predict that high-frequency constructions will undergo this kind of reduction.

3.3.8 Unstable Embracing Negations: A standard/vernacular split?

Standard French is unusual in sustaining an embracing construction as its dominant sentence negator for so many centuries. In a typological study of negation, Dahl (1979) lists French as one of eleven languages to use 'double particles,' out of 108 total, and one of five to use an alternate monomorphemic negation strategy as well. It is clear that in some vernacular varieties, postverbal *pas* alone has been the dominant

26

sentence negator for at least a century, which raises the question of how long this has been the case. Given the data described by Martineau and Mougeon (2003), it is possible that the dominant sentence negation construction in vernacular French actually shifted fairly quickly from preverbal *ne* alone to postverbal *pas* alone and *point* alone, with only a short period of time when the embracing negation was used. If this is the case, then the long-term stability of the *ne ... pas* construction in Standard French can be understood as a top-down imposition by the writers, grammarians and politicians who set the norms for this variety in the seventeenth century.

It is useful to compare this with a similar period in the history of negation in English. Mazzon (2004) conducted a study of negation in a historical corpus of English texts. Although she does not separate declarative sentences from other sentences, her results still show a clear pattern:

Period	Preverbal *ne*	*ne ... not*	Postverbal *not*
Old English	92.2 %	07.8 %	00.0 %
13-14th C.	32.9 %	27.7 %	39.4 %
15th C.	04.9 %	04.2 %	90.9 %
16-18th C.	07.7 %	02.4 %	89.9 %

Table 2. Historical prevalence of negation constructions in English.

We see that, as in Old French, grammaticized embracing constructions were used in the earliest Old English texts. In Middle English, however, the postverbal *not* construction became the dominant sentence negator, displacing both preverbal and embracing constructions. After the 18th Century, preverbal *ne* alone and the embracing *ne ... not* negation disappeared from Standard English. It may be that vernacular English and French underwent the same changes, but that the grammarians who established the rules of Standard English settled on a postverbal construction, while those who set the norms for Standard French used an embracing one.

27

3.4 Non-Sentential Negation in French

Constructions do not exist in a vacuum, and Hopper (1991) and Bybee (1995) argue that language users can form multiple layers of abstraction. I have chosen to focus on the four sentential negation constructions in order to keep this project to a manageable size, but it is likely that there is a significant amount of interaction between these constructions and the many other French constructions that have similar forms and functions. In this section I therefore present a number of other negative constructions, as well as some constructions that are superficially similar to negative ones but do not strictly speaking express a negation. In the next section I will speculate about possible interactions among these constructions that can be explored in future studies.

3.4.1 Negative interjections

The simplest form of negation is the negative interjection, a word or phrase that stands alone as a denial of a preceding proposition. The most common, existing throughout the history of French, is *non*, but Old French had another word, *nennil*:

```
18)   Abel: Iert del meillor!
      Chaïm: Nenil, por veir;
      De cel ferai jo pain al seir.
      'Abel: [Your offering of wheat] will be from the best!
      Chaïm: No, definitely not;
      I'm going to use that to make bread tonight.'
```

Ordo Representacionis Ade, ca. 1160

The word *non* was treated as a sentence in itself and could be a clause in a subordinate construction as in the following example:

```
19)   [E] por quei non?
      'And why not?'
```

Ordo Representacionis Ade, ca. 1160

In Old French, writers occasionally use *non pas du tout* as a more categorical or insistent denial of a proposition. In Classical and contemporary French this is often shortened to *pas du tout*; Classical French writers occasionally use *point* or *point du tout*. In Modern French these are sometimes reduced even further, to just *du tout*.

While *nennil* and *du tout* are semantically very similar to *non*, Modern French is unusual in having a separate interjection *si*, just for denying negative assertions.

20) ```
    Perdican: Vous êtes fâchée de cela ? Oh! que non.
    Camille: Si, j'en suis vraiment fâchée pour vous.
    'Perdican: You are angry about that? I think not.
    Camille: Yes, I am truly angry on your behalf.'
    ```

Alfred de Musset, On ne badine pas avec l'amour, 1834

When *non* and its synonyms are used in response to a negative assertion, they indicate agreement with the assertion. Unlike English, if *oui*, 'yes,' is used in this context it also indicates agreement and is thus synonymous with *non*:

21) ```
    Edgard: N'est-ce pas qu' elle est jolie ? ...
    Le Marquis: Oh ! Oui, quant à ça... oh ! Oui...
    'Edgard: Isn't she pretty?
    Le Marquis: Oh! Yes, now that you mention it ... Oh!
    Yes...'
    ```

Henri Meilhac and Ludovic Halévy, La Cigale, 1877

3.4.2 Negating focused constituents

The particle *non* can also be attached to the beginning of smaller constituents such as noun phrases to deny only that constituent, usually to place focus on it and contrast it with an alternative, or to repeat and reinforce a synonymous constituent for poetic effect, as in the following example:

22) ```
    qui m' est encore ostacle / Et non sceu.
    'which is still unclear / and not known.'
    ```

Miracle de la fille du roy d'Hongrie, ca. 1371.

This fourteenth-century example is the earliest attestation in my corpus. However, prior to that, the phrase *non pas* serves a similar function, as early as the *Ordo Representacionis Ade*:

23) De tei eissit, **non pas** de fors.
     'From you she came, not from anywhere else.'

Ordo Representacionis Ade, ca. 1160

Various compound words were created with this use of *non*, such as *nonpruec*, 'notwithstanding,' *nonporquant*, 'nevertheless' and *sinon*, 'if not,' and it also evolved into a prefix, the same one that was borrowed into English and used in words like *non-smoker*, for example in the following sentence:

24) Il ralume les **non**voians
     'He re-illuminates the sightless'

Jehan Bodel, Jeu de Saint Nicolas, ca. 1200

3.4.3 Negative pronominals

Some of the minimizing idioms evolved not into sentence negators (like *pas, point* and *mie*) but into negative pronominals. The Latin noun *persona* split into the French noun *une personne* 'a person,' and *ne ... personne*, 'nobody,' and *ne ... ame*, *ne ... homme* and *ne ... femme* were used similarly in Old French. The noun *res*, 'thing,' became *ne ... rien*, 'nothing.' In Old French, nouns like *giens*, 'person,' *mot*, 'word,' *goutte*, 'drop,' *chose*, 'thing,' and *ne ... nient* 'a nothing,' also took on senses similar to *ne ... rien* when paired with a preverbal *ne*. While *ne ... chose* and *ne ... nient* had an almost identical usage to *ne ... rien*, *ne ... mot* was more restricted, almost always used with the verbs *dire*, 'say,' and *sonner*, 'ring.' The construction *ne ... goutte* was usually used with the verb *voir*, 'see' (Price 1997).

30

Price (1997) gives evidence that *nient*, *goutte* and *grain* were used 'adverbially,' that is to say, not as direct objects or negative quantifiers, but as general sentence negators, and Gregory (1997) shows similar uses for *riens*, these are relatively rare and were not found in the current study; all of the instances of these four negators were either as objects or as quantifiers.

3.4.4 Negative Quantifiers and Spatial Negation

French inherited a negative quantifier from Latin, in the form of *ne ... nul*. In Classical French, *ne ... nul* was overtaken by a new negative quantifier, *ne ... aucun*. In Old French there was also a construction *ne ... el* meaning 'no other,' which was later replaced by *ne ... autre* and eventually *ne ... pas d'autre*. *Ne ... nul* was combined with the word *endreit*, 'place,' to mean 'nowhere.' Later, *endreit* was supplanted by the noun *part*, which had come to take on a spatial meaning (as in *quelque part*, 'somewhere').

```
25)   Nulle part trouver ne pouons
      Femme pour vous
      'Nowhere can we find
      A wife for you.'
```

Miracle de la fille du roi d'Hongrie, ca. 1371

3.4.5 Temporal negators

A number of temporal adverbs combined with *ne* to form temporal negators. *Ja*, 'now, already' became *ne ... ja*, 'not yet,' and eventually disappeared. *Onc* derives from the Latin *unquam* meaning "sometimes," and became *ne ... oncques*, 'never.' The comparatives *plus* and *mais*, both meaning 'more,' became *ne ... plus* and *ne ... mais*, 'not ... anymore.' Many of these could combine, like *ne ... oncques mais* and *ne ... ja plus*, but the most popular combination was *jamais* 'forever,' and with *ne*, it formed *ne ... jamais*, 'never.' (Foulet 1965)

31

3.4.6 Conjoined negatives and Negative Conjunctions

In Old French, *ne* can function as a conjunction expressing a logical joint denial of more than one constituent, equivalent to English *nor*. Usually the verb was negated by *ne* alone, but it was also negated with an embracing construction like *ne ... pas*. The particle *ne* could also be used to conjoin multiple negated sentences, as with Early Modern English *nor*, but that use gradually disappeared, replaced by constructions like *et ... ne ... non plus*, 'and ... neither.' By the time of Classical French, this conjunction was written *ni*, and the seventeenth-century grammarians established a pattern for its usage where all conjoined elements were required to be preceded by *ni*, as in the following example:

```
26)    Je ne connois ici ni qualités ni sang
       'I recognize here neither character nor blood'
```

Pierre Corneille, Clitandre, 1631

3.4.7 Negations that Aren't Really, Part 2

One of the most striking 'negative' constructions in French is not a negative at all, strictly speaking, but a restrictive construction. In Old French the conjunctions *que* and *fors* could be used in a negative sentence to mark an exception to the negation, as Adam said to Eve after eating the fruit:

```
27)    N'i ad conseil que del morir.
       'There is no solution but to die.'
```

Ordo Representacionis Ade, ca. 1160

By the earliest texts, *ne ... que* and *ne ... fors* had already evolved to the sense of English *only*, and didn't require a direct object:

```
28)    De dis ne remaindront que noef.
       'From ten will remain only nine.'
```

Ordo Representacionis Ade, ca. 1160

The Old French word *guère*, 'much,' and its Walloon and Picard cognate *waires* were borrowed from Frankish *waigaro*. Just like English *much*, *guère* combined with *ne* took on the sense of 'not much,' or 'hardly at all.' While *guère* by itself is used, well, hardly at all in Classical or Modern French, *ne ... guère* is common in Classical French and occasionally used in Modern French.

29) Jo fui dedenz, **n**'en soi **gaires** joïr
     'I was inside, but I hardly had a chance to enjoy it.'

Ordo Representacionis Ade, ca. 1160

3.4.8 Multiple Negations

Some of these negators could be used together in the same clause, as in the following example:

30) Or **ne** porront **pas** dire **aucun** que j'ai antés
     Que d'aler à Paris soie pour nient vantés
     'Now, there is no one that I have known who can say
     That I have prided myself on going to Paris for
     nothing'

de la Halle, Li Jus de la feuillie, ca. 1260

For many years this was perfectly acceptable, but as indicated in the chart in the previous section, Maupas (1607) believed that it was no longer considered to be good French. Later in the seventeenth century, other grammarians began to argue against it, and it gradually disappeared from use.

3.4.9 Non-Sentential Negators in Modern French

When people started to omit the *ne* from *ne ... pas* in Modern vernacular French (yielding postverbal *pas* alone), they also left it out of the non-sentential negations. The constituent negator *non pas* was reduced to *pas*, the negative pronominal constructions

33

became simple postverbal *rien* and *personne*, and so on with the other negative and pseudo-negative constructions that still existed. This has caused some potential ambiguity for *que* and *plus*, since the homophonous conjunction and comparative are still widely used. On closer examination there is very little chance of misunderstanding with *que*, since conjunction *que* is always followed by a subordinate clause (as in example 32) while pseudo-negative *que* is always followed by a lower-level constituent (as in example 31), but if a speaker is concerned about ambiguity they will usually avoid reducing the schwa of pseudo-negative *que*.

31)   Jupiter: Elle **n'**aime **que** son mari.
        'Jupiter: She loves only her husband'

Jean Giraudoux, Amphitryon 38, 1929

32)   Elle eût mieux aimé que son mari fût emporté avec elle
        et l'aimât.
        'She would have preferred that her husband be taken
        with her and love her.'

Alexandre Dumas Père, Les Médicis (history), 1844

There is greater potential for ambiguity with *plus*, since the negative *plus* followed by a partitive (example 33) could sound identical to the comparative of quantity (example 34):

33)   il **n'**y a **plus** de Coca Cola en vente au Zimbabwe
        'there is no more Coca Cola being sold in Zimbabwe'

http://spiritoo.over-blog.com/article-2236021.html

34)   Malgré la saison, il y a plus de Coca que de bières de
        Noël !
        'Despite the season, there are more bottles of Coke
        than of Christmas beer!'

http://www.sans-fils.org/article.php3?id_article=8003

When the *ne* is not pronounced, the two constructions sound the same: *il y a plus*. Even though the syntax of the complete sentence resolves the ambiguity, many French speakers now pronounce the final /s/ in the comparative of quantity, so */plu/ de coca* means 'no more Coke,' but */plus/ de coca* means 'more Coke [than ...]'. This phenomenon was reported by Jespersen (1917: 20) and appears to be the result of a conscious effort to avoid leading interlocutors up a garden path, but more study would be necessary to draw firm conclusions.

As I discuss in Chapter 5, speakers of French appear to have formed a higher-level schema covering many forms of negation and some constructions, like this one, that are not strictly negation. This is a promising avenue for further research, but is outside the scope of this study.

In the next chapter I will give some background for the theoretical framework in which I will be studying the evolution of French negation, and in the following chapter I will discuss previous work on this topic.

4 Sentence Negation in Linguistic Theory

In Chapter 3, I gave an overview of the history of negation in French and its relation to the overall history of the French language. This chapter will discuss the many attempts to understand and model the evolution of sentence negation. I will focus on French, but also discuss work on other languages as necessary.

Negation is a very popular topic, to the extent that it has two recently-published full bibliographies, each with several hundred texts (Seifert and Welte 1987; Brütsch, Nussbaumer and Sitta 1990). French negation has attracted a particularly large share of attention, because the language is so well-known and because it illustrates very clearly certain aspects of language evolution.

I will begin the chapter with a discussion of Jespersen's Cycle, and then cover the semantic and pragmatic changes that are relevant to the Cycle. I will split this into two sections: the evolution from referential nouns to emphatic negation, and that from emphatic negation to unmarked predicate negation. Finally, I will discuss two other likely factors: the creation of a higher-level schema for negations in general in French, and top-down pressure from grammarians and literary trendsetters.

4.1 Jespersen's Cycle

The evolution of sentence negation in French from preverbal *ne* alone to postverbal *pas* alone is similar to changes that have been observed in other languages, and in 1917 Otto Jespersen published a comparative study of the evolution of negation in English, French, German and Danish. In English the process has not only gone further than in French, but it has looped back to the beginning: after preverbal *ne* was

supplemented and eventually supplanted by postverbal *not,* a reduction of *ne a whit*, not a thing, *not* eventually fused with the auxiliary *do* to form *don't*. *Don't* is now almost a preverbal particle, itself ready to be supplemented by some other construction. Because this can form an infinite loop, Dahl (1979) gave it the name of JESPERSEN'S CYCLE.[6]

Versions of Jespersen's Cycle have been found in many other languages, including Piedmontese (Parry 1997) and most of the other Romance varieties (Schwegler 1983, Posner 1985); Welsh (Willis 2005) and Arabic and Berber (Lucas 2005). In fact, Kiparsky and Condoravdi (2006) describe four successive iterations of the Cycle in the history of Greek. In English, German and Danish, the *ne* was repeated in the postverbal negation and incorporated into the new negator, unlike in French, Welsh and Piedmontese.

Jespersen also points out that before it became French, Latin underwent an earlier iteration of the cycle. Ancient Latin texts had an earlier form of *ne*, cognate with the Old English and Old Norse ones, ultimately derived from an Indo-European root 'which I take to be ... a primitive interjection of disgust' (Jespersen 1917: 17). The earlier Latin preverbal *ne* was replaced by *ne oenum*, also preverbal, which was then reduced to *non*, the negator that begins the French iteration of the cycle. Posner (1985) observes that many French-based creoles have completed the French iteration of the cycle by returning *pas* to a preverbal position.

---

[6] Meillet (1912 [1946]: 139-141) discusses the cycle in much the same terms as Jespersen five years before the latter published Negation in English and Other Languages, but it may well be that Jespersen described the issue in his 1909 grammar of English and for some reason Meillet did not cite him.

Many linguists (e.g. Schwenter 2006) refer to the three types of constructions based on their order in Jespersen's Cycle, which I have illustrated with the examples from the Prologue:

Construction	Features	French example	Prologue Example Number
NEG1	preverbal and monomorphemic	de me lever **ne suis** forte.	2
NEG2	two morphemes; can be preverbal, postverbal or "embracing" (one morpheme before the verb and one after)	**je ne suis pas** digne de vivre.	3
NEG3	postverbal and monomorphemic	**je suis pas** doué en dessin.	4

Table 3. Jespersen's Cycle with French examples.

Schwenter (2006) gives a four-stage cycle credited to a number of linguists (Schwegler 1988, 1990; Geurts 2000; Ladusaw 1993; Posner 1985), but discusses a 'Stage ¾' where he places present-day colloquial French. I believe that this stage deserves its own number, and I have accordingly renamed Schwenter's Stage 4 as Stage 5, and Stage 3/4 as Stage 4:

Stage	Constructions used
1	Primarily NEG1 with some NEG2 constructions used for emphasis
2	NEG2 constructions are regular and frequent, but NEG1 still dominates
3	NEG2 constructions dominate, but NEG1 still occurs in some relic contexts
4	NEG2 constructions dominate, but are occasionally reduced to NEG3 constructions
5	NEG3 constructions dominate; some relic NEG2 and NEG1 constructions persist

Table 4. Five stages in Jespersen's Cycle.

NEG1 and NEG2 constructions are in fact used in all five stages; what is distinctive about Stages 3-4 is that the construction(s) with two morphemes are the dominant constructions used in sentence negation. In the next few sections I will discuss theories about the semantic and pragmatic evolution of these constructions.

38

4.2 From Nouns to 'Emphatic' Negation

In the chapter on Negation in the History of French, I showed the morphosyntactic evolution of negation constructions in French from idioms containing nouns and adverbs to complex embracing constructions, and ultimately to monomorphemic particles. In the last section I showed that this is an instance of a general pattern, Jespersen's Cycle, that has been found in many languages. In this section and the next I will discuss the semantic/pragmatic side of this evolution.

The semantics and pragmatics of these constructions clearly change during the course of the cycle, from referential noun phrases to predicate negation. However, the semantics of noun phrases and negations are far enough apart that we would expect to find some intermediate stages. Traugott (1989) and others argue that semantic change typically proceeds gradually through polysemy and ambiguity, driven by the conventionalization of pragmatic inferences. She identifies three tendencies of semantic change based on her investigation of English modal auxiliaries (1989: 34-35):

> Tendency I: Meanings based in the external described situation > meanings based in the internal (evaluative/perceptual/cognitive) described situation.
>
> Tendency II: Meanings based in the external or internal described situation > meanings based in the textual and metalinguistic situation.
>
> Tendency III: Meanings tend to become increasingly based in the speaker's subjective belief state/attitude toward the proposition.

In the specific area of negation, building on Horn (1989), Geurts (1998: 275) sets out a semantic classification of sentence negation, where most sentence negations are either simple descriptive negations or else fall into one of four categories of denial, illustrated by constructed English examples as follows:

PROPOSITION DENIALS serve to simply reject an earlier utterance:

35)   A: The cook is guilty.
       B: The cook is **not** guilty.

Geurts 1998, example 4

PRESUPPOSITION DENIALS aim at the presuppositions of the prevous utterance:

36)   Kurt **DOESN'T** realize that his camels have been
       kidnapped, because they HAVEN'T been kidnapped.

Geurts 1998, example 1b

IMPLICATURE DENIALS remove or block a scalar implicature:

37)   The room **wasn't** WARM: it was SWELTERING.

Geurts 1998, example 2b

FORM DENIALS are concerned with the form, style, or register of an utterance:

38)   I **didn't** manage to trap two monGEESE – I managed to
       trap two monGOOSES.

Horn 1989: 371

Geurts acknowledges (Geurts 1998: 276) that "in practice the line between descriptive negation and denial may be hard to draw." Descriptive negation can be illustrated by Geurts' example 1a, reproduced below, but he makes the point that virtually any descriptive negation can be read as a form of denial if the context permits such a reading.

39)   Kurt **doesn't** realize that his camels have been
       kidnapped.

Geurts 1998, example 1a

Keeping Traugott's three Tendencies and Geurts' five categories in mind, I will discuss the recent contributions of four authors to our understanding of the semantic and pragmatic changes that affect negation constructions. These theories are all compatible,

and each contributes a different piece to our understanding of the evolution of negation. Eckardt (2007: chapter 5) draws on data from French and German, Kiparsky and Condoravdi (2006) from Greek, and Schwenter (2006) from Catalan, Italian and Portuguese, to clarify particular details of this evolution. I will be using their insights to illuminate the path taken by French *ne ... pas*, which is also the focus of Eckardt's chapter.

4.2.1 Stage 0: Literal Reference

What could be the starting point along this path? Kiparsky and Condoravdi (2006) use the following constructed English example to demonstrate one use of this type of construction:

40)   I **haven't** eaten the porridge.

Kiparsky and Condoravdi (2006): 8, example 7.

This sentence is ambiguous between a telic reading ("I haven't eaten any of the porridge") and an atelic reading ("I haven't eaten the entire porridge.") Kiparsky and Condoravdi point out, "Adverbial emphatic negation disambiguates the sentence in favor of the telic reading."

41)   I **haven't** eaten the porridge at all.

Kiparsky and Condoravdi (2006): 8, example 8.

As Eckardt (2007: 249) writes, this use of the construction is TRANSPARENT in that it conforms to the literal meaning of its component words, and is thus likely to be the original construction in this *cline*. It is even more transparent if we modify Kiparsky and Condoravdi's example to use a less opaque minimizing construction:

41

42)　*I **didn't** eat one bite of that porridge.*

4.2.2 Stage 1: Scalar Implicature Denial

Now let us re-examine this Latin expression from Plautus:

43)　**Non** ego **nunc** emam vitam tuam vitiosa **nuce.**
　　'I wouldn't buy your life for a rotten nut.'

Plautus, Miles Gloriosus II, 3, 45

Note that the literal connection is gone: while people are documented to have bought other people's lives throughout history, even in barter economies we are unlikely to find anyone buying anyone else's life with a rotten nut. This is thus hyperbole, an exaggeration to indicate the intensity of the speaker's feeling, in accordance with Traugott's Tendency I.

The noun phrase *one bite* in the previous example, and *nux vitiosa* in this one, are a specific class of noun phrase: an ITEM OF MINIMAL VALUE. Recall that Schweighäuser listed a number of these for Latin, mostly money and food, but also various materials that could be given, sold or bartered. These items of minimal value form the largest subset of NEGATIVE POLARITY ITEMS, a well-known class of entities that are used primarily in negative and interrogative contexts (Jespersen 1917, Ladusaw 1979, van der Wouden 1995). This fits in Geurts' (1998) category of implicature denial described above.

Following Fauconnier (1975), Geurts (1998) and Israel (2001) argue that constructions convey implicature denial by invoking SCALAR MODELS in the mind of the language user, setting up an equivalence between some well-known scale of values and the immediate context of their use. In our example from Plautus, the rotten nut is an item of minimal value in barter. Its use in this quote actually doesn't map the scale of value of life to the scale of value of barter items, but rather invokes the mapping only to *deny* it,

42

thus implying that the value of the life in question is really zero. Constructions like the porridge example, being ambiguous between the literal reference and the minimal value reading, form the bridge between the two interpretations.

4.2.3 Stage 2: What Is Emphatic Negation?

Schwenter (2006) points out that many discussions of Jespersen's Cycle posit a stage where embracing constructions are used for emphatic negation, but are very vague on exactly what 'emphasis' means in the context of negation. He draws an important distinction between 'the possibilities for emphasis (in intuitive terms) that constitute the precursors or input to the Cycle proper, and the emphatic post-verbal elements that become incorporated as GRAMMATICAL elements into the Cycle' (Schwenter 2006: 430). This is essentially a refinement of Palsgrave's (1530: 406) interpretation:

> So that pas, poynt, or mye be used for a more clere expressying of
> negacion, and as though the speker wolde byde by the thing hiche be
> denyeth: in so moche that, if the speker do but fayntly denye a thyng, they
> use than to leave out pas, poynt, or mye...

In intuitive terms - and in the terms of the Oxford English Dictionary (1989) - EMPHASIS includes the following senses: special or important significance, vigor or intensity of expression, intensity or force of feeling, and stress of voice. This fits with the example of the rotten nut, as well as that of the bite of porridge; the shift to subjective belief states is consistent with Traugott's Tendency III. They are both ambiguous between the minimal value reading and the emphatic reading, and thus constructions of these types form the bridge between the two readings.

For a more grammatical sense of EMPHASIS, Schwenter turns to Israel (2001), who argues (following van der Wouden 1994 and Geurts 1998) that by taking a position on

the scalar model in contrast to the scalar norm, these negations count as highly informative, which he further describes as 'saying something which goes beyond some default explanation.' Israel continues, 'the informativity of a proposition determines its rhetorical force: highly informative propositions as rhetorically EMPHATIC' (Israel 2001: 307; emphasis in original, so to speak). In other words, these negatives were called 'emphatic' because they were highly informative.

Eckardt (2007: 235) takes a similar tactic to Israel (2001) and Schwenter (2006), but relates emphasis to focus. Following Rooth's (1992) theory of 'alternative semantics' and Krifka's (1995) discussion of the adverb *even*, she argues that a focused negative presupposes that its referent is 'more surprising than any of its alternatives.' She writes, 'The next step for *mie, goutte, pas* and *point* must have been that their potential to single out minimal (or subminimal) events, and the effects that arise under emphatic focus, were taken as the basis of a use in a new sense. Their new semantic contribution was that of an adverbial modifier "to the least degree," "in a minimal degree".'

Much of Eckardt's discussion is based on her analysis (2007: section 5.2) of the 'puzzling' premodern instances of constructions with postverbal *pas, point*, *mie* and *goutte*, as well as *rien* and *personne*. Here are examples of these 'puzzling' instances for *pas*, *point* and *mie*:

44) Te sera-il **pas** bien sauvage
     Garder désormais la maison ?
     'Will it PAS be a real beast for you
     To keep your house in order from now on?'

                                        Farce nouvelle a six personnages, ca. 1425

45) E sout bien que li reis en sereit mult blasmez, Se
    Thomas l' arcevesque i fust **point** mesmenez
    'And he knew well that it would be a big shame for the
    king if Archbishop Thomas were POINT harmed'

   Eckardt (2007), Example 5.16 (Anonymous, La Vie de saint Thomas, line 4730)

46) Por vos sui en prison misse ..., Mais ... longement
    n'i serai prise, se jel puis **mie**, ...
    'For you I am put in prison, but not for long will I
    be kept there, if I can MIE'

   Eckardt (2007), Example 5.14 (Anonymous, Aucassin et Nicolette 5, 25)

Eckardt observes that these 'puzzling' instances are in the same contexts that are generally associated with negative polarity items, prompting her to analyze them as such in both canonical and 'puzzling' contexts. She also notes (Eckardt 2007: 224) that 'puzzling' postverbal *pas* occurs in a somewhat different set of contexts from the others, being largely restricted to rhetorical questions. She points out (Eckardt 2007: 252-260) that this is consistent with the behavior of STRONG NEGATIVE POLARITY INDICATORS (Zwarts 1986), a set of constructions that express 'subminimal events', and suggests that this may explain how *ne ... pas* ultimately became the primary sentence negator in French.

Interestingly, Eckardt (2007: 257-260) observes that in her data, *point* 'displays the most varied sample of "puzzling" uses.' On this basis, she suggests that by the time of Classical French, *ne ... point* may have been the last grammaticized negator to lose its 'emphatic' status, consistent with the findings of Catalani (2001). As I will discuss in the Results chapter, this is confirmed in the present study.

Eckardt's interpretation suggests that the 'puzzling' uses of these negative particles evolved not from the nouns directly but from the two-morpheme 'emphatic negation'

constructions. This is consistent with the data from the corpus collected for this study, where the first 'puzzling' uses are found well after the appearance of 'emphatic negation' uses.

4.2.4 Stage 3: Proposition Denial

For the next stage, Kiparsky and Condoravdi (2006) state that another 'function of emphatic negation is to mark contradiction of a (possibly implicit) assertion,' in accordance with Traugott's Tendency II, but this is a relatively minor function. In Geurts' (1998) taxonomy this is called PRESUPPOSITION DENIAL, which fits within Horn's (1989) category of 'metalinguistic negation.' Here is an example from the corpus used in this study, where Adam directly contradicts Eve:

```
47)   EVA: Il te ferra changier saveir.
      ADAM: Nel fera pas, car nel crerai.
      'EVA: It will change the way you know things.
      ADAM: No it won't, because I won't believe it.'
```

Ordo Representacionis Ade, ca. 1160, lines 285-286.

In this case, the proposition 'it will change the way you know things' that is denied by Adam is articulated explicitly by Eve on the preceding line. Adam refers to that proposition with a pronoun (the clitic *-l* in *nel*) and a pro-verb *faire*, 'to do.' The denied proposition is not presupposed, but directly articulated.

4.2.5 Stage 4: Presupposition Denial

To understand the uses of 'emphatic' negation, Schwenter (2006) examines use of preverbal, embracing and postverbal constructions in Catalan, standard Italian and Brazilian Portuguese, such as the following (slightly modified) constructed examples:

48)   [speaker B sees interlocutor A putting on a heavy
      coat]
      Avui **no** fa **(pas)** fred.

      'Today it's not cold.'

<div align="right">*Schwenter, Catalan example 2b*</div>

49)   A: Chi viene a prenderti, Gianni?
      B: Non so. Ma Gianni **non** ha **(mica)** la macchina.

      'A: Who's coming to pick you up, Gianni?
      B: I don't know. But Gianni doesn't have the car.'

<div align="right">*Schwenter, Italian example 6b*</div>

50)   A: Você viu esse programa?
      B: **Não** vi (**não**).

      'A: Did you watch that program?'
      B: I didn't watch it.'

<div align="right">*Schwenter, Portuguese example 9*</div>

What he finds, in Geurts' (1998) terms, is not proposition denial or scalar implicature denial, but PRESUPPOSITION DENIAL. 'The difference between proposition and presupposition denials is that the former are directed at the asserted content of the previous utterance, while the latter aim at its presuppositions' (Geurts 1998: 276).

Based on native speaker reactions to the examples above, Schwenter reports that while the preverbal version of each sentence does not require the denied proposition to be active in discourse, the embracing versions are only felicitous where they are used to deny propositions that are already active in discourse. In order to be active this way, the propositions can be implicit and do not have to be overtly stated, believed or accepted by anyone. The embracing constructions can be used to agree with a proposition that has previously been negated, or even implied: Schwenter reports that the Catalan example is felicitous if the other person has said nothing, but simply put on a heavy coat.

Schwenter's study aimed to cast light on the question of French negation through examination of data from other languages (Schwenter 2006: 328). One of the goals of this study will be to test Schwenter's theory against a corpus and measure the extent to which it fits the French situation. Here is an example from Old French that fits the definition of presupposition denial:

```
51)   FIGURA: Or me mostre ton frere vif?
      [...]
      CHAÏM: [E] jo por quei le dei trover?
      Ja nel deveie pas garder.
      'FIGURA: Now show me your brother alive.
      [...]
      CAIN: And why is it me that has to find him?
      I wasn't required to watch him.'
```

Ordo Representacionis Ade, ca. 1160, lines 726, 729, 730

This is the famous line translated in the King James version of the Bible as 'Am I my brother's keeper?' God has made no explicit statement that Cain is required to watch Abel. He asks the question because he knows that Cain has killed his brother, but Cain tries to interpret God's question as presupposing that he had a responsibility to watch Abel. He denies this presupposition, but it is no use since God replies that he knows Cain killed his brother.

Mosegaard Hansen and Visconti (2009) investigated a corpus consisting of five Old French texts (three epic poems, one prose epic and one hagiography) written between 1090 and 1309. They only looked at negations with *ne ... pas* and *ne ... mie*, and they found that all of the tokens in the texts could be characterized with one of the following four pragmatic descriptions: denial of a part of the preceding text, repetition or paraphrase, denial of a presupposition and denial of an inference. They found nothing that was unambiguously predicate negation.

Kiparsky and Condoravdi also observe (2006: 20) that in all the negations they have studied, 'negation is strengthened *only* by combining a simple negative with an indefinite [noun phrase]. A simple negative, or a simple indefinite, never becomes an emphatic negative on its own.' However, Schwenter identifies the Portuguese construction *não ... não* as an embracing construction, which does not fit Kiparsky and Condoravdi's generalization, and there are other examples of adverbials like English *totally not* and French *ne ... pas du tout*. A weaker generalization that is still valid might be that negation is strengthened only by combining a monomorphemic negator with a negative polarity item, and negative polarity items are only formed from indefinites, adverbials and other negators.

4.3 From 'Emphatic' Negation to Predicate Negation

I would describe the primary use of embracing *ne ... pas* in present-day Standard French as PREDICATE NEGATION, the simple statement of a negative predicate without any presupposition required. One of the clearest indications of this is its free use in subordinate clauses:

```
52)   puisqu'ils ne sont pas venus, je m'en vay chez le
      greffier
      'Since they're not here yet, I'm going to the clerk's
      office'
```

Désidério Descombes, La Farce des bossus, 1623

Using Geurts' (1998) terminology, it can be distinguished from various forms of DENIAL (proposition, presupposition, implicature and form). As we will see in the corpus, before the reanalysis, *ne ... pas* could not easily express simple predicate negation, but was almost always restricted to presupposition denial. In the last section we saw how that

sense may have evolved over several steps from the noun sense of *pas*, and in this section we turn to the evolution from presupposition denial into predicate negation.

4.3.1 Jespersen's Strengthening Hypothesis

While Jespersen's observations and his generalizations about the cycle of negation are widely accepted, he also hypothesized a motivation for the transition from monomorphemic negation (NEG1) to bimorphemic negation (NEG2) (1917: 4): 'the original negative adverb is first weakened, then found insufficient and therefore strengthened.' He describes two competing motivations for the placement of negating morphemes: putting them earlier in the clause alerts the listener that the sentence is being negated, while putting it later allows the speaker to place contrastive stress on it. This is particularly important for French, because in the Old French period the prosodic system of the language shifted from the old Latin variable stress to an invariant clause-final stress, eliminating the ability to place contrastive stress on a constituent; the only way to focus a particular item was to place it at the end of a clause (Wenk and Wioland 1982). In addition, Jespersen describes a related phonetic process he calls prosiopesis, where initial syllables are devocalized, making it much harder for the hearer to notice critical information.

There is plenty of evidence of the weakening of Latin *non* to Old French *ne*, and adding phonetic material to it necessarily strengthens it, but Jespersen does not give any evidence for anyone finding *ne* insufficient. Because the postverbal morphemes had the effect of strengthening a weakened preverbal negator, he seems to assume that this is the goal of speakers in using embracing constructions more frequently. Kiparsky and Condoravdi (2006: 4) are among many skeptics of this hypothesis:

The role of phonetic weakening in this hypothetical scenario, however plausible it might seem, is not backed up by any data as far as we know. Our analysis of Greek turned up no support for Jespersen's assumption that phonological weakening triggers the strengthening of negation. There are also some general reasons to doubt it. For one thing, phonetic weakening is too general a phenomenon to explain the specific properties of this unusual pattern of change.

Posner (1985) points out that the cycle has also progressed to varying degrees in many other Romance varieties, particularly in Southeastern France, Switzerland and northern Italy, languages that had very similar structures to French in the Middle Ages, but many of which did not undergo the same prosodic changes as French.

4.3.2 The Rhetorical Inflation Hypothesis

Detges and Waltereit (2002) and Kiparsky and Condoravdi (2006) propose theories that involve an impulse to RHETORICAL INFLATION and subsequent DEVALUATION, following Dahl (2001), who also makes explicit reference to the case of French negation (Dahl 2001: 3). Here is Kiparsky and Condoravdi's explanation (2006: 5):

> Emphatic negation tends to increase in frequency due to pragmatically motivated overuse which is characteristic of inherently bounded evaluative scales. This rise in frequency at the expense of plain negation has an "inflationary" effect, well attested also in politeness systems, hypocoristics, pejoratives, and scalar adjectives of all kinds (Dahl 2001). Uncontroversially, an obligatory element cannot be emphatic, for to emphasizing [sic] everything is to emphasize nothing. Therefore, when emphatic negation rises in frequency to the point where it approaches obligatoriness, it necessarily weakens to regular negation.

This hypothesis is plausible, but as discussed in the next chapter, Croft (2007) argues that it is not necessary to explain the change. He shows that in the normal course of language use, a sufficient number of ambiguous instances of the construction will occur to enable metanalysis, as discussed below.

51

4.3.3 Loss of Emphasis through Metanalysis

Eckardt (2007: 257) has this to say about the final stage in Jespersen's Cycle:

'Loss of emphasis' in the present analysis is tantamount to the following:

- speaker and hearer no longer understand alternative propositions to be salient, and
- speaker and hearer no longer understand that the asserted proposition is the most surprising one among alternatives.

This is consistent with Traugott's Tendency III. However, Eckardt does not indicate what motivates this loss of emphasis. Croft (2000: 130) ascribes it to METANALYSIS, a kind of form-function reanalysis, which I will discuss further in the next chapter. Metanalysis is the simultaneous occurrence of hyperanalysis and hypoanalysis, so that 'the listener swaps contextual and inherent semantic values of a syntactic unit.' In the case of French negation, Croft writes, 'the process by which the emphatic forms become obligatory negative markers is caused by metanalysis. … There is a high degree of correlation between negation and emphasis. … There is a swapping of the two functions: the negative function is attributed to the emphatic element, while the emphatic function is attributed to the nonlinguistic context.'

Eckardt (2007: 257) gives a good overview of the post-reanalysis 'relic' nature of the distribution of these forms:

The failure to understand the respective constructions as emphatic (both in perception and production) will have the following further long-term effects:

- hence, items cease to be subject to pragmatically motivated licensing conditions (because pragmatic side messages cease to be understood)
- the fact that these items only occur in restricted contexts, however, remains part of the collective linguistic knowledge.

52

- the restrictions are no longer contentful but follow simpler surface patterns. Specifically, negative particles in French were used in combination with negation.

Eckardt's prediction that the distribution is no longer contentful (i.e. does not reflect a functional distinction) after the reanalysis is borne out by the descriptive and proscriptive grammars that begin to appear at about the same time as the reanalysis - and in fact, all grammars since then. As I described in the preceding chapter, they are unable to give any principled sense of the distinction between *ne* alone and the embracing constructions; instead they simply list contexts where *ne* alone can be used. Descriptions of the distinction between *ne ... pas* and *ne ... point* are infamous for their vagueness and lack of clarity; some of the earlier ones - and reactions to them - are summarized by Brunot (1891: 466-469).

4.3.4 Summary of the Semantic Evolution of French Negation

Synthesizing the work of Croft (2000), Eckardt (2007), Detges and Waltereit (2002), Geurts (1998), Israel (2001), Kiparsky and Condoravdi (2006) Schwenter (2006), and others, we arrive at this picture of the evolution of French negation:

Semantic Stage	Bridge Construction	Example
0. Literal reference to a thing	A. Transitive sentence	I ate one bite of the porridge
	B. Negative transitive sentence	I didn't eat one bite of that porridge.
1. Scalar denial		
	C. Hyperbolic negative transitive sentence	Non ego nunc emam vitam tuam vitiosa nuce.
2. Emphatic denial		
	D. Common negative idiom	ne punctum quidem temporis ... oppugnatio respiravit
3. Proposition denial		
	E. Grammaticized NEG2 construction	Nel fera pas, car nel crerai.
4. Presupposition denial		Avui no fa pas fred!
5. Predicate negation		puisqu'ils ne sont pas venus, je m'en vay chez le greffier
	F. Reduced NEG3 construction	Je suis pas doué en dessin.

Table 5. Stages in the semantic/pragmatic evolution of embracing constructions

Note that the first column (semantic content) and the second column (syntactic construction) do not match up exactly. This overlap represents BRIDGE CONSTRUCTIONS that are ambiguous between the two semantic readings, in accord with Traugott's (1989) theory as described in the previous section. One of the hypotheses that I will test in this study is whether the corpus shows evidence for a semantic evolution along these lines.

4.4 General Schemas of French Negation

Detailed studies have yet to be done, but there are is evidence suggesting that the evolution of constructions like *ne ... pas* may have been influenced by higher-level schemas that French speakers formed across many kinds of negation. From a formalist perspective, Kroch (1989) argues that the roughly simultaneous decline of both verb-second structures and pro-drop in French were manifestations of a single underlying

change. From a usage-based perspective, as I mentioned in the previous section, Hopper (1991) and Bybee (1995) model the interactions between closely related constructions by postulating that language users can form 'layered' schemas at multiple levels of abstraction, and that these schemas can undergo the same changes as any other schema.

Examining the evolution of all of the French negation constructions, not just those used in sentences, indicates that there was a general trend from the 'negative concord' pattern in Latin (resembling that of present-day Russian) towards a pattern where all negations (and some quasi-negations) were marked by preverbal *ne* alone in combination with another morpheme (*pas, point, rien, que, plus ...*) and now, in present-day French, where these morphemes signify negation by themselves. Confirming this would require conducting similarly detailed studies for these other forms of negation, but it seems clear that there is a general tendency in this direction.

Function	Latin	Old French	Middle French	Classical French	Modern French	Advanced French
Proposition denial	non	nennil	non	non	non	non
Predicate negation	non	nen	ne	ne ... pas	pas	
Presupposition denial	nequequam	ne ... mie	ne ... pas	ne ... pas du tout	pas du tout	du tout
No (partitive)	nullus	ne ... nul	ne ... point de	ne ... pas de	pas de	
Nothing	nihil	ne ... nient	ne ... chose	ne ... rien	rien	rien du tout
Nobody	nemo	ne ... ame	ne ... homme	ne ... personne	personne	
Only	solus	ne ... fors	ne ... que	ne ... que	que	
No more	nec magis	ne ... mais	ne ... mais	ne ... plus	plus	
Hardly	vix	ne ... gaires	ne ... gaires	ne ... guère	guère	presque pas
Never	nunquam	ne ... oncques	ne ... oncques	ne ... jamais	jamais	
Not right now	nec iam	ne ... ja	or ne	ne ... pas maintenant	pas maintenant	
Neither	nec ... nec	ne ... ne ... ne	ne ... ne ... ne	ne ... ni ... ni	ni ... ni	
Nowhere	nusquam	ne ... nul endreit	ne ... nul endroit	ne ... nulle part	nulle part	
No (quantifier)	nullus	ne ... nul	ne ... nul	ne ... aucun	aucun	
No other	nec alter	ne ... el	ne ... autre	ne ... point d'autre	pas d'autre	
Emphatic neither	neque ... neque	ne ... mie ... ne	ne ... pas ... ni	ne ... ni ... non plus	ni ... non plus	
No word	nullum verbum	ne ... mot	ne ... mot	ne ... pas un mot	pas un mot	
Nothing (seen)		ne ... goutte	ne ... goutte	ne ... rien	rien	

Table 6. The evolution of several negation constructions in French.

5 Theories of language change

Broadly speaking, I will be analyzing the syntactic history of negation in French from a usage-based perspective, paying particular attention to pragmatic factors and frequency effects. Since Jespersen's Cycle is widely agreed to be an example of grammaticization and reanalysis, I will discuss how these phenomena have been thought of in recent work.

I will then discuss some of the models that have been applied to the propagation of syntactic changes, including analogy, extension and lexical diffusion. I will then discuss the usage-based model of grammatical change developed by Bybee (1995) and others, that identifies the effects of type and token frequency on language change. Finally, I will discuss quantitative applications of this model, particularly the logistic model developed by Altmann et al. (1983) and Kroch (1989), and propose an extension to this model that takes into account both type and token frequency effects and the competition among equivalent constructions. Many of the linguists cited here use the evolution of French negation in their examples; their claims about this phenomenon will be discussed in the next chapter.

5.1 Grammaticization and Reanalysis

As mentioned in the Introduction, the evolution of French negation has often been held up as a defining example of GRAMMATICIZATION (sometimes known as 'grammaticalization'). This process was identified as early as 1825 by Wilhelm von Humboldt, and first named by Antoine Meillet in 1912; more detail about the history of the concept can be found in Lehmann (1995) and Heine (2003).

57

Meillet's original concept of *grammaticalisation* referred to the 'attribution of grammatical characteristics to a previously autonomous word' (my translation). He gave as his prime example the French word *suis*, 'am,' as in the following examples:

53)   Je suis celui que je suis.
      'I am what I am.'

<div align="right">*Meillet (1912 [1982]: 131)*</div>

54)   Je me suis promené.
      'I went for a walk.'

<div align="right">*Meillet (1912 [1982]: 131)*</div>

In Meillet's description, the word *suis* is at its most *autonome* in the first example, and at its most *grammatical* in the second. My English translation illustrates this: in the first example, the meaning of the French word *suis* and that of the English word *am* are very close. In the second example, *suis* has taken on a function as the auxiliary for compound past tense conjugations of reflexive verbs, a function that *am* does not have in English, so there is no easy way to translate that sentence using *am*. Meillet also discusses the development of other tense markers in French, including the evolution of *avoir* from a verb meaning 'have' into the compound past tense marker and the change of *aller* from a verb meaning 'go' into the compound future marker. Among Meillet's other examples was the subject of the present study: the change of *pas* from a noun to a marker of negation (Meillet 1912 [1982]: 140). All of these examples contain words that have a descriptive or referential function in the earlier stage, but at the later stage have a function that is purely grammatical, such as a tense marker, a negator or a preposition. Although Meillet discussed a continuum from 'most autonomous' to 'most grammatical,' with at

least two steps in between the examples I have quoted, he focused on the fact that the word changed status somewhere along the way.

Interest in the concept of grammaticization was revived by Givón (1979) and has since been further elaborated by a group of *cognitive* and *functional* linguists including Li and Thompson (1976); Lehmann (1982 [2002b]); Bybee (1985, 1995); Langacker (1988); Traugott (1989); Hopper (1991); and Heine, Claudi and Hünnemeyer (1991). Many of these authors investigated Meillet's observation that grammaticization is often accompanied by phonological *reduction* and semantic simplification (*bleaching*). They also placed more emphasis on the continuum of change, and treated grammaticization as a process that moves items further in the 'more grammaticalized' direction (Lehmann 1982 [2002b]: 9). However, others (e.g. Campbell 2001) have observed that reduction and bleaching can be found when grammaticization is not present, and that constructions can become more grammatical without reduction or bleaching. One of the most contentious points of debate is the notion of grammaticization as a UNIDIRECTIONAL process, in that it does not involve grammatical morphemes becoming more lexical (Haspelmath 1999). Some linguists have pointed to examples of grammatical morphemes becoming lexical items (see again Campbell 2001), but others have argued that these are not instances of grammaticization in reverse, but a wholly different process (Haspelmath 2004).

Another concept that has been used to model this change is REANALYSIS, defined by Langacker (1977: 58) as a 'change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation.' A significant debate in the study of language change is the relationship

59

between *grammaticization* and reanalysis; a summary can be found in Croft's *Typology and Universals*, second edition (2003: 253-272). The extreme positions in this debate were taken by Haspelmath (1998: 315) who claimed that 'the large majority of syntactic changes are instances of "pure" grammaticalization and should be explained within the framework of grammaticalization, without reference to reanalysis,' and Campbell (2001: 158) who responded that the notion of grammaticization is 'derivative and therefore lacking a status of its own.' In between are linguists such as Hopper and Traugott (2003: 59) who 'regard grammaticalization as a subset of the changes involved in reanalysis.'

Croft (2003: 268) notes that 'it appears that the role of syntactic reanalysis in grammaticalization depends on one's theory of syntactic representation more than on grammaticalization itself.' In this vein, McDaniels (2003) argues that the standard notion of reanalysis presupposes a level of syntactic structure separate from surface manifestation, a model that many grammaticization theorists are uncomfortable with. In fact, McDaniels and Croft themselves, as well as Detges and Waltereit (2002), offer explanations of syntactic change that rely on pragmatic intentions and interpretations instead of assuming an underlying structure.

Croft (2000) offers a similar model of grammatical change, with a twist, in the context of a broader theory of language change based on Hull's (1988) generalizations of the genetic principles adapted by Dawkins (1976). In Chapter 5 of that book he focuses on the notion of FORM-FUNCTION REANALYSIS, which (unlike 'structural reanalysis') presumes no underlying syntactic structure but instead relies on the interactions of various constructions that the language users rely on for constructing their utterances and interpreting each others' utterances. I will use this model, treating grammaticization as a

60

type of form-function reanalysis that often leads to extension, as discussed in the next section.

Croft gives four different types of form-function reanalysis. In HYPERANALYSIS (Croft 2000: 122), 'the listener reanalyzes an inherent semantic/functional property of a syntactic unit as a contextual property (usually a property of another syntactic unit of the construction). In the reanalysis … the syntactic unit in question loses some of its meaning or function.' In Chapter 3 I described how the postverbal *pas* alone construction emerged as a sentence negator between the sixteenth and nineteenth centuries, and how postverbal *rien* and *personne* took on the negative quantifier functions previously associated with the embracing constructions *ne ... rien* and *ne ... personne*. Croft identifies all three of these as hyperanalysis: 'the inherent negative value of the preverbal particle is attributed to the postverbal element alone, since its (newly acquired) negative value overlaps with that of the preverbal particle; and the preverbal particle is dropped' (Croft 2000: 130).

We can illustrate this by looking back at example (52), which shows predicate negation in a subordinate clause:

55) `puisqu'ils `**`ne`**` sont `**`pas`**` venus, je m'en vay chez le`
     `greffier`
     `'Since they're not here yet, I'm going to the clerk's`
     `office'`

Désidério Descombes, La Farce des bossus, 1623

The following diagram illustrates Croft's explanation applied to one of the examples from Chapter 4.

61
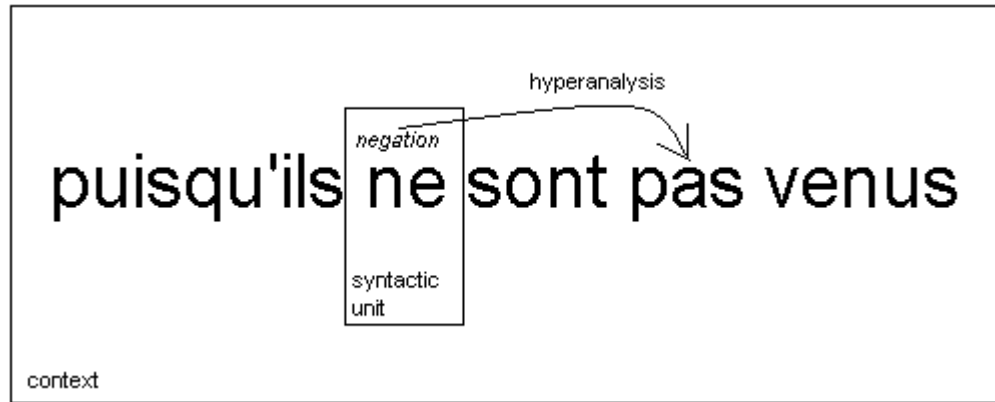
Figure 1. Hyperanalysis of negation from the syntactic unit to the context.

In HYPOANALYSIS (Croft 2000: 126), 'the listener reanalyzes a contextual semantic/functional property as an inherent property of the syntactic unit … and so the syntactic unit gains a new meaning or function.' For examples, he gives the reinterpretation of umlaut as a plural marker in Yiddish and the reanalysis of -*s* as a marker of 'specific or singular events' in Somerset and Dorset English (Croft 2000: 129). It can also be used to explain the change of *est-ce que* from a sequence of words frequently used in yes-no questions to a grammatical marker of yes-know questions, as in the following example:

56)   Est–ce que l'argument
        De ceste fable encore n'avez sceu?
        'Is it that you have not yet understood
        The moral of this fable?'

Etienne Jodelle, L'Eugène, 1552
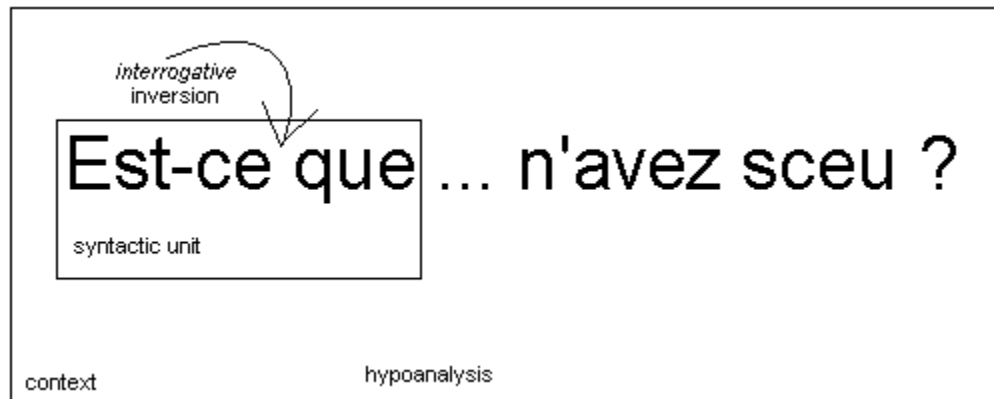
Figure 2. Hypoanalysis of interrogation from the context to the syntactic unit.

This example is a little complicated, in that the interrogative context is marked by the inversion of *est-ce*. In subsequent productions, the phrase is used in contexts where it is incompatible with the reading 'is it that':

57)   Est-ce que l'on consulte au moment qu'on s'enflamme?
      'Do we seek advice at the moment when our passion
      ignites?'

Molière, Psyché, 1671

In metanalysis (Croft 2000: 130), 'the listener swaps contextual and inherent semantic values of a syntactic unit. It appears that these two events occur simultaneously; that is, there is no stage in which hypoanalysis has occurred but not yet hyperanalysis or vice versa.' Croft's metanalysis 'appears to be the mechanism for innovation of the sort of grammatical changes that have gone under the name of invited or pragmatic inference' (Croft 2000: 130). He argues that it is the same process as one called 'context-induced reinterpretation' by Heine, Claudi and Hünnemeyer (1991). Croft specifically refers to the reanalysis of *ne ... pas* as an example of metanalysis:

63
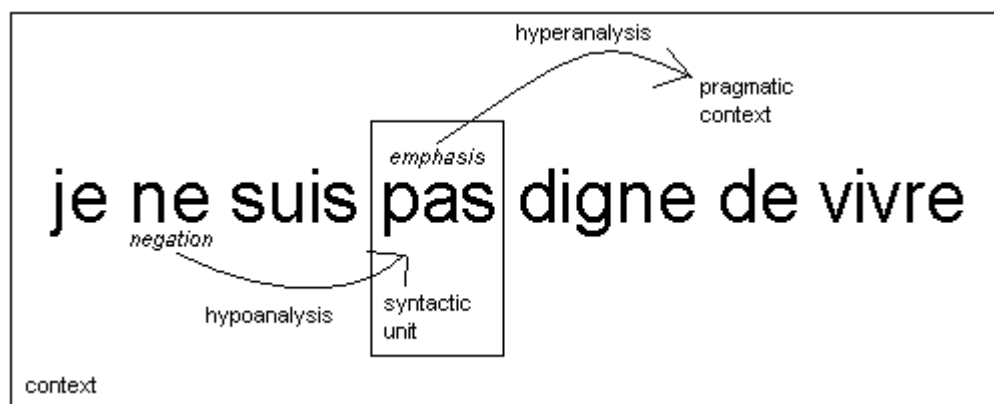
Figure 3. Metanalysis: simultaneous hyperanalysis of emphasis to the pragmatic context and hypoanalysis of negation from the preverbal particle to the postverbal element.

Before metanalysis, the negative meaning is attributed to the preverbal particle *ne*, and the emphatic function is attributed to the postverbal element (*pas, personne*, etc.). From the point of view of the postverbal element, negation is a contextual feature while emphasis is an inherent feature. However, as has been argued by a number of linguists (e.g. Schwegler 1988: 36, Givón 1979, ch. 3), negative utterances are more likely to be emphatic in actual use than positive ones; that is, there is a high degree of correlation between negation and emphasis. This correlation sets the condition for metanalysis: since the emphatic element is found frequently in negative contexts, and negative contexts are frequently emphatic, there is a swapping of the two functions: the negative function is attributed to the emphatic element, while the emphatic function is attributed to the nonlinguistic context.

Finally, in CRYPTANALYSIS (Croft 2000: 134), the language user 'analyzes a covert semantic/functional property of a syntactic unit as not grammatically marked, and inserts an overt marker expressing its semantic value.' Croft interprets PARATACTIC NEGATION as cryptanalysis (Croft 2000: 135-136). This is where negation constructions are used in the complements of verbs such as *deny, prevent, forbid, hinder* and *doubt*, and sometimes after prepositions such as *before, until* and *unless*; he discusses this phenomenon in Latin, Basque and Catalan in addition to the French instance that I discussed in Chapter 3.

64
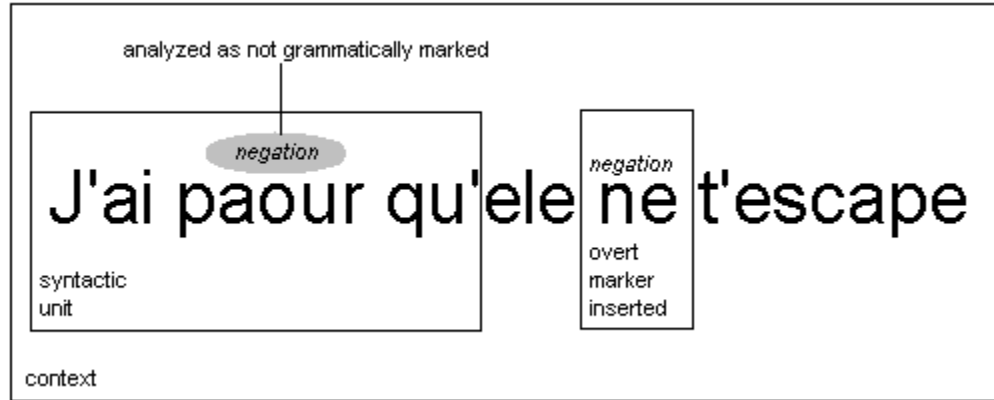
Figure 4. Cryptanalysis where the negation implied by the matrix phrase is analyzed as insufficiently marked.

In order for an utterance to be reanalyzed, inferenced, reinterpreted or metanalyzed, it must first be uttered with sufficient frequency in contexts that favor that reinterpretation. Croft (2000: 159) refers to this use as PERIPHRASIS, a translation of Lüdtke's (1986) *amplification sémantique*. Many linguists believe that this amount of use requires a particular intention, such as expressiveness or avoiding misunderstanding; Croft (2000: 74) cites Lehmann (1985), Lightfoot (1991), Heine (1994) and Harris and Campbell (1995) as using this expressiveness explanation, but builds his work primarily on Keller (1990) and Lüdtke (1986). Detges and Waltereit (2002) go further and assert that this periphrasis is the essence of grammaticization itself. Dahl (2001) makes a connection between the general phenomenon of inflation throughout society including the original economic sense of inflation, and linguistic inflation, such as where Chinese speakers using the expression *hěn kuài* meaning "very fast" in contexts where the speed is not remarkable results in the phrase being devalued to simply mean "fast," and speakers needing to use another word to draw attention to the speed.

In a study of elicited Pear Stories data, however, Croft (2007) demonstrates that in the normal course of describing events, 'innovation [...] is pervasive,' and thus

65

explanations based on expressiveness or avoiding misunderstanding are not necessary. Detges and Waltereit used the expressivity hypothesis as an explanation for the widely observed tendency of grammaticization to be unidirectional. Instead of this, Croft attributes the unidirectionality of grammaticization to the entrenchment and reduction effects of token frequency, which will be discussed in more detail later in this chapter. However, nothing in Croft's explanation precludes the kind of inflation that Dahl and others observe from being part of the innovation he describes, or even an additional source of innovation.

5.2 Propagation and Contexts of Use

In the evolution of constructions, we observe more than just changes in the way a word or construction is intended, or the way it is understood. We also see corresponding changes in the contexts in which it is used. We find new constructions increasing in use at the expense of competing constructions. There are many different ways that this problem has been approached, and we will take a tour of them.

Some older theories of syntactic change (such as Jespersen's 1917 theory) worked on a 'pull-chain model,' where the weakening of one construction created a semantic and pragmatic void that was filled by newly grammaticized constructions. In the example of negation, Jespersen argues that the phonetic weakening of the *ne* alone construction left French speakers without a clear way of expressing predicate negation, and they used the phonologically unambiguous embracing negation constructions more frequently because of this. In section 4.3 I discussed studies by several linguists that found no evidence supporting Jespersen's model. The pull-chain model is not widely used any more, leaving us in search of a new explanation for the increases in frequency.

Since we have found no evidence that the grammar evolves to make it difficult to express a given concept, as long as the concept is not a categorically new one (usually related to a technological advance such as farming or computers), any reanalysis will throw the reanalyzed construction into COMPETITION with existing constructions. While every word and construction has a slightly different meaning and connotation, the two constructions are sufficiently close to be considered 'the same' in some way by the language user.

Croft (2000) integrates this idea into his evolutionary model of language change: just as genetic innovations can be selected to be further replicated, innovations in language can be selected to be replicated and eventually become new conventions; he combines theories from David Lewis (1969) and Herbert Clark (1996) to form a unified definition, which he divides into multiple clauses. His first law of propagation (Croft 2000: 176) states that over the long term, situations where a single community uses two different variants are unstable, and language users tend to shift to one variant or another. This can happen in one of three ways: the first is that the alternative forms can be reassigned to different functions so that they are no longer in competition. The second is that the variation can be reinterpreted as corresponding to a division of the community. The third is to gradually shift the community towards the use of one variant or the other.

One of the oldest models of this shift is known as ANALOGY. According to the overview given by Hock (2003), analogy goes back to the ancient Greek philosophers, and has been used to explain a wide range of language phenomena, including 'any observed regularity' (Hock 2003: 443). The prototype case is called 'four-part analogy' by Hock, who gives the following English example:

67

58) *dog : dogs*
    cow : x

The Old English value for *x* (i.e. plural of *cow*) was *kine*, but it has been replaced by *cows* on the model of *dogs, cats*, etc. Significantly, it often replaces an older, irregular form which is no longer productive, although it can be used to express entirely new concepts, as in *webcams*. In describing this form of analogy (Hock 2003: 441), Hock writes,

> Certain conditions increase the success of this type of analogical change. These include that the 'x-' side of the equation should be a synchronically derived form (such as 'plural' versus 'singular') and that the pattern being generalized should be productive.

There are two kinds of analogical change that are related to four-part analogy: analogical leveling fills out paradigms like the singular/plural paradigm described above; however, it is not restricted to the comparison of single pairs of forms. In Old French, for example, while most of the future forms of *être* followed the regular future paradigm with a stem in *ser-* (*jo serai*, *tu seras*), the third person singular used the suppletive form *iert*, as in the following example:

59) Figura: Tuz jors serras, ta vie n' **iert** pas paie;
    'Figura: Forever you will live, your life will not be
    short;'

Ordo Representacionis Ade, ca. 1160, line 58

By the beginning of the thirteenth century, *iert* had been replaced by *sera* in many contexts:

60) Theophiles: Bien **sera** m'ame devorée
    'Theophile: My soul will be totally devoured'

Miracle de Theophile, ca. 1200

68

ANALOGICAL EXTENSION (sometimes reduced to EXTENSION) is an adaptation of

the concept to patterns that are not tight paradigms, for example to syntactic schemas

such as the following constructed examples from Bybee (2003a):

61)   We are going to Windsor to see the King.

Bybee (2003a: 147), example 1

62)   We are going to get married in June.

Bybee (2003a: 147), example 2

Bybee argues that after the construction *be going to x* was reanalyzed from

movement to intention, it was extended from movement verbs to a wider class of verbs

(which had the effect of increasing its type frequency). Meillet (1912 [1982]) describes

analogie and *grammaticalisation* as two separate processes, but his examples make it

clear that by *analogie* he was referring to analogical leveling, so his notion of

grammaticalisation is not incompatible with analogical extension. Many

grammaticization theories include extension or propagation as a step in grammaticization.

Following Croft (2000) and Bybee (2003b), I prefer to treat it as a separate process that is

a frequent consequence of the reanalysis that constitutes grammaticization.

The nineteenth-century Neogrammarians made a distinction between analogy and

regular sound change. Hock cites a number of linguists who argued that regular sound

change was a form of analogy (including Schuchardt 1885 and Postal 1968), and then

discusses the concept of LEXICAL DIFFUSION (Wang 1969), which was claimed to

subsume both kinds of language change. Contrary to the Neogrammarian assumption that

regular sound change covered the entire lexicon at once, Wang showed that both analogy

and regular sound change tended to start in one part of the lexicon and propagate

throughout. Joan Bybee, in a series of studies (Hooper 1976, Bybee 2002 and others), observed that regular sound change affects high token frequency words first, but analogical extension affects words with low token frequency first.

The shift from competing forms to a single one has also been described by Hopper (1991) as SPECIALIZATION: 'the narrowing of choices that characterizes an emergent grammatical construction.' He introduces this term as a substitute for Lehmann's (1985) 'obligatorification' because the form that 'wins' does not necessarily become obligatory, but I find the term 'specialization' to be less intuitive than 'competition.' It is quite likely the same process, but both Hopper's and Lehmann's terms focus the reader's attention on the narrowing of choices instead of the extension of one form at the expense of the others.

After mentioning the role of productivity in analogical extension, Hock writes (2003: 446), 'there is no ready-made answer [...] to the question of what makes a particular type of formation productive.' Others have attempted to provide an answer, and this will be seen in the next section.

5.3 The Usage-Based Model

The notions of grammaticization, form-function reanalysis and propagation described in the previous section are compatible with several models of language, but they seem to work particularly well with USAGE-BASED or NETWORK models. Although the foundations of this approach go back to early linguistics (Lehmann 1995), the first substantial models were created by Bybee (1985, 1995) and Langacker (1987, 1988) At their core is the notion that patterns of usage affect patterns of language. They were partly

a result of the application of certain models of cognitive science, particularly

connectionist models based on the understanding of neural structure at the time.

In a summary of usage-based models, Croft and Cruse (2004) contrast them with

'structuralist' and 'generativist' models, where 'only the structure of the grammatical forms

determines their representation in a speaker's mind.' Usage-based models rest on the

assumption that the two usage-based properties of frequency and semantics also affect

grammatical representation. Four hypotheses are derived from this assumption (Croft and

Cruse 2004: 291-327, emphasis added):

1. The storage of a word form, regular or irregular, is a function of its token
   frequency.
2. The productivity of a schema is a function of the type frequency of the
   instances of the schema.
3. In addition to source-oriented morphological rules/schemas, there also exist
   product-oriented schemas, which cannot be easily represented by derivational
   rules.
4. Strength of connection between word forms, and thus forces influencing their
   phonological shape (among other things) is a function of similarity. Similarity
   is measurable by comparing words to each other in both meaning and form;
   similarity in meaning is much stronger than similarity in form.

5.3.1 Token Frequency and Type Frequency

Before discussing this issue any further, it is important to focus on two of the

terms used in the hypotheses. TOKEN FREQUENCY is defined by Croft and Cruse (2004:

292) as 'the frequency of occurrence of the word form in language use,' and TYPE

FREQUENCY as 'the number of word forms that are instances of a particular schema.' For

purposes of comparison, token frequency is usually expressed relative the number of

occurrences per thousand words, although when comparing frequencies within a single

text it is sometimes expressed simply as the number of occurrences. For example, as a

measure of the token frequency of English verb conjugation classes, Bybee (1995: 434)

71

refs to Francis and Kučera's (1982) counts of the occurrence of these nouns in the

Brown corpus. For the type frequency of various uses of the English verb *can* (Bybee

2003: 611), she counts the number of different verbs that are listed as co-occurring with

can in a published concordance (Tatlock and Kennedy 1927).

Type and token frequencies are commonly measured by consulting published

sources in this way, but Gernsbacher (1984) observes that published token frequency

tables can be inaccurate measures of experiential familiarity. For type frequency, Poplack

(2001: 410) notes that dictionary counts can be misleading:

> Due to discrepancies between the traditional accounts of [the indicative,
> subjunctive and conditional moods in French] and the facts detailed here, I
> distinguish *prescribed* type frequency or lexical schema (the class of
> lexical items to which the phenomenon is prescribed to apply) from
> *observed* type frequency or schema strength, i.e. the items (or contexts) in
> which a variant form actually occurred.

Because of this, I will rely only on frequency counts taken from the corpus itself

for this study.

Bybee and Thompson (1997) refer to the effect of token frequency as the

CONSERVING EFFECT of lexical strength, or ENTRENCHMENT: in their first example,

> high frequency forms with alternations resist analogical leveling: while
> English *weep / wept, creep / crept* and *leap / lept* have a tendency to
> regularize to *weeped, creeped* and *leaped* respectivey, the high frequency
> verbs with the same pattern, *keep / kept, sleep / slept* show no such
> tendency.

We have a likely example of this from French negation: as described above in

Section 3.3, constructions like *non ai* 'I don't have [it]' and *non ferai* 'I won't do [it]'

retained their forms without subject pronouns for several centuries after subject pronouns

were obligatory in most other sentences. We do not have token frequency data from the

most relevant period, but texts from this period give the reader the impression that these are among the highest-frequency verb constructions.

Lieberman et al. (2007) confirmed this effect with a study of pluralization rules from grammars of Old, Middle and Modern English. They correlated this with token frequency data from the CELEX corpus. It is open to the criticism voiced by Gernsbacher (1984), but demonstrated an effect of token frequency despite this weakness. In Chapter 3 I discussed the observation by Moignet (1965) that the old Latin sentence negator *non* persisted for several centuries in fixed expressions of presupposition denial like *non ai* and *non ferai*; these expressions are relatively high-frequency and thus provide additional evidence in favor of entrenchment. Tottie (1991) observed a similar distribution in contemporary English negation, where older forms such as quantifier *no* and *nothing* were used only in the most frequent contexts. In this study I will test whether there is an effect of token frequency on the entrenchment of lexical items with the preverbal *ne* constructions.

5.3.2 Type Frequency in Morphology

Croft and Cruse's Hypothesis 2 relating type frequency and productivity is based on Bybee and Thompson's (1997: 71) observation that 'the type of change that is resisted by words or phrases of high token frequency is change on the basis of combinatorial patterns or constructions that are productive. [...] But frequency also plays a role in the determination of productivity, where productivity is defined as the likelihood that a pattern will apply to new forms. However, in this case it is type frequency: the type frequency of a pattern determines its degree of productivity.' As mentioned in the quote above, Bybee and Thompson (1997) consider this change to be analogical leveling, or at

least analogical change in general, answering Hock's (2003: 446) question of 'what makes a particular type of formation productive.'

Bybee (1995: 433) discusses a study by Guillaume (1927 [1973]) on nursery school children's overgeneralizations of French conjugations during play. While the 'third conjugation' verbs ending in -re (actually a catch-all category composed mostly of low-productivity irregular forms) had the highest number of tokens (58% of the total), a large majority of the verbs used (76.0%) were of the first conjugation ending in -er, the most productive one. This conjugation was the one that children tended to overgeneralize. Later in the article, Bybee (1995: 438-443) draws a distinction between productive forms and 'default' or 'emergency' forms used mostly for borrowed words that do not fit the native morphology. She shows that while the German plural construction in -s is used in emergencies and for some borrowed forms, the most productive pluralization construction is -(e)n, which also has the highest type frequency.

Although Guillaume (1927) identifies type frequency as a factor in children's overregularizations (and MacWhinney 1978 confirms this), Bybee and Slobin (1982) caution against interpreting this as evidence that analogical change happens through overgeneralization. In spontaneous speech and a series of elicitation tasks, they compared the overregularizations of English past tense forms made by pre-school children (aged 1.5-5.0 years), third graders (8.5-10.0 years) and university-aged adults (under fatigue conditions to induce more overregularizations). They found that the third graders and adults extended the regular past tense *more* than the pre-school children for verbs ending in -t such as *hit* and *cut*, while the pre-school children had internalized the 'affix-checking' rule and tended not to add past tense suffixes to these words. Because the adult

74

and third-grade usage is more in keeping with the historical pattern of English, Bybee and Slobin argue that extension is not primarily driven by small children, but by adult 'errors.'

Bybee and Thompson (1997: 71) explain the effect of type frequency with the network model of language use, where language users store lexical items in a network based on their perceived similarity, and generate new categories (also known as SCHEMAS) to make sense of the items and know how to apply them in the future.

> The more lexical items that are heard in a certain position in a construction, the less likely it is that the construction will be associated with a particular lexical item and the more likely it is that a general category will be formed over the items that occur in that position. The more items the category must cover, the more general will be its criterial features and the more likely it will be to extend to new items.

In an earlier work, Bybee (1995) describes the relationship of type frequency and productivity in more detail: 'Degrees of productivity are highly correlated with the type frequency of a pattern within a language.' In prototypical 'four-part' analogical extension of the kind described by Hock (2003), a 'new' productive form is extended over an old form which has lost its productivity. In these circumstances, language users tend to use the one that feels more general and all-purpose. This feeling of generality and versatility is determined by the type frequency of the alternatives.

It is important to note that the concept of type frequency that emerges from Guillaume (1927) and Bybee (1995) is a *relative* one related to morphological paradigms. The children in Guillaume's study had four patterns to choose from, and they overgeneralized forms from the pattern associated with the most lexical items. The German speakers in Bybee's discussion of German plurals were choosing which plural construction to use for loanwords and nonce forms.

75

In later works, Bybee has expanded this concept to an absolute characteristic of the storage of a construction. For example, in her chapter in the *Handbook of Historical Linguistics* (Bybee 2003b) she draws on examples from the evolution of English *can* which she tracks through at least three distinct semantic stages: *mental ability*, *ability* and *root possibility*. She observes that in each successive stage, *can* occurs with more types of noun phrases in subject position and as an auxiliary modifying more types of verbs. However, she is very clear (Bybee 2003b: 605) that this notion of type frequency is an *effect*, not a cause of the increase in generality: 'it is the high token frequency of grammaticizing phrases which provides the triggering device for many of the changes that occur in the form and function of the grammaticizing construction.' The effect of high type frequency is only indirect, in that each new type brings with it at least one token.

The term *type frequency* is confusing for many people, in part because the term *frequency* suggests events occurring across several equal time periods. Type frequency is not that kind of frequency; the type frequency of a construction is a measure of the status of the construction in the language user's memory. Every time the language user perceives the construction in a new context (where 'new' can mean different things for different people), the type frequency of that construction is increased by one increment in the user's memory. However, memory is not perfect, so it is possible for a language user to forget an utterance that he or she has heard, or for the strength of a memory to fade over time (Goldberg 2006). The type frequency of a construction is thus an approximate measure of the total number of distinct contexts that the language user remembers the

construction occurring in. It can be thought of as the 'market share' or 'mindshare' that the construction commands relative to its competitors.

Bybee and Thompson (1997: 71) write, 'type frequency can range from one to a very large number,' and in many cases it is written as a raw number. However, type frequency as a cause of productivity is only discussed in terms of the productivity of one construction relative to others that are in some sense 'the same.' In other words, it is only used in the context of variation and change, specifically of competition (Hopper 1991's 'specialization,' Lehmann 1985's 'obligatorification'). As Poplack (2001) writes, 'The network model requires that the (putative) replacement form have the same meaning or function as the pre-existing form.'

Type frequency is thus used to determine and predict the productivity of variants *relative to each other*. This is very different from the role of token frequency, which represents the effect of exposure over time and can operate on a single form independent of any others, relative only to the language user's moments of consciousness. I will therefore study it primarily as the proportion of types that a variant is associated with, to highlight this notion of 'mindshare.'

5.3.3 Type Frequency in Syntax

Bybee and Thompson (1997: 72), following Goldberg (1995), take the established concept of type frequency from morphology and apply it in the domain of syntax, to the problem of the DITRANSITIVE ALTERNATION in English. In this case, the ditransitive construction and the prepositional-object construction are considered to 'mean the same thing' – that one entity is in the semantic role of patient, and the other of goal – as in the following constructed examples:

63)   He told the woman the news.

64)   He told the news to the woman.

However, Bybee and Thompson point out that while the prepositional-object construction can be used with any verb, the ditransitive construction is only considered grammatical when applied to a relatively small number of verbs, and ungrammatical with all other verbs, as in the examples below:

65)   * He whispered the woman the news.

66)   He whispered the news to the woman.

Although Bybee and Thompson do not give specific type frequency counts, they observe that the prepositional-object construction has a practical type frequency of 100%, 'occurring with all verbs that can take a patient and recipient argument,' but the ditransitive 'occurs only with specific lexical verbs, and most of these are of Germanic origin and of very high frequency.' They observe that 'this case is made complex by the fact that the ditransitive construction does have some limited productivity, as evidenced by its occurrence with certain new verbs such as *to telephone*, *to e-mail* and *to fed-ex*.' They argue that within the class of 'verbs of sending and communicating,' the ditransitive construction has relatively high type frequency, which 'allows the construction to be used with new verbs with related meanings.'

Smith (2001) investigates the role of type and token frequency in the shift from *have* to be in English compound past tense constructions fulfilling the resultative and

anterior functions. His data broadly confirm Bybee and Thompson's (1997) theory: with regard to type frequency, the *have* constructions already are used with 84% of the types in his earliest Old English data. The verbs that resist the shift the longest are *go* and *come*, which are the highest-frequency verbs used with *be* constructions in *Wuthering Heights*, his nineteenth century sample, appearing in 17 and 11 tokens respectively.

Poplack (2001) found no significant role for type frequency in the relative productivity of any of three competing pairs of irrealis verb constructions in a corpus of adult spoken Quebec French. However, she did not use any concrete measures of type frequency, instead (Poplack 2001: 423) characterizing the type frequency of the older construction in each pair as either 'very low' or 'very high.' Because this was a synchronic study, she also did not have an adequate measure of productivity, relying on 'the weaker requirement of occurring at a substantial, and relatively *homogeneous*, rate across all lexical items and contexts forming its domain' (Poplack 2001: 411). As with type frequency, she relied on subjective, verbal ratings of the productivity of the older construction in each pair (Poplack 2001: 423): the Subjunctive was 'Low and unchanging,' the Inflected Future 'Restricted and decreasing slowly' and the Imperfect 'Restricted and decreasing rapidly.' Poplack did not give details about how she entered this data into VARBRUL, but it is not surprising that the output was unsatisfying.

5.3.4 Autonomous Types
It is important to note Bybee's claim (1995: 433-434) that 'forms with high individual token frequency can be learned by rote, and they can be autonomous, even from other members of their own paradigms.' This implies that they do not count as types for the purpose of measuring type frequency. Building on this, Hay (2001) showed that if

a derived lexical item is more frequent than its base, it tends not to be included in schema formulation.

Hay and Baayen (2001) argue that neither token frequency nor the frequency of a derived item relative to its base are an accurate predictor of whether an item is 'stored' (as a unit) or 'computed' (generated). They plotted a number of items on a graph where the axes were the frequency of the base form and the frequency of the derived form, and drew a 'parsing line' dividing the stored items from the computed ones. They found that the parsing line predicted by an algorithm in a program called *Matcheck* is a more accurate predictor of the observed parsing line than the frequency of the derived item relative to its base. Matcheck was originally developed by Baayen and Schreuder (2000) to model morphological segmentation based on the frequency and similarity of lexical items. It was not possible to test the predictions of Matcheck or the base/derived frequency ratio in this study, but such a test would be valuable in future work.

5.4 Modeling Progression

Describing language change with words allows us to better understand the processes involved, but mathematical models can allow us to describe it more closely and even predict the future course of these changes. The math involved may get complicated, but the increased understanding is worth it.

The first step I took for this investigation of French negation was a pilot study based on a smaller corpus. In that study, as in this one, the number of embracing negations as a proportion of all sentence negations rose very slowly until the seventeenth century, when there was a dramatic increase, and after that increased slowly again for the rest of the study period, forming an 'S' shape. This S-CURVE has been observed by many

investigators of language change, beginning with Osgood and Sebeok (1954: 155) and repeated by Weinreich, Labov and Herzog (1968) and Bailey (1973). However, this was observation without explanation; in these works they did not go beyond the simple statement that the data resembled an S.

According to Altmann et al. (1983), the first to propose a mathematical model underlying the S-curve were Piotrovskaja and Piotrovskij (1974), but they proposed an arc tangent function that did not fit other data or explain the phenomenon. Altmann et al. themselves proposed a logistic function to model observations detailing the acceptance of Arabic loanwords in Persian prose, and draw parallels with other phenomena: 'The same dependence is observed in other fields of science as in the theory of growth, in epidemics, etc. [...] We consider this analogy to be a strong argument for the validity of the model' (Altmann et al. 1974: 111). Kroch (1982, 1989) put forth the logistic function for similar reasons: 'its use is generally considered appropriate in statistical studies of changing percentages of alternating forms over time' (Kroch 1989: 4). Several linguists have applied his model with similar focus, notably Kallel (2007) to the evolution of negation in English. Kallel's data posed challenges for other aspects of Kroch's theory, but provided strong confirmation for the use of the logistic function.

The logistic function, and the exponential function on which it is based, were first developed to describe and explain compound interest and population growth. Altmann et al (1983) and Kroch's (1989) choice of the logistic implies a conjecture that the use of certain constructions grows like an investment or the population of a species. In order to fully explore the implications of this analogy, I will discuss the growth of investments and populations, as well as other useful cases like the spread of ideas. More detail on the

81

history of the logistic function in general can be found in Cramer's (2003) excellent overview, and Denison (2003) provides a good critical review of the studies relating to S-curves in linguistics.

5.4.1 Exponential Growth

There is a famous story used to illustrate exponential growth. When Sissa, the legendary inventor of chess, demonstrated his invention to the Emperor of India, the Emperor was so impressed that he asked Sissa to name his reward.

Sissa replied, 'I would like to have one grain of wheat for the first square of the chessboard, two for the second square, four for the third, and so on, doubling at each square.'

The Emperor agreed, but he didn't know what he was in for. The amounts got bigger and bigger. The number of grains on each chessboard square turned out to be two raised to the power of the number of the square. After eight grains of wheat it doubled to sixteen, then thirty-two. By the second row it was up to 256 grains, and by the fourth row it was 278 metric tons. But it increased faster and faster: by the sixth row it was up to eighteen million metric tons. On the last square alone it would have been nine quintilion grains (2^{63}), more than a trillion metric tons - fourteen thousand times the amount of wheat grown in the Indian subcontinent *in 2007*.

This function has applications way beyond rewarding inventions, and when graphed it looks something like this:
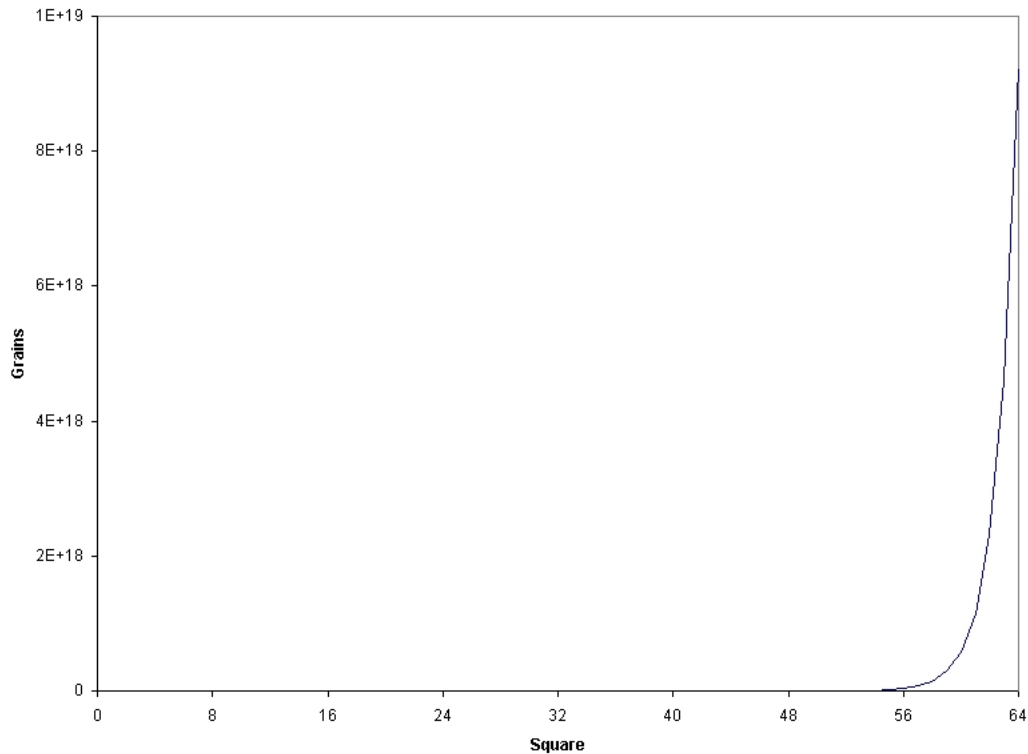
Figure 5. Idealized exponential function.

This graph is quite different from a LINEAR function, which forms a straight line. The numbers it reaches are so large that they in turn have to be written in scientific notation, using powers of ten. Note that the first steps in the progression are so small in comparison to the later stages that they don't even show up on the chart. To the extent that people like Sissa are able to trick others, it works because the initial stages seem manageable.

The deceptively simple word 'doubling' describes a common property of many natural phenomena: at any given time *the rate of growth is proportional to the present size*. The Swiss mathematician Leonhard Euler (1736) developed an equation to model this proportionality:

       1) dx/dt = rt

83

Where x is the amount of grains of wheat, t is the time elapsed (number of the chess square), r is a constant and *dx/dt* represents the rate of growth in wheat. Another example is the multiplication of single-celled organisms such as amoebas, where each organism splits into two. Just as with the grains of wheat, the number of amoebas will double at each generation and grow exponentially.

This proportionality is present in any phenomenon where the items being counted are themselves factors in the change. In many cases, the rate of growth is proportional to the present size, but instead of being equal to it, it is a multiple of it. It may be larger, as in the case of a contagious disease like measles where each carrier can infect ten to fifteen people in just a few days (so r would have a larger value). It may be smaller, as in the case of compound interest, where the borrower gives the lender a fixed percentage of their investments in interest every year (at 5% interest, and a relatively low value of r), adding that interest to the principal. Because the rate of increase of the principal is a fixed fraction or multiple of the principal itself, we get the same effect; all that changes is how long it takes.

Based on the theory of calculus that had just been developed by Newton and Leibnitz, Euler was able to solve the equation for *x(t)* (the amount of wheat on a given square t, or the amount of money in the bank at time t) by integration, obtaining the EXPONENTIAL FUNCTION:

2) $x(t) = x_0 e^{rt}$

where x_0 is the initial value of x. I will discuss e shortly.

5.4.2 Logarithms

Just what does this constant *r*, the 'rate of change' represent in the real world? In order to determine that we would need to be able to undo the exponential operation. Even before Euler worked out the equation and the graph, the Scottish mathematician John Napier (1614) was working on this task. He developed a method of calculating LOGARITHMS, which are the inverse of exponents. If we have an exponential formula as follows:

$$3) \quad y = a^x$$

The inverse of that formula is the following logarithmic formula:

$$4) \quad x = log_a(y)$$

Logarithmic scales have been used for millenia to keep track of values that are naturally distributed along an exponential scale, like the stellar magnitude scale for the brightness of stars and the Richter scale for the intensity of earthquakes. In the case of the Richter scale, for example, even though the difference between a 1.0 earthquake and a 2.0 earthquake is much smaller in terms of how far it moves objects than that between a a 5.0 earthquake and a 6.0 earthquake, the differences in the effect on humans are relatively comparable.

Napier found that logarithmic scales could be used to simplify multiplication problems, because you can get the product of two numbers by adding the logarithms of those two numbers and raising it back to the original power. This is the principle that was used for the slide rule, the most common calculating tool until the invention of the pocket calculator.

5.4.3 Euler's Constant *e* and Natural Logarithms

The use of logarithms to perform multiplication is made much easier, Napier found, if you use a particular constant for the base. Euler refined this constant to approximately 2.718 and called it *e* (Euler 1736: 68). Logarithms to the base *e* are called NATURAL LOGARITHMS and abbreviated *ln*.

This is the *e* that appears in our function representing exponential growth.

$$5)\quad x(t) = x_0 e^{kt}$$

So let's get back to our rate of change, represented by *r*. What would we expect *r* to be for Sissa's chessboard? Since the additional amount of wheat is equal to the amoount on the previous square, we might expect *r* to be 1; since it doubles, maybe *r* would be 2. Neither of those is correct.

We can start with x_0, the amount on the "zeroth square." If x_1 is 1, then we can work backwards, dividing by two to get an x_0 of 1/2. The amount of wheat on each square is 2 to the power of the number of the square, as in the following equation:

$$6)\quad x(t) = (1/2)\, 2^t$$

A little algebra tells us that for Sissa's chessboard - and for any case where the value doubles in exactly one period - *r* is the natural logarithm of 2, which is roughly 0.693.

5.4.4 Exponential Growth Applied to Language Change

Recall that Bybee (1995) observed, 'Degrees of productivity are highly correlated with the type frequency of a pattern within a language.' Productivity is also affected by the degree of openness of the schema - the morphophonological compatibility of it with

various items (Bybee 1995: 430), but Bybee's claim is that type frequency is the dominant factor, so for this study I will assume that openness plays no role, and leave it for a later investigation.

For our purposes, what is productivity? Bybee and Thompson (1997) wrote, 'productivity is defined as the likelihood that a pattern will apply to new forms,' but new coinages – the most obvious kind of 'new form' – are relatively rare. What is more common is when a language user forgets which construction to use with a given lexical item. They then assign that item to one of the various competing constructions based on their type frequencies.

Productivity is thus the likelihood of increasing type frequency, so when Bybee observed that the degree of productivity is correlated with the existing type frequency, she was saying that for type frequency *the rate of change is proportional to the present size*. In other words, it grows exponentially, just like the grains of wheat on Sissa's chessboard, like measles infections, and like compound interest, and it behaves like them.

5.4.5 The Exponential Function - With Limits

You may have noticed that the examples I gave earlier in this section for exponential functions were not very realistic. The Emperor of India ran out of wheat long before he got to the end of the chessboard. If measles keeps infecting people at an exponential rate, pretty soon everyone in the area is either dead or immune. Interest on an investment is eventually spent on something. Earthquakes can only move the earth so far.

The demographer Thomas Robert Malthus (1789) observed that while the human population was capable of increasing exponentially under some circumstances, it was often limited. These limits were slightly different from unchanging limits like the size of

87

a chessboard or the maturity date on an investment. It is possible, Malthus wrote, to increase the amount of subsistence (food, shelter and other necessities) that allow people to live and reproduce, but it is not possible to increase them exponentially (Malthus 1789: 4):

A thousand millions are just as easily doubled every twenty-five years by the power of population as a thousand. But the food to support the increase from the greater number will by no means be obtained with the same facility. Man is necessarily confiried in room. When acre has been added to acre till all the fertile land is occupied, the yearly increase of food must depend upon the melioration of the land already in possession.

Malthus further observed that there were CHECKS on the growth of populations (Malthus 1789: 7):

> The ultimate check to population appears then to be a want of food, arising
> necessarily from the different ratios according to which population and
> food increase. But this ultimate check is never the immediate check,
> except in cases of actual famine.
>
> The immediate check may be stated to consist in all those customs, and all
> those diseases, which seem to be generated by a scarcity of the means of
> subsistence ; and all those causes, independent of this scarcity, whether of
> a moral or physical nature, which tend prematurely to weaken and destroy
> the human frame.

Building on Malthus's work, the Belgian mathematician Pierre-François Verhulst observed that the checks got stronger as the population approached the limits of its environment. He wrote, 'la vitesse d'accroissement de la population est retardée par l'augmentation même du nombre des habitans'[1] (Verhulst 1838: 115). In order to capture this regularity, he developed the LOGISTIC EQUATION, as follows:

---

[1] The rate of growth is checked by the population increase itself.

7) dP/dt = rP(1-P/K)

In the above equation, P is the size of the population and r is the rate of growth. There is a new term, however: K represents the CARRYING CAPACITY of the environment, the maximum population that the environment is capable of supporting.

The solution to the equation is as follows:

8) P(t) = KP$_0$e^{rt}/(K + P$_0$(e^{rt} - 1)) where lim$_{t->8}$P(t) = K

Where P_0 is the initial population level. If we represent the population P as a percentage of the carrying capacity K, represented by x and set our initial time so thatP_0 is 1, we get a much more manageable equation:

9) dx/dt = rx(1-x)

10) x(t) = ert/1+ert

Verhulst tested that equation against population data from France, Belgium, Essex and Russia, finding that it predicted the population dynamics well. The pattern that he found - and that Malthus and Benjamin Franklin had observed before him - was that the population changes started slow, then moved quickly for a while, and then slowed again as they approached the limits. The graph of the equation is the S-curve described at the beginning of this section:
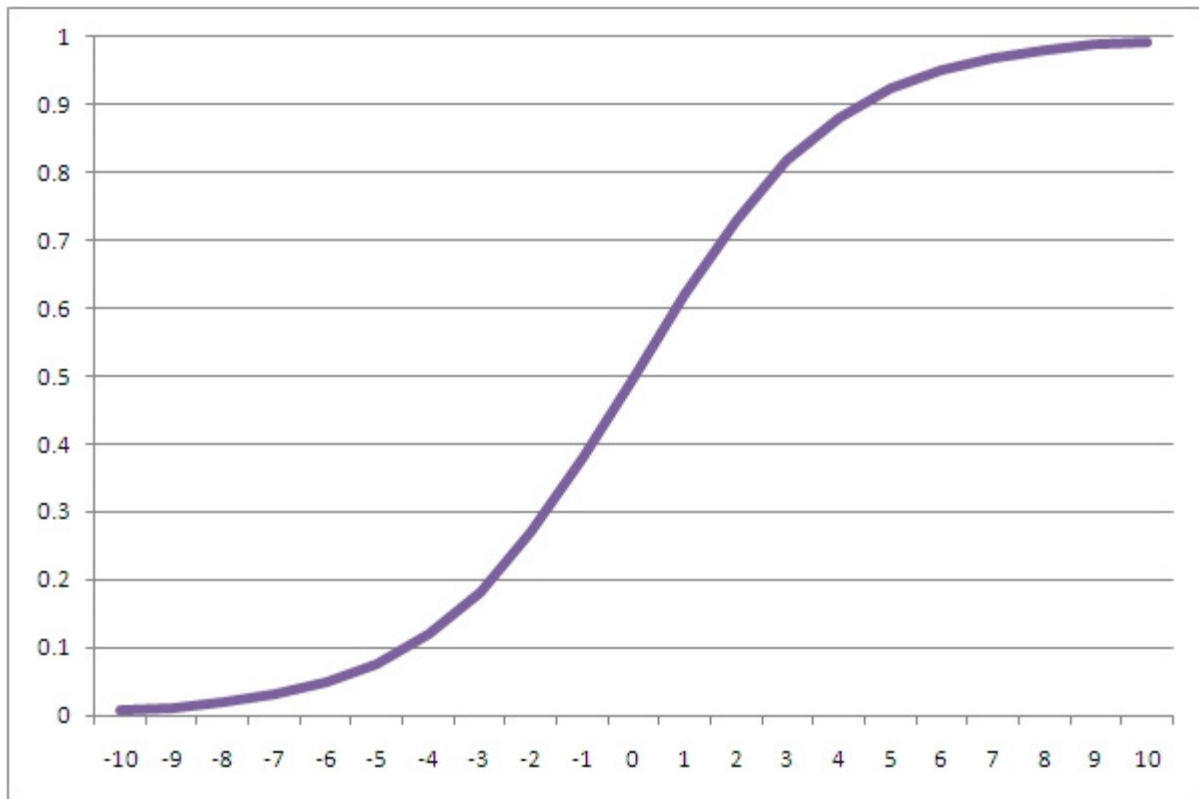
Figure 6. Idealized logistic function.

According to Cramer (2003: 6-8), the logistic model has also been applied to autocatalytic reactions in chemistry and bio-assay, with striking parallels to population growth. In both cases there is something that increases in quantity at an exponential rate, which is facilitated early on by the available space, but limited as it later fills that space. In social sciences, it has been applied to the adoption of social innovations, technologies and products (Manski 2001). This is a step removed from population growth, in that their use implies that there is a limited area of social 'mind share' that is colonized by the innovations, technologies or products.

This is the same s-curve that Osgood and Sebeok (1954) found in language change, and that Altmann et al. (1983) and Kroch (1989) confirmed as an accurate model of many changes they had observed. We would thus expect to find an analogue to the open – but bounded – environments that populations and chemical reactions expand into,

90

and an analogue to the reproductive drive that pushes populations to expand and to the conditions that push chemical reactions to expand.

5.4.6 The Inverse of the Logistic Function: the Logit

Just as the logarithmic function is the inverse of the exponential function, the inverse of the logistic function is the LOGIT FUNCTION; the two functions can be used as TRANSFORMATIONS to convert linear functions to logistic functions and back, in order to compare them and test hypotheses.

$$11)\ t = ln(x/1\text{-}x)\ /\ r$$

The logit transformation allows us to test whether a set of data follows a logistic pattern. If we expect a function to follow a logistic pattern, then we expect its logit to follow a linear pattern. Linear regression is a way of testing the prediction that a given set of data will appear as a straight line on a graph. The practice of LOGISTIC REGRESSION consists of performing a logit transformation on a data set and using linear regression to test how well the resulting data approximates a line.

5.4.7 The Mechanics of Logistic Progressions in Language

Kroch (1989) is very clear in his interpretation that the S-curves he observes reflect the aggregate of the gradual shift from one construction to another in the usage of many individuals:

> The study of language use is the study of the choices that people make among alternative forms in their repertoire of grammatical knowledge in formulating utterances. The usability of grammatical options is sometimes strictly determined by features of extra-sentential context and, to that extent, variation in use may reflect underlying competence extended to the discourse level (Prince, 1988). More germane to our concerns here, however, is another fact, that variation often reflects choices that are not

91

categorically determined by linguistic principles at any level but instead are only probabilistically influenced by features of context and situation. In the case of replacement of one form by another, this is the expected circumstance, at least so long as the change is moving forward and does not turn into a stable alternation. To study such replacement is to determine the nature and weight of these probabilistic factors and to trace their temporal evolution, necessarily using statistical methods applied to corpora of data.

Kroch used the following chart to illustrate the application of the logistic function to language change, specifically to Oliveira e Silva's (1982) data concerning the prevalence of definite articles in possessive adjectives, as in the following example:

67)   Maria conhece o meu irmão.
       'Maria knows my brother.'
                        Kroch (1989: 207), example 7a, literal translation by Kroch.

RISE IN THE USE OF THE DEFINITE ARTICLE BEFORE
POSSESSIVES IN CONTINENTAL PORTUGUESE
(adapted from Oliveira y Silva 1982)
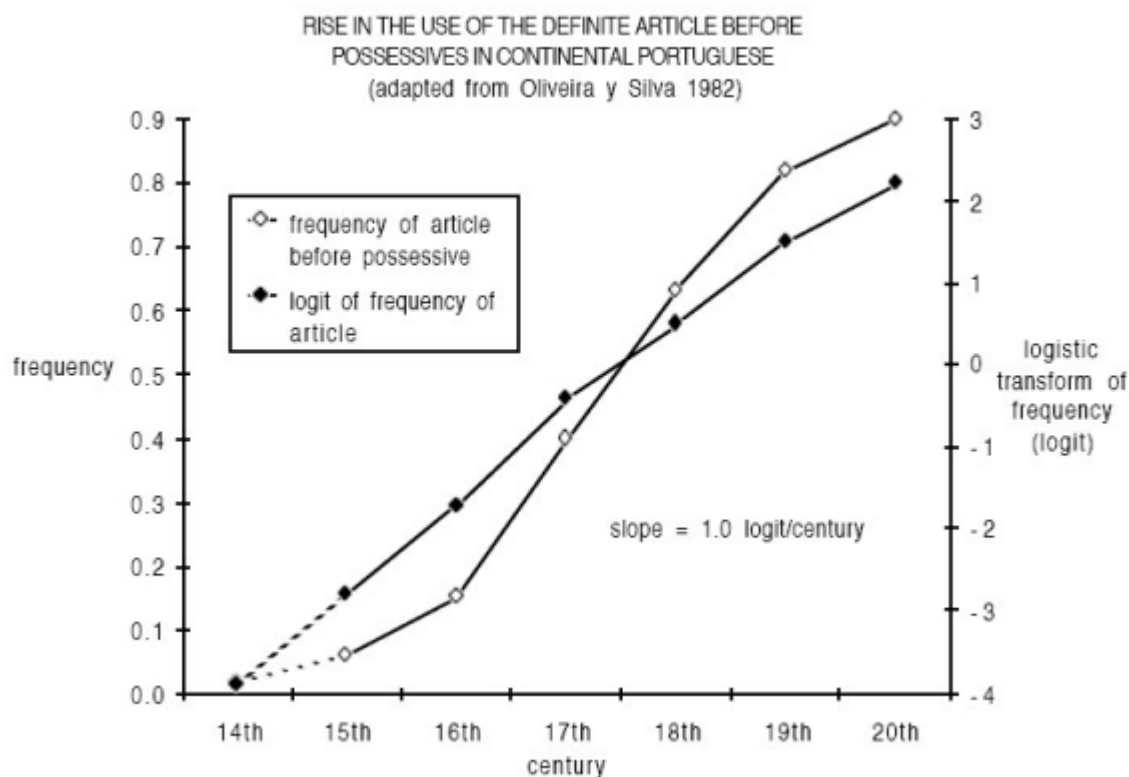
Figure 7. Chart from Kroch (1989: 208) using the logistic function to model the increase in possessive adjective constructions containing definite articles in Portuguese.

According to the frequency curve on Kroch's chart (marked with white diamonds), in the fifteenth century only about five percent of possessive adjectives were

produced with definite articles. That percentage (indicated on the left-hand *y*-axis) rose

slowly in the sixteenth century, and then much quicker until it was above eighty percent

in the nineteenth century, when the change slowed down again. The logit of this

frequency (marked with black diamonds) almost forms a straight line, as predicted for

any logistic distribution. The logit values actually run from about -3 to +2, but Kroch has

plotted them on the right-hand axis and overlapped the two curves to show their

relationship.

Niyogi and Berwick (1995) argue that many linguistic changes follow patterns

that do not fit the S-curve model very well. They propose a 'dynamical systems' model to

account for these changes, and criticize Kroch (1989) for 'imposing an S-shaped logistic

change by *assumption*' (emphasis in the original), and offering an alternative that relies

on learning algorithms and Markov chains. They claim to test the model on data from the

loss of verb-second constraints in French, but it is impossible to evaluate it because there

is no discussion of where the data comes from. Briscoe (2000) offers a stochastic model

that he claims fits the data better than Niyogi and Berwick's, but both models are

complicated, and it is not clear what is gained by this increased complexity.

5.4.8 The Usage-Based Model and Logistic Progression

Although Kroch (1989), as quoted above, believed that the logistic progression

reflected a change 'moving forward,' he did not suggest a mechanism by which these

changes could move forward. The usage-based model of Bybee and Thompson (1997)

and Croft (2000) offers such a mechanism in the form of type frequency and token

frequency.

The critical link between productivity and analogical extension has to do with storage in the lexicon - that is, in long-term memory. Recall that Bybee (1995) wrote, 'Degrees of productivity are highly correlated with the type frequency of a pattern within a language.' Since forms that have high token frequency are stored as units in the lexicon, this implies that sequences without high token frequency are not stored as units in long-term memory. Every time they need to be produced they are 'new forms,' and thus productivity comes into play.

When a construction is perceived as being 'the same' as another, it is in competition with that other construction in that context. Each time the language user needs to use a new or not-separately-stored word in that context, the user chooses one or the other construction, and the likelihood that a construction will be chosen is determined by that construction's productivity relative to the other choices, and thus its type frequency. Recall from earlier in this section that type frequency represents the 'mind share' of a construction, and thus is only used to compare one variant with others.

Assuming for simplicity that the context where a given construction can be used occurs at a constant rate over time, and that separately-stored constructions are a relatively constant subset of the contexts, there is a constant need for new forms to be generated, and a constant rate at which a language user is presented with the choice between these constructions. The fact that the different constructions have different productivities, and thus different likelihoods of being chosen means that *propagation is proportional to type frequency*.

However, propagation is just the relative increase in use of a construction, which is itself an increase in type frequency. As Bybee and Thompson (1997: 71) wrote, 'high

94

type frequency ensures that a construction will be used frequently, which will strengthen its representational schema, making it more accessible for further use, possibly with new items.' In other words, the more types you have, the more you get, just as with measles and compound interest, and the result is exponential growth.

Obviously, the usage of, say, English regular past tense forms, does not continue to grow exponentially; there are analogues to the checks on population discussed by Malthus and Verhulst (1838), in the form of entrenchment due to token frequency. The primary resource used up by analogical extension is the amount of lexical memory allocated to this function, usually at the expense of a competing construction.

Malthus claimed that the availability of natural resources acts as both an enabler and encouragement to reproduction when high, and a restraint and discouragement when low. Similarly, the existence of low-token-frequency, non-entrenched lexical items acts as an enabler and encouragement to extension and to the formation of schemas through type frequency, while the existence of high-frequency, entrenched lexical items acts as a restraint and discouragement to extension. We can represent this with Verhulst's equations:

$$12)\ dx/dt = rx(1-x)$$

$$13)\ x(t) = e^{rt}/1+e^{rt}$$

In the case of language change, dx/dt is the rate of propagation, t is the time, r is a constant, and x is type frequency expressed as a percentage. Since the entrenched types tend to have higher token frequencies, we should expect the per-token prevalence of these constructions to increase at an even higher rate, but still in the form of an S-curve. In this

95

study I will test whether the logistic model can predict the course of a linguistic

propagation by type frequency used as well as by the proportion of tokens.

5.4.9 Actuation, Relics and Asymptotes

There is a particular aspect of the logistic equation that is worth discussing here, if

only because it has confused so many in the past. As Kroch (1989: 5) puts it:

> Note that because the logistic transform [...] varies between -∞ and +∞ as
> p varies between 0 and 1, the logistic, like other functions used in
> statistics, idealizes the empirical situation. Under the model, there is no
> time t for which p=0, nor any for which p=1, although as t approaches -∞,
> p approaches 0 from above and as t approaches +∞, p approaches 1 from
> below. Of course, actual linguistic changes have starting and ending
> points, so the model can only approximate real data; and this
> approximation falsifies the change process precisely at the beginnings and
> ends of changes. In particular, at the beginning of a change p jumps from
> zero to some small positive value in a temporal discontinuity which
> Weinreich, Labov and Herzog (1968) dub the "actuation" of the change.

Another look at the data behind any instance of extension shows that Kroch's *p* is

never 0 or 1. The initial forms do not appear from nowhere, and they are usually the

product of reanalysis. The French pronoun *il*, for example, was present in a significant

proportion of definite noun phrases when it was reanalyzed out of the Latin

demonstrative *ille*, just as the noun *pas* was used in negation contexts before its various

reanalyses. At the end of the extension, there are always a few relic forms hanging

around, some of them with limited productivity. For example, the Middle English plural

forms in -*en* and pronoun *thou* are occasionally heard in relic and archaizing contexts

(both showed up in blogs published the day I wrote this). In his discussion of the tails of

the S-curve, Denison (2003) points to a thorough study by Allen (2000) showing that the

decline of verb-final word order in English was very gradual. This is the phenomenon

described as 'layering' by Hopper (1991). While *no* model can do more than approximate

real data, the asymptotes of the logistic function approximate the real data just as well as any other part of the curve.

Even though the asymptotic nature of the logistic function does correspond better to the reality than Weinreich, Labov and Herzog (1968) claim, their notion of ACTUATION seems to correspond fairly well to the notion of innovation in Croft's (2000) theory, and in the cases we're interested in, to the innovation that occurs through form-function reanalysis. The propagation that is modeled by the S-curve has no reason to exist if the two constructions are not in competition, and they cannot be in competition unless language users consider them to be 'the same' on some level. The momentum in the system suggests that up to a certain point before the propagation begins in earnest the two constructions were thought of as 'different,' and that the form-function reanalysis has the effect of removing that difference from the language users' understanding and throwing the constructions into competition with each other. It thus makes no sense to attempt to apply propagation models before the hypothesized date of reanalysis, since we would not expect propagation to occur then.

5.4.10   Lotka and Volterra's Models of Resource Competition

As I mentioned in Chapter 4, Kroch (1989: 209-214) applies the logistic model to declining uses of two constructions in French, verb-second and pro-drop. However, a closer reading of Verhulst's model reveals that while it can be applied to the decline of a construction, such an application does not offer us any insight into the decline, because populations do not decline in that way. The verb-second and pro-drop constructions declined because of competition from the strict SVO construction, but the logistic model does not take into account this type of competition between constructions.

Since the field of population dynamics produced the logistic model that has proved useful in describing language change, we can look to it for a more detailed model for competition. Verhulst's model was extended, more or less simultaneously, by Lotka (1925) and Volterra (1926), to model a situation where two species compete in a given environment for a limited set of resources. For each pair of species represented by the letters i and j, they added an INTERACTION COEFFICIENT, α, to represent the marginal effect that every member of species i has on a given member of species j, as follows:

$$14)\ dx_i/dt = r_i x_i (1 - \Sigma \alpha_{ij} x_j)$$

Recall that r is the rate of change. The capital sigma Σ indicates the sum of all possible pairs of species i and j, including the interactions of members of the same species on each other.

This equation has been confirmed in various experiments as a model for interspecies competition for resources; for example, Gause (1934) found that it described the competitions between two species of Paramecium protozoa for a constant food source, 'which always resulted in a complete driving out of *P. caudatum* by *P. aurelia*.' (They also developed equations to model the situation where members of one species eat members of another, but this has no known analog in linguistics.) Recall that the simplest version of the Verhulst equation is as follows:

$$15)\ dx/dt = x\ (\ 1 - x\ )$$

For a slightly more complex system where two species *a* and *b* compete for resources which again total to 1, the Lotka-Volterra equation for species *a* would be as follows:

$$16) \; dx_a/dt = x_a \left(1 - x_a - \alpha_{ab} x_b \right)$$

The corresponding equation for species *b* would be as follows:

$$17) \; dx_b/dt = x_b \left(1 - x_b - \alpha_{ba} x_a \right)$$

To my knowledge, while some (e.g. Pinasco and Romanelli 2005) have applied the Lotka-Volterra model to the competition between two languages for speakers, I am unaware of any prior use of this model to represent competition between two *constructions* for usage. In this study, I will be testing the applicability of the Lotka-Volterra model to the competitions between various negation constructions in French.

5.5 Propagation through Society

Until now I have treated the linguistic community as a monolithic entity that adopts changes wholesale. Of course communities are not just made up of individuals but of multiple overlapping sugroups of individuals based on locations, cultural practices and family relationships, and there is a large body of sociolinguistic research showing that communities are far from monolithic in their adoption of language changes. Wolfram and Schilling-Estes (2003) give an overview of the best-known work in this area. They write, 'Language change is typically initiated by a group of speakers in a particular locle at a give point in time, spreading from that locus outward in successive stages ...' (Wolfram and Schilling-Estes 2003: 713)

We might expect all people to be equally involved in the initiation of changes, but Wolfram and Schilling-Estes quote Kroch (1978) as saying that the contemporary sociolinguistic consensus 'universally points to the working class and lower middle class as originators of sound change in contemporary American English.' By 'sound change,' Kroch appears to be referring to regular, reductive sound change; it is not clear to what extent the consensus view holds for reductive sound changes or to analogical changes applied to syntax. Wolfram and Schilling-Estes (2003: 729-732) cite a study by Bailey et al. (1993) that shows the English *don-dawn* merger spreading from higher-income urban areas in Oklahoma to lower-income rural areas, while the quasi-modal *fixin' to* appears to have spread from rural parts of the state to the urban areas.

The S-curves reported by Altmann et al. (1982), Kroch (1989) and others describe the rate of increase in multiauthor corpora, and it is assumed by many linguists that the increase is exclusively due to propagation of the change from one language user to another, possibly from one social class to another, and not due to the propagation across contexts described by Kroch. Labov (1994) makes this assumption, and also observed that the logistic function is the cumulative function of the normal 'bell curve' distribution. It may be true that the social propagation of a change produces an S-curve, but as quoted above, Kroch argues that the S-curves also reflect the increase of use within a text, or within an individual's production. Bailey (1973) and Altman et al. (1983) leave open the possibility that the S-curve may describe propagation across both individuals and contexts. Shen (1990) suggests explicitly that the change spreads simultaneously across the population of language users and across the lexicon, and this is confirmed by Chambers (1992) and Ogura and Wang (1996).

The sociolinguistic consensus described by Kroch (1978) and Schilling-Estes (2003) refers to spontaneous conversational registers, and it is possible that edited, normative genres like theater may not particiate in these changes to the same degree. At times, playwrights are also known to use forms that are associated with particular language communities, sometimes for realism, sometimes in an exaggerated way for humorous effect. In Chapter 6 I will test the corpus for evidence of these aspects of the spread of language change.

6 Hypotheses for the Study

There are several predictions that arise from the theories discussed in the previous chapters. In its original form, this study was designed to use data from the evolution of French negation to test the predictions of the usage-based model as laid out by Bybee and Thompson (1997), specifically the roles of type and token frequency. To these I have added the logistic model of propagation laid out by Kroch (1989), an application of the Lotka-Volterra model of competition, and Schwenter's (2006) application of Geurts' (1998) model of 'emphatic' negation as presupposition denial. Some of these will have sub-hypotheses to reflect particular related questions.

1. **Semantic/Pragmatic Evolution**: The embracing negation constructions evolved semantically and pragmatically from literal reference through scalar denial, emphatic denial, proposition denial and presupposition denial, to end with unmarked predicate negation.

2. **Logistic Propagation of Types**: The prevalence of the increasing negation constructions can be modeled with an S-shaped logistic progression.

   a. **Tokens per text**: This will be true for the proportion of tokens.

   b. **Types per text**: This will be true for the proportion of types (type frequency).

   c. **Rates for types vs. tokens**: The rate of change (as measured by the absolute value of the slope of the logistic curve fitted to the data) will be higher for tokens than for types.

3. **Type Frequency**: When constructions compete for a function, the change in type frequency for each construction in a given period will be predicted by the Lotka-Volterra model based on that construction's type frequency in the preceding period.

   a. **Conjugated verbs**: The correlation will be greater when conjugated verbs are counted as types as opposed to when main verbs are counted as types.

   b. **Autonomy of high-token-frequency items**: Because high-token-frequency items do not take part in schema formation, type frequency counts that exclude high-token-frequency items will correlate better with the increase in type frequency than counts that do not exclude high-token-frequency items.

   c. **Reliability of counts including hapaxes** (items that appear only once in a corpus): Because rare items in a small corpus may not be representative, type frequency counts that exclude hapaxes will correlate better than type frequency counts that include hapaxes.

   d. **Accuracy**: The accuracy of the Lotka-Volterra model will be greater than that of the logistic model, as measured by the R^2 coefficient of determination.

4. **Entrenchment of high-token-frequency items**: As the embracing negation constructions are extended to more and more lexical items, the items that resist the change tend to have higher token frequencies.

6.1 Testing the Hypotheses

I will now discuss each hypothesis in detail, including what theory it comes from,

how to test it against the corpus, and some of the challenges presented by the data.

6.1.1 Semantic/Pragmatic Evolution

The embracing negation constructions evolved semantically and pragmatically

from literal reference through scalar denial, emphatic denial, proposition denial and

presupposition denial, to end with unmarked predicate negation.

In Chapter 4 I discussed a number of theories about the semantic and pragmatic

evolution of negation constructions, including general principles laid out by Fauconnier

(1975), Ladusaw (1993), Traugott (1989), Horn (1989), van der Wouden (1994), Geurts

(1998), Croft (2000), Dahl (2001) and Israel (2001); and specific predictions made by

Jespersen (1917), Detges and Waltereit (2002), Schwenter (2006), Eckhardt (2006) and

Kiparsky and Condoravdi (2006). We have come up with the following general

semantic/pragmatic stages:

68) ```
    Literal reference > scalar denial > emphatic denial >
    proposition denial > presupposition denial > predicate
    negation.
    ```
From this, we can hypothesize that a meaning shift took place at some point in the

history of French. Before that shift, the preverbal construction was used to express basic

predicate negation while embracing constructions were used to deny active

presuppositions. After that shift, some of the embracing constructions were used to

express basic predicate negations, thus becoming essentially synonymous with the

preverbal construction.

How can we test this hypothesis? I have collected a CORPUS of texts covering the

period in question, and I will use it to test all the hypotheses in the study. The

Semantic/Pragmatic Evolution hypothesis is the most difficult to test, because it requires guessing the meaning and function of every token under study (Brinton 2005, 2008). It is difficult for an individual to know what they themselves meant by something they said recently, more difficult for another user of that language variety to know what that individual meant, and even more difficult for a person from another culture and time. One way to compensate for these difficulties is to have multiple coders cross-checking each other's work; unfortunately I did not have the resources to hire additional coders, so the results should be regarded as preliminary and subjective.

The first play in the corpus, the *Ordo Representacionis Ade*, is the oldest complete play[2] in French that is known to have survived to this day. In it and the thirteenth-century plays, there are already many uses of embracing negation constructions for proposition denial and presupposition denial. There are no instances where *ne ... pas* or *ne ... mie* are used to indicate scalar denial and no unambiguous instances of emphatic denial. Because of this, the corpus cannot tell us anything about scalar denial and the constructions we are studying. I therefore tagged every token of declarative sentence negation for whether it was an example of *predicate negation*, *presupposition denial* or *proposition denial*. If the hypothesis is true, I would expect to find a period where there were very few instances of embracing negations that were unambiguously tagged as predicate negation, followed by a period that was more mixed. There were many instances where I could not make a firm decision; the tokens were ambiguous, usually

---

[2] It is missing a bit at the end, but it is substantially intact, unlike all the others from the twelfth century.

between predicate negation and presupposition denial. In those cases, I checked both boxes to indicate the ambiguity.

It is important to note, however, that as I discussed in Chapter 4, Traugott (1989) and others claim that a significant amount of ambiguity is always present in language, and this ambiguity is one of the driving forces of semantic change. Much of the ambiguity I found is therefore predicted by Traugott's theory and an important part of the evolution of negation in French.

I will repeat here some of the examples from the Prelude and Chapter 4 to give the reader an idea of the standards I used for tagging. This is an example of proposition denial:

69)   EVA: Il te ferra changier saveir.
     ADAM: **Nel** fera **pas,** car nel crerai.
     'EVA: It will change the way you know things.
     ADAM: No it won't, because I won't believe it.'

Ordo Representacionis Ade, ca. 1160, lines 285-286.

Unambiguous presupposition denial:

70)   FIGURA: Or me mostre ton frere vif?
     [...]
     CHAÏM: [E] jo por quei le dei trover?
     Ja nel deveie pas garder.
     'FIGURA: Now show me your brother alive.
     [...]
     CAIN: And why is it me that has to find him?
     I wasn't required to watch him.'

Ordo Representacionis Ade, ca. 1160, lines 726, 729, 730

Negation ambiguous between presupposition denial and predicate negation:

71) Cain: J'ay trop peché villainement; je **ne** suis **pas**
    digne de vivre.
    'Cain: Wickedly, I have sinned too much; I am not fit
    to live.'

Mistère du Viel Testament, 1542.

Unambiguous predicate negation:

72) puisqu'ils **ne** sont **pas** venus, je m'en vay chez le
    greffier
    'Since they're not here yet, I'm going to the clerk's
    office'

Désidério Descombes, La Farce des bossus, 1623

The authors of the plays had other ways of marking emphasis and the forms of
denial mentioned above. As discussed in Chapter 3, in Old French putting the
"postverbal" item before the *ne* allowed the authors to place the focus of the sentence on
the negation. Sometimes the subject and the verb of a clause were inverted after an
adverb or conjugation, in order to conform to the verb-second constraint that was present
in French at the time. These adverbs, especially *si*, 'thus' and *or*, 'now, therefore,' could
also serve to mark emphasis, as in the following example, where the King of Hungary's
wife is dead.

73) Or **ne** la peut on trouver tele
    'And so we cannot find one like her'

Miracle de la fille du roi d'Hongrie, 1371

Questions and commands can also add emphasis independently of any of the other
strategies. I have thus excluded tokens with the fronted negators, with subject-verb
inversion and with interrogative or imperative contexts from tests for all the hypotheses.

There are other issues associated with corpus work. One of the major difficulties
is obtaining a representative sample of the texts in the language (Lee 1999; Grieve-Smith

107

2006), because without a representative sample it is impossible to generalize the results to the language as a whole. Language use varies widely, and there is a vast literature discussing variation according to social class (e.g. Labov 1966), region (Chambers and Trudgill 1988), age (Chambers 2003), gender (Tannen 1990), register and genre (Biber 1988). Without controlling for this variation, there is no way to know whether a difference in values for two different time periods is evidence of change over time or some other kind of variation.

It is especially important to account for variation according to register and genre. In addition to the physical and cognitive limitations that result from particular situations such as real-time interaction, various genres and registers have their own sets of norms that control what forms of language are acceptable. Forms from different registers or genres are not directly comparable, and showing two different forms (or frequencies) from different genres and times does not prove that there was change over time. Valli (1984: 142) criticized a historical study by Ashby (1981) for confusing normative and non-normative styles from different periods.

Another difficulty with studying the language of the past is that while there is widespread agreement (Labov 1972, Ochs 1979) that informal, spontaneous conversation is more directly affected than other registers and genres by the kinds of changes discussed in Chapter 5, and thus a more appropriate object for study, records of such conversation are very rare for periods before the invention of audio tape. I have chosen to restrict the corpus to theatrical texts, because they are the closest genre to spontaneous conversation that is available throughout the period under study. Unfortunately, because they are subject to multiple revisions to bring them in line with contemporary language norms and

audience expectations, theatrical texts are subject to TOP-DOWN pressures as well as the

bottom-up changes described in Chapter 5 (Caillet 1857, Degaine 1992). This requires

that we pay special attention to potential top-down pressures and limits the validity of

generalizing the results. The most we can say is that they can sometimes be generalized

to theatrical norms, but even then with extreme caution.

6.1.2 Logistic Propagation

The prevalence of the increasing negation constructions can be adequately

modeled with an S-shaped logistic progression. This will be true for both tokens and

types, but the logistic curve will be steeper for tokens than for types.

This hypothesis is based on the logistic model of Altmann et al. (1983) and Kroch

(1989). The logistic model is simpler than the Lotka-Volterra model, but it may be

accurate enough for many purposes. It can be tested by tagging all the tokens of

declarative sentence negation in the corpus, and marking the types used. It suffers from

the same sampling challenges as the Semantic/Pragmatic Evolution hypothesis, as do all

the other hypotheses. Unlike that hypothesis, it does not suffer from the difficulty of

interpreting semantic and pragmatic content.

It seems clear that published frequency counts will not be adequate

approximations to type or token frequency for this study. As has been established by

Gernsbacher (1984) and Poplack (2001), the corpora used to compile these type and

token frequency tables are not based on representative samples and thus do not yield

reliable measures of experiential familiarity or perceived generality. We also expect the

type and token frequencies of the various features to change over time, and thus require

counts specific to a time period of a century or less, but most of the published counts and

dictionaries are from the twentieth century, or for a longer period (such as 'ancien français,' e.g. Greimas 1969). For this reason I have chosen to rely on frequency counts taken directly from the study corpus.

As I discussed in Chapter 4, it makes no sense to apply propagation models to time periods where we do not expect any propagation to occur, so we can estimate dates for the beginning and end of periods of competition based on the results from the semantic/pragmatic evolution hypothesis. In addition to the difficulties of measuring language change, a precise measure of type frequency for a given individual (the speaker or author) requires knowledge of exactly what constructions that individual was exposed to over the course of their lives. We will never know exactly what type frequencies a given individual has, but we can find ways to approximate this experience. If we do not know much about the author's life, or if the text is anonymous, we can estimate that the author was the average age of an author during the relevant period.

The test poses another difficulty, that of counting types. The literature on type frequency is vague on the method of counting types; some investigators simply count adjectives, verbs or nouns (grouping together singular and plural instances of a noun, or different person and tense conjugations of a verb, for example). Bybee (1995) counted nouns that occurred with various pluralizing suffixes. Bybee and Eddington (2006) counted adjectives that appeared in predicate-adjective constructions. Bybee (1995: 437) argues that Clahsen and Rothweiler (1992) erred in counting verbs with separable prefixes (such as *ausschreiben* 'to write out') as distinct types from their roots instead of counting the root and the compound as a single type (in contrast, she counted *ausschreiben* as a variant of *schreiben* 'to write' and found a stronger effect of type

110

frequency on productivity). I have thus chosen to group all of the uses of a particular verb together, regardless of person, number, tense or reflexivity, but I have separated two high-frequency impersonal constructions, *il faut/il fallait/il faudra/etc.* 'to be necessary' and *il y a/il y avait/il y aura/etc.* 'to exist', and counted them as separate verbs from the full verbs *faillir* and *avoir* that evolved from the same sources.

To minimize the risk of inaccuracy associated with a relatively small number of tokens, I have endeavored to make my corpus as large as possible given the constraints of time and manpower. My goal was to have at least 500 tokens of negation per century, and I was able to exceed that value for every century after 1300. Because some plays were longer than others, some centuries had very large numbers of tokens, but since all of the measurements are independent of the overall number of tokens, this variation should not be an issue. The following table shows the number of tokens I was able to acquire for each century.

Century	Number of tokens
12th	66
13th	397
14th	438
15th	554
16th	1245
17th	700
18th	741
19th	1079
20th	487
Total	5707

Figure 8. Tokens of declarative sentence negation found in the corpus.

Even with this many tokens of negation, however, the number of negations measured during a given author's lifetime is still too low to provide accurate measurements. Because of this, I have extended the periods of measurement for the Type

111

Frequency Hypothesis and the Entrenchment Hypothesis to entire centuries. This is not ideal, but we can hope to find some usable data that way.

I will measure the degrees to which both type frequency and the proportion of tokens correlate with the values predicted by the logistic model. The logistic model can also be used to assign a rate to the linguistic change under study; I will compare the rates given for types and tokens, and test whether the rate for tokens is faster as predicted in Chapter 5.

6.1.3 Type Frequency

When constructions compete for a function, the change in type frequency for each construction in a given period will be consistent with the Lotka-Volterra model based on that construction's type frequency in the preceding period. The accuracy of the Lotka-Volterra model will be greater than that of the logistic model, as measured by the R^2 coefficient of determination.

In Chapter 5 I discussed the theories of grammaticization, reanalysis and analogical extension. I showed how the usage-based model of Bybee (1985, 1995), Langacker (1987, 1988), Bybee and Slobin (1982) and Bybee and Thompson (1997) gave us a satisfying explanation of the processes underlying analogical extension based on the effects of type frequency. Finally, I discussed the logistic models proposed by Altmann et al. (1983) and Kroch (1989), and my proposal to apply the Lotka-Volterra model to analogical extension.

This hypothesis is based on the resource competition model of Lotka (1925) and Volterra (1926). It can be tested by tagging all the tokens of declarative sentence negation in the corpus, and marking the types used. The type frequency can then be calculated for

112

each period by counting the total number of types used during that period, and taking the proportion of types that occurred with the construction in question.We would expect the values predicted by the Lotka-Volterra model to be at least as close to the observed values as the predictions of the logistic model. This test suffers from the same sampling challenges as the tests for all the other hypotheses. Like the Logistic Progression hypothesis, it is subject to the difficulty of counting types. Unlike the Semantic/Pragmatic Evolution hypothesis, it does not suffer from the difficulty of interpreting semantic and pragmatic content.

As with the type frequency measures used in testing the Logistic Propagation hypothesis, I will also test whether to count main verbs or conjugated verbs as types, whether to exclude high-token-frequency items, and whether to exclude hapaxes, I will compile the results for each of the eight possible combinations and compare the correlations.

6.1.4 Measuring Type Frequency

In the past, type frequency has not always been measured the same way, and several researchers have pointed out particular factors that may affect the reliability of type frequency counts. Because of this, in testing the Type Frequency hypothesis I will also test three of these factors to determine how much of a role they play, and then use the combination with the highest correlation in testing the Logistic Progression hypothesis.

One well-known feature of recent Romance languages, and Western European languages in general, is the use of compound verb tenses, particularly to mark past tense and the passive voice. Here is an example:

113

74) FIGURA: En vostre chois vus met e bien e mal!
    Ki ad tel dun, **n'**est **pas** lïez a pal.
    'FIGURA: In your choice I offer both good and evil!
    A person who has this privilege is not chained.'

Ordo Representacionis Ade, ca. 1160, lines 66-67

In this example of the passive voice, the embracing negation *ne ... pas* only embraces the conjugated auxiliary *être*, not the main verb *lier*, which appears as a past participle. There are also modal auxiliaries like *pouvoir*, 'to be able' and *vouloir*, 'to want,' as in this example:

75) FIGURA: Tu es mon serf, e jo ton sire.
    ADAM: [Jo] **ne** te puis **pas** contredire.
    'FIGURA: You are my serf, and I your lord.
    ADAM: I can't contradict you there.'

Ordo Representacionis Ade, ca. 1160, lines 405-406

Here the embracing negation embraces both the clitic object pronoun *te* and the modal auxiliary *pouvoir*, but not the infinitive *contredire*. Later in the evolution of this construction, people began to attach object clitics to the infinitive, but otherwise this construction and the passive construction are still widely used in present-day French.

It is not clear what to do with compound constructions like these in type frequency: are the conjugated verbs the ones we should be looking at, or the main verbs? For this study, in the case of past tense and passive constructions I have chosen to test this question by tagging both conjugated verbs and main verbs and testing which one fits better with the theory. In the case of modal auxiliaries I have chosen to count them and not the infinitives as main verbs. This is a question that deserves to be investigated in more depth in the future.

114

As I discussed in Chapter 5, Bybee (1995) claimed that high token frequency items were stored as chunks and thus do not contribute to the productivity of constructions. Hay (2001) argued that the ratio of the token frequency of the derived form (the morphological equivalent of the lexical item used in the construction) to the token frequency of the base form (the equivalent of the lexical item in all constructions) is a better predictor than type frequency for determining productivity. Hay and Baayen (2001) in turn suggest that the parsing ratio predicted by Matcheck is an even better predictor. It would be nice to be able to test all three predictions, but the scope of the project did not allow me to measure the non-negated frequencies of the verbs in question, or to construct an implementation of Matcheck. I will therefore generate a type frequency count that excludes the verbs with the highest frequency, and compare that count with the count that does not exclude the most frequent verbs.

An accurate measure of type frequency also requires a large number of tokens, and when there is not much data available for a given period it can distort type frequency figures. Baayen and Lieber (1991) drew attention to the number of *hapaxes* (forms that only appear once in a corpus) relative to a particular construction as indicators of the productivity of that construction. However, Baayen and Lieber clearly specify that this measure works best for corpora that contain thousands of tokens of the construction. Goldberg (2006) points out that if a particular verb can be used with either of two different constructions, but only occurs once in the corpus for a given period of study, we run the risk of counting that verb as a type occurring with that one construction instead of with both constructions. If we count all verbs regardless of their token frequencies, we wind up with type frequency percentages that closely parallel the percentages of tokens.

On the other hand, we have established that extension affects low-frequency items first, so if we eliminate too many low-frequency verbs we may wind up with a conservative sample. Because of this, I will make two counts of type frequency, one that includes the verbs that occur only once in negative contexts (hapaxes), and one that excludes them, and I will test the correlation of the model using both counts.

6.1.5 Entrenchment of High-Token-Frequency Items

As the embracing negation constructions are extended to more and more lexical items, the items that resist the change tend to have higher token frequencies. The token frequency of a given lexical item during a given period can be calculated by counting the number of tokens of that item during that period, and dividing by the number of words in the entire corpus during that period; the word count gives us a single measure to equalize all the token frequency measures, so as not to overemphasize the role of a particular playwright; they are an acceptable stand-in for the number of conscious moments that the average French playwright would have lived through while having opportunities to hear the lexical items in question.

This test also suffers from the relatively small corpus size. Although the prediction is that items with high token frequency will resist analogical extension, the token frequencies of individual verbs are still not very high, which allows small variations in the use of individual verbs to overwhelm the larger pattern. I have accordingly taken the same steps in counting token frequency as for the previous hypothesis, i.e. maximize the size of the corpus and count data by century instead of shorter periods.

116

Another challenge in this test is that the high-token-frequency items are not always single words. Later in this chapter I will discuss a number of other constructions that have been associated with preverbal *ne* alone. I will then try to assign a date for each verb as it shifts to one of the embracing negation constructions, and examine the connection between token frequency and the date of the shift.

6.2 The Corpus

Theatrical texts have much in common with spontaneous conversation. In a study of the evolution of the opposition between *be* and *have* in English past tense forms, Rydén (1991) writes, 'the comedies largely revealed themselves as more genuine reflexes of non-conservative language ...' They usually consist of conversation, and are designed to represent spontaneous interaction. Often times the lines were even composed extemporaneously and then written down later from memory (Degaine 1992). Of course, the playwright had time to edit the text to ensure that it conformed to (possibly artificial) standards, and the actors had time to memorize it. Until the seventeenth century, the dialogue in almost all French plays followed strong rules covering rhyme and meter, and many of them were set to music. Molière believed in using a *ton naturel* on stage, and in *Les précieuses ridicules* (scene X), he made fun of people who preferred the artificial diction used in the *Hôtel de Bourgogne* to comediens who *récitent comme l'on parle*. According to Degaine (1992: 216), this position did not contribute to Molière's success: although his comedies were popular, he longed to write tragedies, but his tragedies were commercial and critical failures.

It is periodically said of one playwright or another that they have succeeded in reproducing conversation in a way that other playwrights had not been able to before,

117

which implies that the dialogue of their predecessors was not very realistic. Even in the most poetic, formal play, however, there is still the intent to represent, on some level, people conversing with each other.

6.2.1 A Short History of French Theater

Crucially, there are samples of French theatrical texts available for the most important part of the period under study, from the *Ordo Representationis Ade* in the mid-twelfth century right up to the present day. Symes (2007) gives a good overview of the difficulties involved in studying the history of medieval French theater. Many performances were either improvised or committed to memory without being written down, and others were versions of religious texts, epic poems or other texts that were dramatized according to unwritten conventions, leaving us little or no evidence of how they actually were performed on stage. Many of the texts that were written were not preserved, and of those that were preserved, many continue to languish in the rare text collections of libraries, and have never been distributed in print or on the Internet.

The earliest theatrical texts that exist today are in the tradition of the DRAME LITURGIQUE, a dramatization of a biblical story that was performed by priests or monks inside a church. Other medieval genres were the MIRACLE and the MYSTÈRE, religious performances that became longer and larger over the centuries, until they lasted several days and included dozens of characters performed by a significant portion of the population of the town. There were also more light-hearted genres: the JEU, the SOTIE, the MORALITÉ and the FARCE. While the *jeu* had its own performances, the other genres were shorter and were often performed during intermissions of the *mystères*.

During the middle ages, theatrical performances were often organized by craft guilds; for example, all but one of the 52 fourteenth-century plays that have been preserved were written and performed by the Parisian goldsmiths' guild. From 1402 to 1548 they held a royal monopoly on mystery plays in Paris; similar monopolies were granted for the farces, sotties and morality plays. In the sixteenth century a number of factors conspired to transform the medieval theater. The wars of religion made it difficult to perform religious plays without offending large numbers of people, and a system of traveling theater troupes (who charged for admission) evolved to replace the sedentary annual festivals.

In the middle of the sixteenth century, the *Pléiade* led a revival of interest in classical theater forms. These intellectuals rejected the old medieval genres in favor of the classical genres of COMÉDIE and TRAGÉDIE, and their follower Robert Garnier added TRAGI-COMÉDIE. Unlike the medieval genres that catered to the masses with action-packed stories and quick dialogue, the new renaissance theater featured characters standing on stage delivering long declamations filled with Classical allusions. Joseph (1987: 143-144) argues that the long sentences and turns were developed in imitation of Ciceronian style; even though many of the Renaissance playwrights admired Seneca, it was not until the seventeenth century that the Senecan sentence structures were adopted. The *Pléiade* writers were students and professors, and the plays were performed by teenage students in the *collèges*. Also in this tradition, the first plays in prose were adapted from the Italian by Pierre de Larivey, but prose plays were not really successful until the following century.

The seventeenth century is known as the Golden Age of French Theater. Encouraged by the patronage of Cardinal Richelieu, by King Louis XIV and by the literary *salons* that emerged at the time, playwrights like Pierre Corneille, Molière and Jean Racine combined the Renaissance genres of comedy, tragedy and tragicomedy with elements of the farce and the mystery play to produce a distinctly French tradition. They were also strongly influenced by translations of older Spanish plays, and by a troupe of Italian performers brought to France in the 1570s by Queen Mother Catherine de Médicis.

Although there were two theaters in Paris for a short period in the early seventeenth century, Louis XIV merged them both into the company of the Comédie-Française, which retained a monopoly on theater in the city until the Revolution at the end of the eighteenth century. After the Revolution, theaters proliferated, particularly along the Boulevard du Temple, and a new genre, the melodrama, joined the other theatrical genres. Towards the end of the nineteenth and the beginning of the twentieth century, other genres emerged, such as the *opéra-bouffe* and the *theater of the absurd*.

6.2.2 The Effect of Versification on Negation

As mentioned above, the majority of the texts written before 1700 were in verse; after that, most of them are in prose. To check whether this affects the results, the following table compares the proportion of declarative sentences that used preverbal *ne* alone instead of an embracing construction in three contemporary pairs of plays, one in prose and one in verse:

Date	Title	Author	Versification	Percent of *ne* alone	Total number of negations
1578	Bradamante	Garnier	Verse	66.7%	226
1579	Le Laquais	Larivey	Prose	75.6%	289
1622	Œvres de Tabarin	Tabarin	Prose	27.3%	68
1631	Clitandre	Corneille	Verse	42.0%	123
1701	Atrée et Thyeste	Crébillon	Verse	20.8%	102
1728	On ne badine pas avec l'amour	Marivaux	Prose	12.1%	250

Table 7: Proportion of negative declarative sentences produced with preverbal *ne* alone in pairs of contemporary prose and verse plays, 1550-1750.

Note that in each pair, the use of *ne* alone was relatively comparable whether the plays were written in prose or verse. Because of this, I decided to treat prose and verse plays as equivalent for this study.

6.2.3 Controlling for Regional and Class Variation

Regional variation poses certain problems for this study. By some criteria, many of the Romance varieties spoken in France could be considered separate languages, and the varieties that are strongly differentiated in general also show significant differences from standard French in the field of negation. After 1600, almost all plays were written in a national standard based on Parisian French, although some individual characters spoke in particular dialects if it was relevant to the plot. The difficulty for studying the language before 1600 is that there was so little written down, and restricting the corpus to a particular language variety would result in a corpus that was too small. For example, only ten plays written before 1300 are known to exist now. Of these, four are written in Picard, two in Walloon, two in Parisian, one in Norman and one in Limousin. Since there are many more differences between *langue d'oc* and *langue d'oïl* varieties than within each group, I have chosen to exclude the one *langue d'oc* play, the *Sponsus* of Limoges, from the study but to use plays from any *langue d'oïl* variety written before 1500.

Variation according to social class is a fact of language, and has been documented for other changes in French negation (Ashby 1981). Playwrights, especially good playwrights, can be expected to write dialogue that reflects this variation (Lodge 1991). On the other hand, accurate theatrical representations of lower-class speech have been publicly criticized (this was one of the criticisms of Molière mentioned above), so there has been a certain amount of pressure to either not write about lower class characters, or to modify their speech.

To check whether the social class of characters played a role in determining the negation that they used, I extracted information about the use of negation for the character that produced the most negations in each play, discarding those plays where no character produced more than nine negations. Based on information in the play and external information about historical characters, I assigned each character to class 1 (peasants, prostitutes, beggars, professional gamblers, acrobats, etc.), 2 (merchants, artisans, gentlemen, ladies, clergy and generic characters) or 3 (royalty, nobility, gods and choruses). By subtracting the prevalence of these proportions from an idealized progression function (see the Results section), I isolated the variation that was not due to change in progress. I found no meaningful correlation between that variation and the social class values I had assigned ($p = 0.11$). The full data is available in Appendix C.

Despite this lack of correlation, there are occasional characters who speak non-representative language varieties. In the farces attributed to Tabarin, for example, the character Piphagne speaks something between Italian and Italian-accented French, and Captain Rodomont speaks a similar Spanish-influenced pastiche. In Pierre de Larivey's *Le Laquais* (an adaptation of Lodovico Dolce's *Il Ragazzo*), the pedant character Lucian,

when not speaking in Latin, uses preverbal *ne* alone in 94% of his declarative sentences (46 out of 49), compared to the 77% rate for the entire play. I have thus excluded the pseudo-foreign negations made by characters like Piphagne and Rodomont from tagging, and removed the tags of Lucian's speech from analysis, bringing the overall rate of preverbal *ne* alone down to 74%.

6.2.4 Sampling Challenges

Any introduction to quantitative social science (e.g. Hopkins, Hopkins and Glass 1996, chapter 9) will stress the importance of population sampling. In order to make generalizations that apply to the population in question, the measurements have to be taken from data that are equivalent to the population, in other words, either the entire population has to be measured, or the sample taken has to be representative. The most common way to obtain a representative sample is through random selection, but some researchers use stratified random sampling to ensure that all subsets of the data are studied in proportion to their occurrence in the population.

I set the following targets for the corpus:

- representation from at least every century, and preferably every twenty-five year period
- representation from every subgenre of theater, ideally in every quarter-century when plays were written in that subgenre
- a minimum of half a million words total in the corpus
- a minimum of 500 tokens of negation per century

For the twelfth and thirteenth centuries, sampling is not an issue at all for this corpus: the twelfth century has only one surviving play in *langue d'oïl* (the *Ordo*

123

Representacionis Ade), and the thirteenth century has only ten surviving plays. It is only in the subsequent centuries that there is more data available than can be analyzed in the time period available for this study. The course of history has instituted its own bias; we don't know how many plays were performed but not written down, or written down but not preserved.

In the fourteenth century, sampling is very straightforward. As discussed in the background chapter, there is one mystery play and forty miracle plays from that century. The texts of all the miracle plays and their estimated dates of creation are available from the *Laboratoire de Français Ancien* at the University of Ottawa. I compiled a list and took random samples from each quarter century. Since the first play in the collection was composed no earlier than 1339, I sampled a fourth play from the period 1325-1349 to ensure that I had at least five hundred tokens of sentence negation from that century.

Sampling from subsequent centuries is much more difficult. There is no standard, complete list of plays published or performed in the fifteenth, sixteenth, twentieth or twenty-first centuries. The CESAR database from Oxford Brookes University is a compilation of all performances recorded between 1600 and 1800, but it does not contain complete information about when the plays were first performed or published or whether the text of a given play is available in any form. The nineteenth century possesses an invaluable resource in the five-volume *Parisian Stage* (Wicks 1950, 1953, 1961, 1967, 1979), but it is only available in print form, and suffers from the same interface difficulties as the CESAR database.

A more limited resource was the ARTFL/FRANTEXT corpus. This database was compiled in the 1950s, 60s and 70s for the *Trésor de la langue française* dictionary. The

project involved entering the full texts of thousands of classic works of French literature spanning roughly the period 1550-1950, and saving them on punch cards and paper tape. Beginning in the 1980s, the University of Chicago[3] and the French *Centre National de la Recherche Scientifique* collaborated to convert this corpus to magnetic media, proofread and clean up the texts and make the information available on the Internet.

The *Trésor de la langue française* editors did not sample texts at random, but rather selected them carefully based on what they felt were good models of French usage. Because of this, it is likely biased towards normative usage, and should not be taken as representative of written usage in general, let alone wider usage throughout the language. However, because all of the texts are available on the World Wide Web (except for some texts that are still under copyright), it was the most usable resource. The following chart gives an idea of the uneven distribution of plays in the database. Note, for example, that the period 1650-1674 contains 59 plays by only six authors, while the period 1750-1774 contains 39 plays by twenty authors.

---

[3] In 1993 and 1994, I worked as a research assistant on this project, proofreading texts for later use, although I do not recall proofreading any of the texts used in this corpus.
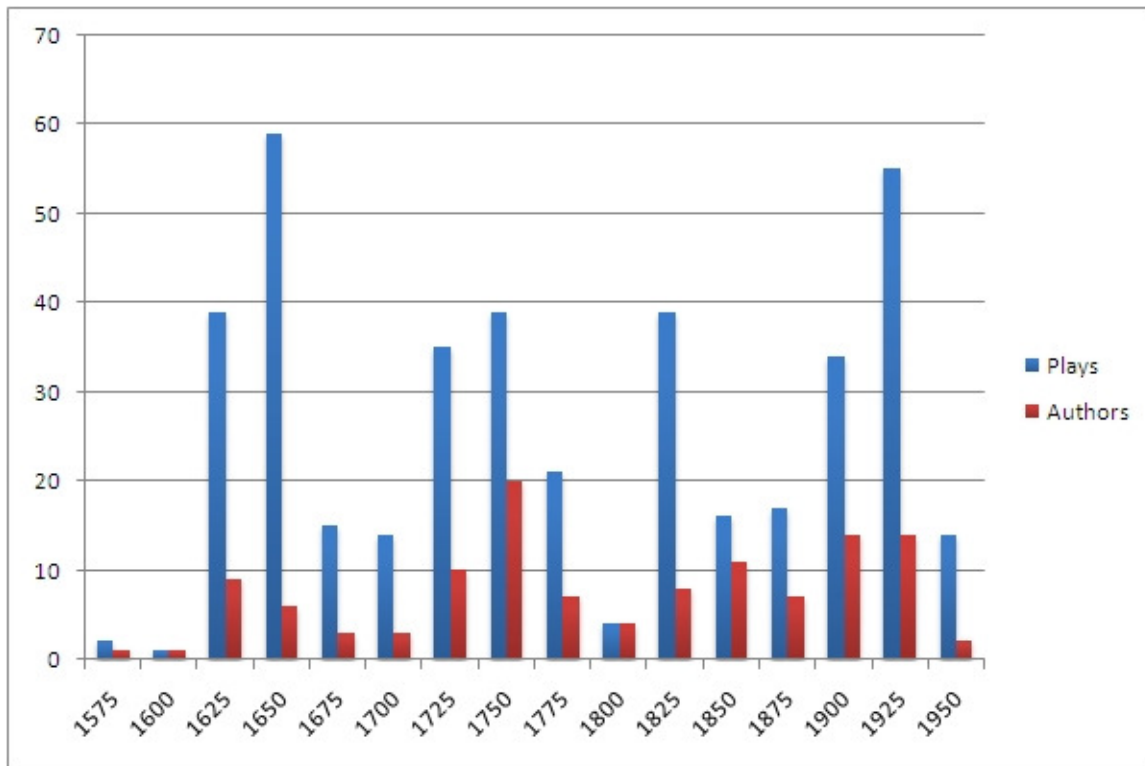
Figure 9: Distribution of plays and authors in the ARTFL/Frantext corpus.

As with the list of fourteenth-century miracle plays, I selected one play at random from the ARTFL database for every twenty-five year period between 1575 and 1949. Almost all of these plays were available in full text form from Gallica, the website of the *Bibliothèque Nationale de France*, but if one was not, I chose another play at random from the same time period. As I will discuss in more detail in the results chapter, the results from three of these plays were discarded because they were found to be unrepresentative of the corpus for various reasons. One play discarded from the 1700-1724 period was replaced with another play chosen at random, and the plays discarded from the 1575-1599 and 1600-1624 periods were replaced with plays from outside the ARTFL corpus chosen for their subgenre membership.

For the fifteenth and early sixteenth centuries, I performed an extensive web search for authoritative lists of plays written during those periods, but was not able to find

126

anything complete. Instead, I compiled a list from a number of lists found in various books and web pages. I searched for full-text and scanned versions of all of these lays, by title and author, where known. I used almost all of the plays that were available in full-text formats, but because they contain more, and longer, mystery and miracle plays than farces, soties and moralities, and almost no Renaissance comedies or tragedies, I supplemented them with scans of printed plays that were available online. This period is probably the most vulnerable to sampling bias, because in order to be usable to me, they had to have been preserved, reprinted in modern type, scanned, in some cases converted to full text, and hosted online. Each of those steps requires time, effort and money, and each allows various people to exercise choice in selecting or rejecting particular plays. A detailed list of the plays used and their sources is in Appendix A.

6.2.5 Preparing plays for tagging

When I acquired the texts for use in the corpus, they were in one of three formats: printed, scanned images or full text. For this study I required each text to be in full text, readable by both humans and computers, and reasonably accurate. I also required the French dialogue, the object of interest, to be clearly separated from all stage directions, letters and books read aloud, quotations, and dialogue in other languages. Finally, to check for effects of dialogue and social class I required the speech of each character to be separately identified.

There was one text that I knew would be indispensable for this corpus, the *Ordo Representacionis Ade*. This is the only theatrical text available from the twelfth century in *langue d'oïl* (apart from a few other liturgical dramas that are primarily in Latin with a few lines in the vernacular). Although it is now available through Google Book Search, at

the time I initially collected the corpus, it was not available online at all. This play was so important that I bought a modern academic edition and scanned it myself.

Many of the other texts I used were not full-text, but distributed on the Internet as scanned images of each page compiled into a Portable Document Format (PDF) file. My two primary sources for these documents were the BNF's Gallica and Google Book Search. I converted each of these scanned documents to full text using a commercial optical character recognition (OCR) program. The accuracy of the OCR varied depending on the clarity of the initial printing and the quality of the scanner and the scanning process, but all of the texts needed a certain amount of cleanup and proofreading.

The OCR program that I used was set to preserve paragraph breaks, but did not automatically recognize character turns or stage directions. Some of the full-text plays available on the Internet did mark these features consistently, but others were formatted for visual presentation and required a significant amount of work to fit the requirements of this study.

6.2.6 Tagging and Tabulation

Once a play was in an adequate full-text format, I loaded it into a custom-made tagging program written in the Hypertext Preprocessor language (PHP). This program identified all instances of *non, ne, n', né, nel* and *nen*, and placed a set of menus next to each one where I was able to specify various features, including what negation construction this word was part of (or if it was identified in error), the main and auxiliary verb and the tense, mood, person and number of the clause, as seen in the following screenshot. Once the play was tagged and double-checked (even the shortest plays required multiple rounds), another custom program imported the tags into a MySQL

128

database. I wrote a suite of PHP scripts that could be used to query the database in order

to count the use of various negation constructions and tabulate the results that I will

discuss in the next chapter.

Figure 10. The PHP-generated form used for tagging the corpus.

In my counts, I included both main and subordinate clauses in declarative

sentences that were negated with any of these four constructions: preverbal *ne* alone,

ne ... pas, *ne ... point* and *ne ... mie*. I did not include imperative and interrogative

sentences, since their semantics and pragmatics are so different from declaratives. I

excluded clauses where *pas, point* or *mie* were used with the negative conjunction

ne ... ni. If there were multiple negators in a clause (such as *ne ... pas rien*), I tagged the

first one. I excluded all instances of "expletive" negation as described in the Background

section, as well as the following:

- *ne ... nient*
- *ne ... mot*
- *ne ... goutte*
- *ne ... chose*
- *ne ... unques*
- *ne ... jamais*
- *ne ... ja*
- *ne ... mais*
- *ne ... plus*
- *ne ... que*
- *ne ... personne*
- *ne ... guère*
- *ne ... ni ... ni*
- *ne ... ni*
- *ne ... nul*
- *ne ... aucun*
- *ne ... rien*
- *ne ... nulle part*
- *ne ... homme*
- *ne ... autre*

To test the Semantic/Pragmatic Evolution Hypothesis, I tagged each instance of

sentence negation for the features **predicate negation**, **proposition denial** and

presupposition denial. As I mentioned earlier in this chapter, even when a native

speaker is tagging for semantic or pragmatic information there can be significant amounts

of ambiguity. For a non-native student of a language it is much more difficult to know the

intent of the speaker or writer, but that non-native student may at least have had

experience using the language in interactive contexts, and can usually fall back on the

intuitions of native speakers. When the period of study is several centuries in the past, the

intuitions of living native speakers cannot be considered authoritative, if they are even

applicable at all. As Brinton (2005, 2008) observes, this limits our ability to judge the

pragmatics of such texts with confidence.

Because of this, when tagging the instances of negation for pragmatic features

such as *predicate negation* or *presupposition denial*, I designed the script to allow me to

mark instances of negation as ambiguous between the two readings. I tagged the

unambiguous ones where possible, but wherever I had any doubts I marked it as

ambiguous.

Semantic tagging of corpora is often done by multiple individuals in order to

guard against individual bias. For this study I did all the tagging by myself and did not

have the resources to have another person tag any of the features. For that reason, any

tagging that depends on individual judgment should be considered very preliminary.

When interpreting the results, it is important to keep in mind that the corpus offers

only limited generalizability. Results from the texts in the ARTFL corpus could

conceivably be generalized to the theatrical subset of the ARTFL corpus as a whole, but

since that is not representative they cannot be relied on to tell us much about theater as a

whole, let alone any other genre. At most, they should be taken as promising suggestions

about what we might find if we are able to get a larger, more representative sample.

6.2.7 A Tagging Example

To illustrate the tagging procedure, I have chosen one example from the corpus to

discuss in detail.

76) FIGURA: De ta côste l'ai [jo] formee,
     N'est **pas** estrange, de tei est nee.
     'From your rib I have made her;
     She is not a stranger, from you she is born.'

Ordo Representacionis Ade, Lines 17-18, page 11

This is the first instance of an embracing negation construction, *ne ... pas* in the text. It is part of a poetic speech about how Eve is to be Adam's companion, and how the two should treat each other well. Since there are only two people in the world and this one is not Adam, it is reasonable for God to assume that Adam could be thinking that Eve is a stranger, so I classified this as *presupposition denial*. The negated verb is *être*, 'to be', and it is in the present indicative, so I marked the tense and mood and typed "être" into both the conjugated and main verb fields. It is third person singular, but the pronoun is dropped, so the subject is not a pronoun and the subject and verb could not be inverted. The *pas* is not fronted or topicalized, and there is no *adversative* construction such as *mais* or *ains*. This is not a transitive sentence, so there is no object pronoun, and it does not use a partitive or passive construction. The negative is not an expletive, quantifier, conjunction or interjection. It is also not a fixed locution, or one of the constructions discussed by grammarians such as Maupas (1607) or Ewert (1943).

7 Results

The data analyzed for this study provide substantial support to the hypotheses within the particular corpus collected. All of the major hypotheses are borne out by the data: the hypothesis of Semantic/Pragmatic Evolution, the hypothesis of logistic propagation of types, the type frequency hypothesis and the entrenchment hypothesis. The application of the Lotka-Volterra model was shown to be a good predictor of the prevalence of types whether conjugated verbs or main verbs were counted as types. It was also a better predictor when high-token-frequency verbs and hapaxes were included in the counts. The logistic propagation of the changes was demonstrated for both the prevalence of types (verbs) and the prevalence of tokens, and the rate of change was shown to be higher for tokens than for types. There were several high-token-frequency verbs that were slow to change to the embracing negation construction.

Tests of statistical significance are common in quantitative studies like this; their function is to determine whether it is appropriate to generalize results. Unfortunately, I have not been able to show that the corpus itself is representative of the genre of French theater, much less the language as a whole. Because of this, the results are only applicable to the specific texts analyzed, and any generalizations beyond that are extremely tentative. I will therefore dispense with significance tests, and only provide tests of effect size where available.

7.1 Semantic/Pragmatic Evolution

In Chapter 4 I discussed a number of hypothesized changes in the semantic and/or pragmatic roles filled by the various constructions under study. The main hypothesized

change is that **the embracing negation constructions evolved semantically and pragmatically from literal reference through scalar denial, emphatic denial, proposition denial and presupposition denial, to end with unmarked predicate negation.**

As described in the previous chapter, the corpus that we have does not include examples of *ne ... pas*, *ne ... point* or *ne ... mie* used for literal reference, scalar denial or emphatic denial. I therefore tagged each instance of sentence negation for the features **predicate negation**, **proposition denial** and **presupposition denial**.

If the hypothesis is true, I would expect to find a period where there were very few instances of embracing negations that were unambiguously tagged as *predicate negation*, followed by a period that was more mixed. The following chart and table show the ways in which the semantics and pragmatic functions fulfilled by *ne ... pas* changed over time in the corpus. I have focused on the embracing constructions with canonical sentence structure, excluding questions, commands, clauses where the particle appears before the verb, and clauses where the subject and verb were inverted, because all of these strategies can be used to add further "emphasis" to a negation. The following table gives the raw results for this test.

Negation Semantics/Pragmatics by Type								
Century	Type	Predn	Predn + Propd	Propd	Predn + Prsupd	Propd+Prsupd	Prsupd	Total
1100	alone	35	1	1	16		1	54
1100	pas			6	2		7	15
1100	mie							
1100	point							
1200	alone	109	1	3	148	1	27	289
1200	pas		1		18		20	39
1200	mie			1	23		39	63
1200	point				4		10	14
1300	alone	150		1	132		13	296
1300	pas	1		1	19		54	75
1300	mie				6		22	28
1300	point	1		1	5		22	29
1400	alone	167		3	132		40	342
1400	pas	2			20		92	114
1400	mie			1	2		7	10
1400	point			1	9		65	75
1500	alone	458		1	302	1	110	872
1500	pas	5			26		167	198
1500	mie				1		13	14
1500	point	5		1	40		111	157
1600	alone	165			128		23	316
1600	pas	12			91		108	211
1600	mie							
1600	point	11			82	1	106	200
1700	alone	58			82		5	145
1700	pas	15		8	153	1	219	396
1700	mie							
1700	point	3		2	69	1	133	208
1800	alone	32		1	59		7	99
1800	pas	77		35	282	2	366	762
1800	mie							
1800	point	2		1	16		18	37
1900	alone	3			27		3	33
1900	pas	13		12	226	3	194	448
1900	mie							
1900	point	1			4		5	10

Table 8. Semantic and pragmatic functions of the four predicate negators in declarative sentences over time.

Note how few of the negations were tagged as expressing proposition denial, even

ambiguously. After the *Ordo representacionis ade*, where the eight tokens constitute

11.6% of all negations, they never make up more than five percent of the negations in a century. This indicates that it is not very likely that proposition denial played a direct role in the evolution of embracing negation constructions. The following chart shows the proportion of tokens of *ne ... pas* that were used for the various pragmatic functions.
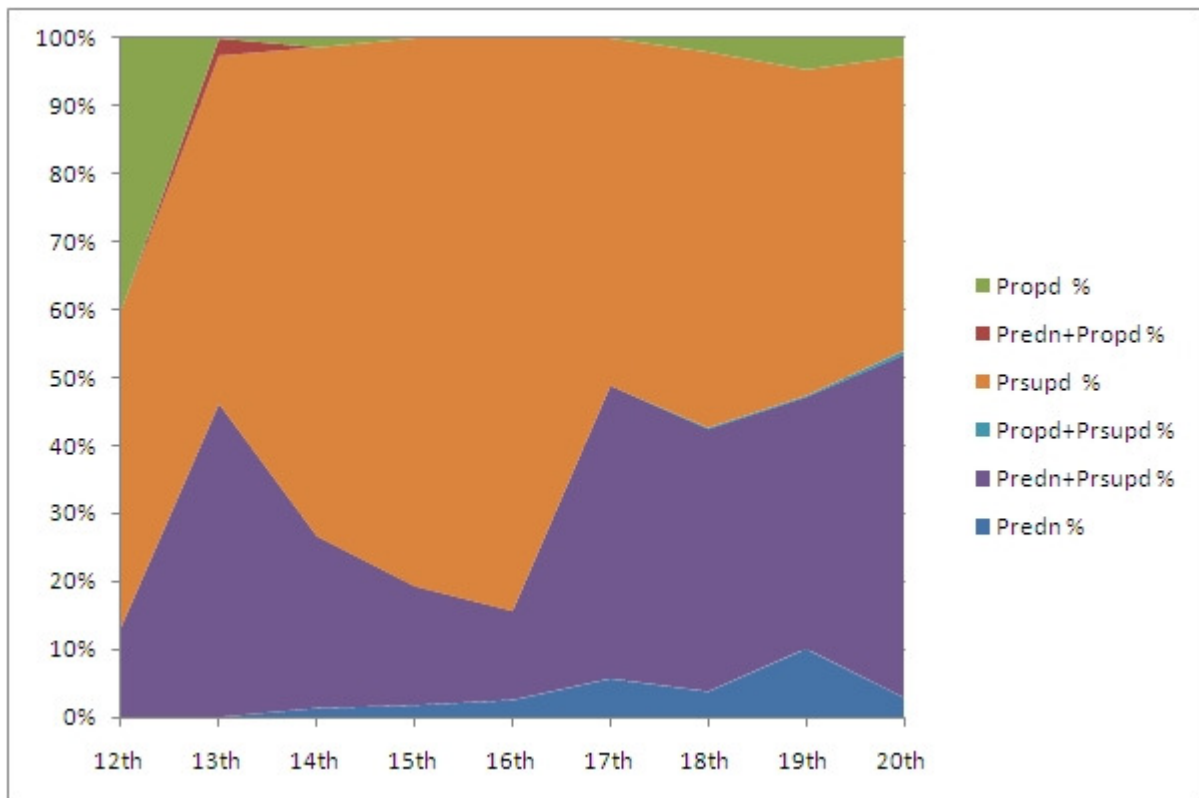

Figure 11. Semantic and pragmatic functions of *ne .. pas* in declarative sentences over time.

This data does not contradict our hypothesis, but it does not overwhelmingly support it either. The unambiguous instances of predicate negation are too few (less than ten percent) to say for sure that *ne ... pas* made inroads in that area. There is a relative increase in tokens of *ne ... pas* in contexts that are ambiguous between predicate negation and presupposition denial between the sixteenth and seventeenth centuries, and a corresponding decrease in unambiguous presupposition denials. This could just be an increase in ambiguity, but it may be evidence of an increase in predicate negations. There

is also a short-term increase in presupposition denials in the 13th century, but that could

be an artifact of the dialect variation and the small, unrepresentative sample.

A clearer picture emerges if we look at the constructions used to represent

predicate negation.
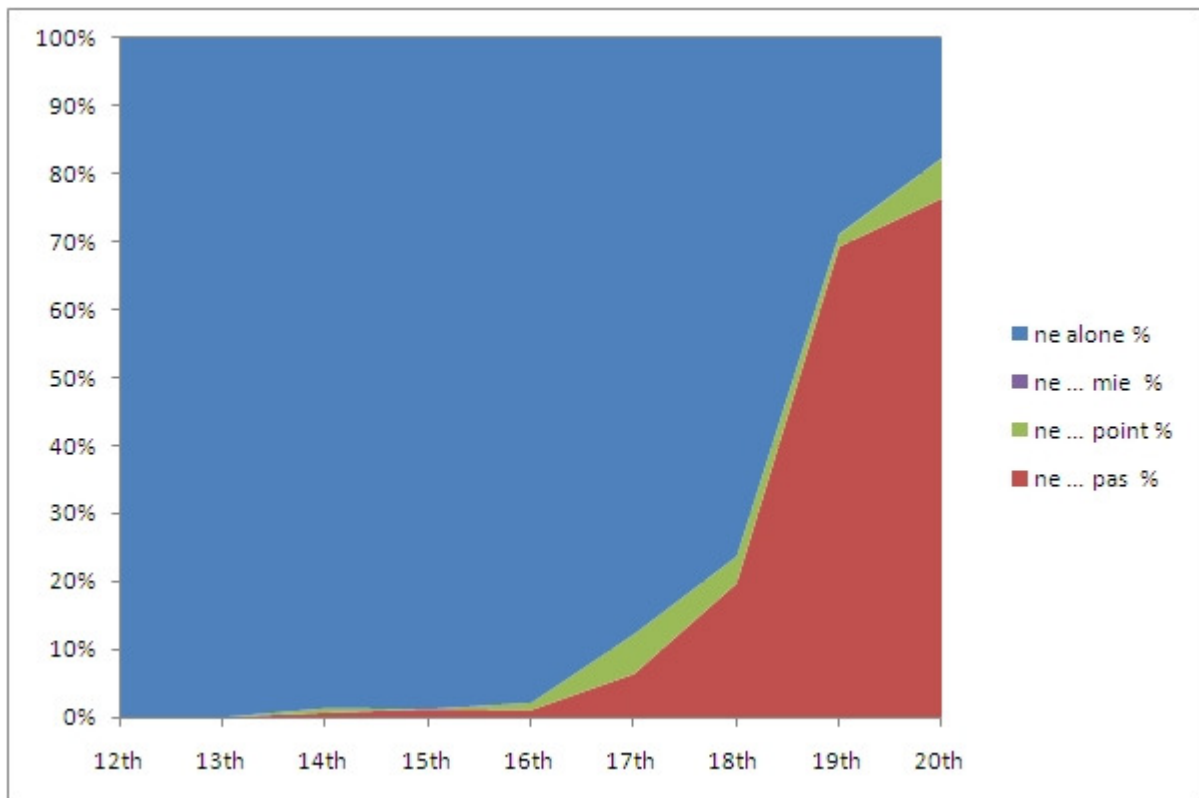
**Figure 12. Syntactic realization of declarative sentences that unambiguously express predicate negation, over time.**

In unambiguous contexts, predicate negations are overwhelmingly expressed with

preverbal *ne* alone until the beginning of the seventeenth century, when playwrights

begin to express them with *ne ... pas*. A small number are expressed with *ne ... point*, but

this is never more than ten percent.
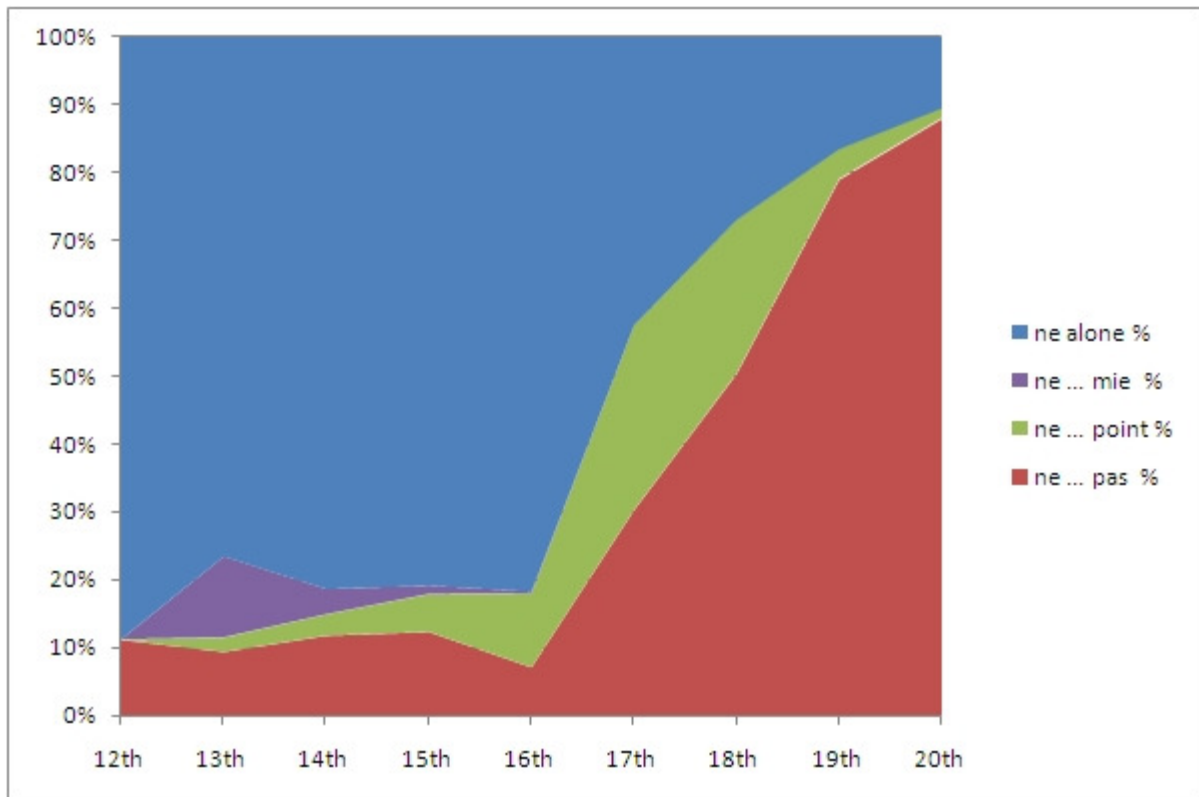
Figure 13. Syntactic realization of declarative sentences in contexts that are ambiguous between predicate negation and presupposition denial, over time.

With the sentences that are ambiguous between predicate negation and presupposition denial interpretations, before 1600 there is always a small minority of tokens that are expressed with *ne ... pas*, *ne ... point* and *ne ... mie*. Beginning in the seventeenth century, there is a large increase in the use of *ne ... pas* and *ne ... point* for these uses, much more dramatic than the seventeenth-century increase for the unambiguous predicate negations.

As I discussed in the chapter on negation in linguistic theory, Traugott (1989) argues that semantic change proceeds through ambiguity and polysemy. It is thus not surprising that the increase in use of *ne ... pas* and *ne ... point* would be used in ambiguous contexts first. This cannot be taken as definitely supporting Traugott's

138

hypothesis, because of the unreliability of the tagging methods I used, but it shows

promise for more rigorous approaches in this area.
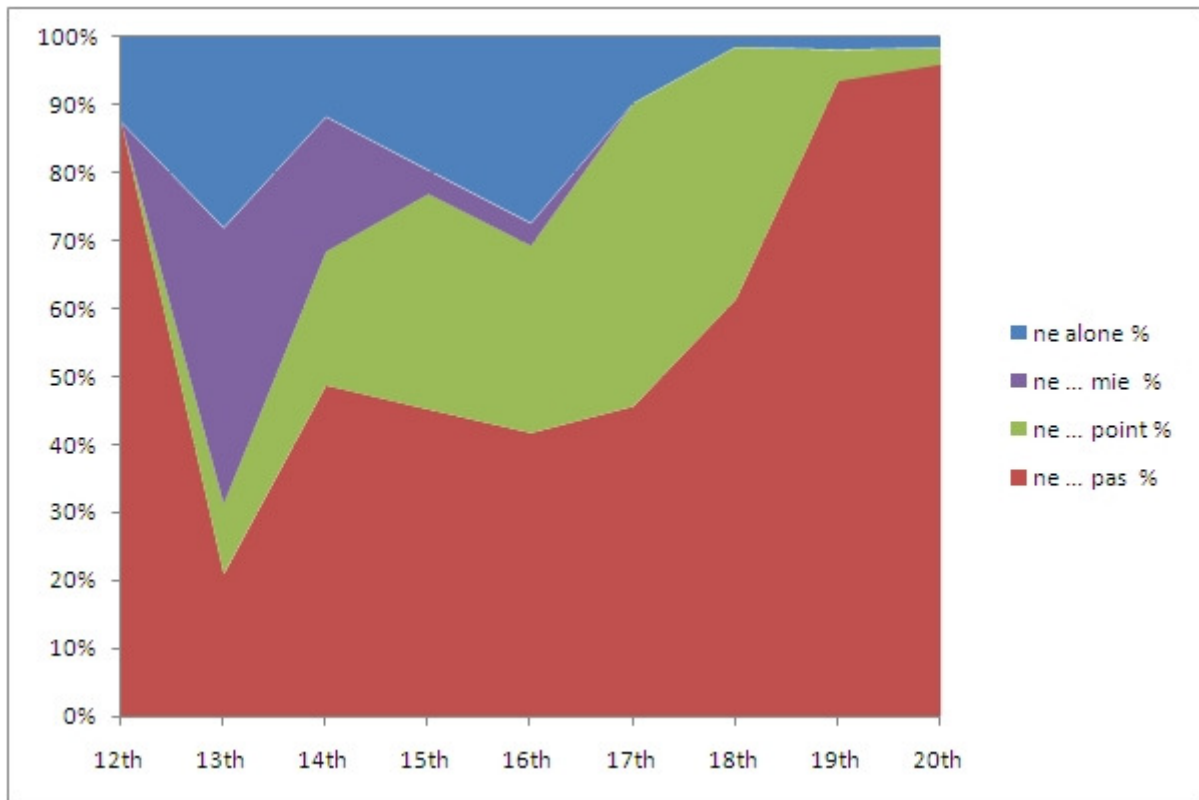
Figure 14. Syntactic realization of declarative sentences in contexts that unambiguously represent presupposition denial, over time.

This chart shows *ne ... pas* and *ne ... point* gaining in usage at the expense of

ne ... mie in the fourteenth and fifteenth centuries, and *ne ... pas* taking over usage from

ne ... point in the eighteenth and nineteenth centuries.

It is also important to note that up through the sixteenth century, there is a

significant number of presupposition denials that were expressed using a variety of

constructions, including preverbal *non* alone, *non ... pas*, various inversion constructions,

and topicalized constructions of the form *pas ne*, *point ne* and *mie ne*. Most of these seem

to have shifted to *ne ... point*.

139

In sum, before the seventeenth century predicate negations and ambiguous sentences were overwhelmingly expressed by preverbal *ne* alone, while unambiguous presupposition denials were expressed with one of the embracing negation constructions. Beginning in the seventeenth century, writers began to use *ne ... pas* more often, first in the ambiguous sentences and then in the unambiguous predicate negations. Based on this, we can estimate that the metanalysis of *ne ... pas* from presupposition denial to predicate negation took place some time around 1600 for theatrical texts. The two constructions *ne ... mie* and *ne ... pas* may have been considered "the same" as early as the fourteenth century, although this data may be clouded by dialect mixtures.

The embracing construction *ne ... point* seems to have been considered "the same" as *ne ... pas* beginning in the eighteenth century. It began to be used more and more frequently in contexts where *ne ... point* had been common. This is consistent with the theories of Traugott (1989), Eckardt (2007), Kiparsky and Condoravdi (2006) and Schwenter (2006) that postulate an "emphatic" stage of *presupposition denial*.

Based on this data, we can estimate the following dates for the hypothesized competitions in the use of negation:

Pragmatic Function	Constructions	Start Century	End Century
Presupposition denial	*ne ... pas*, *ne ... point*, *ne ... mie*	12th	16th
Presupposition denial	*ne ... pas*, *ne ... point*	17th	19th
Predicate negation	*ne* alone, *ne ... pas*	17th	20th

Table 9. Estimated competitions among constructions for negation functions.

These date ranges will allow us to formulate the tests of the Type Frequency, Logistic Propagation and Entrenchment hypotheses.

7.2 Logistic Progression

In the previous section, we concluded that there was likely a metanalysis of the embracing construction *ne ... pas* from "emphatic negation" to general negation, some time around the year 1600. Croft's (2000) theory predicts that after this metanalysis, *ne ... pas* covers the same function as preverbal *ne* alone; the two constructions are thus in competition with each other, and the speech community gradually shifts to the one with higher type frequency, even as its type frequency grows.

The second hypothesis of this study was thus that **the prevalence the increasing negation constructions will follow a logistic progression**, following the predictions of Kroch (1989), modified to include propagation across the lexicon. There are three sub-hypotheses, that the logistic progression would be found for both tokens (instances of use, regardless of which verb) and types (distinct verbs, each counted once), and that the rate of change would be higher for tokens.

7.2.1 Progression by Proportion of Tokens

The aim of this hypothesis is to replicate the findings described by Kroch, for example in this chart that Kroch (1989: 208) adapted from Oliveira e Silva (1982), showing an S-curve for the prevalence of tokens of possessive adjective constructions containing definite articles:

RISE IN THE USE OF THE DEFINITE ARTICLE BEFORE
POSSESSIVES IN CONTINENTAL PORTUGUESE
(adapted from Oliveira y Silva 1982)
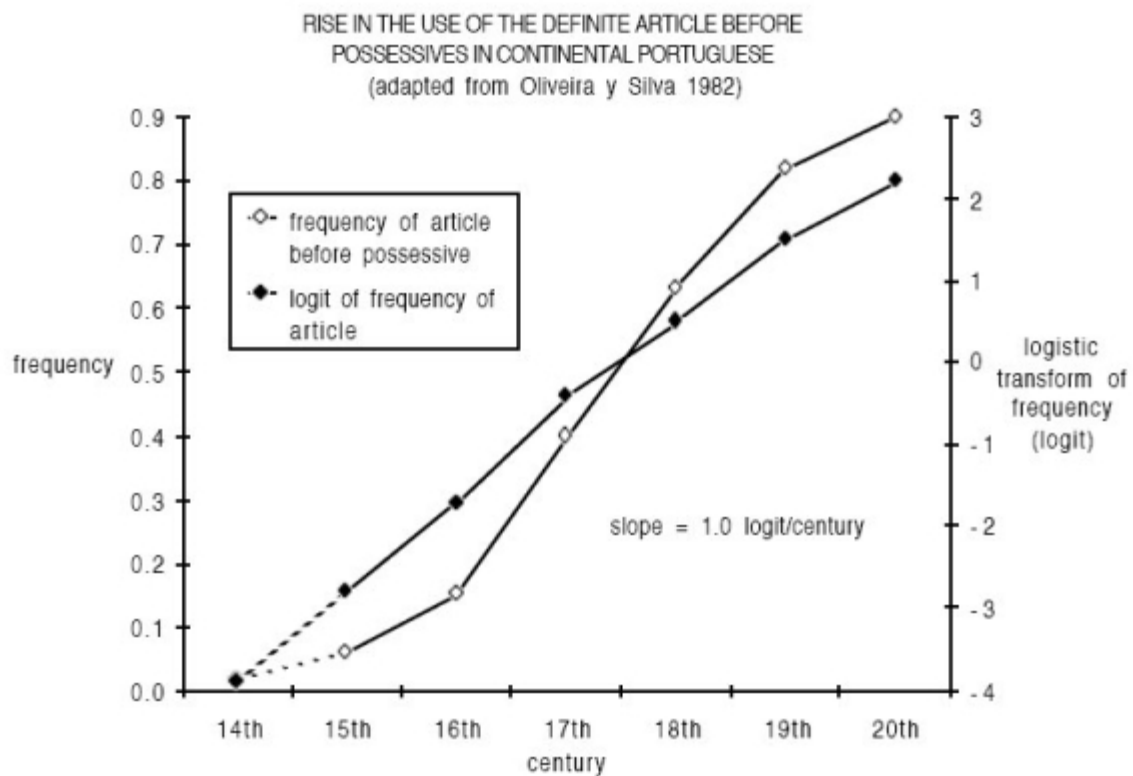
Figure 15. Chart from Kroch (1989: 208) using the logistic function to model the increase in possessive adjective constructions containing definite articles in Portuguese.

The S-curve in Kroch's chart is accompanied by a straighter line repreesenting the

logit of the frequency data. Remember that the logit transformation is performed by

applying the following function to the data:

$$18)\ x(t) = ln(x/1\text{-}x)$$

The logit is the inverse of the logistic transformation, and allows us to assign a

slope to the data, and to use linear regression to measure how well the logistic model fits

the data. A similar S-curve appears in the following chart representing the prevalence of

the four negators we are looking at in our corpus of French:

142
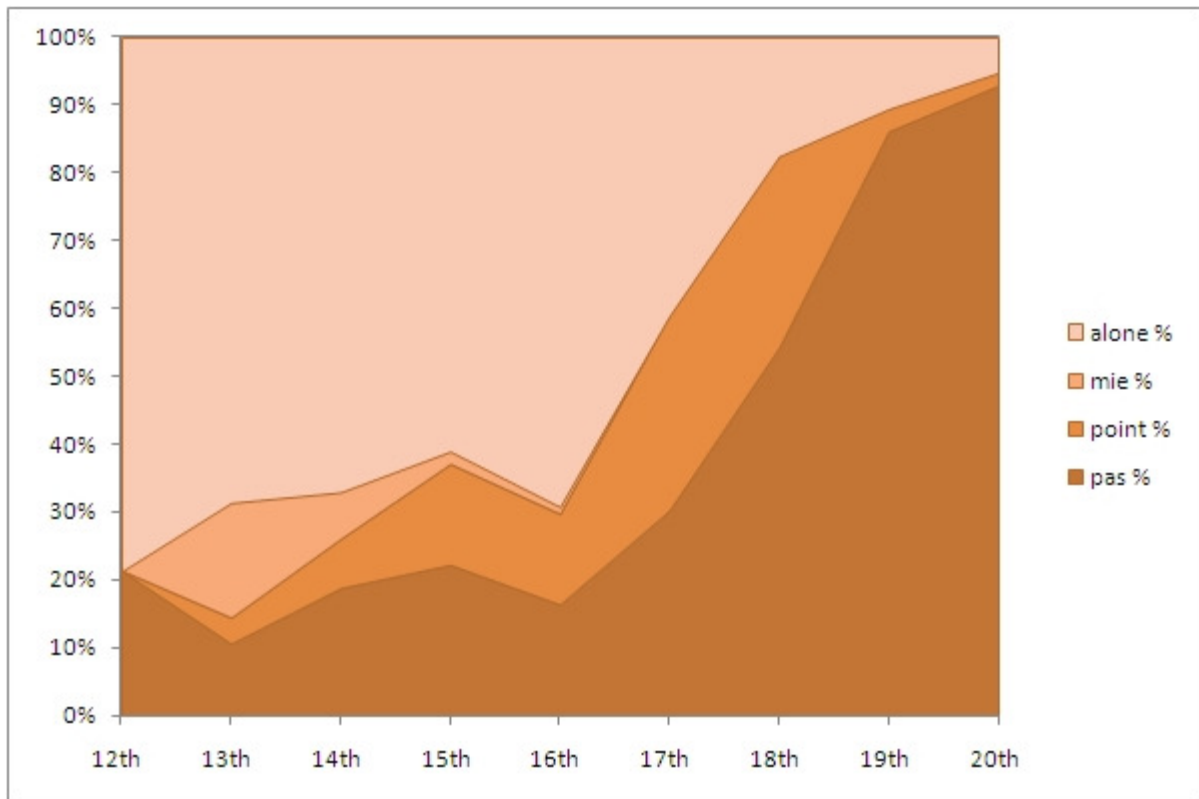
Figure 16. Negation in declarative sentences by century.

Here is the raw data in terms of the number of tokens used in each construction

per century:

Century	alone	pas	point	mie	Total
12th	48	13			61
13th	253	39	14	63	369
14th	268	75	29	28	400
15th	308	112	75	10	505
16th	825	195	159	14	1193
17th	285	210	200		695
18th	128	396	207		731
19th	113	921	37		1071
20th	25	446	10		481
Total	2253	2407	731	115	5506

Table 10. Negation in declarative sentences by century.

It is important to note that the chart by century glosses over a significant amount

of variation within the centuries, as shown in Appendix B and in the following chart:
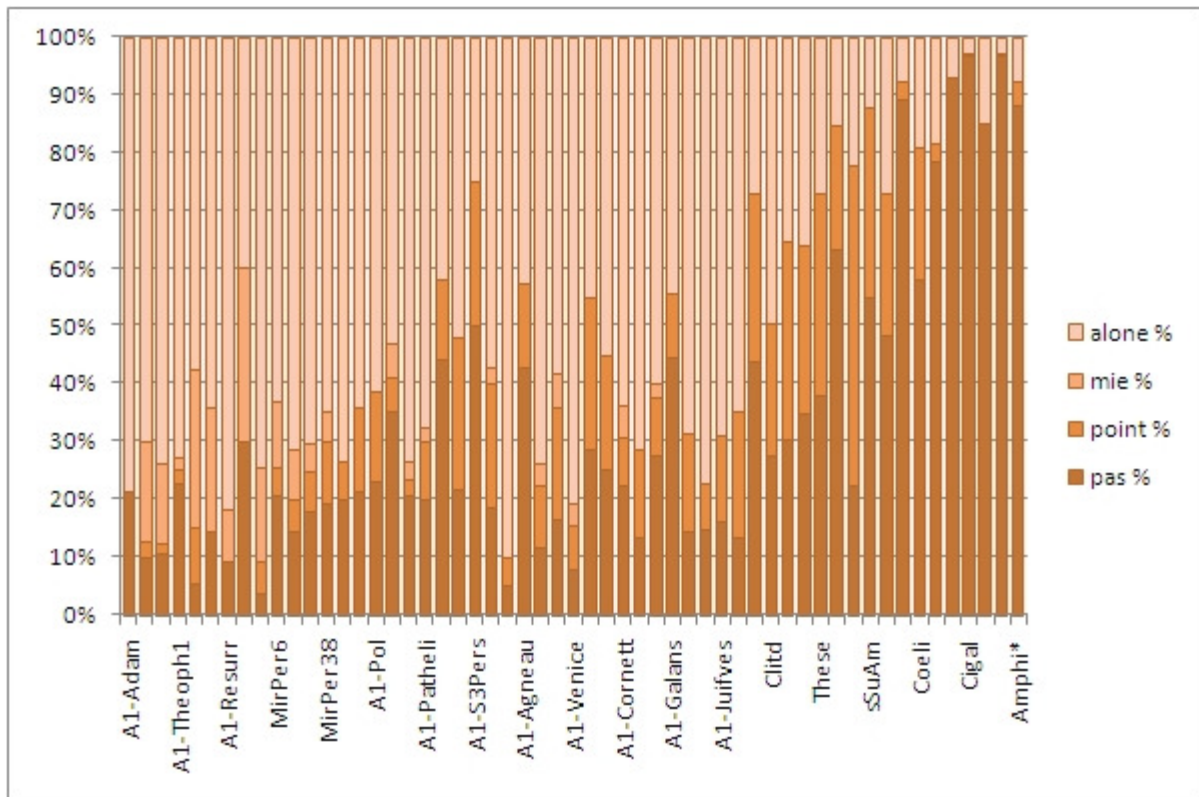
Figure 17. Proportion of tokens of negation in declarative sentences by text.

Even when broken down by text, though, there is a clear S-curve for the rise of

ne ... pas. This is problematic for Labov's (1994) claim that propagation is primarily

social, and that lexical propagation is minor. If we compare the twelve percent of tokens

using *ne ... pas* in sixteenth-century plays such as the *Mystere de Saincte Venice* with the

28% in seventeenth-century plays like Corneille's *Clitandre* (identified on the chart as

"Clitd"), we see that it is not a matter of the new constructions being used 100% of the

time by an increasing number of individual authors, but rather that every author in the

seventeenth-century section of the corpus is using the constructions more frequently (in a

wider number of contexts, as we will see later in this section).

As I discussed in Chapters 4 and 5, it is not justified to apply the Verhulst logistic

model to declining constructions, in this case *ne* alone, *ne ... mie* (the apparent increase

shown on the chart is entirely due to regional variation) and the decline of *ne ... point*. For

144

those, please see the discussion of the Lotka and Volterra models in the next section.

Let's look at *ne ... pas* in more detail:
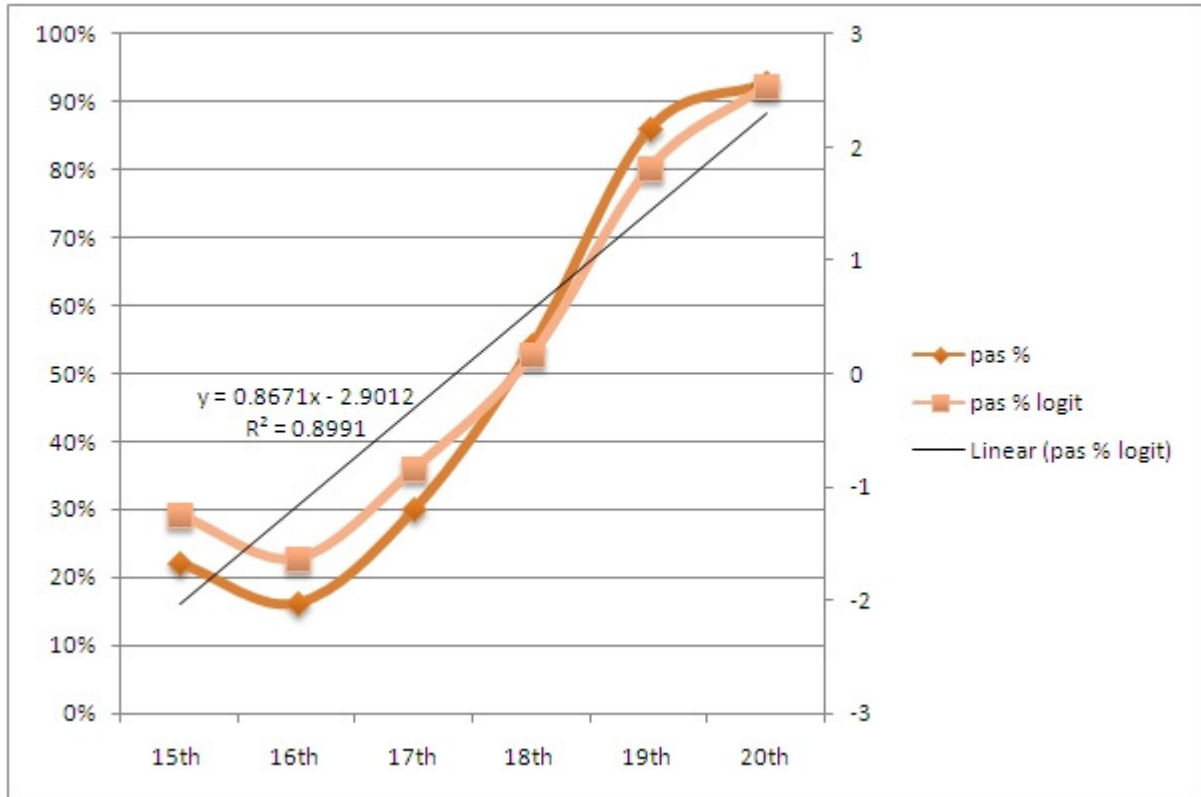
Figure 18. Proportion of tokens of declarative sentence negation with *ne ... pas*, beginning in 1400.

Note that the logit is graphed on a secondary axis, from -3 to 3, which appears on the right side of the chart.

Recall that the semantic and pragmatic analysis suggested that the competition between *ne ... pas* and both *ne* alone and *ne ... point* began in 1600. However, the S-curve begins before the reanalysis, so I have included data about the prevalence of tokens as far back as 1400. Even though the logit curve does not look very linear, the regression indicates that the logistic model fits this data very well, with an R^2 of 0.899. The embracing *ne ... point* construction increases from the thirteenth through eighteenth centuries in this corpus, as seen in the following chart:
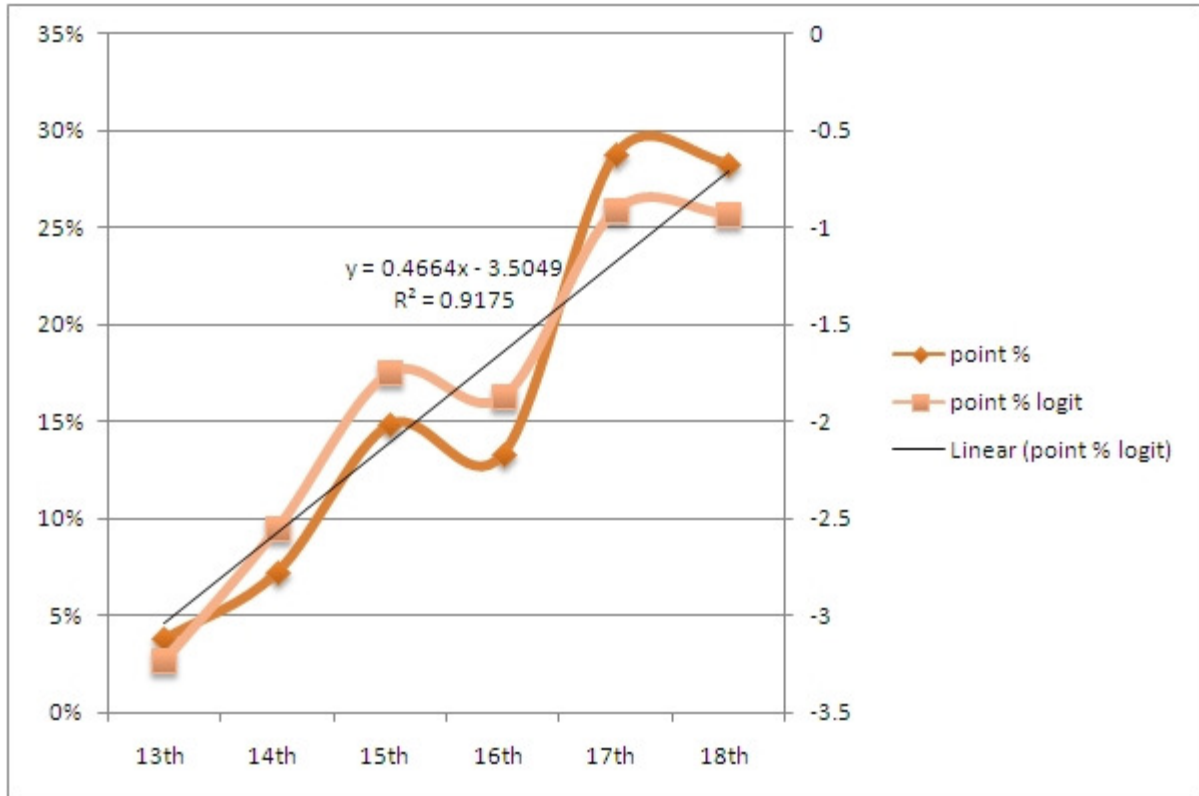
Figure 19. Proportion of tokens of declarative sentence negation with *ne ... point*, from 1200 through 1900.

Despite the squiggle in the middle of the curve, the regression indicates that the logistic function fits the rise of *ne ... point* even better than it fits *ne ... pas*, with an R^2 of 0.918.

As discussed at the end of Chapter 4, there was likely some sort of "embracing negation" schema formed that included *ne ... pas*, *ne ... point* and *ne ... mie*. We can check to see whether the embracing negations as a group increased in a logistic progression.
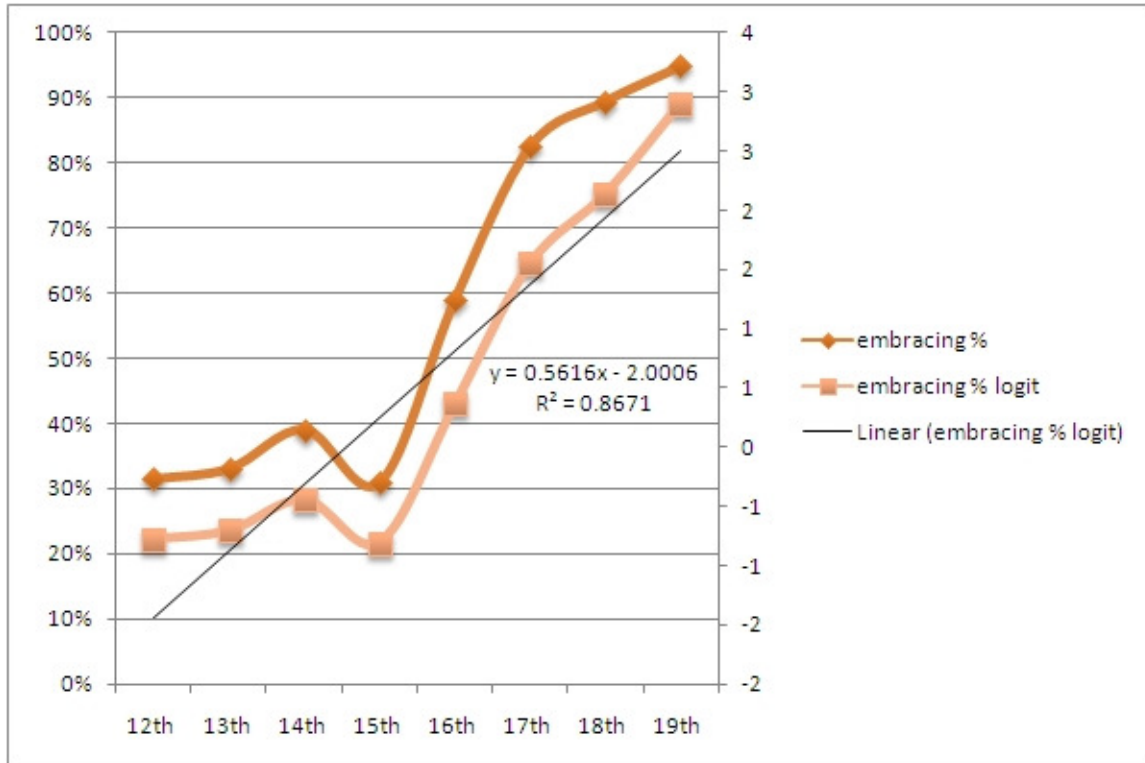
Figure 20. Proportion of tokens of declarative sentence negation with any embracing negation construction, from 1200 through 1939.

This chart shows that when all three embracing negations are combined, the results also fit the logistic model fairly well (R^2 = 0.867). They do not fit it quite as well as the individual changes, but still well enough to support the hypothesis that there was a more general schema that covered all three constructions.

7.2.2 Progression by Type Frequency

Underlying the Logistic Propagation and Type Frequency hypotheses is the idea that an increase in type frequency necessarily results in an increase in tokens as well (assuming that the previously existing types do not lose tokens), because each type is instantiated by at least one token, and often more. Based on Bybee and Thompson's (1997) theory, we would expect changes in type frequency to be one of the driving forces of analogical-type progression. The theory thus predicts that the increases in the

147

prevalence of tokens observed in the previous section will be accompanied by type frequency increases.

As I discussed in the methods chapter, there are several questions about the proper method of counting type frequency: is it better to count main verbs or conjugated verbs, to include or exclude high-frequency verbs or hapaxes? As will be seen in the next section, it does not make much difference whether we count main or conjugated verbs, and the Lotka-Volterra model fits the data best when we include both high-frequency verbs and hapaxes. I will therefore be using type frequency counts of conjugated verbs including high-frequency verbs and hapaxes when testing the question of whether the logistic progression model fits the type frequency data.

Investigating this hypothesis is complicated by the fact that many of the verbs investigated appear in a given century with both preverbal *ne* alone and one or more of the embracing negation constructions, as seen in the following table and chart:

Century	Exclusive preverbal ne	Variable	Exclusive embracing	Total	Exclusive preverbal ne %	Variable %	Exclusive embracing %
12th	19	3	2	24	79%	13%	8%
13th	68	28	18	114	60%	25%	16%
14th	46	18	28	92	50%	20%	30%
15th	56	33	41	130	43%	25%	32%
16th	163	61	44	268	61%	23%	16%
17th	38	39	102	179	21%	22%	57%
18th	11	18	151	180	6%	10%	84%
19th	9	16	185	210	4%	8%	88%
20th	1	4	128	133	1%	3%	96%

Table 11. Type frequencies of preverbal vs. embracing constructions used in negation of declarative sentences by century, based on conjugated verbs and including high-frequency verbs and hapaxes.
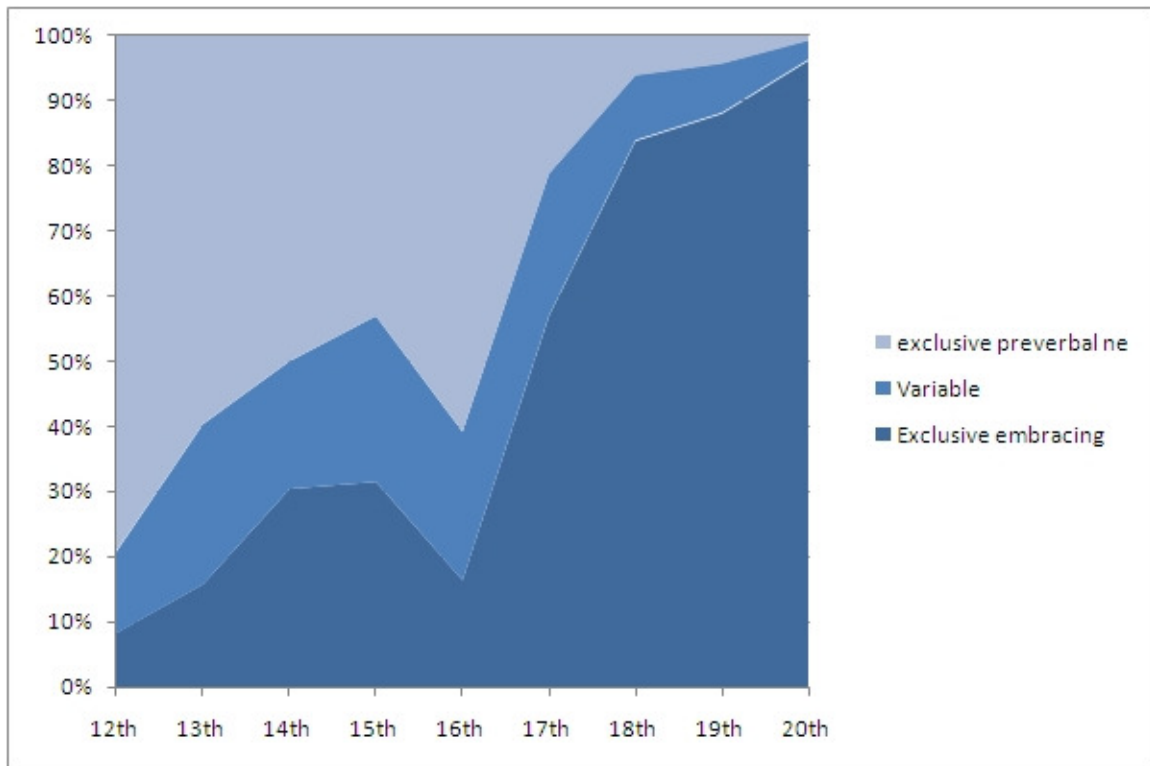
Figure 21. Type frequencies of preverbal vs. embracing constructions used in negation of declarative sentences by century, based on conjugated verbs and including high-frequency verbs and hapaxes.

In the thirteenth, fifteenth and sixteenth centuries these variable verbs constitute

more than 40% of all the verbs. If we separate the embracing negators they add up to

more than 100%, so I have used a non-stacked bar chart to represent them. The data is

presented in more detailed form in Appendix D.
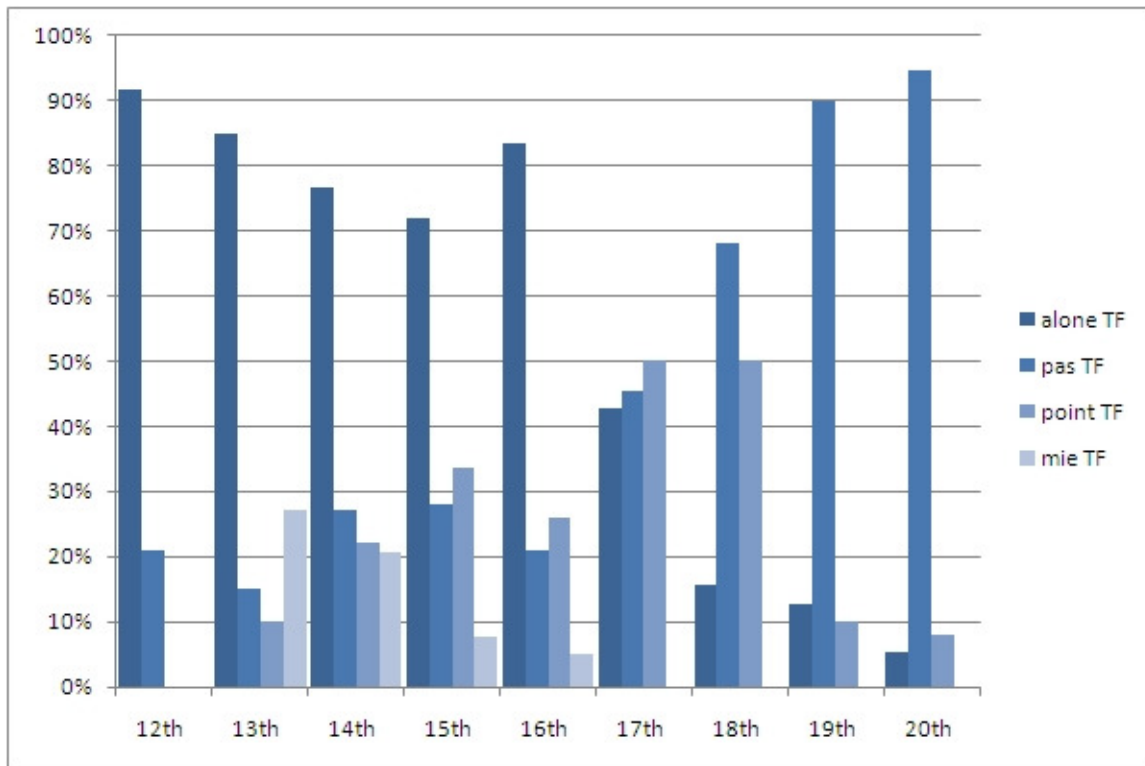
Figure 22. Type frequencies of specific construction used in negation of declarative sentences by century, based on conjugated verbs and including high-frequency verbs and hapaxes.

Let's see how well the model applies to the type frequency of *ne ... pas* and

ne ... point. Recall that from the semantic/pragmatic evolution, we expect the propagation

to begin around 1400. Also recall that the logistic transform is plotted against a different

Y-axis, shown on the right side of the chart.
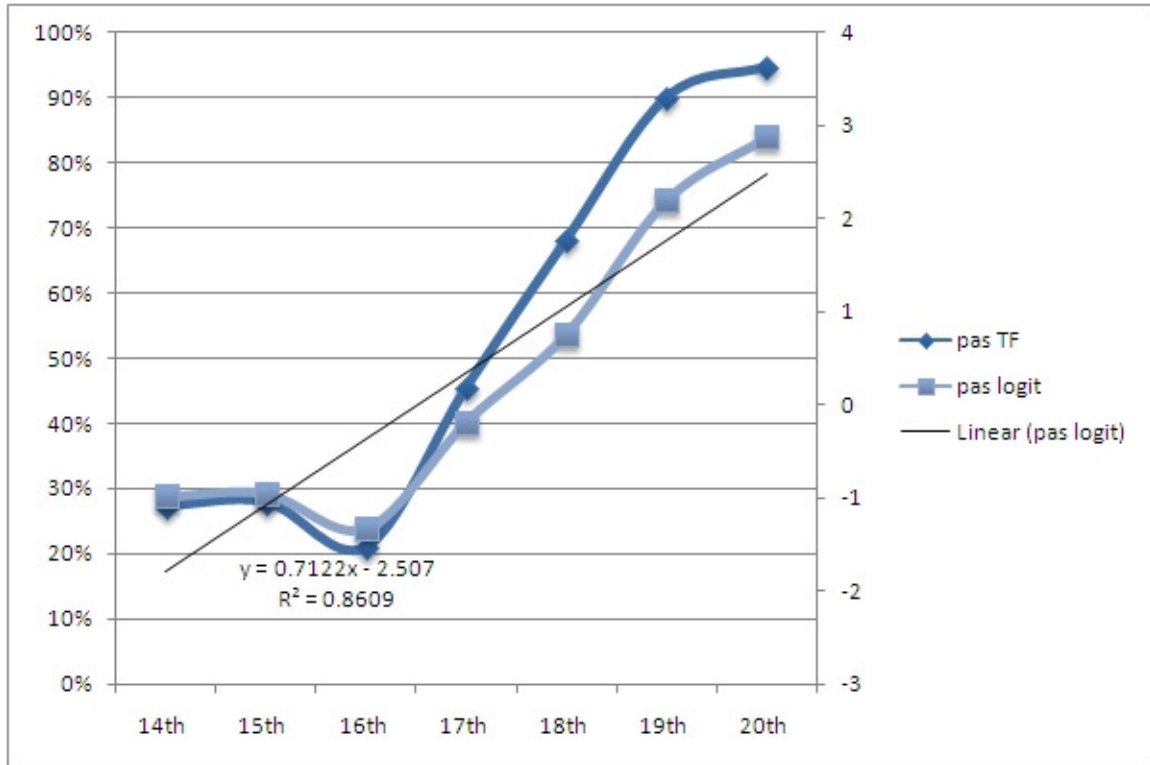
Figure 23. Type frequency of embracing *ne ... pas* in declarative sentences by century, beginning in 1400, based on conjugated verbs and including high-frequency verbs and hapaxes.

The logistic is a good fit, but not quite as good a fit (R^2 = 0.861 instead of 0.899)

as for the prevalence of tokens. The situation is similar for the rise of *ne ... point*:
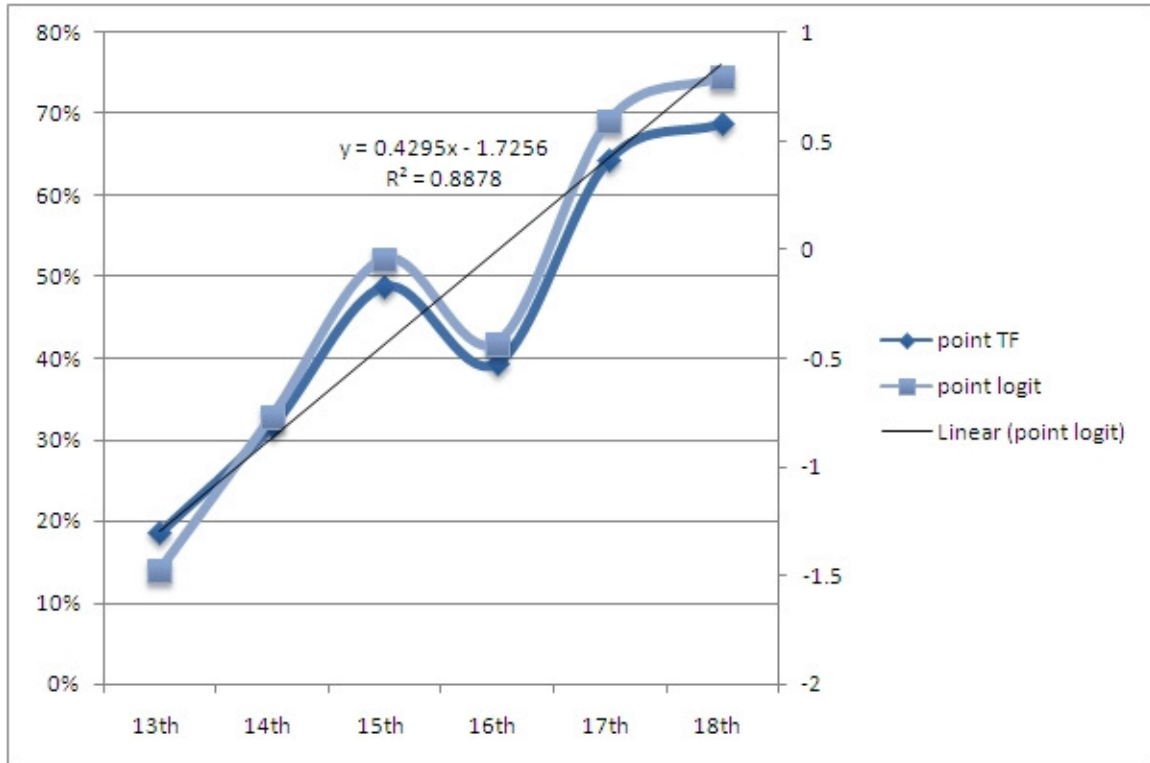
Figure 24. Type frequency of embracing *ne ... point* in declarative sentences by century, from 1200 to 1800, based on conjugated verbs and including high-frequency verbs and hapaxes.

The logistic is also not quite as good a fit for the rise of *ne ... point*, but still very good (R^2 = 0.888 instead of 0.916). As we saw at the beginning of this section, the number of verbs that use embracing negations exclusively follows an S-curve as well. Here is a graph to test how well the logistic model fits it.
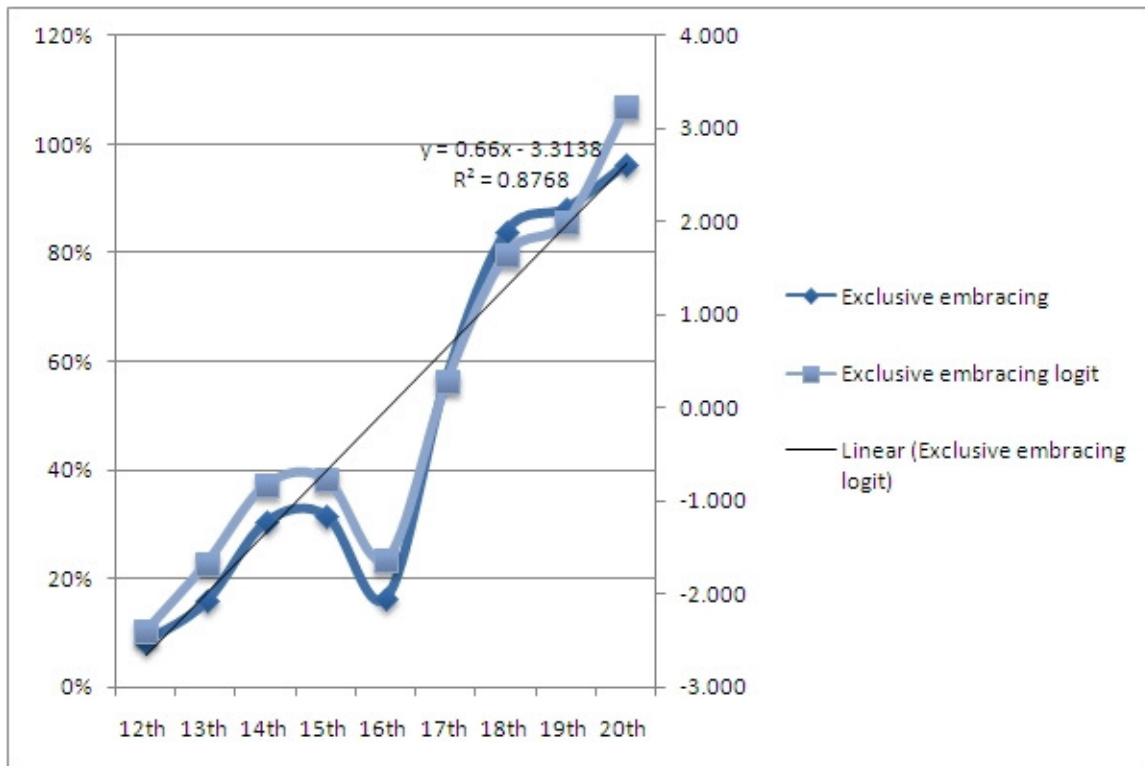
Figure 25. Type frequency of verbs that use embracing negation constructions exclusively, from 1200 through 1939, based on conjugated verbs and including high-frequency verbs and hapaxes.

This change is in the same range as the others ($R^2 = 0.877$) and a slightly better fit than for the proportion of tokens. The high correlations of these models supports the use of the logistic model for analogical-type increases in use.

7.2.3 Comparison of Type Prevalence vs. Token Frequency

The final hypothesis in this section is that the slope values for the logistic regression of the increase in token prevalence will be higher than those for type frequency. This is because we expect many of the additional types to be used with multiple tokens, so each type that shifts over to one of the new forms will be seen in several tokens.

	ne ... pas	ne ... point	all
Tokens Slope	0.867	0.466	0.562
Types Slope	0.712	0.430	0.660
Slope Ratio (token/type)	1.22	1.08	0.852

Table 12. Comparison of the slopes assigned by logistic regression for the increase in token prevalence and type frequency for *ne ... pas*, *ne ... point* and all the embracing negation constructions.

The prevalence among tokens has a higher rate of change than the prevalence among types (type frequency) for the individual changes, which is what we anticipated. For all the exclusive constructions as a general schema, the rate of change among tokens is lower; this may be because the beginning of the change might be earlier than the earliest date for which we have usable data (the thirteenth century). Overall, the logistic model is quite adequate for increases in the use of particular constructions; now let's see what can be done for decreasing constructions.

7.3 Type Frequency

As I discussed earlier, Bybee (1995) predicted that constructions with high type frequency will be more productive, i.e. will tend to increase in type frequency, and those with low type frequency will decrease. Kroch (1989) found that certain grammatical changes will follow a logistic progression. However, he did not take into account the effect of competition from other constructions. Because of this, I hypothesized that **when constructions compete for a function, the change in type frequency for each construction in a given period will be predicted by the Lotka-Volterra model based on that construction's type frequency in the preceding period.** The *Lotka-Volterra model* is necessary to capture the Malthusian idea that the increase in the frequency of an item is limited by the *carrying capacity* of the environment, and the role of *competition*

154

for types as represented by the **competition coefficients**. The function I have used is as described in the section on Malthusian linguistics:

$$19)\ dx_i/dt = r_i x_i (K_i - \Sigma \alpha_{ij} x_j / K_i)$$

K is equal to 1 because we are using percentages to measure type frequency. Since we are measuring correlation, it doesn't matter what r is, and we can set that to 1 as well, yielding the following simplified function:

$$20)\ dx_i/dt = x_i (1 - \Sigma \alpha_{ij} x_j)$$

The full list of verbs found in negative clauses in the corpus can be found in Appendix D. The chart below shows the combination of factors with the most dramatic results.
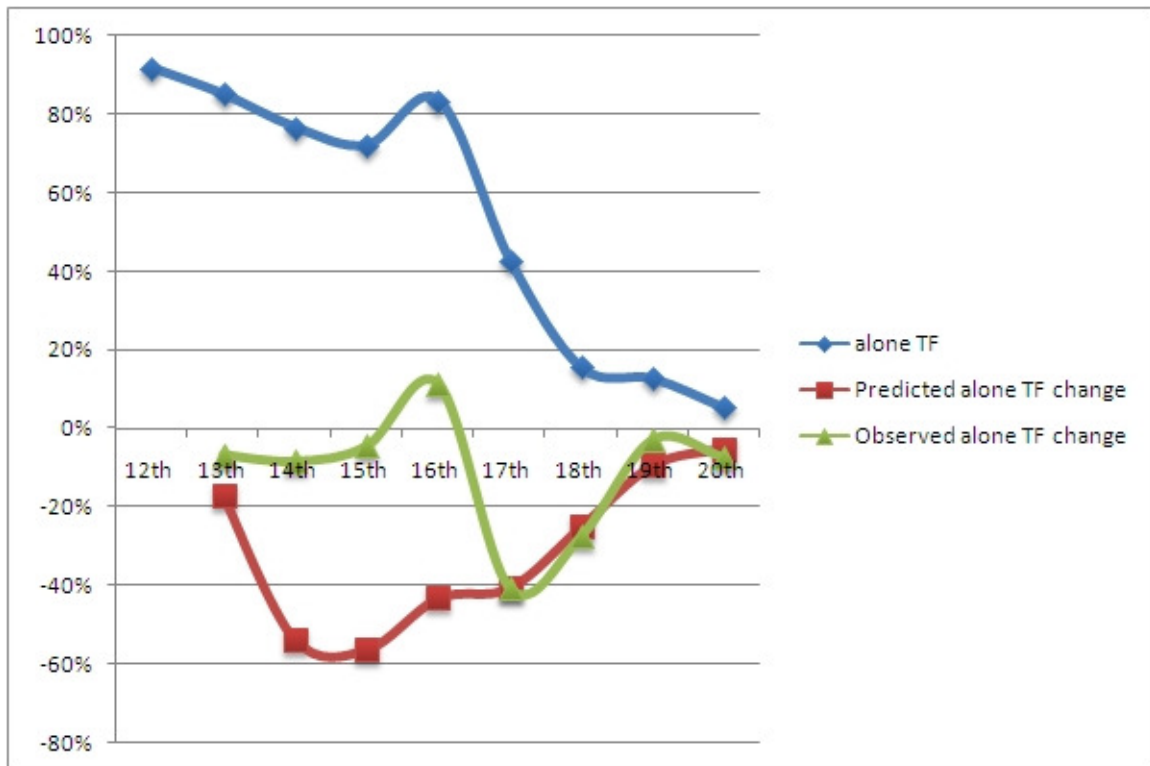

Figure 26. Type frequency, and predicted and observed change in type frequency of preverbal *ne* alone for conjugated verbs, including high-frequency verbs and hapaxes.

155

I had no *a priori* basis for determining the α competition cofficients in this case. They are meant to represent the expected difference between the effect of an average member of one group on another relative to the effect of one average member of the second group on another. I therefore used the least-squared measure to find the α values that best predicted the changes in the negators.

7.3.1 Main vs. Conjugated, High-Frequency and Hapax Verbs

As discussed in the methods chapter, Bybee's work has focused on morphology, and for a case of syntactic evolution like French negation it has not offer a clear expectation as to whether it was more appropriate to count conjugated verbs or main verbs as types. I thus tested another hypothesis, that **the correlation will be greater when conjugated verbs are counted as types as opposed to when main verbs are counted as types.**

Bybee also argued that items with high token frequency do not participate in schema formation, and thus I advanced a third hypothesis, that **type frequency counts that exclude high-token-frequency items will better predict the increase in type frequency than counts that do not exclude high-token-frequency items.**

Finally, Goldberg (2006) warned that observations of constructions that only include a single token (*hapaxes*), especially in a relatively small corpus, may not be reliable indicators of actual usage. To test this I included a fourth hypothesis, that **type frequency counts that exclude hapaxes will correlate better than type frequency counts that include hapaxes**.

Bybee (1995) did not give criteria for deciding whether a verb is "high token frequency," so for this study I examined the frequencies of verbs in the corpus. For every

century, they follow the power-law distribution associated with Zipf's Law (Zipf 1935), as can be seen in this chart from the seventeenth century data found in Appendix D:
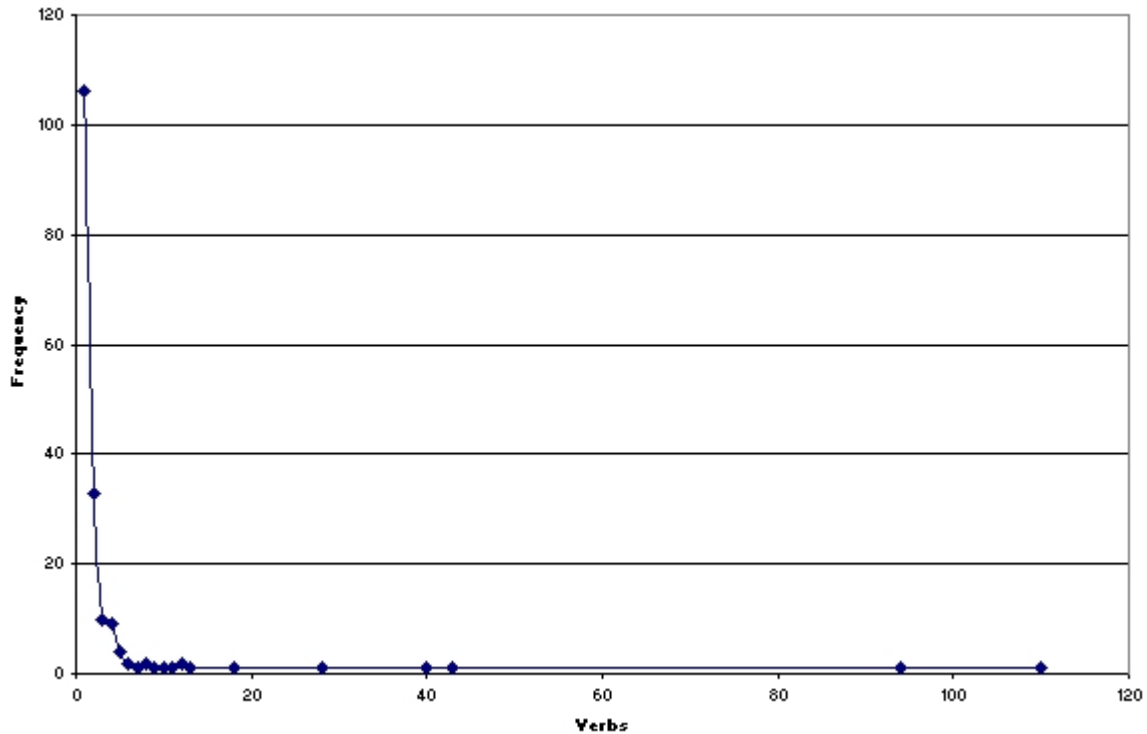
Figure 27. Number of verbs found at each frequency in the seventeenth century.

In the 17th century, the most frequent quartile (25%) of verbs includes all the verbs that occur more than five times. Of these eighteen verbs, all but four (*oser, aimer, mériter* and *importer*) have irregular conjugations. I have therefore chosen that as a starting point for excluding high-token-frequency verbs.

The type frequency counts based on all these strategies are listed in Appendix E. Cohen (1988:110-116) developed a method of estimating the effect size of differences in correlations. The correlation (Pearson product moment) values are converted to z-scores, and if the difference in z-scores (q) is around 0.10, the effect size is small, if it is around 0.30 it is medium, and if it is around .50 it is large. Here are the q values for the differences between the results based on main verbs and those based on conjugated verbs:

157

Competition	+high-frequency -hapaxes	-high-frequency -hapaxes	+high-frequency +hapaxes	-high-frequency +hapaxes
Pas-point-mie, 12th-16th	0.0529	0.0701	0.0106	0.0250
Pas-point, 17th-19th	0.0121	0.0369	0.0308	0.0771
Alone-pas, 17th-20th	0.00442	0.0221	0.0218	0.0141

Figure 28. Effect size indices of differences between correlations based on main verbs and those based on conjugated verbs.

All of these differences are very small by Cohen's (1988) standards, and as we will soon see, the other differences are much larger. I therefore conclude that it does not make much difference whether I use main or conjugated verbs (as I defined them in the methods chapter), and I will therefore arbitrarily choose to use conjugated verbs. Here are the q values for the differences between the results based on including high-frequency verbs and those based on excluding them:

Competition	conjugated verbs +hapaxes	conjugated verbs -hapaxes
Pas-point-mie, 12th-16th	0.113	0.0182
Pas-point, 17th-19th	0.0383	0.197
Alone-pas, 17th-20th	0.222	0.380

Table 13. Effect size indices of differences between correlations based on including high-frequency verbs and those based on excluding them.

Competition	conjugated verbs +high-frequency	conjugated verbs -high-frequency
Pas-point-mie, 12th-16th	0.268	0.399
Pas-point, 17th-19th	0.315	0.0803
Alone-pas, 17th-20th	0.821	0.219

Table 14. Effect size indices of differences between correlations based on including hapaxes and those based on excluding them.

Based on these values, I characterize the differences based on including or excluding high-frequency verbs as small to medium, and those based on including or excluding hapaxes as medium to large. Now let's take a look at the correlation values themselves and see which models fit better with the observed values:

Competition	+high-frequency +hapaxes	-high-frequency +hapaxes	+high-frequency -hapaxes	-high-frequency -hapaxes
Pas-point-mie, 12th-16th	0.146	0.043	0.402	0.418
Pas-point, 17th-19th	0.948	0.948	0.776	0.839
Alone-pas, 17th-20th	0.978	0.964	0.940	0.953

Table 15. Pearson product moment (*r*) values for the correlation between observed type frequency values and values predicted by the Lotka-Volterra model, by type frequency measurement technique.

A Pearson value of 1 indicates perfect correlation and an *r* of zero indicates no correlation. In the first competition, the model seems to be a poor fit, matching observations at most 41.8% of the time. This may be for a number of reasons, but one is that *ne ... mie* was not native to the Parisian and Norman plays at all, but used as a poetic device, and conversely that *ne ... pas* was not native to the Picard plays. Under this theory, the decline of *ne ... mie* would not have been an instance of competition/specialization/obligatorification at all, but simply the decline of a poetic device, and thus not necessarily subject to the effects of type frequency. Because of this, I do not feel comfortable using it to judge these hypotheses.

For the two other competitions, the correlations are highest when we include both hapaxes and high-frequency verbs. Results like these may not be found in all cases, but it seems that here it is most appropriate to include them. Under these conditions, the model predicted type frequency values that were very highly correlated for two of the three competitions under study. However, in a future study it may be worth re-examining the choice of quartiles for excluding high-token-frequency verbs.

Based on these results, I have chosen to focus on **conjugated verbs, including both high-frequency verbs and hapaxes**. Here is the table of type frequency values by century for these variables, identical to Table 1 in Appendix E:

Conjugated verbs +high-frequency +hapaxes				
Century	alone types	pas types	point types	mie types	Total
12th	22	5	0	0	24
13th	85	15	10	27	100
14th	59	21	17	16	77
15th	75	29	35	8	104
16th	187	47	58	11	224
17th	66	70	77	0	154
18th	24	105	77	0	154
19th	23	162	18	0	180
20th	6	106	9	0	112
Total	2258	2408	731	115	

Table 16. Raw numbers of types per century for conjugated verbs, including high-frequency verbs and hapaxes.

The following table shows the alpha values that were generated by the least

squares method:

α	effect of			
effect on	*ne* alone	*ne ... pas*	*ne ... point*	*ne ... mie*
ne alone	1	1.29	1.14	1.76
ne ... pas	0.274	1	0	1.53
ne ... point	0	1.67	1	0
ne ... mie	0.451	3.87	0	1

Table 17. Values for competition coefficients for type frequency of conjugated verbs suggested by the least-squares method, including high-frequency verbs and hapaxes.

In the Lotka-Volterra model, α values cannot be negative. A value of 0 indicates

that the first population has no effect on the second; a value of 1 indicates that a member

of the first population has the same effect on the second population as on other members

of the first population. A value between 0 and 1 indicates a relatively small effect, while a

value greater than one indicates a relatively large effect.

In general, the α values suggest that the constructions that ultimately won the

competitions had strong effects on the losers, and that the ones that lost had little or no

effect on the winners. The embracing negation construction *ne ... pas* had the greatest

effect on its competitors. Embracing *ne ... point* and *ne ... mie* had a strong effect on

160

preverbal *ne* alone, and *ne ... mie* also had an effect on *ne ... pas*, but not on *ne ... point*. Uses of preverbal *ne* alone had a weak effect on *ne ... pas* and *ne ... mie*, but not on *ne ... point*.

By varying the alpha values at random, I discovered that the two most influential coefficients are the ones representing the effects of *ne* alone and *ne ... point* on *ne ... pas*. The lower these values, the higher the correlation between the predicted and observed values of type frequency for both the competition between *ne* alone and *ne ... pas* and that between *ne ... pas* and *ne ... point*. Similarly, the higher the coefficients representing the effects of *ne ... pas* on *ne* alone and *ne ... point*, the better the model fit the observed type frequencies. This suggests that type frequency alone is not enough to account for these changes, and that there is at least one other factor encouraging people to shift to it. The model does not give any idea of what those factors may be, but one could be the formation of a higher-level schema of bimorphemic negation constructions including *ne ... rien* and *ne ... que*, as described in Chapter 4.

Earlier in this section I provided a chart showing the predicted and observed type frequencies for preverbal *ne* alone. Here is the same chart again, followed by the charts for the other negators:
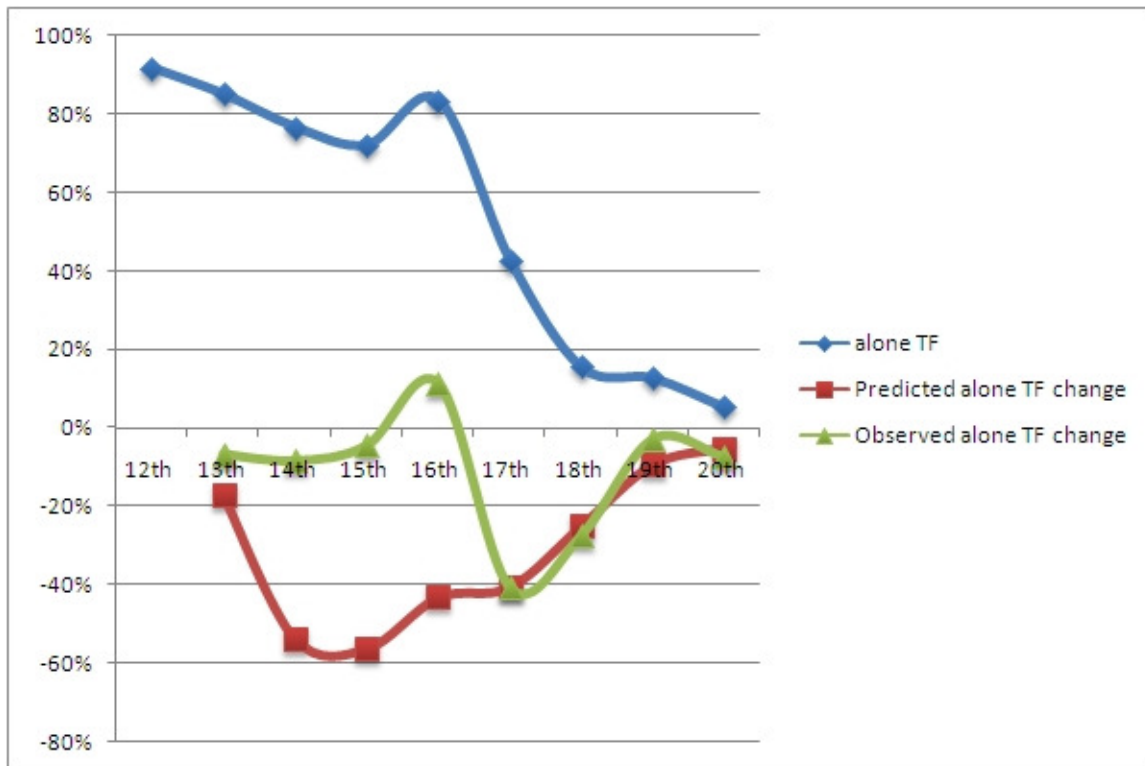
Figure 29. Type frequency, and predicted and observed change in type frequency of preverbal *ne* alone for conjugated verbs, including high-frequency verbs and hapaxes.

This chart provides a visual illustration of the low correlation found for the first competition, between the three embracing constructions for the presupposition denial function from the twelfth through sixteenth centuries. There is essentially no match between predicted and observed type frequency changes during those periods. From the seventeenth century on, however, the match is very close.
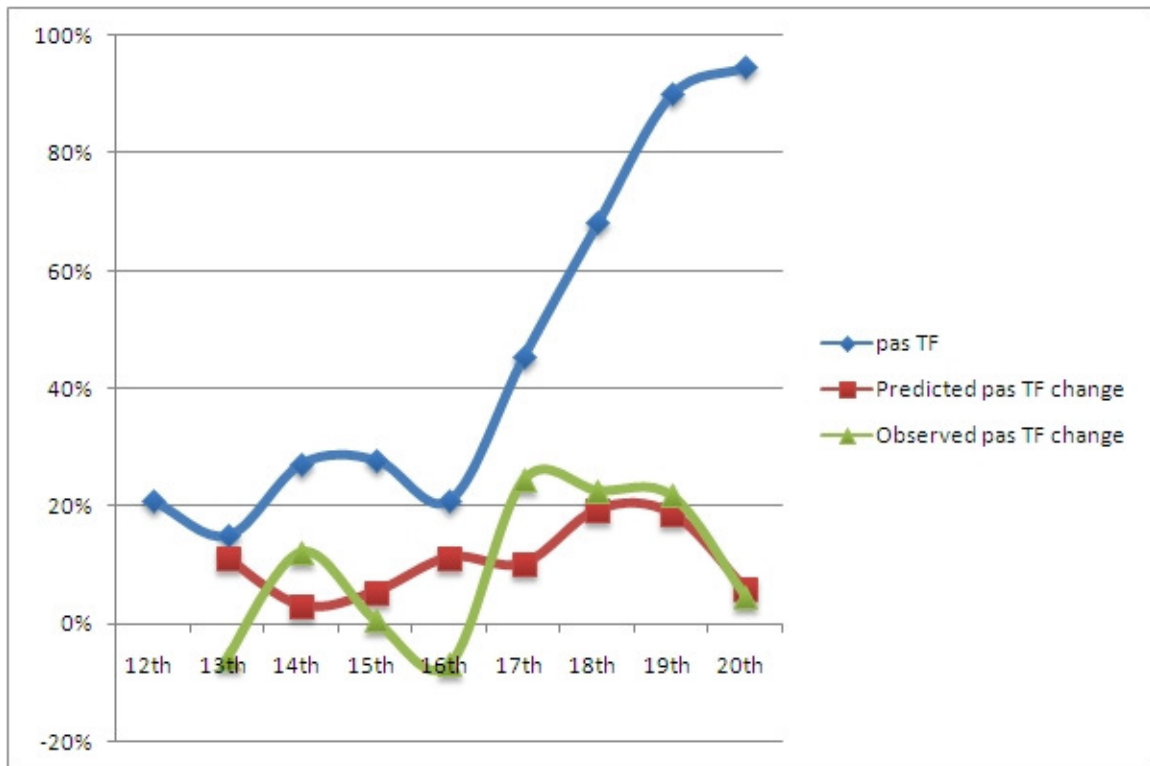
Figure 30. Type frequency, predicted change and measured change in type frequency of embracing *ne ... pas* for main verbs, including high-frequency verbs and hapaxes.

For *ne ... pas* the fit is not quite as bad for the medieval data, but still not particularly accurate. In the seventeenth century the model predicts a greater increase in type frequency than actually occurs in the corpus. From the eighteenth through twentieth centuries, the fit is very close.
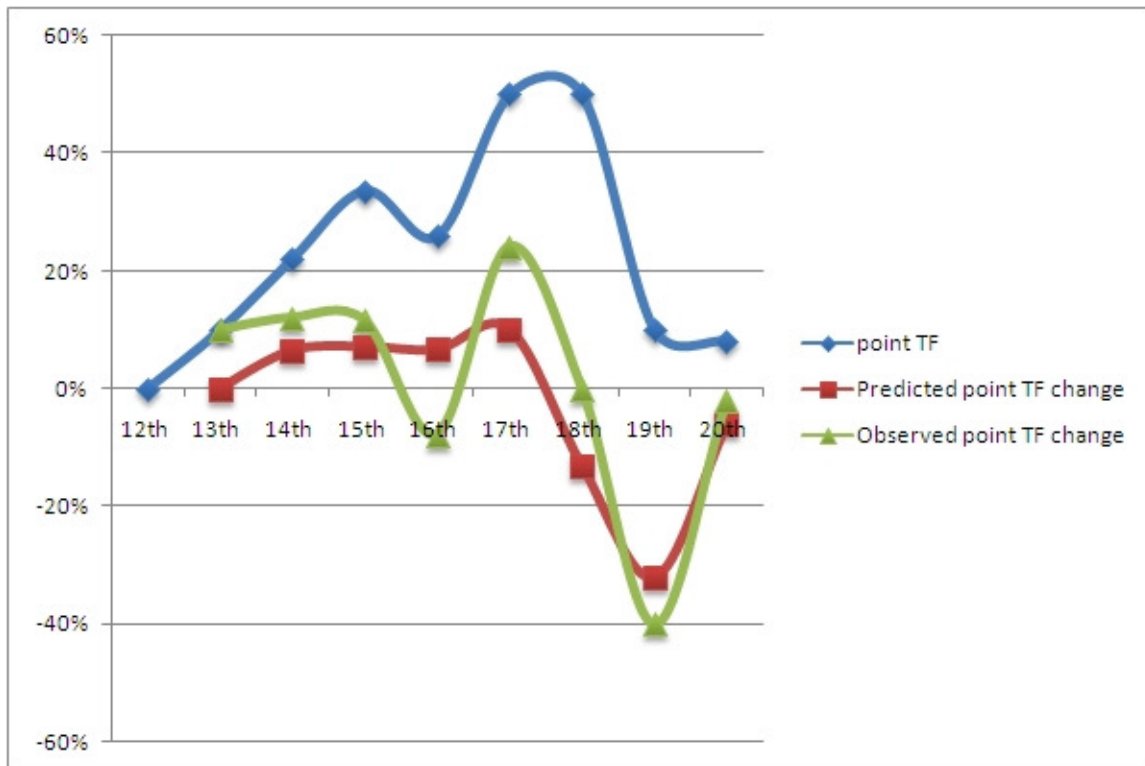
Figure 31. Type frequency, predicted change and measured change in type frequency of embracing *ne ... point* for conjugated verbs, including high-frequency verbs and hapaxes.

For *ne ... point* the fourteenth and fifteenth century predictions are actually fairly close to the observed data. The model does not predict the fall in use of *ne ... point* in the sixteenth century, and predicts a slightly less extreme change than the rapid rise in the seventeenth century and fall in the nineteenth. It does predict the decrease in type frequency in the nineteenth century, and the relatively small amount of change in the eighteenth and twentieth centuries, however.

Overall, the predicted and observed values are very close for the seventeenth through twentieth centuries. In the next section I will look at how these changes propagate through the lexicon, and the role of token frequency.

7.3.2 Lexical Diffusion and Entrenchment

Once a process of analogical extension starts, entrenchment determines the course that it takes through the lexicon. As we saw in the background discussion, contexts with

164

lower token frequency tend to be the most likely to shift to a different construction. My hypothesis was that **as the embracing negation constructions are extended to more and more lexical items, the items that resist the change tend to have higher token frequencies**.

As I discussed in the section on the history of negation in French, beginning in the sixteenth century many grammarians were unable to find a single unifying principle to describe all the contexts where preverbal *ne* alone continued to be used, and settled for lists of contexts. Based on Bybee and Thompson's (1997) theory, we would expect the unifying principle to be one of *token frequency*, which was not much discussed by these grammarians.

In tagging the corpus, I made sure to mark each negation for whether it fit any of these rules. Below is a reprint of the table from Chapter 3 outlining the rules.

Context	Malherbe (1606)	Maupas (1607)	Vaugelas (1667)	Ewert (1930)
with *craindre* and other verbs of fear, uncertainty and doubt (paratactic negation, not counted)	√	√		
with expressions of prevention and caution (paratactic negation)		√		
with causative expressions such as *afin que* and *pour que*		√		
presence of other negators such as *ni, jamais* (not counted)	√	√	√	
with *savoir* and *pouvoir*		√	√	√
subordinate to another negative		√		√
exclamatory or final *que*		√		√
conditional *si*		√		√
comparative (not counted)			√	
time expressions such as *ça fait ... que*			√	
with *oser*			√	√
with *cesser*				√
fixed expressions such as *à Dieu ne plaise*, *n'importe*, *n'avoir garde*, *n'avoir cure*				√
exclamatory *qui*				√
condition by inversion (not counted)				√

Table 18. Negative contexts in which *ne ... pas* and *ne ... point* are not used, as reported by four grammarians.

As discussed in the Background, contexts of fear, uncertainty and doubt and those with comparatives fall under the heading of *expletive negation*. In these contexts, the *ne* alone construction is not in competition with *ne ... pas* or *ne ... point*, but with unmarked verbal constructions, as in the following example:

77)   William: Eh bien, mon ami, j' ai peur que cette
        colonne s' écroule.
        William: Well, my friend, I'm afraid that this pillar
        might crumble.

Robert de Flers and G. A. de Caillavet, M. Brotonneau, 1929

Similarly, contexts with other negators such as *ni* and *jamais* did not take part in the shift to *ne ... pas*. Inversions in general were excluded from the corpus, because they

constituted an alternative form of emphasis. They are important, and should be examined in a later study.

The following chart shows the proportion of tokens that appeared with each negator, but it separates out the tokens that fit one or another of the rules described by the grammarians. The data can be found in Appendix F.
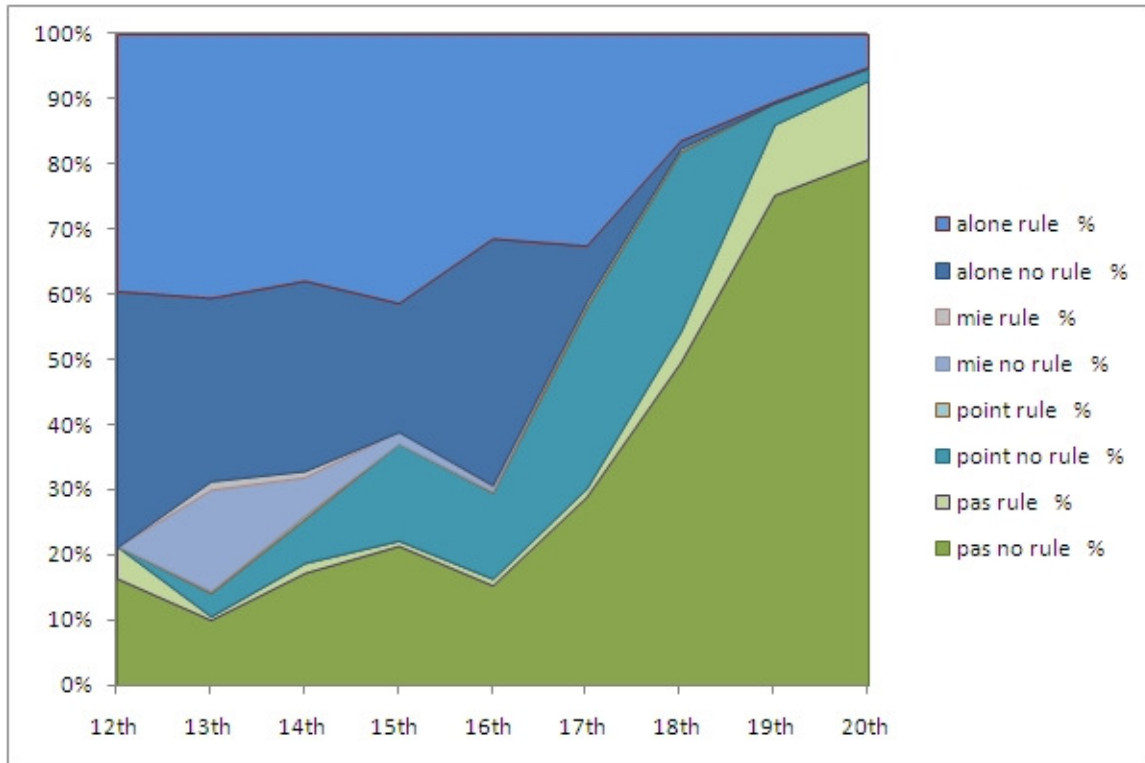

Figure 32. Proportion of tokens that appear with each negator, separated by whether the token fits under one of the grammatical rules described in Chapter 3.

The grammarians appear to have captured the main exceptions to the use of embracing negation with these rules. In the chart above, the dark blue representing the proportion of tokens that did not follow any of the grammarians' rules and were represented with preverbal *ne* alone decreases rapidly in the seventeenth and eighteenth centuries, much faster than those covered by the rules. At the same time there was an increase in the use of *ne ... pas* (dark green) and *ne ... point* (turquoise) in these contexts. In the nineteenth and twentieth centuries, the proportion of tokens where preverbal *ne*

167

alone was used in contexts described by the grammarians (light blue) declined, with a

corresponding increase in the use of *ne ... pas* in these contexts (light green). This shows

that most of the constructions that shifted from preverbal *ne* alone to *ne ... pas* in the

seventeenth and eighteenth centuries did not fit any of the rules, and that the

constructions that did fit the rules continued to shift after that.

Several of the constructions in the rules appeared with *ne* alone less than once in

every ten thousand tokens even in the sixteenth century, which makes them far less

frequent than any of the other rules, so I did not investigate them further. I also excluded

rules that forbid multiple negations and those that described *expletive negation*, because

these are not directly relevant to the question at hand. I will discuss verbs later in this

section. Here are the remaining contexts from Chapter 3, with their token frequencies.

Context	Token frequency (16th C.)	Token frequency (17th C.)	Token frequency (18th C.)	Token frequency (19th C.)	Token frequency (20th C.)
conditional *si*	0.962	0.704	0.427	0.119	0.025
subordinate to another negative	0.351	0.186	0.117	0.022	0
with expressions of prevention and caution, and causative expressions such as *afin que* and *pour que*	0.279	0.080	0.039	0.011	0

Table 19. Negative contexts in which preverbal *ne* alone is used, as reported by four grammarians, with token frequencies per thousand words.

This data is consistent with our hypothesis. The item with the highest token

frequency in the sixteenth century, conditional *si*, is the only one that is still present in the

twentieth century texts, the paratactic expressions consistently have the lowest token

frequencies, and the negative-dependent items are in the middle. However, it may be that

the playwrights in the corpus were simply following the rules set down by Maupas (1607)

and later grammarians.

Several verbs were mentioned by the grammarians, so let us investigate all high-frequency verbs. The following chart shows the proportion of tokens with preverbal *ne* alone for the verbs that appeared at least four times in the corpus in every century from the fifteenth through the twentieth. Since the grammarians seem to have captured the main exceptions to embracing negation, as described above, I have excluded those negations based on rules unrelated to verb choice.

To bring out the patterns more clearly, for the values between the fifteenth and twentieth centuries I have applied a SPAN 3 SMOOTHING TRANSFORMATION, using the median value from each set of three centuries. For example, the proportions of preverbal *ne* alone with the verb *voir* 'to see' in the fifteenth, sixteenth and seventeenth centuries were 58%, 83% and 28%. The smoothing transformation took the median value of 58%, which better captures the overall downward trend.
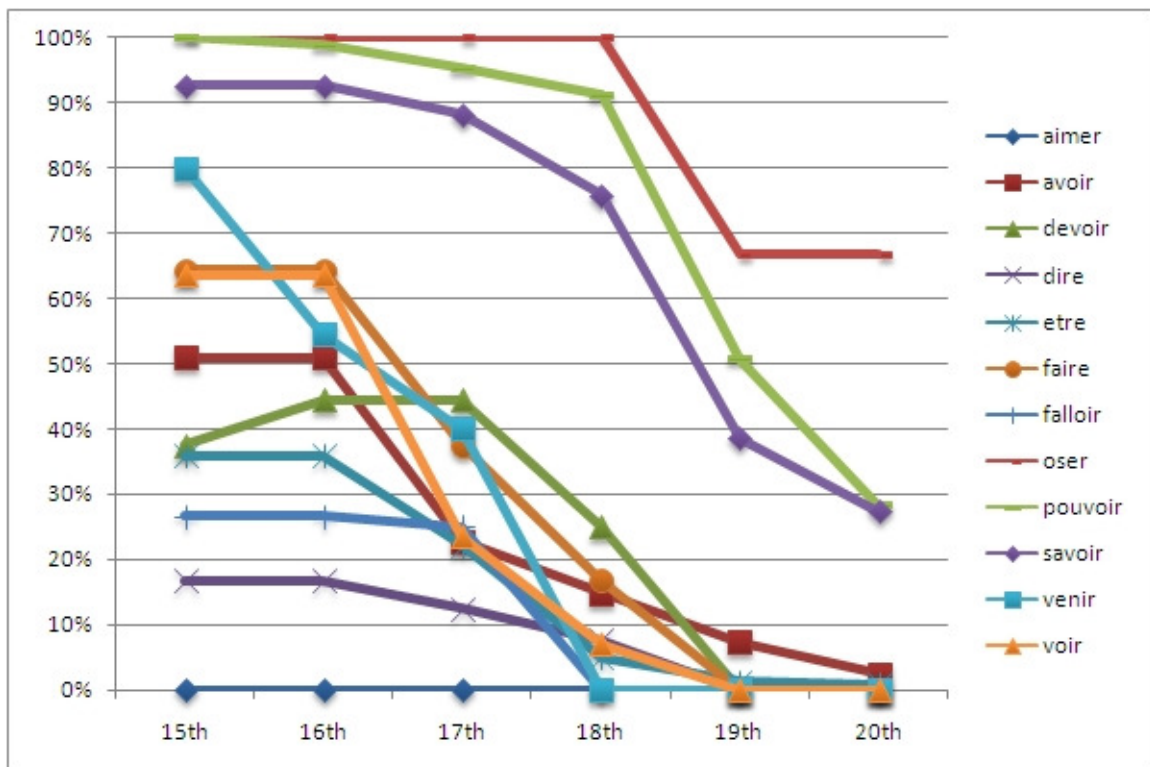

Figure 33. Span 3 smooth of the proportion of tokens with preverbal *ne* alone for the verbs that appeared more than four times in every century from the fifteenth through the twentieth.

169

Notice that the last three verbs to stop being used with *ne* alone are *oser* 'to dare,' *pouvoir* 'to be able' and *savoir* 'to know how'. These are also (with *cesser*, 'to cease') the verbs mentioned by the grammarians. Now let's relate this back to our token frequency hypothesis.

Rank	16th C. Verb	16th C. Token Frequency	16th C. Total N w/ne alone	17th C. Verb	17th C. Token Frequency	17th C. Total N w/ne alone
1	savoir	0.819	91	pouvoir	1.097	90
2	etre	0.792	88	savoir	0.536	44
3	pouvoir	0.720	80	oser	0.158	13
4	avoir	0.666	74	devoir	0.049	4
5	vouloir	0.306	34			
6	devoir	0.144	16			
7	faire	0.144	16			
8	falloir	0.108	12			
9	voir	0.108	12			
10	oser	0.081	9			
11	y avoir	0.081	9			

Rank	18th C. Verb	18th C. Token Frequency	18th C. Total N w/ne alone	19th C. Verb	19th C. Token Frequency	19th C. Total N w/ne alone	20th C. Verb	20th C. Token Frequency	20th C. Total N w/ne alone
1	savoir	0.828	37	pouvoir	0.463	43	pouvoir	0.273	11
2	pouvoir	0.694	31	savoir	0.291	27	savoir	0.149	6
3	oser	0.201	9	oser	0.054	5			

Table 20. Token frequencies per thousand words for the top eleven conjugated verbs most frequently negated with preverbal *ne* alone (in at least 25% of occurrences) by century, excluding tokens covered by one of the rules.

As we see, the verbs that are still used with preverbal *ne* alone with any frequency in the nineteenth century (outside of the constructions described at the beginning of this section), *pouvoir*, *savoir* and *oser*, are the ones that had the highest token frequencies (with the exception of the auxiliaries) in the sixteenth. There are some exceptions: *pouvoir* and *savoir* switch ranks a few times, and *oser* just barely made it into the

category of high-frequency verbs in the sixteenth century. These are likely related to other developments in the language during those periods.

For the less frequent verbs, we can aggregate them into groups based on the number of times they appeared in the sixteenth century. None of these (with the exception of *cesser* and *pardonner* 'to pardon') are found with *ne* alone after 1700.

Tokens	1500	1600	1700	1800	1900
1	0.7310	0.1111	0.0000	0.0183	0.0000
2	0.6512	0.1250	0.0000	0.0000	0.0000
3	0.6000	0.3333	0.0000	0.0000	0.0000
4	0.5000	0.1667	0.0000	0.0000	0.0000
5	0.4167	0.0000	0.0000	0.0000	0.0000
6	0.7500	0.0000	0.0000	0.0000	0.0000
9	0.5806	0.6667	0.5294	0.1786	0.2222

Table 21. Aggregate use of preverbal *ne* alone among low-token-frequency verbs, excluding tokens covered by one of the rules.
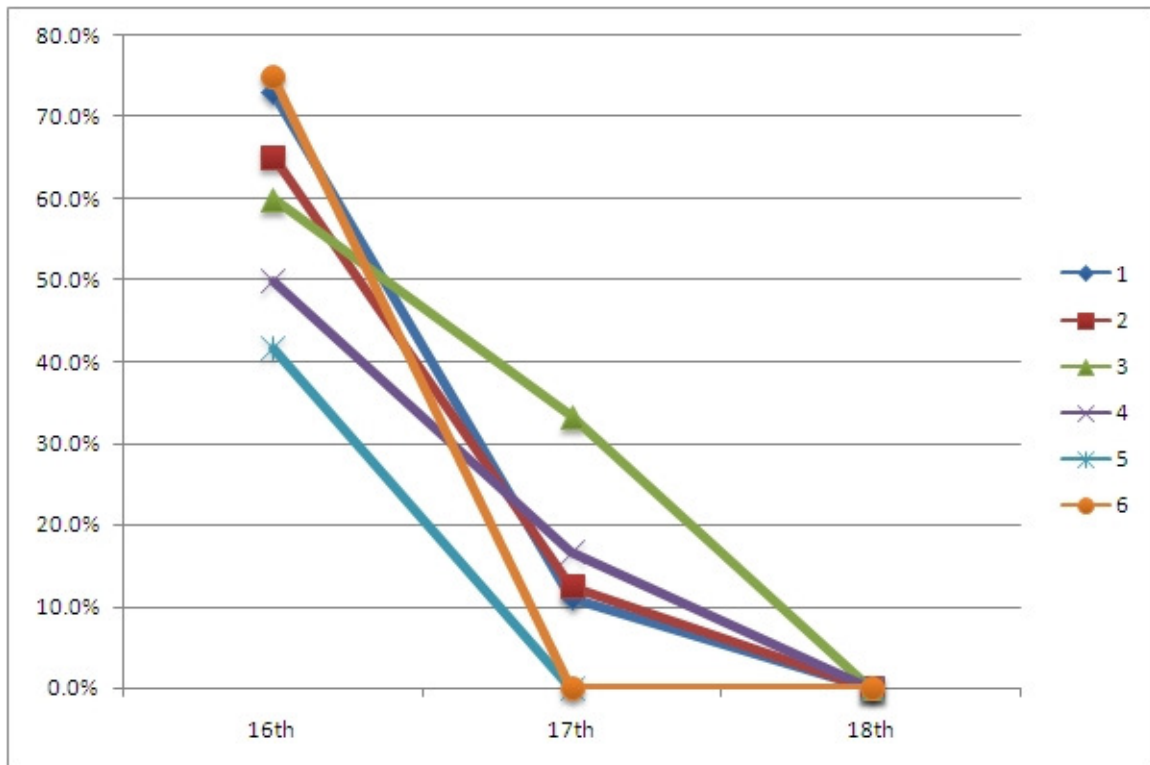

Figure 34. Aggregate use of preverbal *ne* alone among low-token-frequency verbs, excluding tokens covered by one of the rules.

There is an inverse rank-order correlation ($r = 0.8$) between the number of tokens where a verb appeared with preverbal *ne* alone in the sixteenth century and the percentage of tokens of that verb used with *ne* alone in the seventeenth century. The exception is the verbs that appeared three times with *ne* alone in the sixteenth century, but have a higher percentage of *ne* alone in the seventeenth century than the verbs that appeared four times with *ne* alone in the sixteenth. However, we are talking about very small amounts of data (there were four verbs that appeared three times with *ne* alone and two that appeared four times), so it may not be reliable.

As discussed in the Methods chapter, a question worth investigating in the future is how to measure the change in individual verbs, since each of these verbs takes more than a century to shift. In fact, from looking at the chart broken down by verbs it would appear that each verb follows an S-curve. The rise in the use of embracing *ne ... pas* is generally the inverse of that chart:
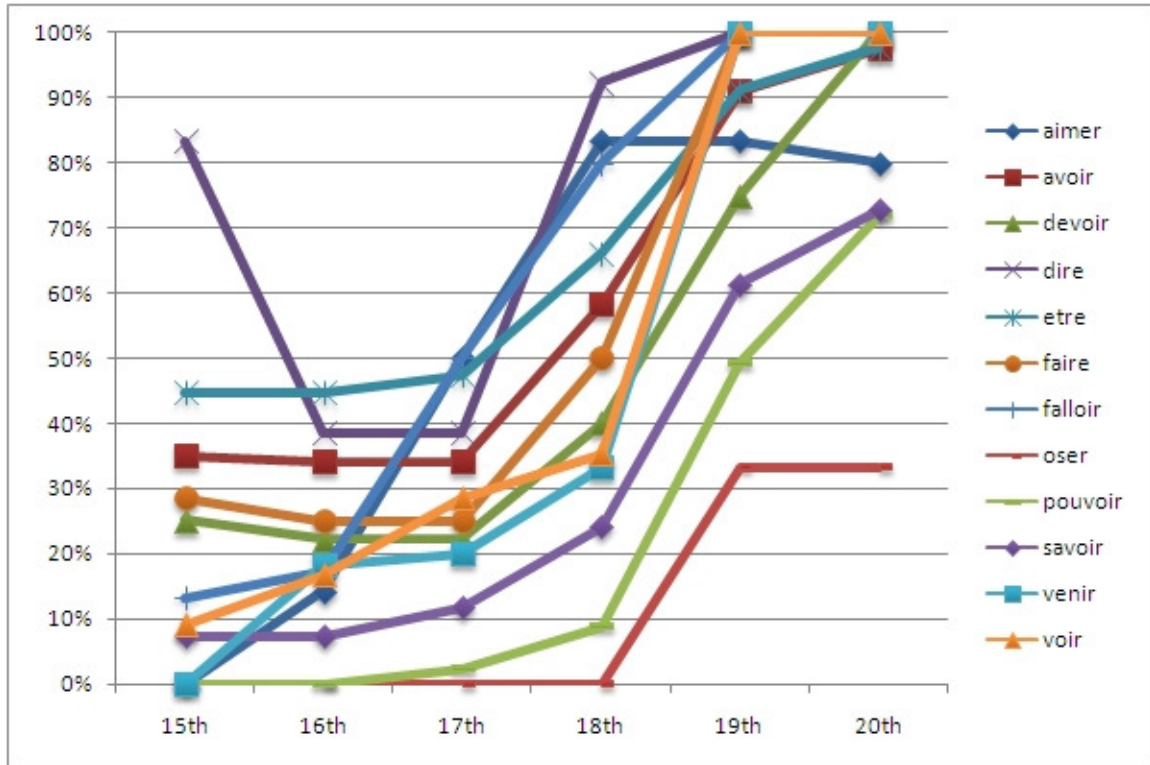
Figure 35. Span 3 smooth of the proportion of tokens with embracing *ne ... pas* for the verbs that appeared more than four times in every century from the fifteenth through the twentieth.

We can test this with the five most common verbs. In the twentieth century *avoir* and *vouloir* are used with *ne ... pas* all the time and in the seventeenth century *pouvoir* is not used with *ne ... pas* at all, which does not allow logistic modeling, so in those cases I replaced 1 with 99.99% and 0 with 0.01.

Rank	16th C. Verb	16th C. Token Frequency	R^2 of logistic regression	Intercept of logistic regression
1	savoir	0.819	0.968	-4.84
2	être	0.792	0.974	-2.04
3	pouvoir	0.720	0.533	-5.75
4	avoir	0.666	0.863	-5.57
5	vouloir	0.306	0.875	-5.51

Table 22. Goodness of fit of logistic regression to the use of embracing *ne ... pas* from the sixteenth through twentieth centuries for the five verbs most frequently used with preverbal *ne* alone in the sixteenth century.

The is clearly a close fit for these verbs. The one exception is *pouvoir*, which is still distorted by the 0.01% in the seventeenth century; if we replace it with a hypothetical

173

1%, the R^2 jumps to 0.912. Another benefit of logistic models is that they yield intercepts based on the curve's position on the x-axis, and in this case we expect those intercepts to correspond to the relative date that the change occurred. This opens up the possibility of testing for correlation between the initial (sixteenth century) token frequencies of the five verbs and their intercepts (as shown in the table). Unfortunately, the correlation is small ($r = 0.426$), but this is a promising avenue for future research.

8 Conclusion

The aim of this study was to investigate the well-known changes that have taken place in the expression of negation in French from a usage-based point of view. The Latin preverbal negator *non* was reduced to *ne*, and then supplemented by a choice of ex-nouns, most notably *pas*, *point* and *mie*. Eventually, *ne* formed "embracing" constructions with these postverbal particles, and *ne ... pas* came to be the dominant sentence negator. Later, people began to omit the preverbal *ne* and just use *pas* to negate sentences.

Specific predictions have been made by Bybee and Thompson (1997) as to the role of type frequency and token frequency in the evolution of grammatical changes. I have used the theories of Lotka (1925), Volterra (1926), Kroch (1989), Hopper (1991), Croft (2000), Detges and Waltereit (2002), Eckardt (2007) and Schwenter (2006) to develop a quantitative framework around the predictions of Bybee and Thompson and apply them to a corpus of French theatrical data, and thereby test Bybee and Thompson's predictions.

The results of the corpus study fit the major hypotheses well: before 1600 playwrights seem to have made a functional distinction between preverbal *ne* alone and the embracing constructions, although it allowed a lot of ambiguity. Some time after 1600 the distinction broke down, and *ne ... pas* was apparently metanalyzed to be equivalent to preverbal *ne* alone, after the manner described by Croft (2000), Eckardt (2007) and Schwenter (2006).

From this point on the two embracing constructions *ne ... pas* and *ne ... point* increased in use at the expense of preverbal *ne* alone. These changes propagated through

the lexicon in S-shaped progressions that reflect the role of type frequency in determining future changes, just as the size of the human population in Malthus's (1789) model determined its rate of increase, and as predicted by Bybee and Thompson's (1997) theory. The prevalence of both tokens and types followed the same logistic model, but the rate of change of the prevalence of tokens was higher, as predicted by Bybee and Thompson's model.

Kroch's (1989) logistic model needed to be extended to deal with the kind of competition among constructions that Hopper (1991) referred to as *specialization*, and I accomplished this by applying Lotka (1925) and Volterra's (1926) equations modeling resource competition. These models turned out to fit the results just as well when main verbs were measured as types as opposed to conjugated verbs, and better when high-token-frequency verbs and hapaxes were included in the type frequency counts. Finally, I found that high-token-frequency forms were resistant to change. These results are satisfying, but they should be generalized outside of the realm of theater only with great caution.

There are a number of avenues for further research. The α values generated by the least squares method for the Lotka-Volterra model can be further interpreted for their significance in describing the competition among constructions. The various catalogs of dramatic performances can be coordinated and adapted to provide a sampling method for gathering texts into a corpus. New texts are being brought online every day, and they can be formatted, tagged and annotated to adapt them for this kind of task. More texts can be examined, especially from the fifteenth and sixteenth centuries. The study can be extended to other genres, such as poetry or letters. The journal of Jean Héroard can be

176

examined in more detail. The question of whether modal auxiliaries count as *types* in type frequency is worth investigating. The present study focuses on declarative sentences, but there is likely some interesting information to be had from imperative and interrogative sentences. To investigate the possibility that the formation of a larger schema motivated some of this change, other negative constructions in parallel changes can be examined, and even completely unrelated morphosyntactic changes, in French or other languages.

This is a promising quantitative extension of Bybee and Thompson's (1997) model. Its predictions are consistent with the data, and suggest that we will find further support in other genres, and when applied to other changes. A few methodological issues about language change emerged from the data, most notably support for Bybee's (1985) claim that high frequency constructions do not participate very much in the formation of broader schemas, and that they should be set aside from the study of any change that fits under the general principle of analogy. Another is that the S-curves observed by sociolinguists represent lexical propagation as well as social propagation.

List of Appendices

Appendix A. Corpus list

This project has been made tremendously easier by several archive projects. First are the reprinting of several medieval plays in modern type by nineteenth and early twentieth-century scholars like Francisque Michel and Gaston Paris. Second are the large-scale scanning projects like the *Gallica* project of the Bibliothèque Nationale Française, and *Google Book Search*, which provided scanned versions of many public-domain books, including these. Then there are corpus projects like the *ARTFL/FRANTEXT* project of the Centre National de la Recherche Scientifique and the University of Chicago, the *Laboratoire de Français Ancien* at the University of Ottawa, and *Editions Gentleman-Cambrioleur* at the University of Indiana, which provided works in proofread, full-text format. Finally there are researchers working independently or in small groups, such as Jeffrey Graf and Graham Runnals, who have made proofread texts available.

The Ordo Representacionis Ade was the only book that I needed to scan myself. The Gallica and Google Book Search texts were not in full-text format. Although Google Book Search now offers raw OCRed text, at the time I was assembling the corpus this was not available. I therefore used ABBYY PDF Transformer to OCR the texts, and proofread them myself. Unless noted, all texts are OCRed and proofread by myself.

Here is the list of plays in chronological order:

Anonymous. ca. 1160 [1967]. Ordo Representacionis Ade (A1-Adam). *Trois pièces médiévales*, ed. by A. Robert Harden. New York: Appleton.

Bodel, Jehan. ca. 1200 [1929]. Jus de Saint Nicholai. (A1-Nichola). *Théâtre français au Moyen Age*, ed. by Louis J. N. Monmerqué and Francisque Michel. Paris: Firmin-Didot. Scanned by the Gallica project.

Anonymous. ca. 1228 [1922]. Li Courtois d'Arras. (A1-Arras). *Courtois d'Arras; Jeu du XIIe siècle,* ed. by Edmond Faral. Paris: Honoré Champion. Second edition revised. Scanned and OCRed by Editions Gentleman-Cambrioleur.

Rutebeuf. ca. 1260 [1929]. Miracle de Theophile. (A1-Theoph1). *Théâtre français au Moyen Age*. Scanned and OCRed by Editions Gentleman-Cambrioleur.

Halle, Adam de la. ca. 1262 [1929]. Li Jus Adan, ou de la Feuillie. (A1-Feuille). *Théâtre français au Moyen Age.* Scanned and OCRed by Editions Gentleman-Cambrioleur.

Anonymous. ca. 1270 [1921]. Le jeu du garçon et de l'aveule. (A1-Aveugle). *Le Garçon et l'aveugle; jeu du XIIIe siècle*, ed. by Mario Roques. Second edition, revised. Paris: Honoré Champion, 1921. Scanned and OCRed by Editions Gentleman-Cambrioleur.

Anonymous, ca. 1275 [1929]. Résurrection du sauveur. (A1-Resurr). *Théâtre français au Moyen Age*. Scanned and OCRed by the Centre d'Études des Textes Médiévaux, Université Rennes 2.

Anonymous, ca. 1278 [1929]. De Pierre de la broche qui dispute a fortune par devant raison. (A1-Broche). *Théâtre français au Moyen Age*.

Halle, Adam de la. ca. 1285 [1998]. *Jeu de Robin et Marion* (A1-Robin), ed. by Olivier Bettens, Geneva: Opéra-Studio de Genève. Virga.org.

Anonymous, 1344 [1883]. Miracle de saint Jehan Crisothomes. (MirPer6). *Miracles de Nostre Dame par personnages*, ed. by Gaston Paris et Ulysse Robert. Paris: Firmin Didot. Scanned and OCRed by the Laboratoire de Français Ancien.

Anonymous, 1348 [1883]. Miracle de l'evesque a qui Nostre Dame s'apparut. (MirPer10). *Miracles de Nostre Dame par personnages.* Scanned and OCRed by the Laboratoire de Français Ancien.

Anonymous, 1371 [1883]. Miracle de la fille du roy de Hongrie. (MirPer29). *Miracles de Nostre Dame par personnages*. Scanned and OCRed by the Laboratoire de Français Ancien.

Anonymous, 1380 [1883]. Miracle de saint Lorens. (MirPer38). *Miracles de Nostre Dame par personnages.* Scanned and OCRed by the Laboratoire de Français Ancien.

Anonymous, ca. 1424 [1949]. Farce française du quinzième siècle. (ABU-farcef). *Recueil de farces françaises inédites du XVe siècle*, ed. by Gustave Cohen. Cambridge, MA: Medieval Academy of America. Scanned and OCRed by Ida Nelson, Association des Bibiliphiles Universels (ABU).

Anonymous, ca. 1440 [1837]. Martyre de Saint Étienne. (A1-Etienne). *Mystères inédits du XVe siècle*, ed. by Achille Jubinal. Paris: Techener. Scanned and OCRed by the Centre d'Études des Textes Médiévaux.

Anonymous, ca. 1440 [1837]. Convercion de S. Pol. (A1-Pol). *Mystères inédits du XVe siècle*, ed. by Achille Jubinal. Paris: Techener. Scanned and OCRed by the Centre d'Études des Textes Médiévaux.

Anonymous, ca. 1440 [1837]. Conversion de S. Denis. (A1-Denis). *Mystères inédits du XVe siècle*, ed. by Achille Jubinal. Paris: Techener. Scanned and OCRed by the Centre d'Études des Textes Médiévaux.

Anonymous, ca. 1440 [1837]. Comment S.Père et S.Pol alèrent à Romme et comment ilz furent martirez. (A1-PerePol). *Mystères inédits du XVe siècle*, ed. by Achille Jubinal. Paris: Techener. Scanned and OCRed by the Centre d'Études des Textes Médiévaux.

Anonymous, 1470 [1859]. Maistre Pierre Pathelin : farce du XVe siècle. (A1-Patheli). *Recueil de farces, sotties et moralités du quinzième siècle*, ed. by Bibliophile Jacob. Paris: Delahays. Scanned by Gallica, N0028126.

Villon, François. 1474 [1859]. Nouveau Pathelin. (A1-NPathel). *Recueil de farces, sotties et moralités du quinzième siècle*. Scanned by Gallica, N0028126.

Anonymous, ca. 1475 [1880]. Farce nouvelle tresbonne et fort joyeuse du cuvier. (A1-Cuvier1). *Nouveau recueil de farces françaises des XVe et XVIe siècles, d'après un volume unique appartenant à la Bibliothèque Royale de Copenhague*, ed. by Emile Picot and Christophe Nyrop. Paris: Damascène Morgand & Charles Fatout. Scanned and OCRed by Editions Gentleman-Cambrioleur.

Anonymous, ca. 1478 [1859]. Sottie a trois personnages. (A1-S3Pers). *Recueil de farces, sotties et moralités du quinzième siècle*.

Anonymous, ca. 1480 [1878]. *Mystère du Viel Testament* (A1-Vielt), ed. by James de Rothschild. Paris: Firmin Didot. Scanned by Gallica, N0005051-N0005056.

Vigne, André de la 1496 [1859]. L'aveugle et le boiteux. (A1-Boiteux). *Recueil de farces, sotties et moralités du quinzième siècle*. Scanned by Gallica, N0028126.

Anonymous, ca. 1520 [1996]. *Vie et hystoire de ma dame saincte Barbe par personnages* (A1-Barbe), ed. by Mario Longtin. Centre d'Études des Textes Médiévaux.

Tasserie, Guillaume, ca. 1520 [1908]. *Triomphe des Normans* (A1-Normans), ed. by Pierre Le Verdier. Rouen: Léon Gy. Scanned and OCRed by the Centre d'Études des Textes Médiévaux.

Thibault, Guillaume, ca. 1520 [1908]. *De la dame à l'agneau* (A1-Agneau), ed. by Pierre Le Verdier. Rouen: Léon Gy. Scanned and OCRed by the Centre d'Études des Textes Médiévaux.

Anonymous, ca. 1530 [1980]. *Mystère de Sainte Venice* (A1-Venice), ed. by Graham A. Runnalls. Exeter: Exeter French Text Studies. Centre d'Études des Textes Médiévaux.

Anonymous, ca. 1540 [1880]. Farce moralisee à quatre personnages. (A1-Farce4). *Nouveau recueil de farces françaises des XVe et XVIe siècles*. Scanned and OCRed by Editions Gentleman-Cambrioleur.

Marot, Clément. 1542 [1880]. Dialogue nouveau fort joyeux. (A1-Dialogu). *Nouveau recueil de farces françaises des XVe et XVIe siècles*. Scanned and OCRed by Editions Gentleman-Cambrioleur.

d'Abondance, Jean. 1545 [1977]. *La Cornette, Farce de Jean d'Abondance* (A1-Cornett), ed. by Michel Rousse, Université de l'Ouest. Scanned and OCRed by Denis Hüe, Centre d'Études des Textes Médiévaux.

Jodelle, Etienne. 1553 [1855]. La Comédie d'Eugène. (A1-Eugene). *Ancien théâtre françois*, v. 4, ed. by Eugène Viollet-le-Duc. Paris: Plon. Scanned by Gallica.

Anonymous. ca. 1556 [1912]. Le Pelerinage de mariage. (A1-PelMar). *Recueil général des sotties*, v. 3, ed. by Emile Picot. Paris: Firmin Didot. Scanned by Gallica.

Anonymous. ca. 1571 [1912]. Farce nouvelle de trois galans et un badin. (A1-Galans). *Recueil général des sotties*, v. 3. Scanned by Gallica, N0005088.

Garnier, Robert. 1578 [1950]. Bradamante. (A1-Bradam). *Bradamante / Les Juifves*, ed. by Marcel Hervier. Paris: Garnier. Scanned and OCRed by Bibliopolis.

de Larivey, Pierre. 1579. Le Laquais. (A1-Laquais). *Ancien théâtre françois*, v. 5, ed. by Eugène Viollet-le-Duc. Paris: Plon. Scanned by Gallica, N0027670.

Garnier, Robert. 1583. Les Juivfes. (A1-Juifves). *Bradamante / Les Juifves*, ed. by Marcel Hervier. Paris: Garnier. Scanned and OCRed by the Bibliopolis society.

Montchrestien, Antoine de. 1601 [1891]. La Reine d'Escosse. (REcos). *Les tragédies de Montchrestien*, ed. by Louis Petit de Julleville. Paris: Plon. Transcribed by G. Mallary Masters, ATHENA.

Tabarin. 1622 [1833]. Farces de Tabarin. (A1-Tabarin). *Les Œuvres de Tabarin*, ed. by Georges d'Harmonville. Paris: Garnier. Scanned by Gallica, N0204373.

Corneille, Pierre. 1631 [1862]. Clitandre. (Clitd). *Œuvres de P. Corneille*, ed. by Charles Marty-Laveaux. Paris: Hachette. Transcribed by ARTFL.

Scudéry, Georges de. 1641. *Eudoxe*. (Eudox). Paris: Courbe. Transcribed by ARTFL.

Molière and Philippe Quinault. 1671. *Psyché: Tragédie-ballet*. (PsychM). Transcribed by Wikisource.

Quinault, Philippe. 1675. *Thesée*. (These). Paris: Ballard. Transcribed by ARTFL.

Regnard, Jean-François. 1700 [1820]. Le Retour Imprévu. (ReImp). *Œuvres completes, v. 3*. Paris: Haut-Coeur. Transcribed by ARTFL.

Crébillon, Prosper Jolyot de, père. 1707 [1812]. Atrée et Thyeste. (AsThy). *Œuvres de Crébillon, v 1*. Paris: Didot l'Ainé. Transcribed by ARTFL.

Marivaux, Pierre de. 1727 [1968]. La seconde surprise de L'amour. (sSuAm). *Théâtre complet de Marivaux*, ed. by Frédéric Deloffre. Paris: Garnier. Transcribed by ARTFL.

Grafigny, Françoise de. 1751. *Cénie*. (Genie). Paris: Cailleau. Transcribed by ARTFL.

Carmontelle, 1781 [1822]. L'Uniforme de Campagne. (UniCa). *Manuel de proverbes dramatiques*, nouvelle édition. Paris: Delongchamps. Transcribed by ARTFL.

Pixérécourt, Guilbert de. 1803. *Coelina ou L'Enfant du Mystere*. (Coeli). Paris: Barba. Transcribed by ARTFL.

Musset, Alfred de. 1834 [1952]. On ne badine pas avec L'amour. (Badin). *Comédies et proverbes*. Paris: Les Belles-Lettres. Transcribed by ARTFL.

Dumas, Alexandre, fils. 1858. Le Fils Naturel. (FilsN). *Théâtre complet, v. 3*. Paris: Calman-Levy. Transcribed by ARTFL.

Meilhac, Henri and Ludovic Halévy. 1877. La Cigale. (Cigal). *Théâtre de Meilhac et Halévy, v. 3*. Paris: Calmann-Levy. Transcribed by ARTFL.

Rostand, Edmond. 1897. *Cyrano de Bergerac*. (CyrBe). Paris: Fasquelle. Transcribed by ARTFL.

Flers, Robert de, and Gaston-Armand de Caillavet. 1914 [1923]. *Monsieur Brotonneau: comédie en trois actes*. (MBrot). Paris: L'Illustration. Transcribed by ARTFL.

Giraudoux, Jean. 1929. *Amphitryon 38*. (Amphi). Paris: Grasset. Transcribed by Ebooksgratuits.com.

Appendix B: Token distribution data

The table in this section gives the distribution of tokens among the various negators in the corpus.

Date	ARTFL	alone	pas	point	mie	Total
1160	A1-Adam	48	13			61
1200	A1-Nichola	56	8	2	14	80
1228	A1-Arras	42	6	1	8	57
1260	A1-Theoph1	32	10	1	1	44
1262	A1-Feuille	42	4	7	20	73
1270	A1-Aveugle	18	4		6	28
1275	A1-Resurr	18	2		2	22
1278	A1-Broche	4	3		3	10
1285	A1-Robin	41	2	3	9	55
1344	MirPer6	67	22	5	12	106
1348	MirPer10	25	5	2	3	35
1371	MirPer29	102	26	10	7	145
1380	MirPer38	74	22	12	6	114
1424	ABU-farcef	11	3	1		15
1440	A1-Etienne	9	3	2		14
1441	A1-Pol	8	3	2		13
1442	A1-Denis	9	6	1	1	17
1443	A1-PerePol	25	7	1	1	34
1470	A1-Patheli	88	26	13	3	130
1474	A1-NPathel	21	22	7		50
1475	A1-Cuvier1	12	5	6		23
1478	A1-S3Pers	1	2	1		4
1480	A1-Vielt	106	34	40	5	185
1496	A1-Boiteux	18	1	1		20
1520	A1-Agneau	6	6	2		14
1520	A1-Barbe	115	18	17	6	156
1520	A1-Normans	39	11	13	4	67
1530	A1-Venice	21	2	2	1	26
1540	A1-Farce4	19	12	11		42
1542	A1-Dialogu	11	5	4		20
1545	A1-Cornett	23	8	3	2	36
1553	A1-Eugene	75	14	16		105
1556	A1-PelMar	24	11	4	1	40
1571	A1-Galans	8	8	2		18
1578	A1-Bradam	148	31	37		216
1579	A1-Laquais	221	42	23		286

Date	ARTFL	alone	pas	point	mie	Total
1583	A1-Juifves	115	27	25		167
1601	REcos	68	14	23		105
1622	A1-Tabarin	18	29	19		66
1631	Clitd	61	34	28		123
1641	Eudox	69	59	66		194
1671	PsychM	51	49	41		141
1675	These	18	25	23		66
1700	ReImp	15	62	21		98
1707	AsThy	22	22	54		98
1728	sSuAm	31	137	81		249
1751	Cenie	53	94	48		195
1781	UniCa	7	81	3		91
1803	Coeli	27	81	32		140
1834	Badin	25	105	4		134
1858	FilsN2	26	335			361
1877	Cigal	7	241	1		249
1897	CyrBe	28	159			187
1914	MBrot	8	254	1		263
1929	Amphi	17	192	9		218

Appendix C: Social class and use of *ne … pas*

This table gives the two characters in each play who produced the most declarative negations, the play code, their name, the social class I assigned to them, the number of declarative negations, the percent that used *ne … pas*, the percent we would expect based on an ideal logistic function, and the difference between those last two. In two of the plays, there was only one character who produced any declarative negations.

Date	Play	Character	Class	N	Pas %	Ideal
1160	A1-Adam	Adam	3	19	21%	3%
1160	A1-Adam	Figura	3	11	18%	3%
1200	A1-Nichola	Clikes	1	27	0%	4%
1200	A1-Nichola	Pincedes	1	33	11%	4%
1228	A1-Arras	Courtois	1	14	7%	5%
1228	A1-Arras	Pourette	1	17	9%	5%
1260	A1-Theoph1	Theophiles	2	4	21%	6%
1260	A1-Theoph1	Salatins	3	4	50%	6%
1262	A1-Feuille	Adans	2	17	21%	6%
1262	A1-Feuille	Maistre Henris	2	24	0%	6%
1270	A1-Aveugle	Li Garçons	1	19	12%	6%
1270	A1-Aveugle	Li Aveules	1	28	18%	6%
1275	A1-Resurr	Miles	1	17	50%	6%
1275	A1-Resurr	Longinus	2	7	0%	6%
1278	A1-Broche	Fortune	3	3	25%	6%
1278	A1-Broche	Pierre	3	3	33%	6%
1285	A1-Robin	Marions	1	8	6%	7%
1285	A1-Robin	Robins	1	6	0%	7%
1344	MirPer6	La Mére Anthure	3	63	21%	9%
1344	MirPer6	Jehan	2	18	33%	9%
1348	MirPer10	L'Evesque	2	16	16%	10%
1348	MirPer10	L'Ermite	2	3	0%	10%
1371	MirPer29	La Fille	3	41	32%	11%
1371	MirPer29	La Mére	3	15	6%	11%
1380	MirPer38	Dacien	1	7	29%	11%
1380	MirPer38	Lorens	2	27	7%	11%
1424	ABU-farcef	Regnault	2	12	0%	15%
1424	ABU-farcef	Godin	2	6	25%	15%

Date	Play	Character	Class	N	Pas %	Ideal
1440	A1-Etienne	Le Premier	2	20	33%	16%
1440	A1-Etienne	Estiene	2	11	33%	16%
1441	A1-Pol	S. Pol	2	15	33%	16%
1442	A1-Denis	S. Pol	2	18	50%	16%
1442	A1-Denis	S. Denis	2	21	33%	16%
1443	A1-PerePol	S. Pierre	1	7	17%	16%
1443	A1-PerePol	Marcel	2	33	20%	16%
1470	A1-Patheli	Pathelin	2	48	24%	19%
1470	A1-Patheli	Le Drappier	2	51	18%	19%
1474	A1-NPathel	Pathelin	2	38	39%	19%
1474	A1-NPathel	Le Pelletier	2	33	39%	19%
1475	A1-Cuvier1	Jacquinot	1	31	25%	19%
1475	A1-Cuvier1	La Femme	3	62	17%	19%
1478	A1-S3Pers	Le Premier	2	47	67%	19%
1480	A1-Vielt	Cayn	3	13	12%	20%
1480	A1-Vielt	Adam	3	31	17%	20%
1496	A1-Boiteux	Le Boiteux	1	32	7%	21%
1496	A1-Boiteux	L'Aveugle	1	91	0%	21%
1520	A1-Agneau	Noble Ceur	3	49	29%	24%
1520	A1-Barbe	Barbe	3	44	22%	24%
1520	A1-Normans	Sarquis	3	34	17%	24%
1520	A1-Agneau	La Dame a l'aigneau	2	52	33%	24%
1520	A1-Barbe	La Folle Femme	1	64	7%	24%
1520	A1-Normans	Le Duc	3	78	40%	24%
1530	A1-Venice	Vaspasien	3	76	0%	25%
1530	A1-Venice	Veronne	2	122	20%	25%
1540	A1-Farce4	Le Premier Mary	2	99	25%	26%
1540	A1-Farce4	Le Second Mari	2	11	27%	26%
1542	A1-Dialogu	Second	2	9	18%	26%
1542	A1-Dialogu	Premier	2	11	25%	26%
1545	A1-Cornett	La Femme	2	4	13%	27%
1545	A1-Cornett	Le Mary	2	9	25%	27%
1553	A1-Eugene	Arnauld	2	11	17%	28%
1553	A1-Eugene	Messire Jean	2	4	6%	28%
1556	A1-PelMar	Le viel Pelerin	1	3	43%	28%
1556	A1-PelMar	Le jeune Pelerin	1	9	25%	28%
1571	A1-Galans	Le Badin	2	18	43%	30%
1571	A1-Galans	Le Deuxième Galant	2	5	50%	30%
1578	A1-Bradam	Bradamante	2	18	3%	31%
1578	A1-Bradam	Beatrix	2	15	7%	31%
1579	A1-Laquais	Valère	1	4	19%	31%
1579	A1-Laquais	Lucian	3	3	4%	31%
1583	A1-Juifves	Amital	3	6	22%	32%
1583	A1-Juifves	Nabuchodonosor	3	5	20%	32%

Date	Play	Character	Class	N	Pas %	Ideal
1601	REcos	Choeur	3	33	18%	34%
1601	REcos	Reine d'Escosse	2	18	14%	34%
1622	A1-Tabarin	Tabarin	2	6	42%	37%
1622	A1-Tabarin	Lucas	2	29	36%	37%
1631	Clitd	Pymante	3	5	16%	39%
1631	Clitd	Dorise	3	3	42%	39%
1641	Eudox	Ursace	2	15	31%	40%
1641	Eudox	L'Imperatrice	3	10	30%	40%
1671	PsychM	Psyché	3	5	28%	45%
1671	PsychM	L'Amour	3	11	36%	45%
1675	These	Medée	3	8	38%	45%
1675	These	Aeglé	3	8	20%	45%
1700	ReImp	Merlin	1	17	74%	49%
1700	ReImp	Géronte	2	8	48%	49%
1707	AsThy	Plisthène	3	6	19%	50%
1707	AsThy	Atrée	3	29	30%	50%
1728	sSuAm	La Marquise	3	47	51%	54%
1728	sSuAm	Le Chevalier	2	35	45%	54%
1751	Cenie	Dorimond	2	21	49%	57%
1751	Cenie	Méricourt	3	11	44%	57%
1781	UniCa	M. Bétassier	2	19	89%	62%
1781	UniCa	M. Gobergeau	2	44	95%	62%
1803	Coeli	Tiennette	1	25	53%	65%
1803	Coeli	Dufour	2	10	39%	65%
1834	Badin	Perdican	3	21	88%	69%
1834	Badin	Camille	3	30	81%	69%
1858	FilsN2	Aristide	2	53	95%	72%
1858	FilsN2	Jacques	2	32	90%	72%
1877	Cigal	Marignan	2	19	95%	75%
1877	Cigal	La Cigale	1	31	97%	75%
1897	CyrBe	Cyrano	3	43	84%	77%
1897	CyrBe	Roxane	3	63	90%	77%
1914	MBrot	Brotonneau	2	69	97%	79%
1914	MBrot	Louise	2	29	100%	79%
1929	Amphi	Alcmène	3	46	92%	80%
1929	Amphi	Jupiter	3	33	85%	80%

Appendix D. Raw type (verb) data

The following tables give the main verbs found in each century. The most frequent quartile of verbs (marked in red) and the hapaxes (verbs that appeared only once, marked in orange) were excluded from the type frequency counts.

1100		
Total number of verbs	8	
Quartile	5.75	
pouvoir	alone-pas	14
etre	alone-pas	11
faire	alone-pas	7
savoir	alone	6
avoir	alone	6
porter	alone	2
croire	alone	2
entendre	alone	2
toucher	alone	1
tenir	alone	1
tarder	alone	1
vouloir	alone	1
songer	pas	1
dire	pas	1
sentir	alone	1
valoir	alone	1
prendre	alone	1
juger	alone	1
devoir	alone	1
offendre	alone	1
oser	alone	1
rendre	alone	1
boire	alone	1
sembler	alone	1
Total number of verbs		24
Number of verbs counted		6

1200		
Total number of verbs	43	
Quartile	32	
etre	alone-mie-pas-point	41
avoir	alone-mie-pas-point	38

savoir	alone-mie	34
pouvoir	alone-mie-pas	23
vouloir	alone-mie-pas-point	13
chaloir	alone	12
faire	alone-mie-pas-point	10
aller	alone-mie-pas	9
voir	alone-mie	9
oser	alone	7
dire	alone-mie	7
prendre	alone-pas	6
venir	alone-mie	6
valoir	alone-mie-pas	5
aimer	alone-mie-pas	5
tenir	alone-mie-point	5
devoir	alone-pas-point	5
oir	alone-pas	4
esmaier	alone-mie-pas	3
trouver	alone-mie	3
y avoir	alone	3
ocire	alone-mie	3
querir	alone-pas	3
crever	alone	3
apprendre	alone-mie-pas	3
baiser	alone	2
perdre	alone-mie	2
croire	alone-pas	2
revenir	alone-mie	2
manger	alone	2
boire	alone	2
falloir	alone-mie	2
donner	alone	2
aferir	alone-point	2
forfaire	mie	2
mesquer	alone	2
douter	alone-mie	2
pendre	alone	2
remuer	alone	2
mettre	alone-point	2
gouter	alone	2
laisser	alone	2
frapper	alone	2
anuir	mie	1
vener	alone	1
atir	mie	1
husser	alone	1

cuider	pas	1
anuire	alone	1
conter	point	1
payer	alone	1
plaire	pas	1
boir	alone	1
appeler	mie	1
aconter	alone	1
penser	alone	1
lier	alone	1
radoter	alone	1
dailler	alone	1
aloir	alone	1
maintenir	pas	1
faindre	pas	1
arreter	alone	1
cailler	alone	1
rejeter	mie	1
reprendre	alone	1
excuser	alone	1
suivre	mie	1
suffire	mie	1
lever	alone	1
surprendre	alone	1
grier	alone	1
mentir	alone	1
renaitre	alone	1
vergonder	alone	1
tailler	alone	1
avioir	alone	1
perir	alone	1
jeter	alone	1
estovoir	alone	1
ennuyer	alone	1
jouer	alone	1
loir	alone	1
ravir	alone	1
abandonner	alone	1
remaindre	alone	1
couper	alone	1
convenir	alone	1
souffrir	alone	1
recroire	mie	1
fendre	alone	1
oublier	pas	1

guetter	point	1
basser	alone	1
accroitre	alone	1
contenter	alone	1
briser	alone	1
battre	alone	1
envoyer	alone	1
entamer	alone	1
finer	alone	1
guarder	alone	1
esrager	alone	1
marier	alone	1
monteployer	alone	1
decevoir	alone	1
sembler	mie	1
durer	mie	1
huer	alone	1
attendre	point	1
aider	alone	1
damner	alone	1
apercevoir	alone	1
remettre	alone	1
Total number of verbs		115
Number of verbs counted		32

1300		
Total number of verbs	41	
Quartile	30.5	
etre	alone-mie-pas-point	65
savoir	alone-pas-point	53
pouvoir	alone-mie-pas	31
avoir	alone-mie-pas-point	31
vouloir	alone-mie-pas-point	20
faire	alone-mie-pas-point	18
finir	alone	12
laisser	alone-mie-pas-point	8
voir	alone-pas-point	7
y avoir	alone-point	7
chaloir	alone	7
mentir	alone-pas-point	6
venir	alone	5
douter	alone-mie	5
oir	alone	5
falloir	alone-mie-pas	4

194

penser	alone-mie	4
devoir	alone-pas	3
mettre	alone	3
deplaire	alone-mie-point	3
reconnaitre	pas-point	3
arder	alone	3
garder	alone	3
ennuyer	alone	3
querir	alone	3
aller	alone-pas	3
appartenir	mie-pas-point	3
accomplir	alone	3
tourner	alone	2
sacrifier	alone	2
demourer	pas-point	2
dire	alone-pas	2
partir	alone	2
refuser	pas	2
oser	alone	2
prendre	alone	2
noyer	alone	2
aconter	alone	2
plaire	alone	2
connaitre	alone	2
decevoir	alone-pas	2
faillir	alone	1
echapper	alone	1
eissir	alone	1
departir	alone	1
tarder	point	1
decheoir	point	1
valoir	pas	1
delayer	point	1
enlasser	alone	1
encommencer	point	1
mostrer	alone	1
passer	alone	1
oter	pas	1
estriver	pas	1
souffrir	pas	1
voulenter	pas	1
defendre	point	1
embesognier	pas	1
guerroyer	mie	1
envoyer	alone	1

perdire	alone	1
sacrificier	alone	1
surmonter	alone	1
delivrer	alone	1
mandre	pas	1
dedire	mie	1
targuer	mie	1
alisser	alone	1
apprendre	alone	1
vois	alone	1
apercevoir	alone	1
perdre	mie	1
oublier	pas	1
convenir	point	1
vendre	mie	1
vivre	alone	1
demeurer	mie	1
boire	alone	1
cesser	alone	1
secourer	alone	1
ressembler	alone	1
courroucer	pas	1
engaigner	pas	1
demander	point	1
errer	alone	1
nier	alone	1
deindre	alone	1
destrier	alone	1
donner	alone	1
entendre	alone	1
quitter	pas	1
Total number of verbs		93
Number of verbs counted		31

1400		
Total number of verbs	45	
Quartile	33.5	
etre	alone-mie-pas-point	83
savoir	alone-pas	54
pouvoir	alone	42
avoir	alone-pas-point	31
faire	alone-mie-pas	19
chaloir	alone	17
falloir	alone-pas-point	15

y avoir	alone-pas-point	14
voir	alone-pas-point	12
vouloir	alone-mie-pas	11
dire	alone-pas-point	9
devoir	alone-pas-point	8
oser	alone	6
venir	alone-point	6
revenir	alone-pas	5
prendre	alone-pas	5
valoir	mie-pas	5
aller	alone-pas	4
tenir	alone-pas-point	4
trouver	alone-point	4
penser	alone-pas-point	4
parler	alone-pas-point	4
cesser	alone	3
croire	alone-pas-point	3
donner	alone-pas	3
manger	alone-pas-point	3
payer	alone-point	3
apprendre	alone-pas-point	3
douter	alone-point	3
connaitre	alone-pas	3
appercevoir	alone-mie	2
accorder	alone-mie	2
secourer	alone	2
offenser	alone-point	2
sembler	point	2
monter	alone-point	2
retourner	alone	2
eissir	alone-pas	2
passer	alone-point	2
deplaire	alone	2
partir	alone-point	2
pourvoir	alone	2
meriter	alone-pas	2
faillir	pas	2
assommer	alone	2
lire	pas	1
decevoir	alone	1
displacer	alone	1
forvoir	alone	1
craindre	alone	1
jeter	alone	1
mourir	alone	1

contraindre	alone	1
hanter	point	1
renter	pas	1
finir	alone	1
entendre	pas	1
souper	point	1
attendre	alone	1
hater	pas	1
souffrir	alone	1
promettre	alone	1
ecrire	alone	1
reprendre	alone	1
soucier	alone	1
porter	point	1
meler	point	1
laisser	point	1
irriter	point	1
reputer	point	1
tromper	alone	1
guerir	alone	1
chier	alone	1
vivre	point	1
user	point	1
esperer	point	1
entremeler	point	1
cuider	pas	1
conduire	pas	1
affoler	pas	1
sentir	alone	1
escondir	pas	1
oublier	pas	1
endormir	point	1
bailler	point	1
avancer	point	1
sortir	alone	1
desrober	alone	1
emprisonner	alone	1
embler	alone	1
punir	pas	1
depecher	alone	1
envoler	alone	1
finer	alone	1
mentir	alone	1
hober	alone	1
garder	alone	1

demeurer	alone	1
suffire	pas	1
rire	alone	1
rappeler	alone	1
meschier	alone	1
oter	alone	1
couper	alone	1
cheminer	alone	1
amener	alone	1
attirer	alone	1
bouger	alone	1
nommer	alone	1
tarder	mie	1
attaindre	point	1
facer	alone	1
oir	point	1
accroitre	point	1
plaider	point	1
recorder	point	1
muire	alone	1
devaler	pas	1
repondre	point	1
vendre	pas	1
tollir	pas	1
pourvoyer	alone	1
pleuvoir	alone	1
commencer	alone	1
souvenir	alone	1
choisir	mie	1
gabber	pas	1
devoyer	pas	1
mettre	alone	1
jouer	alone	1
Total number of verbs		131
Number of verbs counted		34

1500		
Total number of verbs	99	
Quartile	74	
etre	alone-mie-pas-point	201
avoir	alone-pas-point	103
savoir	alone-mie-pas-point	103
pouvoir	alone-mie-pas	88
vouloir	alone-mie-pas-point	64

faire	alone-pas-point	40
devoir	alone-pas-point	24
voir	alone-pas-point	23
falloir	alone-mie-pas-point	23
y avoir	alone-pas-point	22
entendre	alone-pas-point	14
dire	alone-pas-point	14
venir	alone-pas-point	13
prendre	alone-pas-point	12
oser	alone-pas	11
trouver	alone-pas-point	10
tenir	alone-pas	9
craindre	alone-pas-point	8
connaitre	alone-point	7
aimer	alone-pas-point	7
plaire	alone-pas-point	6
parler	alone-pas-point	6
croire	alone-pas-point	6
penser	alone-point	6
mourir	alone-mie-pas	6
deplaire	alone-point	6
chaloir	alone	6
laisser	alone-point	5
entrer	alone-pas	5
punir	alone-pas	5
apprendre	alone-point	5
aller	alone-pas	5
permettre	alone-pas	4
souvenir	alone-point	4
decevoir	alone	4
douter	alone-mie	4
demander	alone-point	4
reprendre	alone-pas	3
celer	alone-mie	3
rendre	alone	3
souffrir	alone-pas-point	3
hasarder	alone	3
manger	alone-pas-point	3
servir	alone-pas	3
toucher	alone-point	3
porter	alone	3
manquer	alone-point	3
adresser	alone-pas	3
hanter	alone	3
tuer	alone-pas	3

faillir	alone	3
oir	alone	3
bouger	alone	3
adoucir	alone	2
oter	alone	2
donner	pas-point	2
appercevoir	pas	2
daindre	alone	2
endurer	alone-pas	2
contredire	alone-point	2
fier	alone	2
perdre	alone	2
lire	alone	2
demeurer	alone	2
commettre	alone	2
invoquer	alone-point	2
lever	alone-pas	2
tirer	alone	2
luir	alone-point	2
mentir	alone-point	2
sembler	alone-pas	2
retourner	alone	2
rire	alone	2
meriter	alone-pas	2
sentir	alone-point	2
dormir	alone-pas	2
decouvrir	alone-pas	2
valoir	alone-pas	2
forcer	alone-pas	2
ecouter	alone	2
esperer	pas-point	2
echapper	mie-pas	2
lasser	alone-pas	2
battre	alone	2
ravir	alone	2
delivrer	alone	2
refroidir	alone	2
regner	alone-pas	2
crier	alone	2
reduire	alone-pas	2
ressentir	alone	2
cesser	alone	2
mordre	alone-pas	2
reprouver	point	2
finer	alone	2

mettre	alone	2
attaindre	alone	2
jouer	alone	2
acoutumer	point	2
respirer	alone	1
emplir	alone	1
deliberer	alone	1
excuser	point	1
brandler	alone	1
consentir	alone	1
accoutumer	alone	1
rider	alone	1
user	alone	1
approuver	alone	1
amender	alone	1
gemir	point	1
reboutter	point	1
rompre	pas	1
acquitter	point	1
chanceler	point	1
vomir	alone	1
violer	alone	1
sortir	alone	1
secher	point	1
vaincre	alone	1
comprendre	point	1
poiser	point	1
moquer	point	1
precher	point	1
employer	alone	1
revoir	point	1
merite	point	1
meurtrier	point	1
domter	point	1
eclaircir	point	1
esbranler	point	1
tracer	point	1
pleurer	alone	1
eclairer	alone	1
devider	alone	1
desaigrir	alone	1
desagreer	alone	1
taire	alone	1
effroyer	alone	1
soucier	alone	1

essuyer	alone	1
survenir	alone	1
tromper	alone	1
empecher	pas	1
creuser	alone	1
appeler	alone	1
appartenir	alone	1
decroire	alone	1
aviser	point	1
dependre	alone	1
sejourner	alone	1
saluer	pas	1
satisfaire	alone	1
foudroyer	alone	1
plaindre	alone	1
importer	alone	1
interpreter	alone	1
passionner	alone	1
exhaler	alone	1
evanoir	alone	1
recourir	alone	1
eprouver	alone	1
estimer	alone	1
pardonner	alone	1
obtenir	alone	1
rejouir	alone	1
nourrir	alone	1
inhumer	alone	1
raconter	alone	1
precipiter	alone	1
ombrager	alone	1
ourdire	alone	1
penetrer	alone	1
entendir	alone	1
former	alone	1
esmoir	point	1
hair	point	1
perir	alone	1
aider	alone	1
encombrer	point	1
comparer	point	1
esbahir	mie	1
nommer	pas	1
soutenir	pas	1
bruler	point	1

exceder	alone	1
preter	alone	1
brider	alone	1
ceder	alone	1
contempler	alone	1
convenir	alone	1
abuser	alone	1
chater	alone	1
tomber	alone	1
chercher	point	1
confesser	point	1
enquerir	point	1
accuser	mie	1
termer	alone	1
embler	alone	1
ennuyer	alone	1
feindre	alone	1
feter	alone	1
defaire	alone	1
concevoir	alone	1
accorder	alone	1
baptiser	alone	1
chanter	alone	1
comforter	alone	1
garder	alone	1
livrer	alone	1
rentrer	alone	1
repondre	alone	1
sacrifier	alone	1
surprendre	alone	1
remuer	alone	1
profiter	alone	1
monter	alone	1
ouvrir	alone	1
pericliter	alone	1
procurer	alone	1
couter	alone	1
crever	alone	1
eloigner	alone	1
entreprendre	alone	1
eprendre	alone	1
escorcher	alone	1
courir	alone	1
contestere	alone	1
arreter	alone	1

bailler	alone	1
choir	alone	1
conformer	alone	1
esmettre	alone	1
gener	alone	1
presser	alone	1
regarder	alone	1
remedier	alone	1
retenir	alone	1
poursuivre	alone	1
nuir	alone	1
gouverner	alone	1
marier	alone	1
montrer	alone	1
nicher	alone	1
affliger	alone	1
abandonner	alone	1
ordonner	alone	1
oublier	alone	1
pourchasser	alone	1
prevenir	alone	1
offenser	alone	1
fouler	alone	1
deceler	alone	1
dedaindre	alone	1
emouvoir	alone	1
eschiner	alone	1
prevoir	alone	1
raller	alone	1
soumettre	point	1
baillir	point	1
cuider	alone	1
murmurer	point	1
imaginer	point	1
reveler	alone	1
vanter	alone	1
quitter	pas	1
rucher	alone	1
Total number of verbs		269
Number of verbs counted		74

1600		
Total number of verbs	73	
Quartile	54.5	
etre	alone-pas-point	110
pouvoir	alone-point	94
savoir	alone-pas	43
avoir	alone-pas-point	40
vouloir	alone-pas-point	28
voir	alone-pas-point	18
oser	alone	13
falloir	alone-pas-point	12
faire	alone-pas-point	12
aimer	pas-point	11
connaitre	pas-point	10
devoir	alone-pas-point	9
meriter	alone-pas	8
dire	alone-pas-point	8
venir	alone-pas-point	7
craindre	alone-pas-point	6
croire	alone-pas-point	6
importer	alone	5
permettre	pas	5
tenir	alone-pas-point	5
mettre	alone-pas-point	5
souffrir	pas-point	4
attendre	pas	4
forcer	pas-point	4
considerer	pas-point	4
entendre	alone-pas-point	4
manquer	alone-pas-point	4
demander	pas-point	4
rendre	alone-pas	4
aller	alone-pas-point	4
toucher	pas-point	3
balancer	point	3
abuser	alone	3
approcher	alone-point	3
menager	alone	3
laisser	alone-point	3
ignorer	pas-point	3
satisfaire	alone-point	3
regner	pas	3
agir	point	3
penser	alone-pas	2
suffire	pas	2

comprendre	alone-point	2
flatter	alone-point	2
disposer	pas-point	2
prendre	alone-pas	2
trouver	pas-point	2
perdre	pas-point	2
voit	point	2
mourir	alone-pas	2
apprendre	alone-point	2
y avoir	point	2
punir	pas-point	2
importe	alone	2
accorder	alone-pas	2
habiter	alone	2
porter	alone	2
sentir	alone-point	2
pretendre	point	2
opposer	alone-pas	2
delivrer	alone-pas	2
ouvrir	alone-pas	2
valoir	alone-pas	2
appercevoir	point	2
user	alone-point	2
celer	point	2
approuver	alone-point	2
recevoir	alone	2
plaindre	alone-point	2
flechir	pas-point	2
retourner	alone	2
offenser	pas	2
entrer	alone-pas	2
promettre	pas	1
vivre	point	1
egaler	point	1
seoir	pas	1
accuser	point	1
cacher	point	1
augmenter	point	1
rester	pas	1
instruire	point	1
reconnaitre	point	1
parler	point	1
murmurer	point	1
sauver	point	1
rougir	pas	1

offrir	point	1
survivre	point	1
irriter	alone	1
obtenir	point	1
mentir	point	1
epuiser	point	1
meler	point	1
defendre	pas	1
epargner	pas	1
soumettre	point	1
troubler	point	1
retenir	point	1
preparer	point	1
plaire	pas	1
combattre	point	1
concevoir	point	1
cesser	point	1
abandonner	pas	1
affliger	pas	1
appartenir	pas	1
venger	alone	1
prevenir	alone	1
effacer	alone	1
enchanter	alone	1
deplaire	pas	1
disconvenir	pas	1
posseder	pas	1
temoins	pas	1
redouter	pas	1
oublier	pas	1
donner	pas	1
interrompre	pas	1
derober	alone	1
tirer	pas	1
relancer	alone	1
surprendre	alone	1
bouger	alone	1
soupir	point	1
poursuivre	point	1
ressortir	point	1
bailler	pas	1
manger	pas	1
chier	point	1
demeurer	point	1
voila	pas	1

pousser	pas	1
marquer	pas	1
oir	point	1
desirer	point	1
preferer	alone	1
regarder	alone	1
languir	alone	1
fourvoir	alone	1
diffamer	alone	1
envoyer	alone	1
retarder	alone	1
reveiller	alone	1
chercher	point	1
depourvoir	point	1
renaitre	pas	1
terminer	alone	1
souhaiter	alone	1
douter	point	1
marier	point	1
importuner	alone	1
resoudre	alone	1
detruire	alone	1
decevoir	alone	1
aigrir	alone	1
conter	alone	1
allonger	pas	1
ceder	pas	1
foudroyer	pas	1
hair	pas	1
eteindre	pas	1
ecouter	pas	1
discerner	pas	1
saisir	point	1
repondre	point	1
nourrir	alone	1
rejoindre	alone	1
autoriser	alone	1
tomber	point	1
moquer	point	1
retablir	alone	1
achever	pas	1
dedire	point	1
envier	point	1
courir	point	1
arreter	point	1

consulter	pas	1
nuir	pas	1
Total number of verbs		180
Number of verbs counted		55

1700		
Total number of verbs	77	
Quartile	57.5	
etre	alone-pas-point	151
avoir	alone-pas-point	62
savoir	alone-pas	58
pouvoir	alone-pas	38
vouloir	pas-point	27
voir	alone-pas-point	16
y avoir	alone-pas-point	15
faire	alone-pas-point	12
connaitre	alone-pas-point	11
falloir	pas-point	10
oser	alone	9
aimer	pas-point	9
croire	pas-point	8
venir	pas-point	8
dire	pas	7
parler	alone-pas-point	6
trouver	alone-pas-point	6
plaire	pas-point	6
laisser	pas-point	5
craindre	pas-point	5
devoir	pas-point	5
quitter	pas-point	5
tromper	alone-pas-point	5
paraitre	pas-point	5
regarder	pas-point	4
mettre	alone-pas-point	4
souffrir	pas	4
partir	point	4
penser	pas	4
manquer	pas	4
repondre	pas-point	4
etonner	pas-point	4
rendre	alone-point	4
convenir	pas	3
agir	pas-point	3
meriter	pas-point	3

210

balancer	pas-point	3
mourir	alone-pas-point	3
aller	pas-point	3
sortir	pas-point	3
deplaire	pas	3
epouser	alone-point	3
permettre	pas	3
attendre	pas-point	3
revenir	pas-point	3
sentir	pas-point	3
survivre	pas	2
rougir	pas	2
vivre	pas	2
reconnaitre	pas	2
naitre	point	2
dementir	point	2
prendre	pas	2
reussir	pas	2
recevoir	point	2
suffire	pas	2
empecher	pas	2
tenir	pas	2
plaindre	pas-point	2
entendre	pas-point	2
inquieter	pas	2
servir	point	2
porter	pas-point	2
toucher	point	2
perdre	pas-point	2
changer	pas-point	2
embarrasser	pas-point	2
accabler	alone-point	2
sauver	alone-point	2
venger	alone-pas	2
donner	pas-point	2
tarder	pas	2
demander	pas-point	2
pretendre	pas-point	2
prevoir	pas	2
oublier	pas	2
meler	point	2
condamner	pas	1
daigner	pas	1
surprendre	pas	1
vor	pas	1

211

cacher	pas	1
daindre	pas	1
ecrire	pas	1
bouger	alone	1
expliquer	pas	1
estimer	pas	1
employer	pas	1
appercevoir	pas	1
acheter	pas	1
songer	point	1
soucier	point	1
durer	pas	1
profiter	point	1
prier	point	1
desapprouver	point	1
tenter	point	1
gouter	pas	1
abandonner	pas	1
opposer	alone	1
nuir	alone	1
acquitter	pas	1
hair	pas	1
accueillir	point	1
charger	pas	1
abuser	point	1
comprendre	pas	1
transpirer	pas	1
confier	point	1
contraindre	point	1
atteler	pas	1
rajeunir	alone	1
renvoyer	point	1
chercher	alone	1
payer	alone	1
soutenir	pas	1
disposer	pas	1
pousser	pas	1
diminer	pas	1
ignorer	pas	1
attraper	pas	1
presenter	point	1
prononcer	pas	1
avertir	pas	1
rester	pas	1
depenser	pas	1

refuser	pas	1
garder	pas	1
dormir	point	1
saviur	alone	1
trahir	point	1
ecarter	pas	1
punir	point	1
presser	point	1
blamer	pas	1
retenir	alone	1
disputer	pas	1
fremir	alone	1
demeurer	pas	1
respecter	alone	1
poursuivre	point	1
epouvanter	pas	1
eclairer	point	1
egaler	point	1
eblouir	point	1
dissiper	point	1
decouvrir	point	1
forcer	point	1
bruler	point	1
fuir	pas	1
outrager	point	1
murmurer	point	1
attendrir	point	1
vendre	point	1
ennuyer	point	1
arriver	point	1
assurer	point	1
amuser	point	1
advenir	point	1
accepter	point	1
cacheter	point	1
entrer	pas	1
occuper	point	1
obstiner	point	1
lire	point	1
juger	point	1
envisager	pas	1
valoir	pas	1
posseder	pas	1
quereller	pas	1
nuire	pas	1

marcher	pas	1
fermer	pas	1
reduire	pas	1
souvenir	pas	1
remettre	pas	1
resister	pas	1
ressembler	pas	1
rejouir	pas	1
emparer	pas	1
Total number of verbs		181
Number of verbs counted		58

1800		
Total number of verbs	97	
Quartile	72.5	
etre	alone-pas-point	206
pouvoir	alone-pas	93
avoir	alone-pas	81
savoir	alone-pas	71
vouloir	pas-point	52
falloir	pas	24
aimer	pas	24
voir	alone-pas	23
faire	pas	23
dire	alone-pas	21
connaitre	pas	18
y avoir	pas	17
croire	pas-point	15
demander	pas	13
oser	alone-pas	12
trouver	pas-point	11
aller	pas	10
comprendre	pas	9
venir	pas	8
reconnaitre	pas	8
donner	pas	7
manger	pas	6
epouser	pas	6
douter	pas	6
cesser	alone-pas	5
tromper	pas	5
laisser	pas	5
attendre	pas	4
devoir	alone-pas	4

gener	pas	4
suffire	pas	4
compter	pas	4
entendre	alone-pas	4
tenir	pas	4
importer	alone	3
mourir	pas	3
ecouter	pas	3
parler	pas	3
regarder	pas	3
penser	pas	3
travailler	pas	3
pardonner	alone-pas	3
sortir	alone-pas	3
paraitre	pas	3
expliquer	pas	3
ajouter	pas-point	3
porter	pas-point	3
quitter	pas	3
perdre	pas	3
partir	pas	3
arriver	alone-pas	3
manquer	pas-point	3
rester	pas	3
sentir	pas	3
effrayer	pas-point	3
deviner	pas	3
souvenir	pas	3
appartenir	alone-pas	3
disparaitre	pas	2
courir	pas	2
mettre	pas	2
interroger	pas	2
admettre	pas	2
cacher	pas	2
detester	pas	2
inviter	pas	2
insister	pas	2
diner	pas	2
amener	pas	2
soucier	pas	2
tolerer	pas	2
accepter	pas	2
elever	pas	2
appeler	pas	2

payer	pas	2
tarder	point	2
accompagner	pas-point	2
plaire	alone	2
retenir	alone-pas	2
songer	pas	2
recevoir	pas-point	2
lire	pas-point	2
plaisanter	pas	2
plaindre	pas	2
permettre	pas	2
estimer	pas-point	2
oublier	pas	2
adresser	pas	2
rappeler	pas	2
abandonner	pas	2
rentrer	alone-pas	2
agir	pas	2
empecher	pas	2
souffrir	pas	2
entrer	pas	2
changer	pas	2
embarrasser	pas	2
reprendre	pas	1
retirer	pas	1
rougir	pas	1
soupconner	pas	1
durer	pas	1
aviser	pas	1
danser	pas	1
hesiter	pas	1
embrouiller	pas	1
chercher	pas	1
essayer	pas	1
facher	pas	1
articuler	pas	1
monter	pas	1
parle	pas	1
parvenir	pas	1
meriter	pas	1
mentir	pas	1
tomber	pas	1
grogner	pas	1
interdire	pas	1
pas	pas	1

216

precher	pas	1
rire	pas	1
secourir	pas	1
souscrire	pas	1
raconter	pas	1
pretendre	pas	1
prendre	pas	1
terminer	pas	1
troubler	pas	1
etouffer	pas	1
abdiquer	pas	1
taquiner	pas	1
infliger	alone	1
deplaire	alone	1
revenir	pas	1
sauver	pas	1
...	alone	1
attaquer	pas	1
atteindre	pas	1
devenir	pas	1
remontrer	pas	1
eprouver	pas	1
demeurer	pas	1
causer	pas	1
attifer	pas	1
battre	pas	1
mefier	pas	1
connatire	pas	1
consoler	pas	1
epouxer	pas	1
briser	pas	1
accorder	pas	1
prouver	alone	1
saisir	alone	1
excluer	pas	1
feindre	pas	1
offenser	pas	1
repeter	pas	1
marier	pas	1
livrer	pas	1
hasarder	pas	1
ignorer	pas	1
trahir	point	1
tenter	point	1
garder	pas	1

juger	pas	1
fremir	pas	1
dormir	pas	1
amuser	pas	1
assassiner	pas	1
renoncer	pas	1
repondre	pas	1
opposer	point	1
persister	point	1
naitre	point	1
entreprendre	point	1
ressembler	pas	1
eloigner	point	1
resister	pas	1
rever	pas	1
mouiller	pas	1
nommer	pas	1
montrer	pas	1
franchir	pas	1
finir	pas	1
forcer	pas	1
occuper	pas	1
offrir	pas	1
promettre	pas	1
reflechir	pas	1
profiter	pas	1
pousser	pas	1
passer	pas	1
pleurer	pas	1
exposer	pas	1
enterrer	pas	1
regretter	alone	1
appercevoir	pas	1
importe	alone	1
interceder	point	1
tendre	pas	1
violer	pas	1
apprecier	pas	1
consentir	pas	1
ecrire	pas	1
ennuyer	pas	1
donnner	pas	1
decider	pas	1
contenter	pas	1
convaincre	pas	1

remettre	pas	1
Total number of verbs		211
Number of verbs counted		73

1900		
Total number of verbs	53	
Quartile	39.5	
etre	alone-pas-point	112
pouvoir	alone-pas	42
avoir	pas	40
savoir	alone-pas	22
vouloir	pas	22
faire	pas	15
falloir	pas	7
devoir	pas	7
comprendre	pas	6
croire	pas	6
y avoir	pas	6
voir	pas	6
aimer	pas-point	5
perdre	pas	5
parler	pas-point	5
souffrir	pas	4
trouver	pas	4
oser	alone-pas	4
bouger	pas	4
repondre	pas	3
songer	pas	3
tuer	pas	3
sentir	pas	3
supporter	pas	3
laisser	pas	3
mourir	pas	3
demander	pas	3
dire	pas	3
venir	pas	3
plaire	pas	2
hesiter	pas	2
offenser	pas	2
aller	pas	2
regarder	pas	2
arriver	pas	2
plaisanter	pas	2
attendre	pas	2

naitre	pas	2
recevoir	pas	2
ouvrir	pas-point	2
agir	pas	2
donner	pas	2
douter	pas	2
connaitre	pas	2
convaincre	pas	2
importer	alone	2
revenir	pas	2
exagerer	pas	2
rentrer	pas	2
reflechir	pas	2
abandonner	pas	2
exister	pas	2
entendre	pas	2
forcer	pas	1
distinguer	pas	1
dissimuler	pas	1
devenir	pas	1
deteindre	pas	1
egarer	pas	1
habiller	pas	1
illuminer	pas	1
nommer	point	1
parvenir	point	1
imposer	pas	1
penser	pas	1
reprendre	pas	1
rougir	pas	1
recommencer	pas	1
reclamer	pas	1
quitter	pas	1
sembler	pas	1
tromper	pas	1
tendre	pas	1
tomber	pas	1
suivre	pas	1
souvenir	pas	1
prouver	pas	1
prevoir	pas	1
percevoir	pas	1
tenir	pas	1
obeir	pas	1
menacer	pas	1

hair	point	1
pleurer	pas	1
craindre	pas	1
prendre	pas	1
porter	pas	1
eventrer	point	1
lire	pas	1
sortir	pas	1
ignorer	pas	1
fatiguer	pas	1
etre	pas	1
ecrire	pas	1
insister	pas	1
interesser	pas	1
offrir	pas	1
occuper	pas	1
montrer	pas	1
mefier	pas	1
durer	pas	1
divorcer	pas	1
avoi	pas	1
attendrir	pas	1
arreter	pas	1
allumer	pas	1
chasser	pas	1
conduire	pas	1
depasser	pas	1
dejeuner	pas	1
decerner	pas	1
paraitre	pas	1
pardonner	pas	1
accepter	pas	1
consentir	point	1
vivre	pas	1
affecter	pas	1
agacer	pas	1
compter	pas	1
choisir	pas	1
apprecier	pas	1
valoir	pas	1
soigner	pas	1
prononcer	pas	1
plaindre	pas	1
permettre	pas	1
partager	pas	1

rappeler	pas	1
recompenser	pas	1
sauter	pas	1
reveiller	pas	1
reconnaitre	pas	1
convenir	pas	1
Total number of verbs		134
Number of verbs counted		40

Appendix E. Raw type frequency data

The following tables give the type frequency counts for each century, for each combination of the various sub-hypotheses.

Century	alone types	pas types	point types	mie types	Total
12th	22	5	0	0	24
13th	85	15	10	27	100
14th	59	21	17	16	77
15th	75	29	35	8	104
16th	187	47	58	11	224
17th	66	70	77	0	154
18th	24	105	77	0	154
19th	23	162	18	0	180
20th	6	106	9	0	112
Total	2258	2408	731	115	

Table 23 Raw numbers of types per century for conjugated verbs, including high-frequency verbs and hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	8	3			8
13th	38	12	7	18	38
14th	35	17	12	11	39
15th	37	22	17	6	39
16th	75	39	34	7	76
17th	41	42	42		66
18th	14	60	47		69
19th	14	79	12		82
20th	6	41	4		42
Total	2258	2408	731	115	

Table 24. Raw numbers of types per century for conjugated verbs, including high-frequency verbs and excluding hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	20	3	0	0	22
13th	76	9	6	19	91
14th	49	13	10	10	67
15th	65	21	30	5	94
16th	168	28	43	6	205
17th	52	56	64	0	138
18th	15	89	63	0	137
19th	17	142	14	0	160
20th	2	96	8	0	102
Total	2258	2408	731	115	

Table 25. Raw numbers of types per century for conjugated verbs, excluding high-frequency verbs and including hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	6	1			6
13th	29	6	3	10	29
14th	25	9	5	5	29
15th	27	14	12	3	29
16th	56	20	19	2	57
17th	27	28	29		50
18th	5	44	33		52
19th	8	59	8		62
20th	2	31	3		32
Total	2258	2408	731	115	

Table 26. Raw numbers of types per century for conjugated verbs, excluding high-frequency verbs and hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	22	5	0	0	24
13th	96	20	11	31	114
14th	64	29	20	17	92
15th	89	41	42	8	130
16th	224	55	69	11	268
17th	77	81	89	0	179
18th	29	119	92	0	180
19th	25	192	21	0	210
20th	5	127	9	0	133
Total	2258	2408	731	115	

Table 27. Raw numbers of types per century for main verbs, including high-frequency verbs and hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	8	3			8
13th	42	15	8	22	43
14th	37	18	13	11	41
15th	42	25	22	6	45
16th	93	49	39	9	99
17th	48	46	47		73
18th	19	65	53		77
19th	18	94	13		97
20th	5	52	4		53
Total	2258	2408	731	115	

Table 28. Raw numbers of types per century for main verbs, including high-frequency verbs and excluding hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	20	3	0	0	22
13th	85	14	7	22	103
14th	54	21	12	11	82
15th	78	32	36	5	119
16th	199	32	48	5	243
17th	61	66	75	0	161
18th	18	101	77	0	161
19th	18	168	17	0	186
20th	2	114	7	0	120
Total	2258	2408	731	115	

Table 29. Raw numbers of types per century for main verbs, excluding high-frequency verbs and including hapaxes.

Century	alone types	pas types	point types	mie types	Total
12th	6	1			6
13th	31	9	4	13	32
14th	27	10	5	5	31
15th	31	16	16	3	34
16th	68	26	18	3	74
17th	32	31	33		55
18th	8	47	38		58
19th	11	70	9		73
20th	2	39	2		40
Total	2258	2408	731	115	

Table 30. Raw numbers of types per century for conjugated verbs, excluding high-frequency verbs and hapaxes.

Appendix F. Token frequency

The following table lists raw numbers of tokens found in the corpus, organized by century, by negator and by whether it appears in a context described by one of the grammarians.

Century	alone no rule	alone rule	pas no rule	pas rule	point no rule	point rule	mie no rule	mie rule	*Total*
12th	24	24	10	3		0		0	*61*
13th	104	149	37	2	13	1	58	5	*369*
14th	117	151	69	6	27	2	24	4	*400*
15th	100	208	108	4	75	0	10	0	*505*
16th	452	373	183	12	157	2	11	3	*1193*
17th	60	225	201	9	193	7		0	*695*
18th	9	119	362	34	202	5		0	*731*
19th	3	110	805	116	35	2		0	*1071*
20th	1	24	388	58	10	0		0	*481*
Total		2253		2407		731		115	*5506*

The following table gives unsmoothed values for the prevalence of tokens used with the verbs that appeared in every century from the fifteenth through the twentieth.

Verb	1400	1500	1600	1700	1800	1900	Last Century
aimer	0.000	0.286	0.000	0.000	0.000	0.000	0
avoir	0.645	0.583	0.100	0.097	0.037	0.000	1600
devoir	0.375	0.750	0.444	0.000	0.250	0.000	1600
dire	0.222	0.429	0.125	0.000	0.048	0.000	0
etre	0.313	0.517	0.218	0.046	0.005	0.009	1600
faire	0.526	0.775	0.333	0.167	0.000	0.000	1600
falloir	0.267	0.522	0.250	0.000	0.000	0.000	1600
oser	1.000	0.909	1.000	1.000	0.417	0.500	1800
pouvoir	1.000	0.955	0.989	0.921	0.527	0.310	1900
savoir	0.926	0.951	0.884	0.759	0.380	0.273	1800
venir	0.833	0.538	0.286	0.000	0.000	0.000	1600
voir	0.583	0.826	0.278	0.125	0.130	0.000	1600
vouloir	0.727	0.719	0.286	0.000	0.000	0.000	1600
y avoir	0.643	0.455	0.000	0.067	0.000	0.000	1500

References

A Concordance to the Complete Works of Geoffrey Chaucer and the Romaunt of the Rose. Washington: Carnegie Institution.

Alighieri, Dante. 1304. *De vulgari eloquentia.*

Allen, Cynthia. 2000. Obsolescence and sudden death in syntax: the decline of verb-final order in early Middle English. *Generative theory and corpus studies: a dialogue from 10 ICEHL*, ed. by Ricardo Bermudez-Otero, David Denison, Richard M. Hogg and C. B. McCully, 3-25. Berlin: Mouton de Gruyter.

Altmann, Gabriel, Haro von Buttlar, Walter Rott, and Udo Strauss. 1982. A Law of Change in Language. *Historical Linguistics*, ed. by Barron Brainerd, 104-115. Bochum: Brockmeyer.

Ashby, William J. 1981. The loss of the negative particle ne in French: A syntactic change in progress. *Language* 57.674-87.

Ashby, William J. 2001. Un nouveau regard sur la chute du ne en français parlé tourangeau: S'agit-il d'un changement en cours? *JFLS* 11.1-22.

Baayen, Harald, and Rochelle Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29.801-43.

Baayen, Harald, and Robert Schreuder. 2000. Towards a psycholinguistic computational model for morphological parsing. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 358.1281-93.

Bailey, Charles-James N. 1973. *Variation and Linguistic Theory*. Washington: Center for Applied Linguistics.

Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5.359-90.

Bellay, Joachim du. 1549. Défense et illustration de la langue française. Paris: Arnoul l'Angelier.

Berkson, Joseph. 1951. Why I prefer logits to probits. *Biometrics* 7.327-39.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Blanche-Benveniste, Claire, and Colette Jeanjean. 1987. *Le français parlé: Transcription et édition*. Paris: Didier.

Bleiberg, Edward (ed.). 2005. Northern Renaissance. *Arts and Humanities Through the Eras*, 141-155. Detroit: Gale.

Brinton, Laurel. 2005. Processes underlying the development of pragmatic markers: The case of (I) say. Pragmatics & Beyond New Series; Opening Windows on Texts and Discourses of the Past, ed. by Janne Skaffari, Matti Peikola, Ruth Carroll, Risto Hiltunen and Brita Wårvik, 279-299. Amsterdam: John Benjamins.

Brinton, Laurel. 2008. Historical Pragmatics and Corpus Linguistics: Problems and Strategies. Paper presented at the conference of the American Association for Corpus Linguistics, Provo.

Briscoe, Ted. 2000. Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. *Language* 76.245-96.

Brunot, Ferdinand. 1891. *La doctrine de Malherbe d'après son commentaire sur Desportes*. Paris: Masson.

228

Brunot, Ferdinand. 1901. *Histoire de la langue française des origines à 1900*. Paris: Armand Colin.

Brütsch, Edgar, Markus Nussbaumer, and Horst Sitta. 1990. *Negation*. Heidelberg: Julius Groos.

Bybee, Joan. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10.425-55.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14.261-90.

Bybee, Joan. 2003. Cognitive processes in grammaticalization. *The New Psychology of Language*, ed. by Michael Tomasello, 145-167. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bybee, Joan. 2003. Mechanisms of Change in Grammaticization: The Role of Frequency. *Handbook of Historical Linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 602-623. Oxford: Basil Blackwell.

Bybee, Joan, and David Eddington. 2006. A usage-based approach to Spanish verbs of 'becoming'. *Language* 82.323-355.

Bybee, Joan, and Dan I. Slobin. 1982. Why small children cannot change language on their own: Suggestions from the English past tense. *Papers from the Fifth International Conference on Historical Linguistics*, ed. by Anders Ahlqvist. Amsterdam: John Benjamins.

229

Bybee, Joan, and Sandra A. Thompson. 1997. Three Frequency Effects in Syntax. *Berkeley Linguistics Society (BLS)*.

Caillet, Jules. 1857. *De l'administration en France sous le ministère du cardinal de Richelieu*. Paris: Firmin Didot Frères.

Campbell, Lyle. 2001. What's wrong with grammaticalization? *Language Sciences* 23.113-61.

Catalani, Luigi. 2001. *Die Negation im Mittelfranzösischen*. Frankfurt: Peter Lang.

Chambers, J. K. 1992. Dialect acquisition. *Language* 68.673-705.

Chambers, J. K. 2003. *Sociolinguistic theory: linguistic variation and its social significance*. Oxford: Basil Blackwell.

Chambers, J. K., and Peter Trudgill. 1988. *Dialectology*. Cambridge: Cambridge University Press.

Clahsen, Harald, and Monika Rothweiler. 1992. Inflectional rules in children's grammars: Evidence from the development of participles in German. *Yearbook of Morphology*. 1-34.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Corneille, Pierre, and Charles-Joseph Marty-Laveaux. 1660. *Au Lecteur de Mélite*. Paris: Hachette.

Cramer, J. S. 2003. The origins and development of the logit model. *Logit Models from Economics and Other Fields*. Cambridge: Cambridge University Press.

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. London: Longman.

Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.

Croft, William. 2007. The origins of grammaticalization in the verbalization of experience. *Cognitive Linguistics* 18. 339-82.

Croft, William, and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Dahl, Östen. 1979. Typology of sentence negation. *Linguistics* 17. 79-106.

Dahl, Östen. 2001. Inflationary effects in language and elsewhere. *Frequency and the Emergence of Linguistic Structure*, ed. by Joan Bybee and Paul J. Hopper. Amsterdam: John Benjamins.

Dawkins, Richard L. 1976. *The Selfish gene*. Oxford: Oxford University Press.

Degaine, André. 1992. *Histoire du théâtre dessinée: De la préhistoire à nos jours, tous les temps et tous les pays*. Saint Genouph: A.-G. Nizet.

Denison, David. 2003. Log(ist)ic and simplistic S-curves. *Motives for Language Change*, ed. by Raymond Hickey, 54-70. Cambridge: Cambridge University Press.

Detges, Ulrich, and Richard Waltereit. 2002. Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21.151-95.

Diez, Friedrich. 1882. *Grammatik der romanischen Sprachen*. Bonn: Eduard Weber's.

Downey, Charles T. 2002. Ad imaginem suam: Regional chant variants and the origins of the jeu d'Adam. *Comparative Drama* 36.359-90.

Dubois, Jean. 1989. Lexis*: Dictionnaire de la langue française*. Paris: Larousse.

Eckardt, Regine. 2007. *Meaning Change in Grammaticalization: An inquiry into semantic reanalysis*. Oxford: Oxford University Press.

Emphasis, n. 1989. Oxford English Dictionary, 2nd edn. Oxford: Oxford University Press.

Englebert, Annick. 1984. Esquisse d'une histoire de la négation en français. *Travaux de linguistique* 11.7-25.

Englebert, Annick. 1985. L'opposition ne / ne … pas en ancien français. *Revue de Linguistique Romane* 195-196.365-78.

Ernst, Gerhard. 1985. *Gesprochenes Französisch zu Beginn des 17. Jahrhunderts: direkte Rede in Jean Héroards Histoire particulière de Louis XIII (1605-1610)*. Tübingen: Niemeyer.

Euler, Leonhard. 1736. *Mechanica sive motus scientia analytice exposita*. St. Petersburg: Academia Scientiarum Imperialis Petropolitanae.

Ewert, Alfred. 1943. *The French Language*. London: Faber and Faber.

Fauconnier, Gilles. 1975. Pragmatic scales and logical structure. *Linguistic Inquiry* 6.353-75.

Ferguson, Charles A. 1959. Diglossia. *Word* 15.325-40.

Foulet, Lucien. 1923. *Petite syntaxe de l'ancien français*. Paris: Honoré Champion.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Frisch, Stefan. 1994. Reanalysis precedes syntactic change: Evidence from middle English. *Studies in the Linguistic Sciences* 24.187-201.

Gause, Georgyi F. 1934. *The struggle for existence*. Baltimore: Williams and Wilkins.

Gernsbacher, Morton Ann. 1984. Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, completeness and polysemy. *Journal of Experimental Psychology: General* 113.256-81.

Geurts, Bart. 1998. The mechanisms of denial. *Language* 74.274-307.

Geurts, Bart. 2000. Explaining grammaticalization (the standard way). *Linguistics* 38.781-788.

Givón, Talmy. 1979. From discourse to syntax: grammar as a processing strategy. *Syntax and Semantics; Discourse and Syntax*, ed. by Talmy Givón, 81-112. New York: Academic Press.

Godard, Danièle, and Jean-Marie Marandin. Reinforcing Negation: the case of Italian. Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar (HPSG) 06.174-194.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gregory, Stewart. 1997. Negative particles in French prose in the twelfth century. *De mot en mot, Aspects of medieval Linguistics*, ed. by Stewart Gregory and D. A. Trotter, 37-51. Cardiff: University of Wales Press.

Greimas, Algirdas Julien. *Dictionnaire de l'ancien français*. Paris: Larousse.

Grieve-Smith, Angus B. 2006. The Envelope of Variation in Multidimensional Register and Genre Analyses. *Corpus Linguistics Beyond the Word*, ed. by Eileen Fitzpatrick, 21-42. Amsterdam: Rodopi.

Guerdan, René. 1984. *Corneille: ou la vie méconnue du Shakespeare français*. Lausanne: Pierre-Marcel Favre.

Guillaume, Paul. 1973. Development of Formal Elements in the Child's Speech. *Studies of Child Language Development*, ed. by Charles A. Ferguson and Dan I. Slobin, 240-251. New York: Holt, Rinehart and Winston.

Haiman, John. 1997. Repetition and identity. *Lingua* 100.57-70.

Harris, Alice C., and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in language* 22.315-352.

Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37.1043-1068.

Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. *Up and Down the Cline: The Nature of Grammaticalization*, ed. by Olga Fischer, Muriel Norde and Harry Peridon, 17-44. Amsterdam: John Benjamins.

Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39.1041-70.

Hay, Jennifer, and Harald Baayen. 2001. Parsing and productivity. *Yearbook of Morphology*.

Heine, Bernd. 1994. Grammaticalization as an explanatory parameter. *Perspectives on Grammaticalization*, ed. by William Pagliuca. Amsterdam: John Benjamins.

Heine, Bernd. 2003. Grammaticalization. *Handbook of Historical Linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 575-601. Oxford: Basil Blackwell.

Heine, Bernd, Ulrike Claudi, and Frederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Hock, Hans Heinrich. 2003. *Analogical Change. Handbook of Historical Linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 441-460. Oxford: Basil Blackwell.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. *Current progress in historical linguistics*, ed. by William Christie, 95-105. Amsterdam: North Holland.

Hopkins, Kenneth D., B. R. Hopkins, and Gene V. Glass. 1996. *Basic Statistics for the Behavioral Sciences*. Needham Heights: Allyn & Bacon.

Hopper, Paul J. 1991. On Some Principles of Grammaticization. *Approaches to Grammaticalization*, ed. by Elizabeth Closs Traugott and Bernd Heine, 17-35. Amsterdam: John Benjamins.

Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Horn, Laurence R. 1989. *A natural history of negation*, 452-462. Chicago: University of Chicago Press.

Horne, Alistair. 2004. *La Belle France*. Vintage: New York.

Hull, David L. 1988. *Science as a process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.

Israel, Michael. 1998. Minimizers, maximizers and the rhetoric of scalar reasoning. Journal of *Semantics* 18.297-331.

Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen:

Munksgaard.

Joseph, John Earl. 1987. *Eloquence and Power: The Rise of Language Standards and Standard Languages*. Oxford: Basil Blackwell.

Kallel, Amel. 2007. The loss of negative concord in standard English: Internal factors. *Language Variation and Change* 19.27-49.

川口裕司研究室 (Kawaguchi, Yuji). 2005.

通時的観点からみたフランス語の否定辞. *言語情報学研究報告7*

*コーパス言語学における語彙と文法*7.319-38.

Keller, Rudi. 1990. *On language change: The invisible hand in language*. New York:

Routledge.

Kiparsky, Paul, and Cleo Condoravdi. 2006. Tracking Jespersen's cycle. Second

International Conference of Modern Greek Dialects and Linguistic Theory.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25.1-49.

Kroch, Anthony. 1982. Grammatical Theory and the Quantitative Study of Syntactic

Change. New Ways of Analyzing Variation (NWAVE).

Kroch, Anthony. 1978. Toward a theory of social dialect variation. *Language in Society* 7.17-36.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1.199-244.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1.97-120.

Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57.267-308.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Basil Blackwell.

Ladusaw, William. 1993. Negation, indefinites, and the Jespersen Cycle. Berkeley Linguistics Society (BLS).

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Langacker, Ronald W. 1977. Syntactic reanalysis. *Mechanisms of syntactic change*, ed. by Charles N. Li, 57-139. Austin: University of Texas.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical prerequisites*. Palo Alto: Stanford University Press.

Langacker, Ronald W. 1988. A usage-based model. *Topics in Cognitive Linguistics*, ed. by Brygida Rudzka-Ostyn, 127-161. Amsterdam: John Benjamins.

Lee, David. 1999. *Modelling Variation in Spoken And Written English: The multi-dimensional approach revisited*. Lancaster: Lancaster University Ph.D. dissertation.

Lehmann, Christian. 1982. Thoughts on grammaticalization. *Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt*.

Lehmann, Christian. 1999. New reflections on grammaticalization and lexicalization. Paper presented at New Reflections on Grammaticalization.

Leonard, Sterling. 1929. *The Doctrine of Correctness in English Usage, 1700-1800*. Madison: University of Wisconsin.

Lewis, David. 1969. *Convention*. Cambridge, MA: Harvard University Press.

Lieberman, Erez, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A. Nowak. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449.713-6.

Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge: MIT Press.

Lindsay, Wallace Martin. 1900. The two recensions of Plautus, A and PA. *American Journal of Philology* 21.23-37.

Lodge, Anthony. 1991. Molière's peasants and the norms of spoken French. *Neuphilologische Mitteilungen* 92.485-99.

Lodge, Anthony. 1993. French*: From Dialect to Standard*. New York: Routledge.

Lotka, Alfred J. 1925. *Elements of Physical Biology*. Baltimore: Williams and Wilkins.

Lucas, Christopher. 2007. Jespersen's cycle in Arabic and Berber. *Transactions of the Philological Society* 105.398-431.

Lüdtke, Helmut. 1986. Esquisse d'une théorie du changement langagier. *La Linguistique* 22.3-46.

MacWhinney, Brian. 1978. *Acquisition of Morphophonology*. Ann Arbor: Society for Research in Child Development.

Malherbe, François de. 1606. Commentaire sur des Portes. Œuvres complètes, 249-474. Paris: Hachette.

Malherbe, François de, Urbain Chevreau, and Gilles Ménage. 1666. Les œuvres de François de Malherbe; avec les observations de mr Ménage, et les remarques de mr Chevreau sur les poësies. Paris: Barbout.

Malthus, Thomas. 1789. *An Essay on the Principle of Population*. London: J. Johnson.

Manski, Charles F. 2001. Daniel McFadden and the econometric analysis of discrete choice. *Scandinavian Journal of Economics* 103.217-29.

Marchello-Nizia, Christiane. 1979. *Histoire de la langue française aux XIVe et XVe siècles*. Paris: Dunod.

Martineau, France, and Raymond Mougeon. 2003. A sociolinguistic study of the origins of ne deletion in European and Quebec French. *Language* 79.118-52.

Maupas, Charles, père. 1606. *Grammaire et syntaxe françoise*. Rouen: Cailioue.

Mazzon, Gabriella. 2004. *A history of English negation.* London: Longman.

McDaniels, Todd. 2003. What's wrong with reanalysis? *Toronto Working Papers in Linguistics (TWPL)* 21.81-8.

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* 12.

Meillet, Antoine. 1912. *Linguistique historique et linguistique générale*. Paris: Honoré Champion.

Moignet, Gérard. 1965. L'opposition non / ne en ancien français. *Travaux de linguistique et de littérature* 3.41-65.

Mosegaard Hansen, Maj-Britt. Forthcoming. The Grammaticalization of Negative Reinforcers in Old and Middle French: A Discourse-Functional Approach.

Current Trends in Diachronic Semantics and Pragmatics, ed. by Maj-Britt Mosegaard Hansen and Jacqueline Visconti. Bingley: Emerald.

Mosegaard Hansen, Maj-Britt, and Jacqueline Visconti. 2009. On the diachrony of 'reinforced' negation in French and Italian. *Studies in Pragmatics; Grammaticalization and pragmatics: facts, approaches, theoretical issues*, ed. by Corinne Rossari, Corina Cojocariu, Claudia Ricci and Adriana Spiridon. Bingley: Emerald.

Nadeau, Jean-Benoît, and Julie Barlow. 2006. *The Story of French*. New York: St. Martin's Press.

Napier, John. 1614. *Mirifici logarithmorum canonis descriptio*. Edinburgh: Andrew Hart.

Niyogi, Partha, and Robert C. Berwick. 1997. A dynamical systems model for language change. Paper presented at the Annual Conference of the Cognitive Science Society.

Ochs, Elinor. 1979. Planned and Unplanned Discourse. *Discourse and Syntax*, ed. by Talmy Givón, 51-80. New York: Academic Press.

Offord, Malcolm. 1976. Negation in Berinus: A contribution to the study of negation in fourteenth-century French. *Zeitschrift für romanische Philologie* 92.313-85.

Ogura, Mieko, and William S.-Y. Wang. 1996. Snowball Effect in Lexical Diffusion: The development of -s in the third person singular present indicative in English. *English Historical Linguistics 1994*, ed. by Derek Britton, 119-141. Amsterdam: John Benjamins.

Oliveira e Silva, Giselle. 1982. Estudo da Regularidade na Variação dos Possessivos no Português do Rio de Janeiro. Rio de Janeiro: Universidade Federal do Rio de Janeiro Ph.D. dissertation.

Osgood, Charles Egerton, and Thomas Albert Sebeok. 1954. Psycholinguistics: A survey of theory and research problems. *Journal of abnormal psychology* 49.1-203.

Ostwald, Wilhelm. 1890. *Outlines of General Chemistry*. London: Macmillan and Company.

Palsgrave, John. 1536. *L'Eclaircissement de la langue française*. Paris: Génin.

Parry, Mair. 1997. On negation in the Ligurian hinterland. *Quaderni di lavoro ASIS* 2.

Pawlowski, Adam, and Maciej Eder. 2001. Quantity or stress?: Sequential analysis of Latin prosody. *Journal of Quantitative Linguistics* 1.81-97.

Pearl, Raymond, and Lowell J. Reed. 1920. On the rate of growth of the population of the united states since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences (PNAS)* 6.275-88.

Pinasco, J. P., and L. Romanelli. 2006. Coexistence of languages is possible. *Physica A: Statistical Mechanics and its Applications* 361.355-60.

Пнотровская, А. А., and Р. Г. Пиотровский (Piotrovskaja, A. A., and R. G. Piotrovskij). 1974. Математические модели диахронии и текстообразования. *Статистика речи и автоматический анализ текста*. 361-400.

Poplack, Shana. 2001. Variability, frequency, and productivity in the irrealis domain of French. *Frequency and the Emergence of Linguistic Structure*, ed. by Joan Bybee, 405-428. Amsterdam: John Benjamins.

Posner, Rebecca J. 1985. Post-verbal negation in non-standard French: A historical and comparative view. *Romance Philology* 39.

Posner, Rebecca J. 1997. *Linguistic change in French*. Oxford: Clarendon Press.

Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper and Row.

Price, Glanville. 1997. Negative Particles in French. *De mot en mot, Aspects of medieval Linguistics*, ed. by Stewart Gregory and D. A. Trotter, 173-190. Cardiff: University of Wales Press.

Prince, Ellen F. 1988. Discourse Analysis: a Part of the Study of Linguistic Competence. *Linguistic Theory: Extensions and Implications*, ed. by Frederick J. Newmeyer, 164-182. Cambridge: Cambridge University Press.

Rickard, Peter. 1974. *A History of the French Language*. New York: Routledge.

Rigal, Eugène Pierre Marie. 1901. *Théâtre français avant la période classique: fin du XVIe et commencement du XVIIe siècle*. Paris: Hachette.

Roosbroeck, Gustave L. van.. 1920. Corneille's early friends and surroundings. Modern Philology 18.361-81.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1.75-116.

Rostila, Jouni. 2006. Storage as a way to grammaticalization. *Constructions* 1.

Rydén, Mats. 1991. The be/have variation with intransitives in its crucial phases. *Historical English Syntax*, ed. by Dieter Kastovsky. New York: Walter de Gruyter.

Schuchardt, Hugo. 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Verlag von Robert Oppenheim.

Schwegler, Armin. 1983. Predicate negation and word-order change: A problem of
multiple causation. *Lingua* 61.297-334.

Schwegler, Armin. 1988. Word-order changes in predicate negation strategies in romance
languages. *Diachronica* 5.21-58.

Schwegler, Armin. 1990. *Analyticity and syntheticity: A diachronic perspective with
special reference to Romance languages*. Berlin: Mouton de Gruyter.

Schweighäuser, Alfred. 1852. De la négation dans les langues romanes: Du midi et du
nord de la France. *Bibliothèque de l'Ecole des Chartes*. 203-47.

Schwenter, Scott A. 2006. Fine-Tuning Jespersen's Cycle. *Drawing the Boundaries of
Meaning*, ed. by Betty J. Birner and Gregory Ward. 327-344. Amsterdam: John
Benjamins.

Seifert, Stephan, and Werner Welte. 1987. *Basic Bibliography on Negation in Natural
Language*. Tübingen: Gunter Narr.

Shen, Zhongwei. 1990. Lexical diffusion: A population perspective and a numerical
model. *Journal of Chinese Linguistics* 18.159-200.

Smith, Aaron K. The role of frequency in the specialization of the English anterior.
Frequency and the Emergence of Linguistic Structure, ed. by Joan Bybee, 361-
382. Amsterdam: John Benjamins.

Smith, John Charles. 2002. Middle French: When? what? why? *Language Sciences*
24.423-45.

Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford: Basil Blackwell.

Stein, Achim. 2009. Histoire de la langue française: Du latin aux premiers textes. http://www.uni-stuttgart.de/lingrom/stein/kurs/vlhist/folien01/siframes.html (May 3, 2009).

Symes, Carol. 2007. Towards a New History of Medieval Theatre: Assessing the Written and Unwritten Evidence for Indigenous Performance Practices. Paper presented at the Twelfth Colloquium of the Société Internationale pour l'Etude du Théâtre Médiéval (SITM).

Tannen, Deborah. 1990. *You just don't understand: Women and men in conversation*. New York: William Morrow.

Tatlock, John S. P., and Arthur G. Kennedy. 1927. *A concordance to the complete work of Geoffrey Chaucer and to the Romaunt of the rose*. Washington: Carnegie Institution.

Tottie, Gunnel. 1991. Lexical diffusion in syntactic change: frequency as a determinant of linguistic conservatism in the development of negation in English. Historical English Syntax, ed. by Dieter Kastovsky, 439-468. New York: Walter de Gruyter.

Traugott, Elizabeth Closs. Constructions in grammaticalization. *Handbook of Historical Linguistics,* ed. by Brian D. Joseph and Richard D. Janda, 624-647. Oxford: Basil Blackwell.

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65. 31-55.

Vaananen, Veikko. 1967. *Introduction au latin vulgaire*. Paris: Klincksieck.

Valli, André. 1983. Exemple d'approche du problème des variantes syntaxiques en linguistique diachronique. *Recherches sur le français parlé* 5.125-46.

244

Valli, André. 1984. Changements de norme, décalages grammaticaux et représentations du français parlé: L'exemple du *Télémaque* travesti de Marivaux. *Recherches sur le français parlé* 6.7-21.

Vaugelas, Claude Favre de. 1647. *Remarques sur la langue françoise*. Paris: Augustin Courbe.

Verhulst, Pierre-François. 1838. Notice sur la loi que la population suit dans son accroissement. *Correspondence Mathématiques et Physiques* 10.113.

Volterra, Vito. 1926. Fluctuations in the abundance of a species considered mathematically. *Nature* 118.558-60.

Walter, Henriette. 1994. *French Inside Out: The Worldwide Development of the French Language in the Past, the Present and the Future*. New York: Routledge.

Wang, William S.-Y. 1969. Competing changes as a cause of residue. *Language* 45. 9-25.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. *Empirical foundations for a theory of language change*. Austin: University of Texas.

Wenk, Brian J., and François Wioland. 1982. Is French really syllable-timed? *Journal of Phonetics* 10.193-216.

Wicks, Charles Beaumont. 1950. *The Parisian Stage, vol. 1*. Birmingham: University of Alabama Press.

Wicks, Charles Beaumont. 1953. *The Parisian Stage, vol. 2*. Birmingham: University of Alabama Press.

Wicks, Charles Beaumont. 1961. *The Parisian Stage, vol. 3*. Birmingham: University of Alabama Press.

Wicks, Charles Beaumont. 1967. *The Parisian Stage, vol. 4*. Birmingham: University of Alabama Press.

Wicks, Charles Beaumont. 1979. *The Parisian Stage, vol. 5.* Birmingham: University of Alabama Press.

Willis, David. 2004. A minimalist approach to Jespersen's Cycle in Welsh. Proceedings of the Eighth Diachronic Generative Syntax (DiGS) Conference.

Wolfram, Walt, and Natalie Schilling-Estes. 2003. *Handbook of Historical Linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 713-735. Oxford: Basil Blackwell.

Wouden, Ton van der. 1995. On the development of marked negation systems. Paper presented at the Twelfth International Conference on Historical Linguistics.

Yvon, Henri. 1959. Les expressions négatives dans la Queste del saint graal. *Romania* 80.63-78.

Yvon, Henri. 1960. Les expressions négatives dans la Conqueste de Constantinople de Villehardouin. *Romania* 81.286-307.

Yvon, Henri. 1960. Les expressions négatives dans la vie de saint louis de Joinville. Romania 81.99-111.

Yvon, Henri. 1962. Les expressions négatives dans Mon Faust de P. Valéry: Comparaison avec les textes du moyen âge. *Français moderne* 30.13-34.

Zipf, George K. 1935. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Boston: Houghton Mifflin.

Zupko, Ronald E. 1985. *Dictionary of Weights and Measures for the British Isles: The Middle Ages to the Twentieth Century*. Philadelphia: American Philosophical Society.

Zwarts, Franciscus. 1986. Categoriale grammatica en algebraïsche semantiek: een onderzoek naar negatie en polariteit in het Nederlands. Groningen: University of Groningen Ph.D. dissertation.