# Variation as a Testing Ground for Grammatical Theory: Variable Negative Concord in Montréal French

**Abstract**

This paper addresses the contribution that corpus-based studies of syntactic variation can make to the construction, elaboration and testing of formal syntactic theories, with a particular focus on the testing dimension. In particular, I present a new empirical study of obligatory and optional asymmetric negative concord phenomena, and I show how an influential analysis for obligatory concord patterns (de Swart, 2010) can be tested using variation data through looking at the predictions that its natural probabilistic extension makes for the forms, interpretations and frequency distributions of expressions in languages in which asymmetric concord is optional. In obligatory negative concord languages like Spanish, negative indefinites, such as *nadie* 'no one', appear bare in preverbal position (i.e. in an expression like ***Nadie*** *ha venido* '**No one** came'), but they co-occur with the negative marker *no* in postverbal negative concord structures such as ***No*** *he visto a* ***nadie*** 'I did not see anyone.' (lit. 'I did **not** see **no one**.'). Furthermore, in this language, co-occurrence between a negative marker and an n-word is either prohibited (**Nadie **no** ha venido*), or it is obligatory (**He visto a **nadie***). Québec French shows a variable version of the Spanish pattern in which the negation marker optionally co-occurs with postverbal negative indefinites (*J'ai **(pas)** vu **personne*** 'I saw no one') but is prohibited with preverbal negative indefinites *****Personne*** *est* ***pas*** *venu* (Ok: ***Personne*** *est venu.* 'No one came'). I show how the predictions for Montréal French of de Swart's analysis of Spanish can be tested (and, in this case, mostly verified) using a quantitative study of the distribution of bare and concord structures in the *Montréal 84* corpus of spoken Montréal French (Thibault and Vincent, 1990) through looking at its natural extension within Boersma (1998)'s stochastic generalization of the Optimality Theory framework, which is the framework in which de Swart's proposal is set.

Keywords: syntactic variation; negative concord; probabilistic grammar; Montreal French

# 1 Introduction

This paper addresses the contribution that language variation and change (LVC) studies can make to the **construction**, **elaboration** and **testing** of formal syntactic and semantic theories, with a particular focus on the testing dimension.

One of the fundamental scientific hypotheses characterizing the field of linguistics in the 20th and beginning of the 21st centuries is that the principles and mechanisms that underly the construction and interpretation of natural language expressions are different from those principles/mechanisms that determine the patterns of use of these expressions. As a consequence of this hypothesis, formal linguists, be they working in the Generative (Chomsky, 1957, 1965, et seq.) or other traditions, have taken their central object of study to be contrasts in **grammaticality** (i.e. whether or not a sequence of words or morphemes is a well-formed expression in the language) and **interpretation** (i.e. which meanings are assigned to grammatical expressions with which forms). Thus, the main goal of formal linguistic theory has been to model grammaticality and interpretative contrasts found in the languages of the world through the use of **deterministic** grammars: formal descriptions of a language consisting of a set of primitive words/morphemes and rules that combine (or 'generate') complex grammatical expressions from these primitives while, at the same time, assigning them an appropriate meaning.

On the other hand, LVC researchers have taken their central object of study to be patterns of linguistic **variation**; that is, contrasts in the distributions of synonymous grammatical forms (or 'variants') in the speech of a speaker (or a population of speakers), as observed in a written or spoken corpus. One of the main aims of scholars working in the Variationist tradition (Labov, 1963, 1966; Weinreich et al., 1968, et seq.) is to identify the set of factors, linguistic and/or social, that influence the use of one variant over another and what the

role of these factors are in both linguistic and social change. In the analysis of these patterns of language use, many LVC researchers make use of **probabilistic** grammars: formal grammars that generate a language along with a probability distribution on its elements. These differences in the classes of grammars used in linguistic analysis in these two different fields have made insights from LVC, which are primarily associated with the source and shape of the probability distribution on the language, difficult to integrate into formal linguistic proposals that concern the form of the language itself.

However, more recently, compelling arguments have been developed that the hypothesis of the total separation between the form of human languages and the form of linguistic usage patterns is not well-founded. For example, there is a large (and growing) body of literature that shows that grammatical factors that determine grammaticality contrasts, also known as **hard** contrasts[1], in some languages (traditionally, the domain of study of formal linguistics) determine preferential, or **soft**, contrasts in other languages (traditionally, the domain of study of LVC and psycholinguistics). A classic example of this hard/soft duality involves person hierarchy effects and their interaction with grammatical voice. In many languages, the set of DPs that can occupy the subject position is restricted by grammatical person. For example, as observed by Jelinek and Demers (1983) (and taken up in Bresnan et al. (2001)), in Lummi, a Salish language spoken in British Columbia, transitive predicates that have third person actors and first or second person patients must appear in the passive voice; that is, in this language, it is impossible to say (the Lummi equivalent) of *The man knows me*, rather one must say (1-a). On the other hand, if the agent is first or second person and the patient is third person, then the active voice is obligatory (1-b); that is, one cannot say the equivalent of *The man is known by me.*

(1) Person-restricted voice in Lummi

[1]See (Bresnan et al., 2001; Sorace and Keller, 2005; Thullier, 2012, among others).

a. xči-t-**ɪŋ** =sən ə cə swəyʔqəʔ
know-tr-**PASS** =1.SING.NOM by the man
'I am known by the man.'

b. xči-t =sən cə swəyʔqəʔ
know-tr =1.SING.NOM the man
'I know the man.'

Cited from (Bresnan et al., 2001, p.1)

The contrast between (1-b) and the corresponding active sentence with a third person subject is one of grammaticality, and, therefore, this is exactly the kind of empirical data that we would like our formal linguistic theories to account for.

Of course, not all languages are like Lummi: in English, having a third person agentive subject with a first person object is perfectly grammatical, and speakers of this language have the option of saying either *The man knows me* or *I am known by the man*. However, as shown by Bresnan et al. (2001) by means of a quantitative study using the parsed *Switchboard* corpus of spoken English (Godfrey et al., 1992), when first and second person actors act on third person patients in the Switchboard corpus, the action is uniformly expressed using the active voice (0/6246 occurrences). On the other hand, when third person actors act on first or second person patients, the action is expressed using the passive voice in 2.9% of the cases (14/486 occurrences), which is a small but highly statistically significant difference. In other words, what appear as grammaticality contrasts in Lummi and other languages appear as preferences in spoken English.

Hard/soft correspondences of the type just described are widespread and found in many different parts of the grammar, including subject definiteness (Givón, 1979), relative clause formation (Keenan and Comrie, 1977; Keenan and Hawkins, 1987; Hawkins, 2004), argument ordering in the noun phrase (Rosenbach, 2002, 2005; Bresnan, 2007), polarity splits in agreement morphology (Chambers, 2004; Tagliamonte, 2011, among many others) and

tense morphology (Deshaies and Laforge, 1981; Poplack and Turpin, 1999; Poplack and Dion, 2009), negative concord patterns (Burnett et al., 2015) and postverbal complement ordering (Bresnan et al., 2007; Thullier, 2012; Tagliamonte, 2014, among others). The coincidences between the shape of grammaticality contrasts in one language and the shape of preferential contrasts in another strongly suggest that these patterns have a common source and that at least some of the mechanisms that underly the construction and interpretation of natural language expressions coincide with at least some of those that determine the patterns of their use. This state of affairs has the important consequence that, if we accept that patterns of variation can shed light on the form of human grammars, we now have access to a whole new empirical domain (variation studies) which can be used to formulate generalizations concerning which kinds of grammatical factors universally condition grammaticality contrasts and/or preferential contrasts in human languages (and which do not).

The hard/soft duality has a further consequence for the form of our linguistic theories: Since we pursue a grammatical explanation for hard contrasts (i.e. we say that, when one expression is grammatical in a language and another is not, it is because of features of that language's grammar), I suggest that the most straightforward explanation for (pertinent) soft contrasts in natural languages is also grammatical. However, in order to integrate 'soft' syntactic patterns into our formal linguistic theories, we need to move to grammars that will allow us to do so[2] . As discussed above, probabilistic grammars can capture preferential

[2]The use of non-deterministic grammars is not the only logically possible way of capturing the hard/soft correspondences discussed in this paper. One could also propose that variation is the result of the speaker having multiple grammars that define multiple languages, in the style of Kroch (2000). Soft syntactic patterns would therefore mirror hard syntactic patterns because the soft patterns are the result of alternation between languages in which these patterns are hard. I view the proposals made in this paper as compatible, in principle, with this view. Indeed, as we will see, the Stochastic OT grammars that are used in the analysis presented in the paper account for variation using alternation between deterministic grammars. Likewise, certain theories of morphological variation, such as Adger and Smith (2010); Adger (2014), use deterministic grammars for syntax and adopt a more complicated theory of the lexicon- morphology interface. The phenomenon analyzed in this paper is located at the syntax-semantics interface, so it is not clear to me how naturally Adger et al.'s model extends to cover something like variable negative concord

distinctions between grammatical expressions in a way that deterministic grammars cannot. Therefore, in this paper, we will explore how (an appropriate subset of) this class of formal systems can be useful in the construction and testing of formal syntactic theories[3].

As an illustration of this proposal, in this paper, I present a new empirical study of obligatory and optional asymmetric negative concord, and I show how an influential analysis for obligatory concord patterns (de Swart, 2010) can be tested through looking at the predictions that its natural probabilistic extension makes for the forms, interpretations and frequency distributions of expressions in languages in which asymmetric concord is optional. In obligatory asymmetric negative concord languages such as Spanish, Italian and European Portuguese, negative indefinites, known in the literature (after Laka (1990)) as *n-words*, may appear in preverbal position without sentential negation; however, when they appear in postverbal position, they must be c-commanded by an appropriate negative operator, usually sentential negation. This asymmetry is exemplified in (2) using the Spanish n-word *nadie* 'no one'.

(2) a. **Nadie** ha venido.
No one has come.
'No one came.'

b. *(**No**) he visto a **nadie**.
**Not** have seen a no one
'I saw no one.'

A further characterizing feature of the Spanish/Italian/Portuguese asymmetric pattern concerns co-occurrence possibilities between preverbal n-words and sentential negation; in

(or even if it is meant to). However, this may be an interesting analytical path to pursue in the future.

[3]Of course, we must be very careful with which kinds of probabilistic grammars that we choose; in particular, we need to avoid systems (such as Markov models or stochastic context-free grammars) which, as discussed already by Chomsky (1957), have certain features that render them inappropriate for the modelization of linguistic competence. See also Stabler (2013) for similar criticisms of this use of certain recent stochastic extensions of multiple context-free grammars and Minimalist grammars.

particular, not only can preverbal n-words in these languages appear bare, they **must** do so. That is, while co-occurrence of *nadie* and *no* is obligatory in (2-b), it is prohibited in (3), under normal intonational patterns.

(3) *__Nadie no__ ha venido.
No one not has come
Intended: 'No one came.'

Another language that shows the preverbal/postverbal asymmetry is the dialect of French spoken in Montréal, Québec. As observed by (Daoust-Blais, 1975; Lemieux, 1985; Muller, 1991; Sciullo and Tremblay, 1996; Déprez, 2002; Labelle, 2010; Larrivée, 2014, among many others), Montréal French (MF) displays a negative concord pattern similar to the one found in Spanish: MF n-words, such as *personne* 'no one', may not co-occur with sentential negation (the VP adverb *pas*) when they appear in preverbal position (4-a)[4]; however, co-occurrence with negation is possible with postverbal n-words (4-b), in which case the unique interpretation created is a negative concord one[5].

(4) Asymmetric Concord in Montréal French

a. **Personne** est (*__pas__) venu.
No one is (*__not__) come
'No one came.'

b. J'ai **pas** vu **personne**.
I have **not** seen no one
'I saw no one.'

[4]As observed by Sankoff and Vincent (1977), bipartite negation (i.e. *ne... pas*) has been (almost) completely eradicated from the language spoken in Québec, with *ne* appearing in only 0.5% of negative sentences in the *Sankoff-Cedergren* (a.k.a. *Montréal 71*) corpus, which is a precursor to the *Montréal 84* corpus used in this paper.

[5]This is an important difference between Montréal French and most varieties of French spoken in Europe, where co-occurrence between postverbal n-words and sentential negation obligatorily gives rise to double negation interpretations, i.e. in most of the French dialects spoken in Europe, *J'ai pas vu personne* only means 'I didn't see no one.'

However, unlike the categorical asymmetric concord systems described above, co-occurrence with negation is optional for postverbal n-words in the French dialect; that is, postverbal n-words can also appear bare (5), and the interpretation is the same the same as (4-b).

(5) J'ai vu **personne**.
I have seen no one
'I saw no one.'

MF thus shows a 'soft' version of the postverbal n-word/negation co-occurence pattern that is 'hard' in Spanish and, in this paper, we will investigate a probabilistic extension of de Swart (2010)'s analysis for the Spanish pattern can capture these patterns of syntactic variation, as observed in a quantitative study of the distribution of negative concord structures in the *Montréal 84* corpus of spoken Montréal French (Thibault and Vincent, 1990).

The paper is laid out as follows: in section 2, I present the main lines of de Swart (2010)'s account of the Spanish pattern, which is set within a Bidirectional Optimality Theory (OT) approach to the syntax-semantics interface. In section 3, I consider extending this analysis within Boersma (1998)'s stochastic generalization of the OT framework (StOT) to account for variable asymmetric concord languages, and I identify what predictions the appropriately extended proposal makes for the distribution of negative concord structures in a corpus of spoken Montréal French. Section 4 presents a quantitative study of the distribution of negation and negative indefinites in the *Montréal 84* oral corpus. I show that the main (non-trivial) predictions of the extension of de Swart's account are borne out; however, I also argue that certain other aspects of the distribution of concord structures in the data are problematic for the naive extension of the analysis presented in section 2. Therefore, in section 5, I modify de Swart's analysis and show how my proposal can

account not only for the range and interpretation of negative structures in this Montréal French, but also for their rate of use in *Montréal 84*. Section 6 gives a summary of the main empirical and theoretical proposals in this work and provides some concluding remarks on the role of both Stochastic grammars and quantitative corpus studies in the construction of syntactic theory.

# 2 Obligatory Asymmetric Concord in Bidirectional OT

This section presents de Swart (2010)'s account of the syntax and semantics of obligatory asymmetric negative concord in languages like Spanish. As mentioned in the introduction, her proposal is set within a Bidirectional OT approach to the syntax-semantics interface. Therefore, I will first lay out a version of the basic OT framework and then show how de Swart derives the preverbal/postverbal asymmetry within it.

## 2.1 Bidirectional OT as a Theory of the Syntax-Semantics Interface

Although Optimality Theory was originally developed to model phenomena associated with the phonological module of the grammar (Prince and Smolensky, 1993); this architecture was quickly extended to the analysis of syntactic phenomena (Grimshaw, 1997; Bresnan, 2000, and much subsequent work), and, more recently, to semantic and pragmatic phenomena (see Hendriks and de Hoop, 2001; de Hoop and de Swart, 2000; Blutner, 2000; Zeevat, 2001; Krifka, 2007, and much subsequent work in the field). In contrast to the vast majority of work done in Interpretative semantics (Montague, 1970) and derivational approaches to the syntax-semantics interface (May, 1985; Chomsky, 1995), in OT approaches to the

syntax-semantics interface, the pairing of linguistic forms with particular meanings is optimized according to both generation (i.e. speaker) and parsing (i.e. hearer) considerations. In order to model this, we use a class of grammars called *Optimality Systems*, which are defined as follows[6]:

**Definition 2.1.** ***OT System.*** *An OT system $\mathcal{O}$ is a pair $\langle GEN, \mathcal{C}\rangle$, where $GEN$ is a relation and $\mathcal{C}$ (the* constraint set*) is a linearly ordered set of functions ($\mathcal{C} = \langle c_1, c_2, c_3 \ldots c_n\rangle$) from $GEN$ into* $\mathbb{N}$.

In phonology and morphology, $GEN$ is usually conceived of as a relation between $\langle$input, output$\rangle$ pairs. In a syntax-semantics interface setting, however, $GEN$ is a relation between a syntactic form $f$ and a meaning $m$ (so we write $\langle f, m\rangle \in GEN$). The OT architecture is extremely versatile, which allows for almost limitless possibilities when it comes to what constitutes a *form* and what constitutes a *meaning*. In order to be as general as possible, in this paper (as in de Swart (2010)), we will consider forms to simply be sequences of words taken from a language-specific lexicon structured into syntactic constituents (although (with some exceptions discussed in section 5) I will remain as agnostic as possible concerning what exactly those constituents are). Furthermore, I assume that the lexical items of Spanish (or English or Montréal French or whatever) that compose these forms are assigned semantic interpretations in the lexicon in a way that consistent with (many versions of) *Generalized Quantifier Theory* (Barwise and Cooper, 1981; Keenan and Stavi, 1986; Keenan and Westerstahl, 1997, among others)[7]. For example, verbal predicates will denote relations of the appropriate arity; that is, intransitive verbs will denote subsets of the domain $D$ (so, for the English verb *arrive*, we write $[\![arrive]\!] \subseteq D$) and transitive verbs will denote binary relations (so, for the English verb *see*, we write $[\![see]\!] \subseteq D \times D$).

[6] All formalizations are taken from Jäger (2002), sometimes recast in my notation.

[7] Thus, this framework is only interested in modelling optimization at the level of compositional semantics, not lexical semantics. This is not necessary, and see Blutner (2000) among others. I will make further remarks on the optimized lexicon in section 5.

Furthermore, expressions that are syntactically determiner phrases (DPs) are analyzed as denoting generalized quantifiers: arity reducing functions which map $n$+1-ary relations to $n$-ary relations in the way specified by the lexical meaning of the DP[8].

In the spirit (if not the letter) of de Swart, we will take meanings to be strings of symbols, for which will give a model theoretic interpretation based on combinations of possible denotations of lexical items. More specifically, we will have a set of (in)transitive predicates such as ARRIVE or SEE. I highlight that, although these symbols bear striking resemblances to English words, they are meant to form part of the abstract semantic representations that will be given for a number of languages in this paper. Symbols like ARRIVE and SEE are interpreted into the same model as the lexical items above, and we write $[ARRIVE] \subseteq D$ or $[SEE] \subseteq D \times D$ for their interpretations. Likewise, we will have symbols that will be interpreted as quantifiers: [NOBODY] will map properties to true just in case they have no human members. Thus, the string (7-a), where NOBODY and ARRIVE are appropriately concatenated based on their semantic type, will be true in a model $M$ just in case the denotation of ARRIVE in $M$ has no human members. Strings with transitive predicates such as (7-b) and (7-c) indicate the order of the composition of arguments: (7-b) is true in a model $M$ just in case [JOHN] maps the set of individuals who saw nobody to true in $M$ (i.e. if John saw nobody), and (7-c) is true in a model just in case [NOBODY] maps the set of humans who saw John to true in that model (i.e. if nobody saw John)[9].

(7) Sample meanings in the range of $GEN$
   a. NOBODY(ARRIVE)

[8]More technically: for a unary generalized quantifier $F^1$ (a function from properties to truth values over a domain $D$), we extend the domain of $F^1$ to include all $n+1$-ary relations $R$ by setting

(6) $F^{n+1}(R) = \{\langle a_1, a_2 \ldots a_n\rangle : F^1(\{b : \langle a_1, a_2 \ldots a_n, b\rangle\}) = 1\}$.

(see Keenan and Westerstahl, 1997, for more details).

[9]Note that just because, in this paper, we use this small language to pick out meanings, this does not mean that the (language independent) meaning denoted by NOBODY(JOHN(SEE)) will necessarily be paired with a form containing a negative quantifier.

b. JOHN(NOBODY(SEE))
c. NOBODY(JOHN(SEE))
d. NOT(JOHN(ARRIVE))

Furthermore, in line with de Swart and Sag (2002) and Peters and Westerstahl (2006), we will represent sentential negation as a 0-ary quantifier which maps truth values to truth values; thus the string in (7-d) is interpreted as true in a model in which John did not arrive.

Thus, $GEN$ will consist of the product of the set of complex forms and the set of meanings, which are strings such as in (7)[10]. So some sample members of $GEN$ for a language like English would be as in (8):

(8) Subset of $GEN$ for English
a. ⟨John saw nobody, JOHN(NOBODY(SEE)⟩
b. ⟨John saw nobody, NOT(JOHN(NOBODY(SEE)))⟩
c. ⟨John saw nobody, FELIX(MEOW)⟩

Of course $GEN$ will contain very many more form-meaning pairs than will form part of the interpreted language, and the way that the final set of pairings between syntactic forms and semantic interpretations is calculated involves the use of the constraint set $\mathcal{C}$ in Def. 2.1. A constraint $c_i \in \mathcal{C}$ is a function from form-meaning pairs ($\langle f, m\rangle \in GEN$) to natural numbers such that $c_i$ maps $\langle f, m\rangle$ to the number of violations of $c_i$ that $\langle f, m\rangle$ incurs. Now, based on the ordering of the constraints in $\mathcal{C}$ and the values of the members of $GEN$ at the $c_i$s, we define an 'optimization' ordering on form-meaning pairs as follows:

**Definition 2.2.** ***Optimization Ordering*** ***($>_{\mathcal{O}}$)*** *Let $\mathcal{O}$ be an OT system and let $\langle f, m\rangle, \langle f', m'\rangle \in GEN$. Then, $\langle f, m\rangle >_{\mathcal{O}} \langle f', m'\rangle$ iff*

[10]Having $GEN$ be quite unrestricted is known as the *free interpretation hypothesis* in OT Semantics (Hendriks and de Hoop, 2001), which can be thought of as the semantic correspondent to OT phonology's *richness of the base* hypothesis.

1. *There is some* $i : 1 \leq i \leq n$ *such that* $c_i(\langle f, m\rangle) < c_i(\langle f', m'\rangle)$, *and*

2. *For all* $j > i, c_j(\langle f, m\rangle) = c_j(\langle f', m'\rangle)$.

According to Def. 2.2, a form-meaning association $x$ is better than another one ($y$) just in case there is some constraint $c_i$ at which $x$ incurs fewer violations than $y$, and $x$ and $y$ have the same values at all the higher ranked constraints. With this definition, it turns out that if $\mathcal{O}$ is an OT system, then $>_\mathcal{O}$ is both transitive and well-founded (cf. Jäger (2002)'s Lemma 2). The acceptable form-meaning associations are only a subset of $GEN$, namely those that are **optimal**[11]:

**Definition 2.3.** ***Optimality****. A form meaning pair* $\langle f, m\rangle$ *is* ***optimal*** *iff*

1. $\langle f, m\rangle \in GEN$.

2. *there is no optimal* $\langle f, m'\rangle$ *such that* $\langle f, m'\rangle >_\mathcal{O} \langle f, m\rangle$.

3. *there is no optimal* $\langle f', m\rangle$ *such that* $\langle f', m\rangle >_\mathcal{O} \langle f, m\rangle$.

According to Def. 2.3, then, grammars pair-up forms and meanings in a way such that the best forms get paired up with the best meanings, the second-best forms get paired up with the second-best meanings, and so on.

Of course, the meat of any OT analysis lies in the exact proposals concerning the inventory and ranking structure of the constraint-set. Although the grammar formalism does not force this, it is generally assumed that the constraint inventory is universal, and, therefore, ideally one would like the constraints to be as general and typologically or functionally well-motivated as possible. The rankings of the constraints, on the other hand, vary from language to language, and this variation is what creates the diverse syntactic and interpre-

[11]The notion of optimality that we use here is Jäger (2002)'s *X-optimality*, which (given the definition of $>_\mathcal{O}$ that we are assuming) is equivalent to and more elegant than Blutner's original formulation, which Jäger calls *Z-optimality*.

tative patterns that we see across languages. As a preview of what is to come, I highlight here that the proposed universality of $\mathcal{C}$ is what allows proposals set in OT frameworks to make clear predictions for linguistic systems that have not yet been studied (such as Montréal French).

First of all, most OT syntax/semantics systems adopt some form of highly-ranked (but ultimately violable) faithfulness constraint that ensures that the elements in the semantic representations largely have correspondences in the syntactic representation (see Bresnan, 2000; Hendriks and de Hoop, 2001; Zeevat, 2001; Blutner et al., 2003, among others). This has the effect of eliminating the most unlikely form-meaning pairings (such as (8-c)) and restoring, in the words of (Zeevat, 2001, p.10), "the important aspects of compositional semantics (not the full principle but essential aspects)". In line with the setup here, the faithfulness constraint that we will adopt is the following, which I call FaithLex.

(9) FaithLex assigns one violation to a pair $\langle f, m\rangle$ for every symbol in $m$ whose interpretation is not the denotation of a lexical item of $f$.

So, for example, FaithLex assigns no violations to (8-a) and 2 violations to (8-c). This constraint concerns general correspondences between forms and meanings; however, in the next section, we will see more specialized ones that create the Spanish asymmetric negative concord pattern, as analyzed by de Swart (2010).

## 2.2 Categorical Asymmetric Concord in Bidirectional OT (de Swart, 2010)

In her 2010 book, de Swart proposes an OT analysis of the main typological patterns associated with the distribution and interpretation of negative morphemes. In this work,

she accounts for an impressive range of data from the various kinds of attested negative concord patterns, the Jespersen cycle, double negation interpretations and the distribution of positive and negative indefinites, among other phenomena. As such, the OT analysis presented in de Swart (2010) is highly articulated and involves many different syntactic and semantic constraints. Since the aim of this paper is much more modest (only to account for categorical and variable asymmetric concord languages), I will only present the subset of her constraint-set that creates the patterns discussed above.

The heart of de Swart's proposal is the claim that the asymmetric pattern is created by the interaction between two markedness constraints: *NEG and NEG FIRST. *NEG is a constraint that considers both members of the form-meaning pair, and assigns violations depending on how many occurrences of negative marking or elements with a negative interpretation are found in the pair.

(10) *NEG assigns one violation to a pair $\langle f, m \rangle$ for every negative morpheme in $f$ and every symbol in $m$ with a negative denotation[12].

As de Swart says (p.78), "the intuition behind *NEG is that negation is marked, both in form and in meaning. Marked forms and meanings should be avoided, so negation should be avoided both in the syntax and the semantics."

Unlike *NEG, which evaluates both the forms and the meanings, NEG FIRST only looks at the form. For de Swart, NEG FIRST is actually conceived of as a family of markedness constraints that govern the placement of negation in the sentence, and we will also take this view in this paper. The version of the constraint that de Swart proposes that is highly ranked for languages like Spanish is shown in (11).

[12] A generalized quantifier $F$ is *negative* iff $F(\emptyset) = 1$ (see Keenan and Stavi, 1986; Peters and Westerstahl, 2006).

(11) NEG FIRST (de Swart, 2010, 96):
Negation precedes the finite verb.

The use of a principle such as NEG FIRST to account for the syntactic distribution of negative morphemes has, in fact, a long history in both functional and formal linguistics. Indeed, de Swart adopts the term *Neg first* from Horn (1989), who describes it as "the strong tendency for negative markers to gravitate leftward so as to precede the finite verb or other possible foci of negation" (Horn, 1989, p. 452). Horn himself traces this principle back to the work of Jespersen (1917, 1933), who suggests that natural languages have "a natural tendency, also for the sake of clearness, to place the negative first, or at any rate as soon as possible, very often immediately before the particular word to be negated (generally the verb)" (Jespersen, 1917, p.5).

In obligatory asymmetric concord languages, de Swart proposes that NEG FIRST is ranked higher than *NEG, since satisfying NEG FIRST allows for two negatively marked elements to be acceptable when the n-words are in postverbal position. As I mentioned, in Bidirectional OT, form-meaning pairs are optimized with respect to both the speaker (meaning → form, as in OT Syntax) and hearer (form → meaning, as in OT Semantics). Thus, it is convention to represent this optimization through the use of two tableaux: one showing the evaluation from the semantic input to the syntactic output (speaker), and then another showing the evaluation from syntactic input to semantic input (hearer). In this way, tableau 1 shows how the meaning NOBODY(ARRIVE) is optimally paired with the form *Nadie ha llegado* (abstracting away from tense), and tableau 2 shows how the form *Nadie ha llegado* is optimally paired with the meaning NOBODY(ARRIVE).

Furthermore, because of the high ranking of NEG FIRST, the optimal output for the semantic form MARIA(NOBODY(SEE)) is *María no ha visto a nadie*, as shown in Tableau 3.

| Input: NOBODY(ARRIVE) | FaithLex | NegFirst | *Neg |
|---|---|---|---|
| a. ☞ Nadie ha llegado | | | * |
| b. Nadie no ha llegado | | | **! |

**Table 1** – Preverbal n-words in Spanish (speaker's perspective)

| Input: Nadie ha llegado | FaithLex | NegFirst | *Neg |
|---|---|---|---|
| a. ☞ NOBODY(ARRIVE) | | | * |
| b. NOT(NOBODY(ARRIVE)) | | | **! |

**Table 2** – Preverbal n-words in Spanish (hearer's perspective)

Since, as shown in tableau 4, MARIA(NOBODY(SEE)) has fewer semantic negations than NOT(MARIA(NOBODY(SEE))), ⟨María no ha visto nadie, MARIA(NOBODY(SEE))⟩ is selected as the optimal form-meaning pair, and we correctly predict that such sentences should only have a negative concord interpretation.

In the next section, I consider applying de Swart's analysis to variable asymmetric concord patterns within a stochastic extension of the bidirectional OT system described above. I show how this account makes straightforward predictions for the patterns of use of negative concord structures in a spoken corpus.

# 3 Optional Asymmetric Concord in StOT

This section examines the prospects for integrating the variable Montréal French asymmetric concord system into the OT theory described in section 2. As mentioned in the introduction, MF shows the preverbal/postverbal contrasts that characterize the asymmetric concord pattern, with the exception that co-occurrence between sentential negation *pas* and an n-word (such as *rien* 'nothing') is optional rather than obligatory (12). That true (i.e. intra-speaker) variation exists in this dialect can be established through the ob-

| Input: MARIA(NOBODY(SEE) | FaithLex | NegFirst | *Neg |
|---|---|---|---|
| a. María ha visto a nadie | | *! | * |
| b. ☞ María no ha visto a nadie | | | ** |

**Table 3** – Postverbal n-words in Spanish (speaker's perspective)

| Input: María no ha visto a nadie | FaithLex | NegFirst | *Neg |
|---|---|---|---|
| a. ☞ MARIA(NOBODY(SEE)) | | | * |
| b. NOT(MARIA(NOBODY(SEE))) | | | **! |

**Table 4** – Postverbal n-words in Spanish (hearer's perspective)

servation that both bare and negative concord variants can be used by the same speakers in the same conversation, as shown, for example, by (12) from the *Montréal 84* corpus. More generally, Burnett et al. (2015) show that education is a significant social factor conditioning the use of concord structures in *Montréal 84*: speakers with a higher level of education are less likely to use concord sentences than speakers with a lower level of education. However, I highlight that there are speakers from every education level that have the variable system described in this paper[13].

(12) a. La loi cent un moi j'ai **rien** contre ça (27 213)
'Loi 101 me I have nothing against that.'

b. C'est pour ça que j'ai **pas rien** contre la loi cent un (27 221)
'It's for that that I have nothing against Loi 101.'

Although de Swart does not discuss this dialect of French (or other systems that show the same variable pattern), she suggests that similar cases of variation in the realization and interpretation of negative indefinites should be handled within a stochastic extension of her

[13]Although very interesting and important, the question of the social meaning of variable negative concord, the style(s) (in the sense of Eckert (2008)) with which it is associated and how its social conditioning should be modelized in an OT framework is out of the scope of this paper.

OT system. Moving to a probabilistic system (or some other kind of significant departure) is necessary to allow for the co-existence of the two synonymous forms in (12-a) and (12-b) for the following reason: if we wish to explain the ungrammaticality of a sentence like ***Rien** *me plaît* **pas**. (Intended: 'Nothing pleases me') through the use of the *NEG constraint (as we did for Spanish), then we will need to find some reason why (12-b) is not ungrammatical, because it presumably incurs the same number of violations. In Spanish, negative concord structures with postverbal n-words were permitted because they satisfied a higher-ranked NEG FIRST constraint; however, again, if we use the ranking NEG FIRST≫*NEG to allow (12-b), we end up predicting that it should win over (12-a) because the negative element appears later in the utterance. In other words, the optimization algorithm that we use to determine which form-meaning pairs are in the language predicts only obligatory patters; there is no place for optionality.

## 3.1 Modelling Grammatical Variation

An influential way through which grammatical optionality is modelled in Optimality Theory is through the use of Boersma (1998)'s stochastic generalization of the framework. Stochastic OT (StOT) shares the generation relation ($GEN$) and the constraint set $\mathcal{C}$ with regular (also known as *ordinal*) OT, and the main differences between the two classes of grammars come in the form of the ordering relation between constraints and the evaluation algorithm. Unlike in the system described above in which constraints are ordinally ordered, in StOT each constraint $c_i$ is assigned a real number on a continuous scale, called its *rank*. Since the constraints are ranked on a continuous scale, we can now talk of the distance between two constraints in a meaningful way. Additionally, at every evaluation event, a small amount of noise (a value chosen from a normal distribution with mean 0 and standard deviation 2 (or some other arbitrary value)) is added to the constraint ranking. The

rank of a constraint after the noise is added is called its *selection point*. If the ranks of two constraints are very far apart, the noise that is added at evaluation time will not change the ordinal ranking of the selection points at each evaluation event. However, if the constraints are ranked very close together, then the ranking of selection points might change from evaluation to evaluation. It is in this way that StOT models linguistic variation: the evaluation algorithm defines a probability distribution over ordinal rankings of selection points, which, in turn, defines a probability distribution over the set of candidates. More specifically, the probability of a particular expression (in our case, particular form-meaning pairs) being optimal will be the sum of the probabilities of all ordinal rankings that make it optimal. As such, a StOT grammar describes an interpreted language just in case it assigns probabilities to the form-meaning pairings that correspond to their relative frequencies in the language.

With these considerations in mind, let's examine the predictions that a stochastic extension of de Swart's analysis applied to variable concord in Montréal French make for the distribution of negative concord structures with *pas* in an oral corpus. As discussed above, MF displays a preverbal/postverbal asymmetry (*Personne est (****pas****) venu.* 'No one came.' vs *Jean a* ***pas*** *vu personne.* 'Jean saw no one.' ). Therefore, in line with de Swart, we would like to account for this pattern using the interplay between the NEG FIRST constraint and *NEG: the extra negation in *Jean a* ***pas*** *vu personne.* is acceptable because it satisfies NEG FIRST; however, it is prohibited in **Personne est* ***pas*** *venu.* because *Personne est venu.* is already optimal for NEG FIRST. This being said, variation in this dialect is possible when the n-word is in postverbal position: n-words in this position can also appear bare (ex. *J'ai vu personne.* 'I saw no one.'). Within the context of the analysis developed above, the most natural account of the acceptability of bare postverbal n-words is that it is due to the ranking of *NEG over NEG FIRST: it is more important to avoid the proliferation of negative markers than to express negation as soon as possible in the sentence. The optimal

pairing of *Jean a vu personne* with JEAN(NOBODY(SEE)) is shown in tableau 5.

| Input: JEAN(NOBODY(SEE) | FAITHLEX | *NEG | NEGFIRST |
|---|---|---|---|
| a. ☞ Jean a vu personne | | * | * |
| b. Jean a pas vu personne | | **! | |

**Table 5** – Bare postverbal n-words in Montréal French (speaker's perspective)

Indeed, the ordering *NEG ≫ NEG FIRST is what de Swart proposes to derive the Spoken European French pattern (13), in which postverbal n-words appear bare and co-occurrence with *pas* creates an obligatory double negation interpretation.

(13) Spoken European French

a. J'ai vu personne.
I have seen no one
'I saw no one.'

b. J'ai **pas** vu personne.
I have not seen no one
'I didn't see no one.', i.e. I saw someone. Double negation only

In StOT, the variation between the European French-style system (with bare n-words) and the Spanish system (with negative concord structures) can be modelled through the construction of an OT grammar in which the *NEG and NEG FIRST constraints are assigned very close values on the continuous ranking scale. The noise that is added at each evaluation event could perturb the initial ranking order between *NEG and NEG FIRST, which will, in turn, define a probability distribution over ordinal rankings: *NEG will be ordered before NEG FIRST (which associates *Jean a vu personne* with JEAN(NOBODY(SEE))) some proportion of the time, while NEG FIRST will be ordered before *NEG (which associates *Jean a pas vu personne* with JEAN(NOBODY(SEE))) the rest of time. In order to determine how close the initial rankings of the two constraints need to be to create this pattern, we will need to look at corpus data, which is what we will do in section 4.

An important way in which MF differs from Spanish (which will affect the version of NEG FIRST family of constraints that we test) is in the morphophonological properties of its sentential negation marker. Unlike Spanish, Italian and European Portuguese in which negation is a preverbal head (sometimes called a *light* negation (Giannakidou, 2006)), MF negation is a phrasal adverb (Pollock (1989); Abeillé and Godard (1997), also known as a *heavy* negation), which generally occupies the third position in the sentence following the finite verb (14-b) or auxiliary (14-c).

(14) a. **No** he visto a Juan. Light negation (Spanish)

b. Je vois **pas** Jean. Heavy negation (Montréal French)

c. J'ai **pas** vu Jean. Heavy negation (Montréal French)

So, unlike in Spanish in which NEG FIRST is satisfied through having negation in preverbal position, the fact that there exists a preverbal/postverbal asymmetry in Montréal French, shows that NEG FIRST in this language must be satisfied with the negation in third position[14]. In other words, de Swart's proposal for the constraint that is highly ranked in Spanish (11) needs to be replaced by a more general version of the constraint.

Furthermore, if we would like to evaluate whether NEG FIRST is active (at the 'soft' level) in Montréal French, we first have to be more precise about its formulation. In the literature (Jespersen, 1933; Horn, 1989; Corblin and Tovena, 2003; de Swart, 2010, among others), NEG FIRST is treated as a constraint that is sensitive to linear order. Therefore, I will devote the majority of the rest of this paper to testing the pertinent extension of de Swart's analysis using linear NEG FIRST, which counts syntactic constituents[15] from the beginning

[14]A similar point could be made about (informal) Welsh, which allows (restricted) variable negative concord with a postverbal adverbial negation marker. See Borsley and Jones (2005) for discussion.

[15]Of course, we could also count the number of 'words' from the beginning of the sentence, which would make similar predictions, but, since we will ultimately revise the linear NEG FIRST constraint in the next section, I will not consider this alternative here.

of the utterance. As shown in (15), a form-meaning pair containing a negative constituent is assigned the number of violations equal to the number of syntactic constituents in the utterance that (properly) linearly precede the constituent.

(15) Linear NEG FIRST (NEG FIRST$_L$):

Let $>_L$ be the linear precedence relation between syntactic constituents $(\psi, \psi_1, \psi_2 \ldots)$, and let $\langle f, m \rangle$ be a form-meaning pair such that $f$ contains a constituent $\phi$ with a negative denotation. Then,

$$\text{NEG FIRST}_L(\langle f, m \rangle) = |\{\psi : \psi >_L \phi\}|$$

Since the set of candidates that we will be considering in the rest of the paper will differ minimally in their number of syntactic constituents, I will only mark the violations of NEG FIRST$_L$ that distinguish two candidates in the tableaux that follow. For example, in a tableau comparing candidates *J'ai pas vu personne* and *J'ai vu personne*, I will mark the concord structure has having no violations of NEG FIRST$_L$ and the bare structure as having a single violation of this constraint. This will greatly increase the readability of the tableaux, and I hope that this will not cause confusion.

In combination with some syntactic constraints forcing the realization of sentential negation as an adverb, the constraint in (15) has a consequence for the predictions of an analysis in which variable asymmetric concord patterns are derived through a 'soft' NEG FIRST constraint. In particular, unlike in Spanish/Italian/Portuguese in which having a bare direct object n-word is always less optimal with respect to NEG FIRST than having a concord structure, this is not necessarily the case in Montréal French. As shown in (16), if the verb is in the present tense, a direct object n-word will be in third position (16-a), which is the same place in the linear order as *pas* in the negative concord version of the sentence

(16-b). Therefore, both structures should satisfy NEG FIRST$_L$ equally well. However, (16-a) and (16-b) do not both satisfy *NEG equally well: (16-b) has an extra negation marker which earns this candidate an extra violation.

(16) a. Je vois **personne**.
I see no one
'I see no one.'

b. Je vois **pas personne**.
I see not no one
'I see no one.'

As a consequence, even at an evaluation event in which the selection point of NEG FIRST$_L$ is higher than that of *NEG, the optimal form for a sentence with a transitive verb in the present tense will be *J'ai vu personne.* This is shown in tableau 6.

| Input: JEAN(NOBODY(SEE) | FAITHLEX | NEGFIRST $_L$ | *NEG |
|---|---|---|---|
| a. ☞ Jean voit personne | | | * |
| b. Jean voit pas personne | | | **! |

**Table 6** – Postverbal third position n-words (speaker's perspective)

In other words, a stochastic extension of de Swart's analysis predicts that negative concord structures where n-words follow the finite verb (modulo *pas*) are harmonically bounded and should not appear in a spoken corpus.

The same prediction is made for fragment answers: because of the extra negation, *pas personne* is harmonically bounded and not predicted to occur.

(17) Qui as-tu vu? 'Who did you see?'

a. **Pas personne**/ **Personne.** 'No one.'

On the other hand, where we do expect to see the contribution of NEG FIRST$_L$ is in struc-

tures in which negation appears in an earlier position in the sentence than the bare n-word would, such as in sentences with composed tenses (18-a), when the n-word is embedded in a prepositional phrase (18-b), when the n-word is in an infinitival construction (18-c), and when the n-word is embedded in subordinate clause (18-d) under a neg-raising verb.

(18) a. J'ai **(pas)** vu **personne**.
I have not seen no one
'I saw no one.'

b. Je parle **(pas)** à **personne**.
I speak not to no one
'I speak to no one.'

c. Je veux **(pas)** voir **personne**.
I want not see no one
'I don't want to see anyone.'

d. Je veux (**pas**) que Jean voie **personne**.
I want **not** that Jean see no one
'I don't want Jean to see anyone.'

## 3.2 Summary

In summary, although de Swart (2010)'s analysis of asymmetric negative concord was designed with invariant systems like Spanish in mind, by virtue of its Optimality Theoretic architecture, it makes clear predictions for the kinds of patterns that should exist in variable systems like Montréal French. In particular, given that, in categorical asymmetric concord languages, NEG FIRST appears as a hard constraint, in variable asymmetric concord languages, it should appear as a soft constraint. Furthermore, taking into account the heavy morphophonological status of MF sentential negation, a straightforward StOT extension of her proposal predicts that variation should be limited to postverbal n-words in 4th position or later in the sentence. These predictions are summarized in Table 7.

| Syntactic Position | Prediction |
|---|---|
| Preverbal | No concord |
| Fragment | Same as 3rd (no concord) |
| 3rd position | Same as fragment (no concord) |
| 4th position | Variable concord |
| 5th position | Same as 4th (variable concord) |
| 6th+ position | Same as 4th (variable concord) |

**Table 7** – Predictions of 'soft' NegFirst analysis for Montréal French

I test these these predictions in the next section using the *Montréal 84* corpus of spoken Montréal French.

# 4 Variable Negative Concord in *Montréal 84*

This section presents a quantitative study of the distribution of negative concord structures in the *Montréal 84* corpus.The *Montréal 84* corpus is composed of sociolinguistic interviews performed in 1984 with 72 speakers of a variety of ages, education levels and professions. From this corpus, we extracted all the occurrences of n-words (*personne* 'no one', *rien* 'nothing', *aucun* 'no', *jamais* 'never', *nulle part* 'nowhere'). From this initial dataset, we excluded clearly idiomatic expressions and structures where the n-words are within the scope of another n-word such as (19); these are instances of what is called (after den Besten (1986)) the *negative spread* construction, which does not allow for co-occurrence with negation.

(19) a. Puis il-y-a **jamais** què: **personne** qui s'est plaint de ma bouffe
Then there is never that no one that refl is complaint of my food
'No one has ever complained about my food.' (126 2114)

b. **personne:** débouchait sur **rien**
No one open up on nothing
'No one ended up with anything.' (85 532)

Additionally, I excluded examples in which the n-words are modified by a maximizing degree adverb such as *absolument* 'absolutely', *presque* 'almost', *pratiquement* 'practically', and *quasiment* 'almost', since modification by this class of adverbs uniformly blocks negative concord in Québec French (Déprez and Martineau, 2004, p.10) and cross-linguistically (Giannakidou, 2006).

(20) a. Nécessairement, il connaît **absolument** rien. (2 293)
'Necessarily, he knows **absolutely** nothing'

b. Oui parce-que: j'ai **presque** jamais été au travail. (91 273)
'Yeah because: I **almost** never went to work.'

c. on les voit **pratiquement** jamais (64 128)
'We **practically** never see them.'

d. mais la j'écoute **quasiment** rien. (131 506)
'but now I listen to **almost** nothing.'

Furthermore, we also excluded expressions in which an n-word co-occurs with negation yielding a double negation interpretation, not a single negation interpretation. Double negation interpretations in the *Montréal 84* corpus are limited to 5 occurrences of the expression *pas pour rien* 'not for nothing' (21), which suggests also that double negation interpretations in Montréal French are largely idiomatized.

(21) Un gars qui parle bien pour moi c'est un gars comme Robert-Charlebois. <hum> Un gars qui: bien Robert-Charlebois a: a déjà sacré comme tout le monde, il doit sacrer encore de toute façon <humhum> mais: un gars qui a: qui prend **pas** des mots: longs comme ça: pour **rien**.

*A guy that talks well for me that's a guy like Robert-Charlebois. <hum> A*

*guy that: well Robert-Charlebois has: has a cursed before like everyone, he must still curse in any case <humhum> but: a guy that has: that does **not** use long words like that: for **nothing**.* (113 606)

Finally, since this paper only addresses the question of the analysis of variable asymmetric concord systems, I limit the investigation to speakers that actually have such systems. In fact, 11 of the 72 speakers never use negative concord structures with *pas*[16]; in other words, these speakers have an invariant grammar along the lines of the Spoken European French grammar in which *NEG≫ NEG FIRST and the ranges of these constraints do not overlap. Thus our final dataset contains 2160 n-words taken from the speech of 61 speakers.

We coded each n-word for its syntactic position (preverbal, postverbal (i.e. following the finite verb/auxiliary) or fragment) and whether it co-occurs with *pas*. Then, within the set of postverbal n-words (n = 1964), I coded for the presence of the NEG FIRST$_L$ constraint, as defined in (15).

## 4.1 Results

The rates of the use of concord structures (i.e. co-occurrence with *pas*) by syntactic position are shown in Table 8.

As expected, there were no occurrences of structures of the form **Personne est pas venu.* Furthermore, when we consider the rate of concord between n-words in postverbal position, we find a significant difference between the use of *pas* with n-words in the 3rd position compared to the 4th, 5th or ≥6th positions ($\chi^2 = 460$; $p < 0.0001$); whereas, we find no significant difference between the rate of negative concord between 4th, 5th and 6th

[16]As mentioned above, although there is social conditioning in the data, the 11 speakers who are invariant in the interviews do not constitute a well-defined socio-economically defined class, according to the stratifications of the corpus.

| Syntactic Postion | Bare n-word | Negative concord | % Concord |
|---|---|---|---|
| Preverbal | 77 | 0 | 0.0% |
| Fragment | 236 | 17 | 6.7 % |
| 3rd position | 1621 | 128 | 7.3% |
| 4th position | 39 | 42 | 51.8% |
| 5th position | 22 | 40 | 64.5% |
| ≥6th position | 27 | 45 | 62.5% |

**Table 8** – Distribution of negative concord structures by syntactic position in *Montréal 84*.

position ($\chi^2 = 2.85$; $p = 0.239$). Since our analysis predicted that the rate of negative concord should be drastically higher with n-words in 4th position and later than in third position (and that we should not find major differences in the use of *pas* with n-words past 4th position), I suggest that the data from *Montréal 84* supports this hypothesis.

Our 'soft' linear NEG FIRST analysis also predicted that fragments and n-words in third position should show the same rate of concord; this was borne out (albeit in a trivial way): there is no significant difference between the rate of negative concord in fragments as in 3rd position ($\chi^2 = 0.118$; $p = 0.73$). However, by virtue of the fact that both fragment and postverbal 3rd position structures with *pas* were predicted to be harmonically bounded, it is therefore surprising to find an (albeit limited) number of concord sentences such as (22) and (23) in the corpus.

(22) Third position negative concord structures.

a. Moi: Je suis **pas aucun programme** anyway
*'Me, I follow no program anyway.'* (27 367)

b. Puis: je connais **pas personne** de parfait.
*'So: I don't know anyone perfect.'* (2 993)

c. J'ai pas de vacances, j'ai **pas rien**.
*'I have no vacation; I have nothing.'* (2 616)

d. mes frères ont **pas jamais** été tu-vois dans le hockey

*'My brothers were never, you know, into hockey'* (54 419)

e. Il-y-a pas de gang.Il y en a **pas nulle part**

*'There aren't any gangs; there aren't any anywhere.'* (4 942)

(23) Fragment negative concord structures.

a. 2. **Pas aucune**. Aucune aucune aucune influence[17].

*'2. None. No no no influence.'* (66 863)

b. . . . de voir: ma mère dans la maison **pas personne** d'autre.

*'. . . to see my mother in the house; no one else.'* (126 831)

c. 5. Qu'est-ce-que tu vas faire avec ça? 2. Bien non, **pas rien**. (rire)

*'5. What are you going to do with that? 2. Well no, nothing.'* (7 126)

d. j'avais bien du stock mais; **pas jamais** ( ).

*'I had a lot of stuff but; never ().'* (90 500)

In summary, although many of the predictions made by the StOT analysis given above for the distribution of negative concord structures in the *Montréal 84* corpus were borne out, we find an unexpected contrast between third position postverbal and fragments contexts on the one hand, and preverbal contexts on the other. These observations are summarized in Table 9.

| Position | Prediction | Observation |
|---|---|---|
| Preverbal | No concord | No concord |
| Fragment | Same as 3rd (**no concord**) | Same as 3rd (**limited concord**) |
| 3rd position | Same as fragment (**no concord**) | Same as fragment (**limited concord**) |
| 4th position | Variable concord | Variable concord |
| 5th position | Same as 4th (variable concord) | Same as 4th (variable concord) |
| ≥6th position | Same as 4th (variable concord) | Same as 4th (variable concord) |

**Table 9** – Negative Concord in *Montréal 84*: Predictions vs Observations

[17]In the final dataset, we also excluded n-words that were part of repetitions, i.e. the last two *aucune*s in this example, since negative concord is uniformly excluded from these contexts.

In the next section, I address the question of what makes third position postverbal n-words and fragment n-words different from preverbal n-words, and I modify the existing StOT analysis to account for the observed patterns of variation.

# 5 A New Probabilistic Analysis of Montréal French

This section modifies the analysis presented in section 3 with an account of the variable use of *pas* in fragments and third postverbal position. Then, with the final constraint-set in place, I show how we can use Boersma (1998); Boersma and Hayes (2001)'s *Gradual Learning Algorithm* (GLA) to assign ranking values to the proposed constraints that will generate the appropriate distribution of concord structures Montréal French.

## 5.1 Structural Neg First

Although the results in the previous section suggest that the use of Neg First has significant potential for explaining both categorical and variable negative concord patterns, our extension of de Swart (2010) did not make quite the right predictions. In addition to predicting categorical exclusion of concord in fragment nwords and third position nwords, closer inspection of the set nwords in third position in *Montréal 84* show that not all utterances where the nword occupies linear third position are created equal. For example, if we restrict our attention to the 1749 occurrences of n-words in third position, we find a significant effect of the presence of (non-)finite clause boundaries in the use of concord structures ($\chi^2 = 146.08; p < 0.0001$). In particular, n-words in third position that, presumably, appear in a lower infinitival or finite clause (i.e. would appear in a lower clause than *pas*) are significantly more likely to appear in a concord structure (49% of cases) than n-words that appear in the same finite clause as *pas* would (6%). This pattern is shown in

Table 10.

| Position of n-word | Bare n-word | Negative concord | % Concord |
|---|---|---|---|
| Upper clause | 1593 | 101 | 6 % |
| Lower clause | 28 | 27 | 49% |

**Table 10** – Negative concord with n-words in third position

In other words, despite the n-word appearing in third position in both cases, examples like (24-a), where there is a (infinitival) clause boundary between where sentential negation is placed and where the direct object n-word appears, are much more frequent than examples like (24-b), where *pas* and *rien* appear in the same clause.

(24) a. Je peux **pas rien** faire face a ça.
I can not nothing do.INF face to that
'I can't do anything faced with that.' (1 770)

b. il-y-a **pas rien** mais ça serait calme
there is not nothing but it be.FUT calm
'there is nothing, but it will be calm' (8 175)

The pattern in Table 10 tells us that our hypothesis that Neg First makes reference to linear order is not fine-grained enough; rather, hierarchical relationships between constituents play an important role in the co-occurrence patterns between sentential negation and postverbal nwords in Montréal French. Therefore, in order to account for differences in the frequencies of utterances like (24-a) and (24-b), I propose that Neg First should be restated with respect to dominance relations between syntactic nodes in a tree, rather than linear order.

(25) Structural Neg First (Neg First$_D$):
Let $>_D$ be the **dominance** relation between syntactic nodes $(\psi, \psi_1, \psi_2 \ldots)$, and let $\langle f, m \rangle$ be a form-meaning pair such that $f$ contains a constituent $\phi$ with a

negative denotation. Then,

$$\textsc{Neg First}_D(\langle f, m\rangle) = |\{\psi : \psi >_D \phi\}|$$

(25), then, would make a distinction between (24-a) and (24-b) because there are presumably more syntactic nodes (for example, infinitival clause structure) dominating *rien* in (24-a) than in (24-b). Additionally, moving to a structural characterization of Neg First can help explain why, in this system, we might get some occurrences of negative concord with nwords in third position. If we assume, following, for example, Pollock (1989), that the tensed verb in French is base-generated in a low syntactic position and raises into a higher tensed position, and that *pas* occupies an adverbial position that is medial between these two positions, then, under $\textsc{Neg First}_D$, fewer nodes dominate the negative element in the concord structure than in the bare structure. So the concord structure would receive fewer violations than the bare structure, and we would no longer predict that *Je vois pas personne* would be harmonically bounded[18]. I highlight that this is only one possible analysis of the syntax of *Je vois (pas) personne*. The point that is pertinent for the paper is simply that the structural relations between *pas* in a concord structure and *personne* in a non-concord structure are different, and so we would like $\textsc{Neg First}_D$ should be sensitive to these differences.

Although this structural revision to Neg First improves on our previous linear version of the constraint, we now need to capture the difference in frequency between n-words that directly follow the finite verb and those that are separated from it by some other constituent. Once again, I suggest that typology can give us some insight into what is driving these patterns. Although the difference in the use of a negative concord structure

[18]Under this account, we would have the same explanation for variation in fragments. Thus, this kind of analysis would have to adopt an analysis of the syntax of fragments in which fragment answers are associated with at least some non-audible syntactic structure (for example, in a theory such as Merchant (2004)).

between *Je vois pas personne* and *J'ai pas vu personne* is a matter of frequency in Montréal French, we see this preferential contrast showing up as a grammaticality contrast in some other languages. For example, Zanuttini (1997) shows that in Piedmontese, an Italian dialect, the postverbal sentential negation marker *nen* cannot co-occur with an nword such as *gnun* 'no one' if the verb is in (what she calls) a 'simple' form (i.e. it does not consist of an auxiliary and a past participle (Zanuttini, 1997, 76) (26-a). However, if the n-word is embedded under a participle (26-b) or within a prepositional phrase (26-c), then negative concord is grammatical[19].

(26) Piedmontese (Zanuttini, 1997, 77)

a. *A veddu **nen gnun**.
   I see not no one
   Intended: 'I don't see anyone'.

b. I l'hai **nen** vist **gnun**.
   I it.have not seen no one
   'I have not seen anyone.'

c. A parla **nen** cun **gnun**.
   he talks not with no one
   'He doesn't talk with anyone.'

Thus, to capture the categorical patterns found in Piedmontese, we need a constraint that makes reference not only to the presence of syntactic structure, but also to particular syntactic domains. In the logic of the OT analysis, then, this would boil down to saying that forms in which the n-word is not in the same unembedded domain as the finite verb

[19]Piedmontese also has a second negation marker *pa* which, according to Zanuttini, is not subject to the same restrictions as *nen*; however, *pa* also has very different syntactic and semantic/pragmatic properties than *nen*, so it would seem that certain meaning-related considerations can override the high rankings of NEG FIRST$_{D_i}$. But I leave a full analysis of the Piedmontese negation system within StOT to future work.

receive an extra violation of NEG FIRST than forms in which the n-word occurs with no embedding next to the finite verb. So, the version of NEG FIRST that is highly ranked in Piedmontese is a more specialized case of the structural NEG FIRST constraint proposed above, one that only looks at a subset of the dominance relations in the tree. I will call this the *domain-relativized* NEG FIRST constraint (and notate it as NEG FIRST$_{D_i}$).

(27) Domain-Relativized NEG FIRST (NEG FIRST$_{D_i}$):

Let $>_{D_i}$ be the dominance relation between syntactic nodes ($\psi, \psi_1, \psi_2 \ldots$), **relativized to the appropriate domain** (i.e. $>_{D_i} \subset >_D$), and let $\langle f, m \rangle$ be a form-meaning pair such that $f$ contains a constituent $\phi$ with a negative denotation. Then,

$$\text{NEG FIRST}_{D_i}(\langle f, m \rangle) = |\{\psi : \psi >_{D_i} \phi\}|$$

In Piedmontese (which also allows optional negative concord with *nen*), NEG FIRST$_{D_i}$ and *NEG would be ranked closely together. The more general NEG FIRST$_D$ would be ranked much lower. In Montréal French, on the other hand, I propose that the three constraints are be ranked very close together such that we could have optional negative concord with n-words in postverbal position (but not with n-words in preverbal position), but still capture the difference in frequency between concord structures with n-words with 'simple' verbs and n-words with some level of embedding[20].

[20]Note that we might need some other explanation for the contrast between *Je vois pas personne* and *J'ai pas vu personne* if we assume, following Abeillé and Godard (1996); Abeillé (2002), that composed tenses in French are associated with a flat structure. However, I leave exploring this possibility to future work.

## 5.2 Structural Neg First and the GLA

Now that we have a full constraint-set, the final step in the analysis is to show that it is possible to assign ranking values to the constraints proposed in this paper (FaithLex, Neg First$_D$, Neg First$_{D_i}$, and *Neg) such that our grammar generates the appropriate distribution of concord and bare forms in spoken Montréal French. In order to show that our StOT grammars describe this language, we need to assign ranking values to the constraints in $\mathcal{C}$ and show that they generate the patterns of variation that we see in the language. In order to do so, I will use the *Gradual Learning Algorithm* (GLA) (Boersma, 1998; Boersma and Hayes, 2001), which is a learner for StOT grammars from categorical or variable data and for which there is an implementation in the Praat system (Boersma and Weenik, 2014). The GLA is given an OT grammar in which all the constraints have the same ranking (as a convention, we set them at 100.00) and learning data which consists of a set of form-meaning pairs[21] with the statistical distribution of the language under study, in this case Montréal French. The learner then assigns ranking values to the constraints in the grammar, modifying its ranking assignment based on the form-meaning pairs it is exposed to. In this study, the GLA was fed with 100 000 observations of negative meanings paired with sentences with n-words in preverbal position (i.e. *Personne est (pas) venu*), non-embedded postverbal position (i.e. *Jean voit (pas) personne.*), and embedded postverbal position (i.e. *Jean parle (pas) à personne.*) according to their distribution in *Montréal 84*. To get an idea of the range of constraint rankings that the GLA will learn on the Montréal French dataset, I ran the simulation 5 times, and the resulting learned grammars are shown in Table 11.

As shown above, Faith Lex does not interact with any of the other constraints that are

[21] There is a bidirectional GLA (Jäger, 2003); however, since the variation that interests us in this paper concerns variation in the form of sentences with negative indefinites, we will only need to look at the unidirectional speaker → hearer system.

| **Grammar** | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| NegFirst$_{D_i}$ | 102.675 | 102.699 | 102.619 | 102.629 | 102.691 |
| *Neg | 102.172 | 102.221 | 102.171 | 102.251 | 102.183 |
| FaithLex | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| NegFirst$_D$ | 97.828 | 97.779 | 97.829 | 97.749 | 97.817 |

**Table 11** – 5 grammars learned by the GLA learner in Praat (Montréal 84 data)

the focus of this paper, so it remains ranked at 100.00 even after exposure to the Montréal 84 data. The other constraints, on the other hand, separate themselves out such that *Neg and Neg First$_{D_i}$ have almost the same ranking, allowing for much variation with embedded postverbal n-words; however, the more general Neg First$_D$ is ranked much lower, only overlapping with *Neg a very small proportion of the time. Thus, we should rarely find concord structures with unembedded postverbal n-words.

Note that if we feed the GLA learner a different dataset, one that is just like Montréal 84 but where there are no occurrences of concord sentences with un-embedded postverbal n-words (i.e. what we might observe in a corpus of spoken Piedmontese), after 100 000 observations the distance between the rankings of *Neg and Neg First$_D$ gets larger and overlap is no longer possible, as shown in Table 12. Thus, I argue that this new analysis can capture not only the typological relationships between Spanish and Montréal French (and Piedmontese) negative concord on the one hand, but also the more subtle connections between Montréal French-style variable systems and the variable systems found in Italian dialects like Piedmontese.

| **Grammar** | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| NegFirst$_{D_i}$ | 107.474 | 107.758 | 107.540 | 107.957 | 107.878 |
| *Neg | 107.070 | 107.164 | 102.171 | 107.527 | 107.354 |
| FaithLex | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| NegFirst$_D$ | 92.930 | 92.836 | 92.964 | 92.473 | 92.646 |

**Table 12** – 5 grammars learned by the GLA learner in Praat (Hypothetical Piedmontese)

# 6 Conclusion

In conclusion, this paper has argued that language variation and change studies have an important role to play in the construction, elaboration and testing of formal syntactic theories. As an illustration of this proposal, I showed how we can test a probabilistic extension of de Swart (2010)'s analysis of obligatory asymmetric negative concord using *Montréal 84* corpus of spoken Montréal French. I argued that the main lines of the predictions of de Swart's analysis were born out; however, looking at variation data showed us that the syntactic patterns associated with negative concord can require a more subtle definition of the NEG FIRST family of constraints. More specifically, I argued that the NEG FIRST constraint family should be conceived of as taking into account hierarchical structure, rather than simple linear order, and that members of this family can differ with respect to the syntactic domains to which they are sensitive.

Of course, the choice of testing de Swart's proposal (rather than other proposals) was not arbitrary: in particular, in order to test the quantitative predictions of a formal analysis that was formulated to account for categorical syntactic patterns, we need for this analysis to be set within a syntactic framework that has an (appropriate) stochastic generalization. Since de Swart's proposal was framed within OT, and we have extensions of this framework such as Boersma's that are equipped to model syntactic variation, her analysis for Spanish made clear predictions for Montréal French. However, there are many other analyses of the forms and interpretations of asymmetric negative concord sentences that are set in frameworks that do not permit variation (Penka and Zeijlstra, 2010, for a recent overview). Although a fair amount of progress has been made in extending other mainstream syntactic frameworks, such as Chomsky (1995)'s *Minimalist* framework, to deal with certain kinds of sociolinguistically conditioned variation such as variable morphological agreement (Adger and Smith, 2010; Comeau, 2011; Adger, 2014, among others), these proposals do not so

clearly extend to phenomena at the syntax-semantics interface such as negative concord. I therefore leave the exploration of the probabilistic analysis of variable negative concord in other frameworks as an open research area.

# References


Abeillé, A. and Godard, D. (1996). La complémentation des auxiliaires en français. *Langages*, 122:32–61.

Abeillé, A. and Godard, D. (1997). The syntax of French negative adverbs. In Forget, D., editor, *Negation and Polarity: syntax and semantics*, pages 1–27. John Benjamins, Amsterdam.

Abeillé, A. and Godard, D. (2002). The syntactic structure of French auxiliaries. *Language*, 78:404–452.

Adger, D. (2014). Variability and grammatical architecture. In Picallo, C., editor, *Linguistic Variation in the Minimalist Framework*. Oxford University Press.

Adger, D. and Smith, J. (2010). Variation in agreement: A lexical feature-based approach. *Lingua*, 120:1109–1134.

Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219.

Blutner, R. (2000). Some aspects of optimality in natural language semantics. *Journal of Semantics*, 17:189–216.

Blutner, R., Hendriks, P., and de Hoop, H. (2003). A new hypothesis on compositionality. *Proceedings of ICCS 2003*, page 53?57.

Boersma, P. (1998). *Functional Phonology. Formalizing the interaction between articulatory and perceptual drives*. PhD thesis, University of Amsterdam.

Boersma, P. and Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32:45–86.

Boersma, P. and Weenik, D. (2014). Praat: Doing phonetics by computer. Version 5.3.80, retrieved 29 June 2014 from http://www.praat.org/.

Borsley, R. and Jones, B. (2005). *Welsh Negation and Grammatical Theory*. University of Wales Press, Cardiff.

Bresnan, J. (2000). Optimal syntax. In Joost Dekkers, F. v. d. L. and van de Weijer, J., editors, *Optimality Theory: Phonology, Syntax and Acquisition*, pages 334–385. Oxford University Press, Oxford.

Bresnan, J. (2007). A few lessons from typology. *Linguistic Typology*, 11:297–306.

Bresnan, J., Cueni, A., Nikitina, T., and Baayen, H. (2007). Predicting the dative alternation. In Boume, G., Kraemer, I., and Zwarts, J., editors, *Cognitive Foundations of Interpretation*, pages 69–94. Royal Netherlands Academy of Science, Amsterdam.

Bresnan, J., Dingare, S., and Manning, C. (2001). Soft constraints mirror hard constraints : Voice and person in English and Lummi. In Butt, M. and King, T., editors, *Proceedings of the LFG01 Conference*, Hong Kong.

Burnett, H., Tremblay, M., and Blondeau, H. (2015). The variable grammar of Montréal French negative concord. In Fisher, S., editor, *Penn working papers in linguistics*, volume 21. University of Pennsylvania.

Chambers, J. (2004). Dynamic typology and vernacular universals. In Kortmann, B.,

editor, *Dialectology meets typology: Dialect grammar from a cross-linguistic perspective*, pages 127–145. Mouton de Gruyter, Berlin/New York.

Chomsky, N. (1957). *Syntactic Structures*. MIT Press, Cambridge.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge.

Comeau, P. (2011). *A window on the past, a move towards the future: Sociolinguistic and formal perspectives on variation in Acadian French*. PhD thesis, York University.

Corblin, F. and Tovena, L. (2003). L'expression de la négation dans les langues romanes. In Godard, D., editor, *Les langues romanes: problèmes de la phrase simple*, pages 279–242. CNRS Publications, Paris.

Daoust-Blais, D. (1975). *L'influence de la négation sur certains indéfinis en français québécois*. PhD thesis, Université du Québec à Montréal.

de Hoop, H. and de Swart, H. (2000). Adjunct clauses in optimality theory. *Revista di Linguistica/Italian Journal of Linguistics*, 12:107–127.

de Swart, H. (2010). *Expression and interpretation of negation: An OT Typology*. Springer, Dordrecht.

de Swart, H. and Sag, I. (2002). Negation and negative concord in Romance. *Linguistics & Philosophy*, 25:373–415.

den Besten, H. (1986). Double negation and the genesis of Afrikaans. In Muysken, P. and Smith, N., editors, *Substrata versus universals in Creole languages*, pages 185–230. John Benjamins, Amsterdam.

Déprez, V. (2002). Concordance négative, syntaxe des mots-N et variation dialectale. *Cahiers de linguistique française*, 25:97–117.

Déprez, V. and Martineau, F. (2004). Pour use analyse microparametrique de la concordance négative. In Corblin, F., Ferrando, S., and Kupferman, L., editors, *Indéfinis et prédications*. Presses Universitaires Paris-Sorbonne.

Deshaies, D. and Laforge, E. (1981). Le futur simple et le futur proche dans le français parlé dans la ville de québec. *Langues et Linguistique*, 7:21–37.

Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12:453–476.

Giannakidou, A. (2006). N-words and negative concord. In Everaert, M., editor, *Blackwell Companion to Syntax (Volume 3)*, pages 327–391. Blackwell.

Givón, T. (1979). *On understanding grammar*. Academic, New York.

Godfrey, J., Holliman, E., and McDaniel, J. (1992). Switchboard : Telephone speech corpus for research and development. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pages 517–520, San Francisco.

Grimshaw, J. (1997). Projection, heads and optimality. *Linguistic Inquiry*, 28:373–422.

Hawkins, J. (2004). *Efficiency and complexity in grammars*. Oxford University Press, Oxford.

Hendriks, P. and de Hoop, H. (2001). Optimality theoretic semantics. *Linguistics and Philosophy*, 24:1–32.

Horn, L. (1989). *A natural history of negation*. CSLI Publications.

Jäger, G. (2002). Some notes on the formal properties of bidirectional Optimality Theory. *Journal of Logic, language and information*, 11:427–451.

Jäger, G. (2003). Learning constraint sub-hierarchies: the bidirectional gradual learning algorithm. In Blutner, R. and Zeevat, H., editors, *Optimality Theory and Pragmatics*, pages 217–242. Palgrave McMillan.

Jelinek, E. and Demers, R. (1983). The agent hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics*, 49:167–185.

Jespersen, O. (1917). *Negation in English and other languages*. Horst, Copenhagen.

Jespersen, O. (1933). *Essentials of English grammar*. Allen & Unwin, London.

Keenan, E. and Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8:63–99.

Keenan, E. and Hawkins, S. (1987). The psychological validity of the accessibility hierarchy. In Keenan, E., editor, *Universal grammar: 15 essays*. Croom Helm, London.

Keenan, E. and Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9:253–326.

Keenan, E. and Westerstahl, D. (1997). Generalized quantifiers in linguistics and logic. In van Benthem, J. and ter Meulen, A., editors, *Handbook of logic and language*, pages 837–893. Elsevier, Amsterdam.

Krifka, M. (2007). Approximate interpretation of number words: A case for strategic communication. In Bouma, G., Kräer, I., and Zwarts, J., editors, *Creative Foundations of Interpretation*, pages 111–126. Koninklijke Nederlandse Akademie van Wetenschapen, Amsterdam.

Kroch, A. (2000). Syntactic change. In Baltin, M. and Collins, C., editors, *The handbook of contemporary syntactic theory*, pages 699–729. Blackwell, Oxford.

Labelle, M. (2010). Negative words and negation in french. In Larrivée, P. and Ingham, R., editors, *The Evolution of Negation: Beyond the Jespersen Cycle*. De Gruyter Mouton.

Labov, W. (1963). The social motivation of a sound change. *Word*, 19:273–309.

Labov, W. (1966). *The social stratification of English in New York City*. Center for Applied Linguistics, Arlington.

Laka, I. (1990). *Negation in Syntax*. PhD thesis, MIT.

Larrivée, P. (2014). The continuity of the vernacular. In Mosegaard Hansen, M.-B. and Visconti, J., editors, *The Diachrony of Negation*, pages 253–256. John Benjamins, Amsterdam.

Lemieux, M. (1985). Pas rien. In Lemieux, M. and Cedegren, H., editors, *Les tendances dynamiques du français parlé à Montréal*, pages 91–140. Office de la langue française.

May, R. (1985). *Logical Form*. MIT Press, Cambridge.

Merchant, J. (2004). Fragments and ellipsis. *Linguistics and Philosophy*, 27:661–738.

Montague, R. (1970). English as a formal language. In Visentini, B., editor, *Linguaggi nella societa e nella tecnica*, pages 189–224. Edizioni di Communita, Milan.

Muller, C. (1991). *La négation en français*. Droz, Geneva.

Penka, D. and Zeijlstra, H. (2010). Negation and polarity: an introduction. *Natural language and linguistic theory*, 28:771–786.

Peters, S. and Westerstahl, D. (2006). *Quantifiers in Language and Logic*. Oxford University Press, Oxford.

Pollock, J.-Y. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, 20:365–424.

Poplack, S. and Dion, N. (2009). Prescription vs. praxis: The evolution of future temporal reference in French. *Language*, 85:557–587.

Poplack, S. and Turpin, D. (1999). Does the futur have a future in (canadian) french? *Probus*, 11:133–64.

Prince, A. and Smolensky, P. (1993). Optimality theory: Constraint interaction in generative grammar. *Rutgers University Center for Cognitive Science Technical Report*, 2.

Rosenbach, A. (2002). *Genitive Variation in English: Conceptual Factors in Synchronic and Diachronic Studies*. Mouton de Gruyter, Berlin.

Rosenbach, A. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language*, 81:613–644.

Sankoff, G. and Vincent, D. (1977). L'emploi productif de ne dans le français parlé à montréal. *Le Français Moderne*, 45:243–256.

Sciullo, A.-M. D. and Tremblay, M. (1996). Configurations et interprétations: les morphèmes de la négation. *Recherches linguistiques de Vincennes*, 25:27–52.

Sorace, A. and Keller, F. (2005). Gradience in linguistic data. *Lingua*, 115:1497–1524.

Stabler, E. (2013). Two models of minimalist, incremental syntactic analysis. *Trends in Cognitive science*, 5:611–633.

Tagliamonte, S. (2011). Variation as a window on universals. In Siemund, P., editor, *Linguistic universals and language variation*. Mouton de Gruyter, Berlin/New York.

Tagliamonte, S. (2014). A comparative sociolinguistic analysis of the dative alternation. In Torres-Cacoullos, R., Dion, N., and Lapierre, A., editors, *Linguistic Variation: Confronting fact and theory*. Routledge, London and New York (to appear).

Thibault, P. and Vincent, D. (1990). *Un corpus de français parlé: Montréal 84*. Université Laval, Québec.

Thullier, J. (2012). *Contraintes préférentielles et ordre des mots en français*. PhD thesis, Université Paris-Diderot.

Weinreich, U., Labov, W., and Herzog, M. (1968). *Empirical foundations for a theory of language change*. University of Texas Press.

Zanuttini, R. (1997). *Negation and clausal structure. A Comparative Study of Romance languages*. Oxford University Press, Oxford.

Zeevat, H. (2001). The asymmetry of optimality theoretic syntax and semantics. *Journal of Semantics*, 17:243–262.