Arbeitspapiere des SFB 340
Bericht Nr. 132, September 1999

A LEXICALIST COLLOCATION ANALYSIS OF SENTENTIAL NEGATION AND NEGATIVE CONCORD IN FRENCH

FRANK RICHTER, MANFRED SAILER

Contents

Many thanks to our informants Olivia Gerig, Agnès Gobbo, Christian Richard and Raphaëlle Rivet, to Michaela Banzhaf and Anne Sailer for additional help with the data, to Christian Richard for help with English, to Mike Calcagno, Francis Corblin and Adam Przepiórkowski for comments and discussion, and to our editor, Valia Kordoni, for her patience and moral support.

ISSN 0947-6954/99

1. Introduction

The study of negation in natural language is attractive to linguists for various reasons. First, it seems to be easily delimited in scope, as natural languages tend to provide an identifiable small set of elements and constructions that are intuitively considered as expressing negation. Second, most formal languages which are likely candidates for semantic representation languages have a straightforward notion of negation[1]. These two facts taken together invite researchers on both sides of the syntax-semantics interface to substantiate the form-meaning correlation of language. But in order to attribute a general relevance for the investigation of language to the study of an empirical area as restricted as negation, it is usually deemed necessary that the empirical domain in question promise at least some potential for a general explanation of the structure of natural languages.

We disagree with that view, much in the spirit of Baker (1991), and suggest that the distribution of negation in natural languages should be reconstructed from particular idiosyncratic properties of single lexical items. A lexical item can only appear in contexts which fulfill certain conditions. These two premises lead us to derive the supposedly general properties of negative constructions from the interaction of the semantic contribution and the distributional restrictions of lexical elements. It is crucial to acknowledge that both factors are extremely idiosyncratic and unpredictable. Data are interesting just by themselves. The level of description at which we expect valid generalizations over linguistic structures is, thus, much more abstract than phrase structure and phrase structural relationships. With the architecture of a collocation theory which expresses the mutual restrictions between syntactic structure, semantic representations, and idiosyncratic lexical requirements, we give a first tentative answer as to where we believe they should be sought. In the present study, we investigate the negation system of French. French is famous for its high number of options for expressing negation, which synchronically preserve differing — and partly contradictory — historical stages. Making the versatility of the French negation system even more intriguing and challenging, a distinction between written and spoken language cuts across the negative constructions. Surprisingly, though, French was evoked as the principal example in Pollock (1989), which is a proposal aiming at a phrase structural explanation of negative constructions. In that paper,

[1] Note that this is the crucial difference between negation and questions (Groenendijk and Stokhof 1997).

the full range of data has not been taken into account. More importantly, it is not clear how a broader range of negation data could be reconciled with that theory.
We will cast our analysis of a broad range of French data in HPSG, because its explicit formal foundations force and enable us to be precise about the principles that are needed to describe the empirical facts. Syntactically, we rely on the results of major HPSG contributions that investigate the syntactic properties of lexical elements which are involved in negative constructions (Kim and Sag 1996; Recourcé 1996; Abeillé and Godard 1997). These syntactic analyses will be amended by a collocation theory that refers to logical forms in a higher order, type-theoretic semantic representation language. Richter and Sailer (1999) present both the semantic framework and the theory of collocations which we are going to adopt. They provide an equivalent in HPSG to the semantic representation language used in the Government and Binding framework of linguistics, an integration of a general collocation theory with HPSG, and an application of these two analytical tools to an analysis of negative concord in Polish. Combining the syntactic work with the collocation theory, we will provide an account of French negation that captures both the observable sub-regularities of negative constructions and their puzzling idiosyncrasies. The emerging synthesis of the analyses will support the decidedly lexicalist perspective of the powerful collocation theory of Richter and Sailer (1999) with empirical data from French.

In section 2, we discuss the main data from Standard French. Section 3 sketches formal aspects of the grammar architecture, and introduces the semantic representation language and the collocation theory. They are then applied to the Standard French data in section 4. In section 5 we present more data from Standard French that strengthen the lexicalist account, before we briefly turn to data from other varieties of French, Canadian French and Haitian Creole, in section 6. A general comparison of our analysis to work in other frameworks and in HPSG is given in section 7. Finally, in section 8, we point out a number of open questions. The appendix contains the formalization of the theory presented here and in Richter and Sailer (1999).

2. The French Data

We start with an investigation of the behavior of the words that are most intimately connected to sentential negation. We will consider the distribution of the pre-verbal negative particle *ne*, glossed as NE, the adverbial negative marker *pas* (*not*), glossed as NM, and so-called n-words such as *personne* (*nobody*), *rien* (*nothing*), and *aucun N* (*no N*).[2] In the context of n-words, we will have to study in which constructions they trigger negative concord and where they do not. With Ladusaw (1995), we perceive negative concord (NC) as a "morphological indication at multiple positions in a clause of the fact that the clause is to be interpreted as negated," since this definition is broad enough to cover the cross-linguistic diversity of NC. The focus of our inquiry lies on the following two questions: (a), which semantic contribution can be attributed to which of the above mentioned words; and (b), how can their particular distributional properties best be characterized.

2.1. **Pre-verbal *ne*.** In contemporary French, three types of occurrences of the pre-verbal negative particle, *ne*, can be distinguished (Recourcé 1995). First, it may occur as the only lexical exponent of sentential negation in a sentence. Second, it may be used together with the adverbial negative marker, *pas*, or with n-words, in sentences that exhibit sentential negation. Third, there is an expletive use of *ne* which does not coincide with negation. We refer to these instances of *ne* as *negative*, *co-negative* and *expletive ne*. Negative *ne* can be considered archaic and is mostly restricted to a number of fixed constructions. Although co-negative *ne* is productive, it tends to disappear from spoken language. It is, however, obligatory in formal registers for expressing sentential negation with *pas* or an n-word.[3] Expletive *ne* is bound to a certain group of constructions which we shall compare to traditional polarity contexts.

We will, in turn, investigate the three types of *ne*, and propose a preliminary classification of the data. For the syntactic properties of *ne*, we build on the HPSG analysis of Recourcé (1995) and Recourcé (1996). They show that, independent of the distinctions we have just drawn, *ne* leads a syntactic double life as a pre-verbal affix on finite verbs and as a syntactic marker in infinitival constructions. In both cases, *ne* contributes a MARKING value, *NE*, to the syntactic verb projection.

[2]The term n-word is attributed to Laka (1990). We use the notion n-word merely to be able to refer to the indicated class of words.
[3]See Sturm (1981) for a corpus-based investigation to the present day distribution of co-negative *ne*.

2.1.1. *Negative ne.* *Ne* may express sentential negation only in a number of frozen constructions.[4]

(1) a. * Jean ne vient.
Jean NE comes
b. Jean ne peut venir.
Jean NE can come 'Jean can't come.'
c. Paul semble ne savoir que faire.
Paul seems NE know what to-do
'Paul doesn't seem to know what to do.'
d. Ne bougez de là!
NE move from here. 'Don't move!'

In (1b) and (1c), negative *ne* appears with particular verbs, namely (certain uses of) *pouvoir* (*can*), *savoir* (*know*), *oser* (*dare*), *cesser* (*stop*), but also *bouger* (*move*). As our examples show, the verbs that allow negative *ne* cannot be characterized by any semantic or syntactic properties as a plausible natural class, but appear to be quite heterogeneous. This indicates that the availability of their combination with negative *ne* should be regarded an idiosyncratic property of these verbs rather than a property of *ne*.[5] Hereafter, we refer to verbs that can combine with negative *ne* as *ne*-verbs.

There is one more argument for putting the burden of the idiosyncrasy on the verb rather than on *ne*. In infinitival constructions as (1c), an adverb may appear between *ne* and the verb. This adverb, then, is obligatorily interpreted outside the scope of the negation:[6]

(2) a. Paul semble ne vraiment savoir que faire.
Paul seems NE really know what to do
'It seems that Paul really doesn't know what to do.'
b. Paul prétend ne fréquemment savoir que faire.
'Paul pretends that he often doesn't know what to do.'
$ 'Paul pretends that he doesn't often know what to do.'

[4]Our examples of negative *ne* are mainly taken from Grevisse and Goosse (1993, §975) and Jones (1996, p. 362f.). They list more contexts for negative *ne* than we discuss.
[5]Rowlett (1998) comes to the same conclusion.
[6]The symbol '$' indicates that a given reading is not available.

We take the data in (2) as an indication that, in (1), it is not *ne* that contributes semantic negation, but the main verb.[7] We assume a separate lexical entry for *ne*-verbs which does not differ in the phonological form from its non-*ne*-variant, but introduces semantic negation in the CONTENT value. Moreover, it specifies the MARKING value as *NE*. Due to the syntactic properties of *ne*, we cannot treat negative *ne* as an element directly selected by *ne*-verbs. To account for the obligatory co-occurrence of verbs that are licensed by that second lexical entry with *ne*, we distinguish two cases. If the verb is finite, the affix, *ne*, is a member of its set of pronominal affixes (Recourcé 1996, p. 70). If the verb is non-finite, a collocation requirement of the verb guarantees that it be dominated by a *head-marker-phrase* whose MARKING value is of sort *NE*.

It might be questioned whether such a syntactic collocation restriction is appropriate. Further evidence that this analysis is on the right track comes from double negation facts. Negative *ne* can co-occur with n-words to, marginally, express a double negation reading; however, it can never co-occur with *pas* under a double negation reading. This is indicated by the inappropriateness ('#') of (3b) in the context provided in (3):

(3) A: J'avais l'impression que Pierre ne savait que faire dans cette classe.
'I had the impression that Pierre didn't know what to do in that class.'

a. B: Ah non, ce n'est pas vrai! PERSONNE ne savait que
No, that's not right. Nobody NE knew what
faire, tous s'amusaient très bien.
to-do, they were all enjoying themselves.
' . . . nobody did not know what to do, . . . '

b. # B: Ah non, ce n'est pas vrai! Il ne savait PAS que faire.
No, that's not right. he NE knew NM what to-do.
' . . . He didn't know what to do.'
$ ' . . . He did not not know what to do.'

---

[7] If *pas* is present, the reading of the sentences depends on the relative linear order of *pas* and the adverbial.

(i) Paul prétend (ne) fréquemment pas savoir quoi faire.
'Paul pretends that he often doesn't know what to do.'

(ii) Paul prétend (ne) pas fréquemment savoir quoi faire.
'Paul pretends that he does not often know what to do.'

We can account for that restriction with appropriate selection properties of the adverb *pas*. With its MOD attribute, it selects a verb that is MARKING *unmarked*.[8] *Ne*-verbs cannot co-occur with *pas*, because they bear the MARKING value *NE*, which is a syntactic reflex of their inherently negative semantic form. They require the syntactic presence of a semantically vacuous *ne* as an overt collocational reflex of negation. Double negation readings of *ne*-verbs with other negative elements are possible, since they are not prohibited by these elements. Also, *ne* and *pas* may co-occur in the absence of a *ne*-verb.

Apart from the lexically determined occurrences, negative *ne* is also attested in some fixed syntactic constructions, as exemplified in (4). The crucial difference to the data in (1) is that the set of verbs to which negative *ne* attaches in these syntactic environments is not restricted.

(4) a. Si je n'avais pris la voiture, l'accident ne serait pas
if I NE had taken the car, the accident NE was NM
arrivé.
happened
'If I hadn't taken the car, the accident wouldn't have happened.'

b. Il y a longtemps que je ne l'ai vu.
it there has long time that I NE him have seen
'I haven't seen him in a long time.'

We see two options to analyze these constructions. Either we stipulate the existence of a *ne* that, like in older stages of French, contributes semantic negation; or we try to assimilate these cases to expletive *ne*, and locate the semantic negation in a different element of the sentence, *si* in (4a), and *il y a longtemps que* in (4b).[9]

If we choose the first option, we must restrict the occurrences of this negative *ne*, ne^-, to the relevant syntactic constructions. Let us spell

[8] As noted below, we follow the syntactic treatment of *pas* of Abeillé and Godard (1997) and Kim and Sag (1996). This means that *pas* can be realized as either a complement or an adverb. For *pas* as complement, we will ensure that its SYNSEM value can only be added to the list of complements of unmarked finite verbs.

[9] Adam Przepiórkowski (p.c.) pointed out to us that a constructional approach provides a further solution. As far as we can see, a constructional analysis would enforce the same properties that we attribute to the lexical contribution of *si*. It seems that the negation would prove to be part of the idiosyncratic meaning of the construction, which would force us to introduce a type of *ne* which would not be inherently negative. In effect, the constructional approach would not differ from our second option with respect to the nature of the particle, *ne*.

out some of the necessary restrictions for sentence (4a). Among other environments which would have to be made explicit, ne^{-} can only occur inside a *si* (*if*) clause. That *si* clause must, in addition, be embedded under a negated matrix clause[10] and may not contain *pas* or an n-word:

(5) a. * Si je n'avais pris la voiture, j'aurais été en retard.
if I NE had taken the car, I would-have been late
'If I hadn't taken the car, I would have been late.'

b. Si je n'avais pas pris la voiture, ...
if I NE had NM taken the car,
'If I hadn't taken the car, ... '
$ 'If I hadn't not taken the car, ... '

c. Si personne n'avait pris la voiture, ...
if nobody NE had taken the car,
'If nobody had taken the car, ... '
$ 'If nobody hadn't taken the car, ... '

The first option has the advantage of directly reflecting the idea that previous historical stages are still conserved in part of the French negation system. There would be a truly negative *ne* in modern French that obeyed strict collocation requirements. On the other hand, this analysis gives rise to an incoherent picture for negative *ne*, as we had dismissed the idea of a truly negative *ne* in the case of *ne*-verbs above.

Choosing the second option, we would attribute all particularities of the construction to lexical properties of *si*. *Si*, then, semantically introduces not only an implication but also semantic negation ($\lambda P \lambda Q[\neg P \rightarrow Q]$). It selects a *NE*-marked finite clause via both, its COMPS and its MOD attribute.[11] The clause on the COMPS list, however, may not already contain a semantic negation[12].

As it stands, there is an apparent problem with the second option. If the above-given semantic contribution of *si* is correct, we would predict an obligatory wide scope of the negation over the *si* clause. This prediction is not borne out. In (6), the quantifier in subject position may take scope over the negation.

---

[10] According to some of our informants, this restricion is not as strict as the others.
[11] For our purposes, it is immaterial whether a single SUBCAT list or a combined SUBJ/COMPS system are responsible for the syntactic selection.
[12] The relation `contains-event-negation`, defined in the appendix (90), can be used to express this restriction formally.

(6) Si la plupart des étudiants ne prenaient la voiture, l'air à Tübingen ne serait pas tellement pollué.
'If most of the students did not use the car, the air in Tübingen wouldn't be that polluted.'
'If it weren't the case that most students take the car, ... '

The scope facts in (6) can be reduced to a more basic problem of sub-lexical semantics, i.e., to the observation that some modifiers can intervene into parts of the semantic contribution of a lexical item.[13] Notice that the suggested sub-lexical solution to the problem of the scope of negation in *si* clauses is compatible with our explanation of the scope facts of adverbials relative to *ne*-verbs as illustrated in (2). With *ne*-verbs, the wide-scope reading of the negation is not available at all, even though it is available if *pas* precedes the adverb (see footnote 7), and we have interpreted the restrictions on possible scopings of adverbials as evidence that *ne*-verbs express negation lexically. In (6), on the other hand, both readings are available.

Lacking an empirical criterion to choose between assigning the semantic contribution of negation to *ne* or other construction-specific lexical elements in (4), we choose the second option, because it will allow us to postulate a single lexical entry for all types of *ne*. All occurrence restrictions of the negative *ne* have been attributed to elements other than *ne*.

2.1.2. *Co-negative ne.* Co-negative *ne* arises whenever *ne* co-occurs with an element that expresses sentential negation. As has been mentioned earlier, *ne* may occur in sentences that contain an n-word or *pas*.[14] These sentences receive a single negation interpretation.[15]

(7) a. Jean ne vient *(pas/ jamais).
Jean NE comes (NM/ never)
'Jean doesn't come/ never comes.'

[13]This very problem is part of the motivation for underspecification-based approaches to semantics, as pursued by Copestake, Flickinger, Malouf, Riehemann, and Sag (1995) and Richter and Sailer (1997). We hope to be able to provide a non-underspecified account for it similar to the one we propose in Richter and Sailer (1999) for scope ambiguities.
[14]We will ignore the *ne ... que* (*only*) construction as well as *ni ... ni* (*neither ... nor*) coordination.
[15]Except for the marginal cases of n-words with lexically *ne*-marked verbs mentioned in (3).

b. Jean affirme de ne *(pas) avoir compris.
Jean affirms to NE (NM) have understood
'Jean affirms that he hasn't understood.'

Although the co-negative *ne* is the most frequent, in spoken language *ne* is optional in sentences like (7). In (8), *ne* has been dropped, and the sentences are still grammatical with the same readings.[16]

(8) a. Jean vient pas/jamais.
b. Jean affirme de pas avoir compris.

The optionality of *ne* is best explained if we assume that the marker (or affix), *ne*, makes no semantic contribution in its co-negative use. In this respect, negative and co-negative *ne* behave the same. The real effect of *ne* consists in introducing a collocation restriction. Co-negative *ne* requires the presence of a negation that has scope over the event or proposition expressed in the clause in which *ne* occurs.[17] Note that the collocation restriction of co-negative *ne* is compatible with the distribution of negative *ne*. Therefore, we do not need a separate lexical entry.

2.1.3. *Expletive ne.* Next, we turn to expletive *ne*. Our aim is to characterize expletive *ne* as a non-negative use of *ne*. If we achieve this goal, co-negative and expletive *ne* can be considered instances of a uniform polarity use of *ne*.

In its expletive variant, *ne* need not occur with a clausemate *pas* or n-word. Expletive *ne* also appears in comparative constructions, with certain negative complementizers and in the complement clauses of some verbs.[18]

[16] Dropping co-negative *ne* in infinitivals is allowed in formal registers of French as well.

[17] In the terminology of Przepiórkowski and Kupść (1999), *ne* is licensed by both eventuality and non-eventuality negation. (i) is an example of co-negative *ne*, (ii) an instance of negative *ne* in non-eventuality contexts.

(i) Tu ne peut pas m'aider?
you NE can NM me help
'Can't you help me?'

(ii) Que ne ferait-il pour réussir?
what NE would-make-he to succeed
'What wouldn't he do to succeed?' (Jones 1996, p. 363)

Due to the restrictions of the semantic representation language we use in the present paper, we only deal with eventuality negation.

[18] Wilmet (1997, p. 512f.) points out that historically as well as synchronically, there are noticable differences between expletive *ne* in comparatives and other syntactic

(9) a. Il est plus riche qu'il ne l'était.
he is more rich than that he it was
'He is richer than he used to be.'
b. Je ne doute pas qu'elle n'arrive$_{\mathrm{subj}}$.
I NE doubt NM that she NE arrives
'I have no doubt that she comes.'
c. Je crains qu'il ne vienne$_{\mathrm{subj}}$.
I fear that he NE comes 'I'm afraid he comes.'
d. (il réussira), à moins qu' il ne fasse$_{\mathrm{subj}}$ une erreur stupide.
(he will succeed), unless he NE makes a mistake stupide
'... unless he makes a stupid mistake.'

Again, parallel to the facts we have seen for co-negative *ne*, the presence of the pre-verbal expletive *ne* is not obligatory in spoken language. Thus, the *ne*-less variants of the sentences in (9) are all grammatical.

(10) a. Il est plus riche qu'il l'était.
b. Je ne doute pas qu'elle arrive.
c. Je crains qu'elle vienne.
d. Il réussira, à moins qu'il fasse une erreure stupide.

Expletive *ne* can best be accounted for if we posit a class of complementizers in French that select an (optionally) *ne*-marked clause that — with the exception of the comparative construction discussed in footnote 18 — is in subjunctive mood. These complementizers must be able to fulfill the licensing conditions for *ne*. It might be possible to formulate their semantic contribution uniformly in such a way that they introduce a semantic negation that scopes over the *NE*-marked clause. This solution is plausible in the case of the comparative complementizer or complementizers

environments. The adverbial negative marker, *pas*, seems marginally possible in comparative constructions. Following most grammars and the intuitions of our informants, we include comparatives as contexts for expletive *ne*, and consider the non-negative interpretation of examples as (i) and (ii) synchronically unavailable.

(i) Il faut avoir l'esprit plus libre que je ne l'ai pas.
it needs have the spirit more free than I NE it have NM
'One must have a freer spirit than I do.' (Racine, after Wilmet (1997, p. 513))

(ii) Il est plus instruit que je suis pas.
he is more instructed than I am NM
'He is better instructed than I am.' (Sturm 1981, p. 24)

like *à moins que* and *sans que*, and it seems to be conceivable for others as well.

We have now shifted the burden of determining when expletive *ne* may appear to the occurrence of some complementizer. The crucial question is now whether the occurrence of the complementizer is determined by syntactic selection alone or by some semantic properties of the surrounding context. To sketch the interaction of selectional and collocation properties of lexical elements, we consider a *NE*-marked clause which occurs as the complement of *douter* (*doubt*):

In the non-negated case, *douter* does not allow a complement clause with expletive *ne*. This indicates that *douter* does not select the special *ne*-licensing complementizer.

(11) Je doute qu'il (*ne) vienne.
I doubt that he (NE) comes
'I doubt he is coming.'

There is, however, a negative polarity version of *douter*, *douter*$_2$. This verb has collocation restrictions that only allow it to appear in the scope of a clausemate negation or interrogative operator. In addition, *douter*$_2$ differs from its positve counterpart in that it selects a complement clause which is introduced by the *ne*-licensing complementizer.[19]

(12) a. Xavier ne doutait pas qu'il ne fît semblant de lire.
Xavier NE doubted NM that he NE made seeming to read
'Xavier didn't doubt that he was pretending to read.'
b. Nierez-vous/ Douterez-vous que Canova et Rossini ne
will-deny you/ will-doubt you that Canova and Rossini NE
soient de grands artistes?
are some great artists
'Will you deny that Canova and Rossini are great artists?'

We have been able to provide an account for pre-verbal *ne* which is able to homogenize negative, co-negative and expletive *ne*. They are licensed by one lexical entry whose most important feature is a collocation requirement on logical forms that must be satisfied within a certain syntactic environment. A final argument for assimilating the three types of *ne* comes from the fact that it is often unclear whether an occurrence of *ne* is an instance of negative or expletive *ne*. As an example, consider

[19]The examples are taken from Grevisse and Goosse (1993, §983c).

the use of *ne* in subjunctive adverbial clauses introduced by *que* (*that*), if synonymous to *sans que* (*without*), as in (13):

(13) Il ne fait point de voyage qu'il ne lui arrive quelque
he NE makes not-at-all any travel that it NE him happens some
accident.
accident
'He doesn't make a journey without having an accident.'

Grevisse and Goosse (1993, §974,5.) interpret *ne* in (13) as an instance of negative *ne*. The synonymous sentence with *sans que* instead of *que*, on the other hand, they classify as an instance of expletive *ne*. Under our approach, it is, in both cases, the complementizer that contributes the semantic negation. The difference between *sans que* and *que* in (13) consists in an additional condition of the latter that the *que* clause may only modify a *NE*-marked matrix clause. For *sans que*, a *NE* value of the matrix clause is optional.

2.2. **The adverbial negative marker *pas*.** As mentioned above and illustrated in (7) and (8), French possesses an adverbial negative marker, *pas*, which can be present — either in combination with *ne* or without *ne* — in sentences that express sentential negation. It can be demonstrated that *pas* contributes semantic negation:[20] (a), it can productively occur as the only conceivable lexical exponent of negation in a negated sentence, and, (b), two occurrences of *pas* in one sentence result in a double negation reading.

(14) a. Jean (ne) vient pas.
Jean (NE) comes NM 'Jean isn't coming.'
b. Jean (ne) peut pas (ne) pas connaître Marie.
Jean (NE) can NM (NE) NM know Marie
'Jean cannot not know Marie. [he surely knows her.]'

We follow Abeillé and Godard (1997), Kim and Sag (1996) and Kim (1996) in assuming that *pas* is an (adverbial) complement, if it is combined with a finite verb, and an adjunct otherwise. Semantically, it contributes a negation in both cases. It should be noted that the syntactic analyses of Recourcé (1996), Abeillé and Godard (1997) and Kim (1996) have compatible accounts for the co-occurrence of *ne* and *pas* in infinitival constructions, which has been problematic for GB-oriented accounts.

[20]See footnote 18 for a putative counterexample.

(15) Il regrettait de ne vraisemblablement pas pouvoir partir avant
he regretted to NE probably NM be-able-to leave before
avril.
april
'He regretted [probably not being able to leave until april.]'
(Recourcé 1996, p. 65)

2.3. **N-words.** We can distinguish three different environments in which n-words might occur. First, there can be exactly one n-word in sentences with a sentential negation reading, where a pre-verbal *ne* is optionally present (16a). Second, an n-word co-occurs with the negative marker, *pas*, and in most varieties of Standard French we observe an obligatory double negation reading (16b)[21]. Finally, sentences with two n-words are usually assigned a single negation reading, but the double negation reading is also available (16c).

(16) a. Personne (n') est venu.
nobody (NE) is come 'Nobody came.'
b. Personne (n') est pas venu.
nobody NE is NM come 'Nobody did not come.'
c. Personne (n') a rien lu.
nobody (NE) has nothing read.
'Nobody read anything.' or 'Nobody read nothing.'

Based on the double negation reading of (16b) and the optionality of *ne*, the only other plausible lexical exponent of negation besides the n-word in (16a), it is plausible to assume that French n-words contribute a negation to the meaning of the sentence. At the same time, the first, and preferred reading given for (16c) indicates a second nature of French n-words: In a sentence with a single negation reading but two n-words and no other plausible source of negation, one of the two n-words must necessarily be semantically non-negative. In that reading, (16c) has the characteristics of a negative concord construction. French n-words may also occur in the negative polarity contexts of questions (17a) and adversative predicates (17b), which are known to license negative polarity items in other languages as well.[22]

[21] Cited from Muller (1991, p. 258). It is important to note the absence of a negative concord reading.
[22] Note that the licensing of n-words is not a general cross-linguistic property of questions and adversative constructions, since they do not license n-words in Polish.

(17) a. As-tu jamais vu ça?
have you never seen this? 'Did you ever see this?'
b. (Muller 1991, p. 263):
J'hésite à rien écrire sur Simone Weil, ...
I hesitate to nothing write about Simone Weil, ...
'I hesitate to write anything about Simone Weil, ... '

Muller (1997, p. 256) provides further examples of n-words in typical negative polarity contexts. In (18), they stand in the scope of *trop* (*too*) and adversative predicates.

(18) a. Il est trop pauvre pour s'acheter jamais de chemise.
he is too poor for himself to-buy never any shirt
'He is too poor to ever buy himself a shirt.'
b. Il est trop désagréable pour qu'aucun d'entre nous lui fasse de cadeau.
he is too unkind for that none of among us him makes any present
'He is too unkind for any of us to make him a present.'
c. Il est impossible qu'aucun d'eux boive d'alcool.
it is impossible that none of them drinks any alcohol
'It is impossible for any of them to drink alcohol.'

From these data, we conclude that there exist two homonymous varieties of n-words in French. There are n-words that contribute negation (n_1-words hereafter) and n-words that do not (n_2-words hereafter). The question of whether n-words contribute negation or not can be separated from the other details of their logical form. We can, thus, investigate the distribution of n_1-words and n_2-words without having to commit ourselves to a specific logical form, but we will turn to that issue in section 2.4 below.

If one accepts the distinction between n_1-words and n_2-words, then it is clear which variant is present in the sentences with negative meaning and exactly one n-word. There, the negative variant, n_1-word, is attested. (17) and (18), on the other hand, contain instances of n_2-words. In cases with an n-word and a clausemate *pas*, a double negation reading arises[23], which is indicative of an n_1-word (16b). If two n-words co-occur in one sentence, as in (16c), an ambiguity arises, and the negative concord reading is preferred. This suggests that both, an n_1-word or an n_2-word, may

[23]See Corblin (1994) and, for some exceptions, Muller (1991) and section 5 below.

co-occur with an n_1-word. There are, however, a number of problems with this explanation.

The uncontentious instances of n_1-words allow modification by *presque* (*almost*) or *pratiquement* (*practically*), whereas such a modification is excluded in unequivocal cases of n_2-words, as illustrated in (19) and (20), respectively.[24]

(19) a. Presque personne n'est venu.
almost no-one NE is come
'Almost no-one has come.'
b. Jean n'a pas parlé à presque personne.
Jean NE has NM spoken to almost no-one
'Jean hasn't spoken to almost nobody.'
(it holds for almost no person that Jean has not spoken to that person)

(20) a. * As-tu presque jamais vu ça?
have-you almost never seen that
b. ?* Il est impossible que presque aucun d'eux boive
it is impossible that almost none of them drinks
d'alcool.
any alcohol

In the light of these data, we would expect that the availability of modification by *presque* (*almost*) is an indication for the presence of an n_1-word, i.e., the variant of n-words that contributes logical negation. This is exactly the way in which modifiability by *presque* is interpreted by Muller (1991) and van der Wouden and Zwarts (1993). As shown for Catalan and Greek by Giannakidou (1998) and for Polish by Richter and Sailer (1999), this reasoning cannot be correct, as it predicts the wrong reading for sentences of the kind illustrated in (21):

(21) Presque personne n'a presque rien vu.
almost no-one NE has almost nothing seen
'Very few people saw anything at all.'
$ 'Most people saw a lot.'

[24]The *almost*-test has been popular since Zanuttini (1991). See Muller (1991, p. 318) for an independent application of the *almost*-test to French. Note that sentence (20b) is, of course, grammatical under a negated interpretation of the embedded clause, but the presence of *ne* is strongly preferred in the negative interpretation.

Since both n-words in (21) are modified by *presque*, they should both be n_1-words. But the meaning of the sentence suggests that only one n-word can contribute a negation. The *presque* diagnostic could be maintained if one assumed, with Déprez (1997b), that n_1-words are treated as numerals, parallel to *zéro* (*zero*)[25]. The NC data in (16c) and (21) could, then, be analyzed as a pair reading. But Haegeman (1997, p.125f.) points to differences between numerals and n-words which seem hard to account for in Déprez's analysis. A clear difference lies in the fact that *zéro* does not license pre-verbal *ne*[26]. Błaszczak (1998a) and Richter and Sailer (1999) propose that Polish n-words do not carry a semantic negation even though they can be modified by *prawie* (*almost*). If that is correct and of cross-linguistic significance, it might be justified to assume that one of the n-words in (21) is an n_2-word. But then the *presque* diagnostic loses its value for distinguishing between the two variants of French n-words.

If there is no reliable test to distinguish between the putative two types of n-words, one might be led to envisage a theory in which there is only one type of n-words. To be consistent with our overall architecture, all n-words would have to be non-negative. French could be assimilated to the account of Polish of Richter and Sailer (1999), and the logical negation would be attributed to the verb and appropriate collocation conditions of verbs that carry negation. The reason we reject this analysis as inappropriate for French are the double negation readings of sentences with more than one n-word (16c). Additionally, if *ne* contributed logical negation, then its optionality would lead us to predict an optional negative meaning for ordinary positive sentences, contrary to the facts. For example, all sentences in (16) preserve their readings no matter whether *ne* is present or not. That means that the sentences without *ne* would have to be able to provide the missing semantic negation by themselves. Since a principle to the effect that they can only do so in the absence of *ne* and the presence of an element (an n-word) that needs to be licensed, seems to be linguistically implausible, that negation could always occur spontaneously in any sentence.

Muller (1991, p. 305,316) observes that the class of n-words that appear in negative concord readings differs from the class of n-words that appear in expletive negation contexts. On the basis of these observations one might posit the existence of not only two but three homonymous

[25]See Brown (1995) for a similar proposal concerning English n-words.
[26]See also Wilmet (1997, p. 168).

forms for n-words, negative n_1-words, co-negative n_2-words and expletive n_3-words. N_2-words appear in combination with n_1-words in negative concord readings. N_1-words and n_2-words allow modification by *almost*. N_3-words are restricted to non-negative contexts such as questions, adversative constructions and embedded constructions with matrix licensers, and they cannot be modified by *almost*. But a closer look at the data reveals that the potential n_3-words form a proper subset of the potential n_2-words. In other words, if an n-word has an expletive use, it also occurs negatively and co-negatively. On the other hand, if a word has a negative and a co-negative use, it might not necessarily have an expletive use. For example, *plus* (*(no) more*) and *que* (*only*) are not allowed in polarity contexts except for negation, but they can be combined with n_1-words. They can occur as negative and as co-negative n-words only. The observed contrast in distribution can, therefore, be accounted for by minor idiosyncratic differences in the collocation restrictions of n-words. Words like *plus* and *que* have stronger collocation restrictions in their lexical entries, which prevents them from occuring in non-negative licensing contexts. Others have more liberal collocation conditions, and they can occur in all licensing contexts for n-words. With a theory of collocations, making a semantic distinction between three types of French n-words is not necessary.

2.4. **Semantic properties of n-words.** We assume that French n-words exhibit a lexical ambiguity between negative (existential) quantifiers and indefinites. In this respect we disagree with Corblin (1994) and Corblin (1995), who rejects the assumption of a lexical ambiguity of n-words. In the following discussion, we will defend the ambiguity approach by adducing differences in the scope behavior of the two different kinds of n-words

2.4.1. *N-words and negation.* For simple sentences with only one n-word, we already know that the n-word contributes negation and belongs to the class of n_1-words. In (22), *personne* is the only n-word:[27]

(22) a. Je n'ai donné quoi que ce soit à personne.
I NE have given anything whatsoever to nobody
'I have given nobody anything at all.'

[27] The data in (22) are from Muller (1991, p. 317). *qui/quoi que ce soit* (*anyone/ anything*) is a negative polarity item (NPI). Some speakers find (a) ungrammatical.

b. Je n'ai rien donné à qui que ce soit.
I NE have nothing given to anyone whatsoever
'I haven't given anything to anyone at all.'

In sentences with more n-words and a single negation reading the situation is more difficult. As argued by Muller (1991, pp. 317–321), any one of several n-words can contribute semantic negation in NC constellations. Therefore, either sentence in (22) can be used to paraphrase (23).

(23) Je n'ai rien donné à personne.
I NE have nothing given to nobody
'I haven't given anything to anyone.'

The following example of Muller (1991, p. 320), (24a), demonstrates that the relative scope of the n-words does not directly reflect their linear order. In the reading (24b), linear order is respected, but in the reading (24c), the adverbial *plus* (*no more*) scopes over the subject. The negation always takes wide scope.

(24) a. Aucun de mes amis divorcés ne s'est plus marié.
none of my friends divorced NE refl. is no-more married
b. 'None of my divorced friends has married again.'
c. 'It was no more the case that any of my divorced friends married.'

These data lead to the conclusion that either one of the n-words can be interpreted as an n_1-word in NC constructions with several n-words. Once we have identified the n_1-word, all remaining n-words can be treated as n_2-words. In cases like (23), whether the first or the second n-word is the n_1-word makes no difference as for which readings are available, but the sentence in (24a) shows that, in principle, there can be a difference.

In fact, the semantic behavior of n_1- and n_2-words differs even if we ignore negation. The following negated sentence with a single n_1-word shows that the negation contributed by the n_1-word, *rien*, and the rest of its lexical meaning cannot be separated (Muller 1997, p. 260):

(25) a. Luc n'a souvent rien dit.
Luc NE has often nothing said
b. 'It is often the case that Luc didn't say anything.'
c. ? 'There is nothing that Luc said frequently.'
d. $ 'It wasn't often the case that Luc said something.'

This is not the case for true n_2-words, as evidenced in the following sentence with a non-negative occurrence of an n-word:

(26) a. ?? J'hésite à souvent rien écrire sur Simone Weil.
I hesitate to often nothing write about Simone Weil
b. $\mathsf{hesitate}'(i, \mathsf{often}'(\exists x.\mathsf{write}'(i, x, simone)))$
c. \$ $\mathsf{hesitate}'(i, \exists x.\mathsf{often}'(\mathsf{write}'(i, x, simone)))$

The sentence in (26a) only has the non-specific reading, in which *often* scopes over the things to be written.

In an NC constellation, an n_2-word must be interpreted with narrow scope relative to the scope of adverbials, even if this leads to a situation in which the adverbial takes intermediate scope between the semantic negation and the n_2-word. Restricting our attention to those readings of (27) in which the first n-word, *personne*, contributes negation, we can construct three logically possible readings, (27b)–(27d), for that sentence:[28]

(27) a. Personne n'a souvent rien dit.
no-one NE has often nothing read
b. 'It is often the case that nobody said anything.'
c. \$ 'Nobody said anything often.'
d. 'For nobody it was often the case that they said anything.'
[= Nobody often talked.]

Astonishingly, the third reading, (27d), whose analogue is unavailable in the case in which *rien* is an n_1-word (25d), is possible with an n_2-word in the same position. On the other hand, the reading (27c), where the second n-word, *rien*, takes scope over the adverbial in oder to be in the immediate scope of the negation, is excluded.

The scoping possibilities of n_2-words are clearly distinct from those of n_1-words. Even though they are licensed in the scope of a negation (or some other affective operator), they need not be in its immediate scope. Regarding their quantificational properties, n_2-words are always interpreted with the narrowest possible scope.

Our analysis receives further support from the interaction of n-words and *pas*. As we have seen earlier in (16b), n-words in French do not exhibit

[28]It is not clear, whether there are readings with *rien* as n_1-word. One of these readings would be logically equivalent to (27b); a second one would be equivalent to the unavailable (27c). In the third hypothetical reading, *rien* would have scope over *souvent*, which in turn scopes over *personne*. Our informants did not get this reading, which suggests that there is at least a preference to assign an n_1-reading to n-words that are pre-verbal subjects.

negative concord with the negative adverbial *pas*.[29] Corblin (1994) has shown that the linear order of an n-word and *pas* determines their relative scope.[30] In sentence (28),[31] the first n-word, *personne*, is treated as an n_1-word, i.e., it contributes semantic negation. The adverbial, *pas*, also contributes negation. Due to the linearity constraint, this negation must be in the scope of the subject. The crucial question now is which scope readings the second n-word can receive. In the most natural reading, it is interpreted as an n_2-word.[32]

(28) Personne n'a pas dit rien.
nobody NE has NM said nothing

(29) a. $\neg\exists x \neg\exists y\ \mathsf{say}'(x, y)$
b. \$ $\neg\neg\exists x \exists y\ \mathsf{say}'(x, y)$
c. \$ $\neg\exists x \exists y \neg \mathsf{say}'(x, y)$

The only available reading, (29a), is one in which the second n-word is interpreted in the scope of the lower negation. At first, this looks as if the second n-word participates in an NC constellation with *pas*. The problem with that explanation is that if this were allowed, we would be unable to explain the unavailability of an NC reading in sentences as (30).

(30) Jean n'a pas dit rien.
Jean NE has NM said nothing
'It is not the case that Jean has said nothing.'
\$ 'Jean hasn't said anything.'

In our analysis, an n_2-word must be interpreted with the narrowest possible scope. For (28) this means that the n_2-word must be in the scope of the lower negation. However, it need not be in the immediate scope

[29]As we will see below, this generalization, although widely believed in the literature, is not entirely correct. Nevertheless, it holds for all constructions discussed in this section.
[30]This might not be true in all cases, since, according to our informants, example (ii) has both readings stated under (i), in contrast to (i).

(i) Jean n'a pas dit rien.
Jean NE has NM said nothing
'It is not the case that Jean said nothing.' [= Jean said something]
\$ 'Nothing was not said by Jean.' [= Jean said everything]
(ii) Jean n'a pas rien dit.

[31] We are grateful to Francis Corblin for discussing this sentence with us.
[32]See section 5 for an example in which the second n-word is interpreted as an n_1-word, which results in a triple negation reading.

of its licenser, which is the n_1-word. Our account correctly predicts that the NC reading of (30) is not available, and the pseudo NC reading is the only possibility in (28).

We found that in French, two distinct kinds of n-words can be observed even in NC constellations, where they show different scopal behavior. It is, therefore, not sufficient to say that n_1-words express negation, whereas n_2-words do not. In the remainder of our discussion of the logical form of n-words, we will further study the semantic similarities and differences of n_1- and n_2-words by applying a few tests from the literature which tell us more about their behavior compared to other quantifiers.

2.4.2. *N-words and quantification.* Richter and Sailer (1999) employ a number of semantic tests to determine the logical form of n-words in Polish. We will use some of these tests for a closer investigation of the two varieties of French n-words.

Scopal interaction with quantifiers. The scope of n_1-words interacts non-trivially with other quantificational elements, whereas, as we have seen, n_2-words are bound to the narrowest possible scope interpretation. In (25) above, the relative scope of the n_1-word and the frequence adverbial is free. The freedom of scope persists if another quantificational complement is added to the construction. Compare the behavior of the n_2-word, *rien*, in (27) to that of a quantifier such as *la plupart* (*most*):

(31) a. Personne n'a raté la plupart des classes.
nobody NE has missed most classes
'Nobody is such that he missed most classes.'
'Most classes weren't missed by anybody.'
b. Jean a souvent raté la plupart des classes.
Jean has often missed most classes
'Often, Jean has missed most classes.'
'Most classes were such that Jean often skipped them.'
c. Personne n'a souvent raté la plupart des classes.

The translations of (31a) and (31b) indicate that there is a scope ambiguity between *la plupart* and the n_1-word and between *la plupart* and the quantificational adverb, respectively. Sentence (31c), which contains all three quantificational elements, has all six logically possible readings.

The observable scopal interactions reveal a clear contrast between n_2-words and quantifiers, whereas n_1-words pattern with quantifiers. To capture this difference, we suggest that n_2-words should be analyzed not as qantifiers but as Heimian indefinites which must be bound with the narrowest possible scope. That means that all other quantificational expressions must have scope over the quantifier that binds them. At the same time, the n_2-words need not be bound in the immediate scope of the licensing operator. In short, French n_1-words are negative existential quantifiers ($\lambda P \neg\exists x.P(x)$), and n_2-words are indefinites ($\lambda P.P(x)$) which must be bound with narrow scope and must be collocationally licensed by an affective operator. In the constructions we have considered so far, they could not be licensed by *pas*.

French n_2-words share the property that their scopal behavior differs from that of quantifiers with Polish n-words, but the manifestation of this difference is subject to cross-linguistic variation. Whereas Polish word order determines the scope of quantifiers but not the scope of indefinites, French word order determines the scope of indefinites but not the scope of quantifiers. The differences in the systems of n-words between French and Polish should ultimately be derived from the fact that in Polish, word order constrains the relative scope of quantifiers whereas in French, word order is not as strictly mapped to scope.

Predicative Constructions. It has been argued in Partee (1987), Giannakidou (1997), Błaszczak (1998a) and Błaszczak (1998b) that universal quantifiers can appear in predicative constructions only under strong contextual restrictions, if at all. As the examples in (32) – (34) illustrate, neither type of n-words allows a predicative use in French.

(32) a. Jean n'est pas médicin.
Jean NE is NM doctor
'Jean isn't a doctor.'
b. * Jean n'est aucun médicin.
Jean NE is no doctor

(33) a. Personne n'est médicin.
nobody NE is doctor
'Nobody is a doctor.'
b. * Personne n'est aucun médicin.
nobody NE is no doctor

(34) a. C'est pour empêcher que Jean (ne) devienne médicin.
that is to prevent that Jean (NE) becomes doctor
'This is to keep Jean from becoming a doctor.'

b. * C'est pour empêcher que Jean devienne aucun médicin.

In (32b), *aucun* is an n_1-word, in (33b) it is ambiguous between an n_2-word licensed by *personne* or an n_1-word which should induce a double negation reading, and in (34b) it is an n_2-word licensed by the complementizer *que*. At first glance, the behavior of n_1-words could be taken as argument in favor of a universal analysis, since predicative constructions are incompatible with quantifiers as *chaque* (*every*).

(35) * Jean est chaque médicin.
Jean is every doctor

The fact that n_2-words are not permitted either, as shown in (33b) and (34b), weakens the force of the argument. In this situation, it is important to notice that in French, contrary to English, there exists an articleless variant of predicative sentences which is clearly preferred:

(36) Jean est (un) médicin.
Jean is (a) doctor
'Jean is a doctor.'

Apparently, the appearance of determiners in French predicative constructions of the type we consider is problematic for independent reasons. If that is so, then we cannot construct an argument in favor of a specific analysis of n-words as quantifiers or indefinites from these data.[33]

Donkey pronouns. Since Roberts (1989) it is known that there exist dynamic binding effects across negation in modal contexts. These effects are attested for indefinites and existentials.[34] As the examples show, both types of French n-words are admissible in these constructions.

---

[33]Determinerless constructions are equally problematic and inconclusive. In (i) and (ii), an n_1-word, *personne*, appears in a predicative construction. It is, however, unclear whether these sentences share all relevant properties with the profession-type predicatives discussed above.

(i) Paul n'est personne d'important.
Paul NE is nobody of important
'Paul isn't anybody important.'

(ii) Paul n'est *(pas) personne.
Paul NE is NM nobody
'Paul isn't nobody.'

[34]See Giannakidou (1998) for observations concerning the logical form of n-words in Greek.

(37) a. Soit qu'il n'y a aucune salle de bains dans cette
be it that it NE there has no bathroom in this
maison, soit qu'elle se trouve dans un endroit bizzare.
house, be it that it itself finds in a place strange
'Either there is no bathroom in this house, or it is in a funny place.'

b. Soit que personne n'a construit aucune salle de bain dans
be it that nobody NE has built no bathroom in
cette maison, soit ...
this house, be it
'Either nobody has built any bathroom in this house, or it is in a funny place.'

(38) a. N'achète aucun livre! Il pourrait être dangereux!
NE buy no book! It might be dangerous!
'Don't buy any book! It might be dangerous!'

b. Je ne doute pas que Marie achète aucun livre, bien qu'il
I NE doubt NM that Marie buys no book, even though it
pourrait être dangereux.
could be dangerous
'I don't doubt that Marie buys [any book]$_i$, even though it$_i$ might be dangerous. '

In the sentences (37a) and (38a), *aucun* is an n_1-word; in the sentences (37b) and (38b), *aucun* is an n_2-word.

In dynamic contexts, both types of n-words show the same behavior which is attested for indefinites and for existentials. In predicative constructions neither type of n-word is allowed. Other French determiners with the notable exception of the indefinite article, *un*, share this behavior with n-words.

The last two tests have shown that the two types of n-words behave similarly in certain constructions. In predicative constructions they seem to pattern with universal quantifiers. As we have seen, this behavior can perhaps be ascribed to independent properties of French determiners. In contrast to universals and in support of our analysis, n_1-words and n_2-words allow dynamic binding effects.

3. Technical Preliminaries

Before we turn to the HPSG analysis of French negation, we need to clarify our assumptions about the major technical ingredients of our theory.

The formalism. The formalism in which we formulate our HPSG theory is RSRL (Richter 1999; Richter, Sailer, and Penn 1999). The most important features which distinguish RSRL from other feature logics are quantification over components of objects and relations as first class citizens of the logical language. Their combination leads to a classical interpretation of relations under the scope of negation that corresponds to the intuitive meaning of relational expressions. The devices of RSRL are capable of expressing current linguistic analyses proposed in the HPSG literature.

An RSRL grammar consists of a signature and a set of principles like the Head Feature Principle and the Subcategorization Principle. Among these principles is the Word Principle (Höhle 1999), which is of special interest to linguistic analyses which emphasize the lexical component of the grammar. The Word Principle contains a disjunctive enumeration of the lexical entries in the consequent of an implication whose antecedent is *word*. The *derivational rules* (DRs) which we introduce below to relate (possibly simple) words to derived words are to be understood as instances of the object-level lexical rules defined in Meurers (1995).

Semantics in HPSG. In Richter and Sailer (1997) and Richter and Sailer (1999), we have shown that it is possible to employ expressions of a traditional logical language such as Ty2 as values of the content attribute of linguistic signs. The syntax of the semantic representation language is defined in terms of a sort hierarchy and principles that ensure the well-formed semantic typing of the Ty2 expressions, and the noncyclicity of the relevant structures. The details are specified in the appendix below.

We assume that the meaning of signs is constructed by simple composition rules along the syntactic phrase structure from the semantic contribution of the lexical elements that occur in the sign. For the treatment of the phenomena under consideration, simple functional application with lambda conversion suffices.[35] Functional application and lambda conversion are

[35]We are open to allow other basic semantic operations at the phrasal level, such as lambda abstraction or existential binding (von Stechow 1993; Bittner 1994). A constructional introduction of negation as proposed by Ladusaw (1995) would, however,

built into the SEMANTICS PRINCIPLE. The SEMANTICS PRINCIPLE says that the CONTENT value of a phrase is the redex-free result of recursively applying lambda conversion to the functional application of one daughter's CONTENT value to the other daughter's CONTENT value.

In order to account for scope ambiguities, we specify a set of derivational rules that encode type-shifting operations on the semantic contribution of words. Our type-shifting rules are borrowed from Hendriks (1993). In addition, we assume DRs for existential closure and for the insertion of *pas* on the COMPS list of finite verbs. The latter DR is our formalization of the corresponding lexical rule proposed in Abeillé and Godard (1997) and Kim and Sag (1996). Finite verbs are input to the existential closure derivational rule, which means that, like *pas*, they can introduce existential quantifiers into the logical form of sentences which ultimately bind the variables of indefinites. To control the lexical introduction of existential closure, two components of our theory of logical forms come into play, one merely technical in nature, the other linguistically motivated. On the technical side, we use a definition of lambda conversion that is not fully meaning-preserving, as it precludes the re-naming of variables (α-conversion). We allow variables that are free in a daughter's CONTENT value to be bound in the phrase's CONTENT value by entering into the scope of a quantifier. The proper binding of all variables in the logical form of a sentence and the prohibition of empty quantification in utterances are enforced by the PRINCIPLE OF FULL INTERPRETATION (PFI). The PFI states that all variables must be bound, all quantifiers must bind distinct variables, and there is no empty quantification. Thus, the PFI ensures that every element in the semantic representation is properly used and the variables are all bound.

Collocations. Building on work of van der Wouden (1992), Richter and Sailer (1999) incorporate a lexical theory of collocations in the grammar of negative concord in Polish. They propose a very general device to account for collocation properties. Each word can specify its own occurrence restrictions. The basic theory postulates two dimensions of occurrence restrictions. First, a licensing context must be specified. Second, the licensing context must appear within a certain domain relative to the word under consideration. For example, Polish n-words are licensed in the context of a negation. The negation context must be found in the

violate our requirement of constructing the meaning of phrases only from the meaning of their words.

same clause in which the n-word occurs. Richter and Sailer (1999) show that it is necessary to allow multiple, simultaneous occurrence restrictions. According to Progovac (1988), certain negation sensitive elements in Serbo-Croatian behave as positive polarity items (PPIs) with respect to the clause in which they appear, but as NPIs with respect to some larger syntactic domain.

A natural formalization in HPSG of conditions that words impose on their semantic environment within some syntactic domain involves making the relevant structure locally available for lexical conditions by adding it under an attribute defined on words. We call this attribute COLL, and extend the signature in such a way that its value is a list of *barrier*s. *barrier*s indicate in which domain the lexical conditions must be satisfied. Therefore, *barrier* has several subsorts which specify the various types of empirically observed barriers for collocations.

(39)
$$\begin{bmatrix} \textit{barrier} \\ \text{LIC } \textit{content} \end{bmatrix}$$

complete-clause (cc) *utterance* ...

According to (39), we expect that some words want their collocation requirements to be satisfied within the minimal complete clause in which they are contained; whereas others want them to be satisfied in the utterance they are part of, that is, the minimal structure that bears illocutionary force. The dots indicate that there might be other relevant environments.

The attribute LIC(ENSER) is defined on all barriers. Its purpose is to make the structure accessible on which words can lexically impose their occurrence restrictions. For our theory of French sentential negation, the licenser will always be the logical form (the CONTENT value) of the minimal complete clause in which the element that must be licensed occurs. The correct identification of the licenser of words with collocation restrictions in signs is guaranteed by the Clausal Content Collocation Principle (CCCP), (40). The CCCP makes sure that the licensing logical form for a word is indeed the CONTENT value of the minimal complete clause that contains that word.

(40) Clausal Content Collocation Principle (CCCP)
In each sign, s, for each complete clause z that is dominated by s,

> for all words with a *cc* object, x, in their COLL list that are dominated by z and not dominated by another intervening *cc* barrier, z's CONTENT value is identical with x's LIC value.

With the crucial components of our theory in place, we can now spell out our analysis of the French data of section 2.

4. Analysis

We are now ready to present our analysis for French in more formal detail. Issues of logical form will be our primary concern. We will not discuss syntactic licensing barriers. For the purpose of this paper we work with the simplistic assumption that the collocation requirements of n-words must be satisfied in the minimal complete clause in which they are contained. French finite sentences are instances of complete clauses in the relevant sense.

Pre-verbal ne. We treat *ne* as a prefix (when combined with a finite verb) or marker (in infinitival constructions) which makes no semantic contribution, but introduces collocation conditions. More specifically, we claim that the clause in which *ne* appears must contain a negation (or a licensing complementizer) that scopes over the main event.[36] The relevant parts of the lexical entry of the marker *ne* and its licensing condition are spelled out under (88) in the appendix. The collocation requirements of the marker *ne* are identical to those of finite verbs with the affix *ne*.

The adverbial negative marker pas. Syntactically, we follow Abeillé and Godard (1997) and Kim and Sag (1996) in assuming that the SYNSEM value of *pas*, like the SYNSEM value of other members of a class of adverbials with similar syntactic behavior, is added to the COMPS list of a finite verb by a derivational rule (*pas*-DR). *Pas* is the syntactically least oblique complement in the structure. The *pas*-DR also puts an additional semantic argument slot onto the CONTENT value:

(41) *pas*-DR

[36]As we do not restrict the distribution of verbs without the prefix *ne*, we correctly predict the optional character of *ne*.

$$\begin{bmatrix} \textit{word} & \\ \text{SYNS L} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \textit{finite} \end{bmatrix} \\ \text{VAL COMPS} & \boxed{1} \\ \text{MARKING} & \textit{unmarked} \end{bmatrix} \\ \text{CONT} & \lambda x_{1} \ldots \lambda x_{n}.\phi \end{bmatrix} \end{bmatrix} \implies$$

$$\begin{bmatrix} \textit{word} & \\ \text{SYNS L} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{VAL COMPS} & \left\langle \begin{bmatrix} \textit{synsem} \\ \text{L CAT H MOD L CAT MARKING } \boxed{2} \end{bmatrix} \middle| \boxed{1} \right\rangle \\ \text{MARKING} & \boxed{2} \end{bmatrix} \\ \text{CONT} & \lambda a \lambda x_{1} \ldots \lambda x_{n}.a(\phi) \end{bmatrix} \end{bmatrix}$$

The semantic contribution of *pas* consists of logical negation and existential closure over an arbitrary finite number of variables. We obtain the logical form $\lambda p_t.\neg\exists x_1 \ldots \exists x_n p$. In (42), the relevant parts of the lexical entry of *pas* are given.

(42) The lexical entry of *pas*:

$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & \langle \textit{pas} \rangle \\ \text{SYNS L} & \begin{bmatrix} \text{CAT HEAD MOD} & \begin{bmatrix} \text{L CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \textit{inf} \end{bmatrix} \\ \text{VAL COMPS} & \textit{elist} \\ \text{MARKING} & \textit{unmarked} \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \lambda p_t.\neg\exists x_1 \ldots \exists x_n p \end{bmatrix} \end{bmatrix}$$

The PFI, (80), guarantees that, in an utterance, each quantification which *pas* introduces binds a variable. Consider the derivation of the CONTENT value of the sentence in (43a):

(43) a. Jean ne lit pas un livre.
Jean NE reads NM a book
$\neg\exists y.\mathsf{book}'(y) \wedge \mathsf{read}'(j, y)$

b. derivation of the finite verb:
1. *lit*: from the base lexical entry: $\lambda y \lambda x.\mathsf{read}'(x, y)$
2. *lit*: raising the quantifier: $\lambda Y \lambda x.Y(\lambda y.\mathsf{read}'(x, y))$
3. *lit*: *pas*-DR: $\lambda a_{\langle t,t\rangle} \lambda Y \lambda x.a(Y(\lambda y.\mathsf{read}'(x, y)))$
4. *ne lit*: adding affixal pre-verbal *ne*: $\lambda a_{\langle t,t\rangle} \lambda Y \lambda x.a(Y(\lambda y.\mathsf{read}'(x, y)))$

c. semantic contribution of the NPs:
1. *Jean*: j
2. *un livre*: $\lambda P \exists y.\mathsf{book}'(y) \wedge P(y)$

d. derivation of the phrasal semantics:
   1. *ne lit pas*: $\lambda Y \lambda x.\neg Y(\lambda y.\mathsf{read}'(x,y))$
   2. *ne lit pas un livre*: $\lambda x.\neg\exists y.\mathsf{book}'(y) \wedge \mathsf{read}'(x,y)$
   3. *Jean ne lit pas un livre*: $\neg\exists y.\mathsf{book}'(y) \wedge \mathsf{read}'(j,y)$

The DR in (41) mentions the MARKING value of the input and of the output word. This is necessary to account for the incompatibility of *ne*-verbs with *pas*. In the lexical entry of *ne*-verbs, the MARKING value is specified as *NE*. *Ne*-verbs can, thus, never serve as input to the *pas*-DR. As illustrated in (43b), the output of the *pas*-DR is a word with the MARKING value *unmarked*. Only the words that bear the affix *ne* are marked *NE*. In the example above, *ne* combines with *lit* one derivation step after the application of the *pas*-DR.

When *pas* combines with a non-finite verb as in the sentence (15), it is syntactically a modifier of a verb projection, and its SYNSEM value does not appear on the COMPS list of the verb. A derivational rule extends the semantic arguments of the verb which *pas* (or another adjunct) modifies to provide the argument slots for adjuncts:

(44) ADJUNCT INTRODUCTION DR

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT} \;\; \lambda x_1 \ldots \lambda x_i \lambda x_{i+1} \ldots \lambda x_n.\phi \end{bmatrix} \Longrightarrow$$

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT} \;\; \lambda x_1 \ldots \lambda x_i \lambda a \lambda x_{i+1} \ldots \lambda x_n.a(\phi) \end{bmatrix}$$

According to (44), an adjunct introduced by this derivational rule can take any scope relative to the other arguments, x_1 to x_n. The actual scope of an adverb in a sentence will usually be restricted by the syntactic structure and word order relationships. Note that despite the different syntactic behavior of *pas* in finite and non-finite constructions, there is only one lexical entry for it.

N-words. There are two lexical entries for each n-word: For *personne*, we need *personne$_1$*, which is a negative existential quantifier, and *personne$_2$*, which is an indefinite. Let us first consider their logical forms:

(45) a. *personne$_1$*: $\lambda P \neg\exists x \exists y_1 \ldots \exists y_n.P(x)$
     b. *personne$_2$*: $\lambda P.P(x)$

In (45a), the first existential quantifier which binds the variable, x, introduced by *personne$_1$*, is followed by a finite sequence of existential quantifications. This sequence of quantifications can informally be understood as an optional existential closure triggered by the negation. If that negation proves to be the only operator in a sentence which can contribute existential closure, and there would otherwise be free variables, then the PFI enforces an existential binding at that point in the logical form.

For *personne$_2$*, stating the logical form is not enough to account for its semantic properties. We still must specify its collocation restrictions. An operator must bind its variable at some node of the semantic derivation. The binding operator must be in the scope of a clausemate negation (or of a special negative complementizer). To exclude the NC reading of *personne* with *pas* in sentences like (16b), which is repeated in (46a), we demand that the quantifier that binds the free variable introduced by the n$_2$-word may not be the first in the scope of the licensing negation. The relevant part of the lexical entry of *personne$_2$*, including its collocation requirement, is formalized in the appendix (87). In (46c), the licensing context for the free variable, x, introduced by an n$_2$-word such as *personne$_2$* is sketched informally. The negation symbol is meant to represent the operator that licenses the n-word[37].

(46) a. Personne$_{1/*2}$ (n')est pas venu.
$ \$ \neg\exists x.\mathsf{come}'(x)$
$\neg\exists x.\neg\mathsf{come}'(x)$
b. Personne$_1$ n'a rien$_{1/2}$ lu.
$\neg\exists x\exists y.\mathsf{read}'(x,y)$
$\neg\exists x\neg\exists y.\mathsf{read}'(x,y)$
c. $\ldots\neg\exists y\ldots\exists x\ldots[\ldots x\ldots]\ldots$

(46a) and (46b) make the readings for (16b) and (16c), respectively, explicit. In (46a), the occurrence of *personne$_2$* is excluded because it would necessarily require that the variable introduced by the n$_2$-word be bound by the first quantifier after the negation. Only the double negation reading, which involves *personne$_1$*, can be derived. For the sentence (46b), we can derive both a single negation reading and a double negation reading. In the NC reading with only one negation, *rien$_2$* is used. The variable introduced by *rien$_2$*, y, is bound by the existential closure under negation

[37]Note that this does not exclude the possibility that a negation symbol precedes $\exists x$, as long as that negation is not the licensing negation of the n$_2$-word. The only reading of sentence (28), given in (29a), is an example for that.

that is contributed by *personne$_1$*. The existential quantifier that binds y follows the existential quantifier that binds the variable, x, introduced by *personne$_1$* itself, which, in turn, follows the negation symbol. The collocation restriction sketched in (46c) is, therefore, satisfied.

It is interesting to observe that there seems to be a possibility to drop the requirement that the binding quantifier not be the first after the licensing negation. Occurrences of single n-words which receive an NC interpretation with *pas* are documented on television, in newspapers and in literature. (47) shows three examples, which we quote from Muller (1991, p. 261):

(47) a. Personne ne sait pas ce qu'il y a derrière.
nobody NE knows NM what that it there has behind
'Nobody knows what is there behind.'
b. Je n'ai pas besoin d'aucune preuve.
I NE have NM need of no proof
'I don't need any proof.'
c. ... il fera pas d'cadeau à personne.
... he will-make NM any present to nobody
' ... he is not going to make a present to anybody.'

Another way to approach the positional condition on the quantifier binding the variable of an n$_2$-word would be to deny its existence as part of the collocation requirements of *personne$_2$*. Instead, one might derive the facts from pragmatic considerations. Under its NC reading, (16b) has the same logical form as (16a). But (16a) has the advantage of being shorter and unambiguous. It should, therefore, be pragmatically preferred.[38]

Even though our informants uniformly reject the NC reading for all sentences in (47), they report noticable differences. They consider an NC reading completely impossible in the case of (a), quite impossible for (b) and, finally, almost available for (c). In section 5, we will provide evidence that, in certain constellations, an n$_2$-word may be licensed by *pas*. Interestingly, (c) resembles these constellations the most, and (a) shows the least resemblance.

As it stands, the status of the restriction on the quantifier binding the variables of indefinite n-words is not entirely clear yet. If, as we have

[38]This is different from the case of (16c). There exists no more compact way to express either reading. But, perhaps, similar pragmatic considerations can explain why, even in the light of the existence of clear cases of a double negation reading for sentences with several n-words, most speakers tend not to mention it in cases as (16c).

suggested, the empirical facts can plausibly be derived on the basis of independent pragmatic considerations, it could be omitted in the collocation requirements of n_2-words. On the other hand, the apparent counterexamples in (47) might receive a different explanation.

5. Further Data

In this section, we consider more French negation data, indicate how they fit into our analysis, and discuss other treatments which they have received in the literature. In particular, we will defend our analysis of n-words as lexically ambiguous expressions against the express criticism brought forward against the ambiguitiy approach by Corblin (1995) and Ladusaw (1995). Along the way, we will point out what we think is problematic about other treatments.

First, we will reconsider the interaction of n_2-words and *pas* as their potential licenser in the light of the predictions of our analysis (section 5.1). Second, we discuss readings with multiple negations (section 5.2). And finally, we will see that there exist other classes of n-words than the ones we have considered so far. While they can also be devided into n_1-words and n_2-words, they are distinguished from other types of n-words by their collocation properties (section 5.3).

5.1. **Licensing of n_2-words by *pas*.** Let us first come back to the widespread claim that n-words cannot be licensed by *pas*. As we have seen above, although n_2-words can receive an NC reading, this only seems to be possible when the licensing negation comes from an n_1-word, but not with *pas* as licenser. According to Corblin (1995, p. 467) this is a "knock down argument against the analysis of *rien* ... as a lexical item having either a negative or a polarity item interpretation." In (48), the relevant data are repeated:

(48) a. Jean ne lit pas rien.
Jean NE reads NM nothing [only double negation]
b. Personne ne lit rien.
no-one NE reads nothing [double negation or NC]

Corblin (1995) affirms that these data show that what we have called *rien*$_2$ cannot be a negative polarity item.[39] In fact, all theories of NPIs

[39] Corblin (1995) explains the NC reading of (48b) as the result of the application of certain construction rules to which we turn below. An NPI analysis has been suggested

predict that any NPI should be licensed by the strong sentential negation marker *pas*. If *rien* exists as an NPI, we would, therefore, expect that it can occur in NC contexts induced by *pas*. To see why the *pas* constructions we considered so far could not license *rien*$_2$, recall the collocation conditions that we have deduced for its lexical entry. We found that n$_2$-words must be bound in the scope of a licensing negation with the narrowest possible scope of their binding quantifier. In addition, the binding quantifier must not immediately follow the licensing negation. These conditions are independent of the relative position of *pas* and the n-word, and they exclude, thus, an NC reading of the sentences with *pas* and a single n-word in (49).

(49) a. Personne n'a pas parlé.
nobody NE has NM spoken
$\neg\exists x \neg\ \mathsf{speak}'(x)$
$\$\ \neg\exists x\ \mathsf{speak}'(x)$
b. Luc n'a pas rien dit.
Luc NE has NM nothing said
$\neg\exists x \neg\ \mathsf{say}'(l, x)$
$\$\ \neg\exists x\ \mathsf{say}'(l, x)$

As a negation introducing element, *pas* can, in principle, license n$_2$-words. The problem with the constructions in (49) is that the licensing potential of *pas* is superseded by the collocation conditions of *rien*$_2$, which exclude the logical forms that would result from the semantic composition rules. Our characterization of the collocation requirements of n$_2$-words makes the availability of n$_2$-words dependent on properties of the logical form of a sentence, but not on the presence or absence of particular words, like *pas*.[40] This leads to the prediction that *pas* licenses n$_2$-words in contexts that differ from the ones in (49) in a way that results in logical forms of the sentences that are compatible with the lexical requirements of n$_2$-words. More specifically, in the logical form there must be a quantifier intervening between the negation and the quantifier that binds the variable introduced by the n$_2$-word. Consider the following data:[41]

---

by Jones (1996) and by Corblin (1994, p. 23). In the latter proposal, French n-words are NPIs exactly in the absence of a clausemate negation.

[40]Recall that in sentence (28) above, we have already seen that *pas* and an n$_2$-word might co-occur in one sentence. But in (28), the n$_2$-word, *rien*, is not licensed by *pas*; it is licensed by an n$_1$-word, *personne*.

[41] According to Francis Corblin (p.c.), sentence (50a) is not fully grammatical. He gave us the example (i) with the NPI *le moindre N* (*the least N*). In (i), there is a preferred NC reading. The difference in grammaticality among speakers might be due

(50) a. Je n'ai pas donné quoi que ce soit à personne$_{1/2}$.
I NE have NM given anything to nobody
$\neg\exists x \exists y$ give$'(I, x, y)$
$\neg\exists x \neg\exists y$ give$'(I, x, y)$
b. Je n'ai donné quoi que ce soit à personne$_{1/*2}$.
$\neg\exists y \exists x$ give$'(I, x, y)$

In (50), an NC reading arises when the quantifier binding the variable, x, of the indefinite, *quoi que ce soit* (*anything*), intervenes between the negation operator and the variable, y, of the n$_2$-word. In (51), the order of the indefinite and the n-word are reversed. This brings the n-word into the immediate scope of the negation contributed by *pas*. As our theory predicts, no NC reading is available, and *rien$_2$* is prohibited.

(51) a. Je n'ai pas donné rien$_{1/*2}$ à qui que ce soit.
$\neg\neg\exists x \exists y$ give$'(I, x, y)$
\$ $\neg\exists x \exists y$ give$'(I, x, y)$
b. Je n'ai donné rien$_{1/*2}$ à qui que ce soit.
$\neg\exists x \exists y$ give$'(I, x, y)$

While our analysis correctly predicts (50), it is problematic for Corblin (1994). There, an n-word can only enter into an NC-like reading with *pas* if the complexity threshold of two negations in a single clause has been reached. But in (50), we observe an NC reading of an n-word with only one negation.[42]

---

to different restrictions on certain NPIs. In particular, even though it is widely used in linguistic literature (Muller 1991; Muller 1997; Déprez 1997b), some speakers do not like the NPI *qui/quoi que ce soit*.

(i) Je n'ai pas donné le moindre renseignement à personne.
I NE have NM give the least information to nobody

'I didn't give any information to anybody.'

[42] Francis Corblin (p.c.) rejects the NC reading for sentences with an intervening universal quantifier like (i):

(i) % Je n'ai pas donné tous les cadeaux à aucun enfant d'ici.
I NE have NM given all the presents to no child from here

Some of our informants agree with this judgment, others find the sentence grammatical under an NC reading. The source of the deviation might be the special behavior of universals with respect to negation in French. Universals tend to be interpreted with narrow scope relative to sentential negation. In (i), then, some speakers might try to construct the universal below the n-word, bringing the latter into the immediate scope of *pas*. Therfore, the licensing requirements of *personne$_2$* cannot be satisfied.

5.2. **Aspects of multiple negations.** In Corblin (1994) and Corblin (1995), it has been argued that simple sentences in French never allow a reading with more than two negations. For example, for sentence (52), there are four hypothetically possible readings under a lexical ambiguity approach, only two of which are actually available.

(52) Personne ne dit rien à personne.
no-one NE says nothing to no-one

(53) a. $\neg\exists x \exists y \exists z.\mathsf{say}'(x, y, z)$
b. $ $\neg\exists x \exists y \neg\exists z.\mathsf{say}'(x, y, z) \equiv \forall x \forall y \exists z.\mathsf{say}'(x, y, z)$
c. $\neg\exists x \neg\exists y \exists z.\mathsf{say}'(x, y, z)$
d. $ $\neg\exists x \neg\exists y \neg\exists z.\mathsf{say}'(x, y, z) \equiv \forall x \exists y \forall z.\mathsf{say}'(x, y, z)$

Corblin's analysis is cast in the DRT framework and investigates the possibilities of deriving exactly the right readings of multiply negated sentences from constraints on the application of the construction rules that build the DRSs. A central constraint imposes a threshold for the number of negations which are allowed to occur in a single clause, and it sets the upper bound to two. In combination with linear order scopal requirements, the readings (53a) and (53c) are the only ones that are derivable. The way the construction rules are set up, it is sufficient for Corblin (1995) to postulate only one type of n-words, which introduces a negation and corresponds to our n_1-words.

We fail to see why, as Corblin claims, the apparent need for adopting a threshold constraint should be a fundamental problem for the lexical ambiguity approach. Any analysis that relies on semantic representations as part of linguistic objects instead of pure entailment notions, such as the semantic representations of DRT, HPSG's CONTENT value and GB's LF, assumes that there exist conditions on the form of its semantic representations. These conditions can easily restrict the complexity of the semantic representations. If it can be convincingly argued that there really is a competence limitation to the number of negations in the semantic representations of clauses, then this limitation could be expressed in HPSG as a constraint on admissible CONTENT values just as it can be phrased as a processing restriction on constructing DRSs.[43]

[43]Przepiórkowski and Kupść (1999) and Przepiórkowski (1999) reorganize HPSG's CONTENT value of sentences in such a way that at most most two negations may occur in a simple sentence. This can be viewed as a direct incorporation of Corblin's threshold into the feature geometry. A crucial difference between the two approaches

However, we do not want to commit ourselves to hypotheses about upper bounds for the number of negations yet. At least for some speakers, the threshold of two negations seems not to be absolute. Consider the following small dialog.[44]

(54) A: Jean n'a rien dit à PERSONNE.
$\neg\exists x \neg\exists y\ \mathsf{say}'(j, x, y)$
B: Mais non, Jean n'a PAS rien dit à personne.
$\neg\neg\exists x \neg\exists y\ \mathsf{say}'(j, x, y)$

The sentence uttered by A is to be taken in its double negation reading. B's reply, then, semantically negates A's utterance. As the logical formulae below the sentences indicate, this results in the presence of three "$\neg$" operators in the semantic representation of a single clause.

A further argument against the ambiguity of n-words comes from data that involve the pre-verbal negative particle, *ne*. Corblin (1994) as well as Recourcé (1996) claim that an n-word in argument position must be interpreted as a negative quantifier if it co-occurs with a clausemate *ne*. In our terms that would mean that those n-words would always have to be n_1-words. While the described empirical observation holds most of the time, Muller (1991, p. 266, 387) cites counterexamples which falsify its general validity:

(55) a. C'était aussi (...) pour empêcher que nul spécialiste ne puisse
it was also to prevent that no specialist NE can
pointer les insuffisances — pourtant notoires — de la
stick the inadequacies — nevertheless well-known — of the
Corporation.
corporation
'It was also to keep any specialist from piercing through the corporation's nevertheless well-known inadequacies.'
b. Nous enterrerons le garçon sans que personne ne le sache.
we will-burry the boy without that nobody NE it knows
'We will burry the boy without anybody knowing it.'

---

is that Corblin assumes that both negations are of the same kind, whereas for P&K, one negation is an eventuality negation while the other is a propositional negation.

[44]Note that *pas* in B's reply could be considered an instance of non-eventuality negation, if one assimilates this notion to the propositional negation of Dik (1997, p. 174–177). The example is a problem for the theory of Przepiórkowski (1999), because it contains three negations in a single clause.

In (55a), *nul spécialiste* is licensed as an n_2-expression by the complementizer *que*; in (55b), *sans que* licenses *personne*$_2$. In our view, what is at stake is determining the exact collocation conditions of *ne* and n_2-words. Not only may we find different classes of n_2-words which exhibit some variation in their licensing contexts. The pre-verbal particle, *ne* and n_2-words could prove to be subject to collocation conditions which might be able to help explain why their co-occurrence is rare.

This behavior of n_2-words and *ne* is not unexpected. It is well known that there is a hierarchy of contexts in which NPIs are licensed. A recent attempt at specifying a hierarchy of licensing contexts can be found in Giannakidou (1998). According to the licensing hierarchy we expect that all NPIs are licensed by *pas*; whereas n_1-words as licensers are further down in the hierarchy. An NPI that illustrates this pattern is *un chat* (taken from Déprez (1997b)) :

(56) a. Je ne vois pas un chat.
I NE see NM a cat 'I don't see a soul.'
b. * Personne ne voit un chat.
no-one NE sees a cat

Being a strong NPI, *un chat* can only be licensed by *pas*. For n_2-words we have seen above how other collocation requirements can interfere with the hierarchy of licensing contexts. In the case of n_2-words, additional restrictions obscured the fact that, in principle, they can be licensed by *pas*, as the hierarchy of licensers predicts. Progovac (1988) demonstrated that elements exist that are simultaneously NPIs and PPIs.

From our survey of French data we conclude that a theory that wants to account for all uses of n-words needs to assume a lexical ambiguity. Even if, as in Corblin (1994), all occurrences of n-words in negated clauses can be explained by a single type of n-words, a truly non-negative homonym must be added to the theory as soon as n-words in other polarity contexts are considered. We claim that an analysis resulting from that approach is less attractive than one that starts with the assumption of lexical ambiguitiy, because the former must (a) stipulate a restriction on the construction rules for semantic representations which is not independently motivated and (b) assume a collocational treatment of the positive uses of n-words in terms of both NPIs and PPIs. In our analysis, n_1-words are not (collocationally) constrained at all, and n_2-words carry a single collocation restriction.

5.3. **N-words with other collocation properties.** We finish our survey of collocations that are related to the French negation system with a short look at two more classes of expressions which illustrate the diversity and degree of lexical idiosyncrasy which can be found among collocations. For the first class of expressions, we take the word *mot* (*lit.: word*) as typical exponent, and for the second class, we discuss the behavior of the expression *âme qui vive* (*lit.: soul that lives*).

Grevisse and Goosse (1993, §977) note that there is a small set of n-words whose members are restricted to the co-occurrence with a particular semantic class of verbs. Furthermore, if such an n-word is the only n-word in a clause, the pre-verbal marker *ne* is obligatory.

(57) a. Je *(ne) vois/ comprends/ entends goutte.
I NE see/ understand/ hear drop
'I see/ ... nothing.'
b. Qui *(ne) dit mot consent.
who NE says word agrees
'Who doesn't say anything, agrees.'

Notice that the verbs in (57) are not *ne*-verbs. The n-words *goutte* and *mot* can enter into an NC reading with other n-words but not into an NC reading with *pas*. Moreover, they do not allow a double negation reading.

(58) a. Personne ??(ne) dit mot. [only NC]
b. * Je ne dit pas mot.

An n-word like *mot* can also occur in non-negative contexts. In these contexts, no *ne* is present.

(59) a. Le capitaine, après avoir désarmé son rifle, le jeta sur son épaule et s'avança d'un pas saccadé, ...
'The captain, after unloading his rifle, slung it over his shoulder and stepped forward in a series of jerks, ... '

sans mot dire.
without word say
'wordless.' (Jules Verne, *De la terre à la lune*)[45]

[45]The sentence has been found in the publicly accessible *French Collection at University of Virginia*, `http://etext.lib.virginia.edu/french.html`

b. Il est impossible qu'il dise mot.
it is impossible that he says word
'It is impossible that he says anything.'

This suggests that our new class of n-words comes in two variants. Following the pattern we have seen previously, there is a mot_1 and a mot_2, but the restrictions which they obey are stronger than the ones for normal n-words. In particular, mot_1 is a positive polarity item. It is excluded in sentences that contain another negation introducing element. For mot_2, the same restrictions as for n_2-words hold.

To capture the general restrictions of both variants of *mot*, we must state in their lexical entries that they occur in a *NE*-marked clause. Their semantic contribution does not only consist of logical negation and the restriction of the argument position; it also restricts the kind of event the main predicate of the sentence denotes. The semantic models on which we base our formalization are very simplistic and do not contain events as individuals. In a more elaborated semantic analysis, we could approximate the semantic contribution of *mot* in (57b) by a logical form like (60). The symbol e is meant to be an event variable. It is restricted to events of verbal expression.

(60) $\lambda P.\neg\exists x[\mathsf{thing}'(x) \wedge \mathsf{participant}'(e,x) \wedge \mathsf{saying}'(e) \wedge P(x)]$

The expression *âme qui vive* (*a living soul*) also requires the presence of pre-verbal *ne* (61a). In addition, its members exhibit an NC reading with both, *pas* (61b) and n-words (61c), and they allow non-negative uses in polarity contexts different from negation (61d).[46]

(61) a. Je *(n')y connaissais âme qui vive.
I NE there knew a living soul
'I didn't know anyone there.'
b. Il ?(n')y a pas âme qui vive.
it NE there has NM a living soul 'There is no-one.'
c. Personne ?(n')a jamais rencontré âme qui vive dans cet
no-one NE has never met a living soul in this
endroit desert.
place deserted
'No-one has ever met a soul on earth in this deserted place.'
d. Tout à l'heure, en traversant le village,

[46]With the exception of (61c), the data in (61) are from Gaatone (1971, p. 200f.).

c'est à peine si j'ai aperçu âme qui vive.
it's hardly if I have noticed a living soul
'Passing through the village a short while ago, I hardly saw a living soul.'

As was the case for *mot*, *âme qui vive* comes in two variants. *Âme qui vive$_1$* is a strong PPI, just as *mot$_1$*. *Âme qui vive$_2$*, on the other hand, is less restricted than other n$_2$-words. It may occur freely in the scope of any element that introduces sentential negation.

In connection with *âme qui vive* we can make an interesting observation that supports our decision to group n-words that co-occur with clausemate negation together with their non-negative uses, instead of putting them in one class with the truly negative n-words. In spoken language, the negation particle, *ne*, can marginally be dropped in (61b) and (61c), but never in (61a).[47] This indicates that whereas the requirement to co-occur with clausemate *ne* is part of the collocation restrictions of *âme qui vive$_1$*, it is not for *âme qui vive$_2$*.

Our list of collocational diversity within the group of n-words could easily be extended by taking a closer look at non-negative contexts that allow some n-words but not others (Muller 1991, p. 265), or by investigating which combinations of negative concord are really available (Muller 1991, p. 269; Cristea 1971, p. 116). From the data which we have discussed, it should already have become clear that a simple general processing constraint of the kind proposed by Corblin (1994) is inadequate and ultimately incapable of expressing the fine-grained distinctions that we have seen. Non-lexical accounts also do not seem to be promising to us, because they strive to push the negative meanings into constructions, thus multiplying the number of constructions and disregarding the importance of the lexical semantic contributions for the empirical distribution of lexical elements. But lexical theories are only suitable to the task at hand if they are furnished with a strong theory of collocations.

6. Canadian French and Haitian Creole

There are several variants of French whose negation system differs significantly from Standard French (SF). The differences which we can observe

[47]The marginal status the sentences (61b) and (61c) without *ne* is due to the fact that the expression *âme qui vive* belongs to a register that normally does not allow the omission of *ne*. Nevertheless, the contrast indicated in (61) is still quite clear for native speakers.

in those dialects can tell us more about the principles that govern the behavior of French negation. In this section, we will consider data from a dialect of Canadian French, Acadien (AC), and from Haitian Creole (HC). We will not try to provide a formal analysis of these languages. Our discussion is meant to illustrate how our collocation approach can account for the variation in the French negation system across different variants in terms of differing collocation restrictions of certain word classes, and in terms of the presence or absence of certain negation sensitive elements in the language.

6.1. **Acadien.** In Acadien, the pre-verbal negative marker, *ne*, is present only in registers that aim to be close to the SF model. On the other hand, the adverbial negative marker, *pas*, can freely co-occur with n-words without giving rise to a double negation reading.[48]

(62) a. No *ne*:

Eux autres te reconaissont pas, mais toi tu les recounais.
they you recognize NM, but you you them recognize

'They dont recognize you, but, you recognize them.' (Maillet, p. 28)

b. Just a single n-word:

Nous autres, parsoune s'en vient frotter chus nous.
we, no-one comes to-scrub by us

'But nobody ever comes scrubbing at our place.' (Maillet, p. 15)

c. Negtive Concord without *ne* and *pas*:

Parsoune s'en vient non plus laver nos hardes.
no-one comes neither to-wash our clothes

'And nobody comes to wash our clothes either.' (Maillet, p. 15)

d. Negative Concord with *pas*:[49]

Ah! Non, apparence que je faisait pas zire à parsoune de
oh! no, apparently that I made NM disgust to nobody of
mon temps.
my time

[48]We are grateful to Christian Richard for providing native judgements for AC. Additional data are taken from a contemporary play by Antoine Maillet, *La Sangouine*, Leméac, 1990, which is a very valuable source of negation data in AC.

[49]The Québecois data given in Muller (1991) suggest that this construction is quite frequent in Québecois. It is, however, rather the exception in the use in Maillet (1990).

'Oh! No, apparently I was far from disgusting in my time.' (Maillet, p. 23)

In certain respects, the data in (62) are parallel to SF; in others, they differ. Just as SF, Acadien has n-words that contribute a logical negation, which we have classified as n_1-words (62b). And, like in SF, *pas* contributes a negation to the meaning of the sentence (62a). So far, we could simply apply our analysis of section 4 to AC, eliminate *ne* from the lexicon, and predict the data correctly. The problem is, then, how to account for sentences as (62d), in which an n-word, *parsoune*, and *pas* co-occur without a double negation effect. (62d) can receive a negative concord interpretation, which is not present in the corresponding SF constructions.

Two possible modifications of our previous analysis suggest themselves: (a) in AC, *pas* is ambiguous between a logical negation and a negative polarity item, or (b) in AC, n-words show an ambiguity similar to the n-words of SF, but they obey slightly different collocation conditions. In other words, what we see in (62d) is an n_2-word, *parsoune*, which is licensed by the negation of *pas.* The second option, (b), seems to be clearly preferrable, since it provides a minimal modification of our analysis that preserves the explanation of the the NC reading of (62c) which we have suggested for the corresponding SF constructions. In (62c), *parsoune* can be analyzed as an n_1-word licensing the n_2-word *non plus.* This analysis is supported by (63), which supplies independent evidence for a non-negative use of n-words in AC:

(63) a. V'là la place la plus crasseuse que j'ai jamais forbie.
here is the place the most filthy that I have never cleaned
'That is the most filthy place I've ever cleaned.' (Maillet, p. 19)

b. Pas que je pouvions acheter rien de ça, nous autres, mais
not that I could buy nothing of that, we, but
je pouvions regarder.
I could look
'Even though I couldn't buy any of those, I could still look.' (Maillet, p. 32)

In (63b), *rien* is an n_2-word licensed by *pas que.* On the other hand, just as in SF, there are no non-negative instances of *pas* in AC. In (64), we have constructed sentences which, in SF, allow expletive *ne* (which does not exist in AC). Even in these contexts, the sentences with *pas* obligatorily receive a negative interpretation.

(64) a. J'ai peur qu'i' va pas venir.
I fear that he will NM come
'I am afraid he will not come.'
$ 'I am afraid he will come.'
b. À moins qu'i' vienne pas ...
unless he comes NM
'Unless he doesn't come ... '
$ 'Unless he comes ... '

To summarize, *pas* in AC behaves exactly like SF *pas*. For n-words, we assume a lexical ambiguity. Like in SF, n_1-words contribute logical negation and are negative existential quantifiers which may lexically contribute quantifiers for existential closure. N_2-words are indefinites which must be bound by existential closure in the scope of a negation. But in AC, in contrast to SF, the relative position in the logical form of sentences of the licensing negation of n_2-words and the existential quantifier that binds their variable is not constrained. The licensing negation might immediately precede the binding existential quantifier. The difference between SF and AC consists, thus, of the lack of the lexical element, *ne*, in AC, and of the collocation requirements of n_2-words. The logical forms of the lexical elements are the same in both varieties of French.

6.2. **Haitian Creole.** Déprez (1997a) and Déprez (1997b) document that HC lacks a pre-verbal negative particle, and the adverbial negative marker, *pa*, is obligatory for marking sentential negation. In complete sentences, n-words do not occur without the negative adverbial *pa* to express sentential negation.[50]

(65) a. Jan *(pa) we pèsòn.
Jan NM see nobody 'Jan did not see anyone.'
b. Pèsòn a pa di pèsòn anyen.
nobody has NM said nobody nothing
'Nobody said anything to anyone.'

These data indicate that *pa* is the lexical element that contributes logical negation. N-words, on the other hand, are best treated like the non-negative n_2-words of AC, since they lack a negative interpretation independent of the presence of *pa*.[51] Déprez (1997b) illustrates that n-words

[50]In the glosses, we do not always follow Déprez (1997a) and Déprez (1997b), where some instances of n-words are glossed as *any*-NPIs.
[51]Even though this is not mentioned explicitely in Déprez (1997a) and Déprez (1997b), a double negation reading seems to be unavailable with n-words in HC.

in HC have the distribution of French negative polarity items. They can occur in contexts where they are licensed by matrix negation. Déprez's data are repeated in (66) and (67).

(66) HC: M'pa te di pèsòn ap vini.
I NM have said no-one would come
'I did not say that anyone will come.'
SF: Je n'ai pas dit que *personne$_2$/ qui que ce soit allait
I NE have NM said that no-one/ anyone would
venir.
come

(67) HC: M'pa jamn konne koman Macoute la te tuye pèsòn.
I NM never knew how Mocoute this had killed no-one
'I never knew how this Macoute had killed anyone.'
SF: Je ne savait jamais comment ce Macoute avait tué
I NE knew never how this Macoute had killed
*personne$_2$/ qui que ce soit.
no-one/ anyone

In (66), the n-word *pèsòn* (*nobody*) appears with a non-negative interpretation in a complement clause of a negated predicate in the matrix clause. In the corresponding SF sentence, the NPI *qui que ce soit* must be used because there the n-word (*personne*) cannot be used in its non-negative meaning.[52] Example (67) illustrates the same point with an n-word in a non-subject position in an embedded interrogative clause. We leave a precise specification of the collocation restrictions of HC n-words to further investigation.

Our brief overview over negation systems that are closely related and yet different from Standard French has indicated that the account developed in this paper is well-suited for an investigation and analysis of cross-linguistic variations. An explanation for the variation of the French negation system is to be sought in the principles governing the logical forms and the collocation restrictions of the words which play a central role in expressing sentential negation.

[52]The corresponding example in Déprez (1997b, p. 110, ex. (10b)) does not contain *pas*. Note that the SF sentences with *personne* are not ungrammatical but have a meaning that is different from that of the HC sentences. In their grammatical readings, *personne* is interpreted as *personne$_1$*.

7. Comparison

In the present section, we point out in which respect our theory of logical form differs from the most widespread assumptions about semantics in HPSG, compare our linguistic theory of negation to previous work in HPSG, and, finally, discuss the conceptual relationship of our analysis to related work in GB.

7.1. **Semantics in HPSG.** The most common option for the description of scope phenomena available in HPSG is a Cooper-storage mechanism with phrasal quantifier retrieval as outlined in Pollard and Sag (1994), with optionally phrasal or lexical quantifier retrieval (Pollard and Yoo 1998), or with obligatorily lexical quantifier retrieval (Przepiórkowski 1998). All varieties employ very simplistic semantic structures inspired by situation semantics and do not provide a model-theoretic interpretation for these structures. The first approach suffers from serious empirical shortcomings since it is not able to account for all possible readings arising from the interaction of raising verbs with the quantifier scope of raised NPs (Pollard and Sag 1994, p. 328). The analysis of Pollard and Yoo (1998) was developed in response to these problems. When quantifier retrieval becomes obligatory at the word level as in Przepiórkowski (1998), the lexical theory has the remarkable property of already representing the meaning of sentences at their lexical heads. Consequently, the syntactic nodes in the phrase structure no longer correspond to applications of semantic composition rules. Instead, the meaning of a sentence originates at its syntactic head and is simply passed up in the syntactic structure. Especially in the context of linearization grammars, this architecture calls into question the purpose of phrase structure.

Our approach to semantics radically departs from the mainstream HPSG tradition. This move is mainly motivated by an attempt to draw on the wealth of linguistic research that has been conducted in semantic frameworks that employ higher order type-theoretic languages. Adopting a classical, model-theoretic semantics and logical forms that are standard in the linguistics literature makes the results and techniques of that research immediately available to HPSG, and vice versa. Moreover, this step establishes a close link to frameworks that posit a tree-configurational level of logical forms. Although there exist significant conceptual differences, our discussion has shown that many analytical insights can be retained. Supporting a fresh look from a new angle, the transparent logical foundations of HPSG should help to shed more light on the nature of conditions on logical forms and their interaction with syntactic well-formedness conditions in general. As HPSG, by its very design, is an eclectic framework, we consider it legitimate to step beyond its boundaries and espouse a semantic theory that hitherto has been neglected, *pace* Nerbonne (1992).

In the same spirit, our treatment of scope ambiguities is inspired by work in Categorial Grammar. Since we adopt the analysis of quantifier raising of Hendriks (1993), we can rely on its mathematical results, which ensure the right readings. As an immediate consequence, we know that our approach does not suffer from the empirical problems of Pollard and Sag (1994) concerning the interaction of raising verbs and quantified NPs. At the same time, we do not give up the idea of deriving the meaning of phrases from the meaning of their syntactic daughters by the application of semantic composition rules at each node in the syntactic structure.

7.2. **Work on negative concord in HPSG.** In a series of papers, Adam Przepiórkowski and Anna Kupść (P&K) developed a syntactic analysis of the distribution of sentential negation and negative concord in Polish and Italian.[53] Their analysis builds on the idea of a lexical treatment of unbounded dependencies presented by Sag (1997). Words amalgamate nonlocal properties of their dependents and pass them on to their phrasal projections as determined by grammatical principles. In P&K's view, NC is an unbounded dependency construction. Whereas earlier versions of the analysis of P&K were entirely syntactic, Przepiórkowski and Kupść (1999) introduce a special attribute, SENSITIVITY, on n-words, whose value characterizes them as looking for a semantic environment which contains an eventuality negation. In combination with principles that characterize certain syntactic projections as islands within which the semantic licensing of n-words must take place, this extension of their theory accounts for what we consider the collocational aspects of n-words and NPIs, but without embedding the analysis in a general theory of collocations.

We have argued that a collocation theory, based on a semantic analysis of certain lexical items that impose contextual restrictions on logical form in certain syntactic domains, can account for the whole range of data on its own. It seems, thus, to be more parsimonious and methodologically preferable to develop an account of the distribution of n-words solely

[53]Przepiórkowski and Kupść (1997a), Przepiórkowski and Kupść (1997), Przepiórkowski and Kupść (1997b), Kupść and Przepiórkowski (1998), Przepiórkowski and Kupść (1999), Przepiórkowski (1999).

on the basis of an independently and cross-linguistically motivated theory of collocations which expresses the mutual contingency of syntax and semantics displayed by NC phenomena. Further support for this architecture comes from the great diversity of French negative constructions and the cross-linguistic considerations of section 6. It is also important to keep in mind that the architecture of our collocation theory is modeled after a theory of unbounded dependencies[54]. It is, thus, well-equiped to formalize structural islands that constitute boundaries within which certain requirements must be satisfied. At the same time, a collocation analysis easily lends itself to a straightforward lexical treatment without necessitating an extension of the structure of signs beyond independently motivated structures.

For a more detailed comparative evalutation of the present analysis with the one pursued by P&K, our analysis would have to be enriched by a more fine-grained investigation of the syntactic licensing domains of n-words and by a semantic treatment of what Przepiórkowski and Kupść (1999) call non-eventuality negation. P&K, on the other hand, would have to show how their analysis can account for the great variety of lexical idiosyncrasies which we have discussed in the present paper.

7.3. **Work on negative concord in GB.** Finally, we briefly compare our HPSG account of negation and NC with two principal representatives of the Chomskyan tradition of GB theory.

Decidedly non-lexicalist analyses such as Zanuttini (1991), Haegeman and Zanuttini (1991) and Haegeman and Zanuttini (1996) posit uniform lexical entries for n-words but define alternative interpretational options at the level of Logical Form. Depending on the structural position of the n-words at some level of representation or derivation, operations are stipulated that allow multiple negations to reduce to a single negation (*negation factorization*). As far as we can see, these approaches have at least three problems. Firstly, it is unclear how some lexical items can be exempt from negation factorization operations. Secondly, dialectal variation must be accounted for in terms of the availability of such an option in each particular case in question and remains, for the lack of a differentiated formal analysis that could shed light on the nature of the underlying

[54]See Richter and Sailer (1999).

structures, essentially mysterious. Finally, the necessary operations invite a distinctly procedural view which we would like to avoid.[55]

The binding-theoretic approach of Progovac (1988) and Progovac (1994) is in certain respects similar to our approach. It uses the terminology of GB's binding theory to describe the distribution of negative polarity items and positive polarity items in Serbo-Croatian, and relies on GB-internal assumptions of domain-extension to express the long-distance licensing of certain negative polarity items. The approach suffers, however, from the fact that it is purely syntactic in nature, which we believe to be ultimately inadequate for a linguistic phenomenon that is closely related to the meaning of expressions.[56] It resembles our analysis, because the binding-theoretic metaphors can be directly expressed in the theory of collocations.

All GB-based approaches are basically syntactic in nature and rely on the existence of some level of Logical Form. We believe that primarily syntactic approaches are inadequate for an analysis of phenomena that involve semantic notions like negation, unless a theory of interpretation for logical forms exists and is explicitly used. For generative theories, a theory of logical forms has been explicated by von Stechow (1993) and Heim and Kratzer (1997). Acquaviva (1997) and Giannakidou (1997) develop a theory of negation within that framework. As we have seen, though, a tree-configurational syntactic level of logical forms is not necessary to account for the facts. In HPSG, it is sufficient to impose conditions on admissible forms of Ty2 formulae in the CONTENT of signs.

8. Open Questions

We have tried to integrate independent strands of current research in syntax and semantics and explored the explanatory potential of their emerging synthesis in a new approach to sentential negation in French. As is natural with such an enterprise, many questions remain to be answered before its merits and difficulties can be fully assessed.

---

[55] See Acquaviva (1997), Déprez (1997b) and Giannakidou (1998) for a detailed criticism from a GB-oriented perspective.

[56] See Hoeksema (1996) for a critical review of Progovac (1994). Błaszczak (1998b) is a recent attempt to express a theory of Polish n-words and NPIs with the descriptive devices of binding theory.

One basic problem with the focus on syntactic aspects of negative sentences in large parts of the literature is the lack of well-established, commonly acknowledged tests to decide on their logical form. In section 2, we discussed various tests to determine the form of the semantic contribution of French n-words. As in the corresponding discussion of Polish n-words in Richter and Sailer (1999), it became clear that no test known from the literature can be used as a decisive diagnostic which strictly excludes alternative solutions. Further research is needed to come to a full understanding of the significance of these tests and, hopefully, to the development of new and clear diagnostics.

Another important semantic issue we did not discuss are the licensing conditions of negative polarity items. Although they are usually phrased in terms of entailment requirements, it could be argued that, at least for some data, the relevant property is not entailment but the presence of negation or some other semantic operator in a certain domain of the logical form. In the present paper, we adopted the latter position.[57] To settle this question, one must find adequate characterizations of the logical forms that constitute so-called affective environments relevant for polarity item licensing. In particular, a fully explicit syntactic and semantic analysis for non-negative contexts such as questions and adversative predicates would be of interest.

A collocation theory in the spirit of van der Wouden (1992) can account for instances of apparent non-compositionality in natural language by appropriately restricting the contexts under which a certain interpretation of a lexical item is possible. Under a broader perspective, a collocation theory also accounts for the markedness of certain constructions. If the collocation requirements of an element are stricter, then it can appear in fewer, more specific contexts.

A collocation theory also gives us a handle on historical language change. Consider the meaning shift in negative constructions in French, as schematized in the Jesperson cycle. If, by their collocation requirements, two lexical elements are bound to appear usually in the same contexts, the individual semantic contribution of the given elements is not transparent in their most frequent contexts. Therefore, language learners are given a choice as to which semantics to attribute to which element. In French,

[57]See Przepiórkowski and Kupść (1999) for interesting observations that are problematic for an entailment-based approach but compatible with a logical form analysis.

this may have led to shifting the negative meaning from the pre-verbal particle, *ne*, to the former indefinite, *pas*.

The powerful collocation theory put forward in this paper might raise questions of learnability. Disregarding the fact that it is an open question how learnability considerations affect theoretical reasoning in HPSG, we would like to suggest tentatively that the solutions to learnability problems investigated in Fodor's (1992) LPSG (Learnable Phrase Structure Grammar) are applicable to our lexical approach. In particular, we might assume that whenever a new word is learned, the context in which it appears is in some form stored as the collocation requirements of that word. After a number of occurrences of the word have been encountered, a generalization procedure may extend the possible contexts of a word by generalizing over the attested occurrences.

9. Appendix: Formalization

We formalize our theory of logical form and the collocation theory of negative concord in Relational Speciate Re-entrant Logic (RSRL, (Richter 1999; Richter, Sailer, and Penn 1999)), a typed feature logic which extends SRL (King 1989; King 1994) by relations and substructural quantification, and provides a model theory for HPSG grammars similar to the one King (1999) presents for SRL. To make the formalization easily accessible, it employs the AVM syntax for RSRL which is defined in Richter (1999).

Sections 9.1 to 9.3 introduce the necessary technical background for the linguistic analysis of negative concord by defining a theory of Ty2 expressions as logical forms of linguistic signs in HPSG. Section 9.1 contains the sort hierarchy and principles which supply Ty2 as the semantic representation language. Section 9.2 presents the sort hierarchy and principles which are needed to formalize the Semantics Principle. The following section, 9.3, provides an explicit version of the Principle of Full Interpretation, which imposes special well-formedness conditions on the logical forms of certain distinguished signs. A list of the most important derivational rules in section 9.4 and the formalization of our collocation theory of negative concord in section 9.5 conclude the appendix.

9.1. **Formalization of Ty2.** The domain of *meaningful expressions* (*me*) of Ty2 is structured by the sort hierarchy and appropriateness conditions in (68) and (69). The subsort relation is indicated by indentation. For example, *me* is the supersort of all other sorts in (68), and *const* is the immediate supersort of the species $const_1$ to $const_n$. Attributes are written in a column behind the highest sort they are appropriate to, and their appropriateness is inherited by all subsorts of that sort. For example, TYPE is appropriate to all sorts in the hierarchy in (68). Moreover, the function denoted by the attribute TYPE maps objects which are of a maximally specific subsort of *me* to objects which are of a maximally specific subsort of *type*, as indicated by the sort symbol *type* in the second column behind *me*.

(68) The sort hierarchy of *me*:

me TYPE *type*
- *var(iable)*
- *const(ant)*
  - $const_1$
  - ⋮
  - $const_n$
- *appl(ication)* FUNC(TOR) *me*
  ARG(UMENT) *me*
- *abstr(action)* VAR *variable*
  ARG *me*
- *equ(ation)* ARG1 *me*
  ARG2 *me*
- *neg(ation)* ARG *me*
- *l(ogical)-const(ant)* AGR1 *me*
  ARG2 *me*
  - *dis(junction)*
  - *con(junction)*
  - *imp(lication)*
- *quant(ifiers)* VAR *variable*
  SCOPE *me*
  - *uni(versal)*
  - *exi(stential)*

(69) The sort hierarchy of *type*:

- *type*
  - *atomic-type*
    - *entity*
    - *truth*
    - *w-index*
  - *complex-type* IN *type* OUT *type*

The theory of Ty2 expressions must guarantee the proper types of the Ty2 expressions in the denotation of the grammar. The finiteness and non-cyclicity of the expressions can either be considered meta-assumptions about interpretable logical forms or expressed in RSRL. In (70), we state the necessary type restrictions for meaningful expressions.

(70) Type restrictions:

$$
\mathit{appl} \rightarrow \begin{bmatrix} \textsc{type} & \boxed{2} \\ \textsc{func type} & \begin{bmatrix} \textsc{in} & \boxed{1} \\ \textsc{out} & \boxed{2} \end{bmatrix} \\ \textsc{arg type} & \boxed{1} \end{bmatrix}, \quad \mathit{abstr} \rightarrow \begin{bmatrix} \textsc{type} & \begin{bmatrix} \textsc{in} & \boxed{1} \\ \textsc{out} & \boxed{2} \end{bmatrix} \\ \textsc{var type} & \boxed{1} \\ \textsc{arg type} & \boxed{2} \end{bmatrix},
$$

$$
\mathit{equ} \rightarrow \begin{bmatrix} \textsc{type} & \mathit{truth} \\ \textsc{arg1 type} & \boxed{1} \\ \textsc{arg2 type} & \boxed{1} \end{bmatrix}, \quad \mathit{neg} \rightarrow \begin{bmatrix} \textsc{type} & \mathit{truth} \\ \textsc{arg type} & \mathit{truth} \end{bmatrix},
$$

$$
\mathit{l\text{-}const} \rightarrow \begin{bmatrix} \textsc{type} & \mathit{truth} \\ \textsc{arg1 type} & \mathit{truth} \\ \textsc{arg2 type} & \mathit{truth} \end{bmatrix}, \quad \mathit{quant} \rightarrow \begin{bmatrix} \textsc{type} & \mathit{truth} \\ \textsc{scope type} & \mathit{truth} \end{bmatrix}
$$

To enforce non-cyclicity, we need a relation between a meaningful expression and its components. In $\texttt{component}(x, y)$, x is a component of y. The Non-Cyclicity Principle, (72), uses that relation for requiring that no complex meaningful expression occur in any proper subexpression of itself.

(71) $\texttt{component}(x, y) \overset{\forall}{\Longleftarrow}$

$$
\begin{array}{l} x = y \lor \\ \left(\begin{array}{l} \left(\begin{array}{l} {}^{y}\!\left[\textsc{arg}\ \boxed{1}\right] \lor {}^{y}\!\left[\textsc{arg1}\ \boxed{1}\right] \lor {}^{y}\!\left[\textsc{arg2}\ \boxed{1}\right] \\ \lor {}^{y}\!\left[\textsc{func}\ \boxed{1}\right] \lor {}^{y}\!\left[\textsc{scope}\ \boxed{1}\right] \lor {}^{y}\!\left[\textsc{var}\ \boxed{1}\right] \end{array} \right) \\ \land\ \texttt{component}(x, \boxed{1}) \end{array} \right) \end{array}
$$

(72) The Non-Cyclicity Principle:

$$me \rightarrow \forall \boxed{1} \left(\left(\left[\textsc{func}\ \boxed{1}\right] \vee \left[\textsc{arg}\ \boxed{1}\right] \vee \left[\textsc{arg1}\ \boxed{1}\right] \vee \left[\textsc{arg2}\ \boxed{1}\right] \vee \left[\textsc{scope}\ \boxed{1}\right] \right) \rightarrow \neg\ \texttt{component}(:,\boxed{1}) \right)$$

9.2. **Ty2 expressions as logical forms.** We want the expressions which represent the meaning of linguistic signs to be redex free. Therefore, a mechanism is needed that handles β-reduction. We choose to implement it as a recursive sort within CONTENT. For a treatment of β-reduction with a recursive sort in the CONTENT of signs, it is necessary to make the structure of the CONTENT value somewhat more complex than mentioned in the text. The meaningful expressions of Ty2 which specify the model theoretic meaning of a linguistic sign are the value of an attribute, LOGICAL-FORM (LF), appropriate to *content*. A second attribute, β-REDUCTION, is also appropriate to *content*, and contains recursively defined structures which specify the β-reduction of complex expressions.

(73) The sort hierarchies of *content* and *β-reduction*:

content	LOGICAL-FORM	*me*
	β-REDUCTION	*β-reduction*
β-reduction	TERM	*me*
	AUX	*me*
no-reduction		
one-reduction	REC	*β-reduction*

The idea of the recursive structure of *β-reduction* is that if no reduction step is performed, the AUX value equals the TERM value. If there is a reduction, REC AUX is deduced from AUX by one reduction step, and the TERM value is passed up from the bottom of the reduction chain deeper down in the structure:

(74) The theory of *β-reduction*:

$$\textit{no-reduction} \rightarrow \begin{bmatrix} \textsc{term} & \boxed{1} \\ \textsc{aux} & \boxed{1} \end{bmatrix}$$

$$\textit{one-reduction} \rightarrow \begin{bmatrix} \text{TERM} & \boxed{1} \\ \text{AUX} & \boxed{2} \\ \text{REC} & \begin{bmatrix} \text{TERM} & \boxed{1} \\ \text{AUX} & \boxed{3} \end{bmatrix} \end{bmatrix}$$
$$\wedge\ \texttt{find-redex}\left(\boxed{2}, {}^{y}\begin{bmatrix} \text{FUNC} & \begin{bmatrix} \text{VAR} & \boxed{5} \\ \text{ARG} & \boxed{4} \end{bmatrix} \\ \text{ARG} & \boxed{6} \end{bmatrix}, \boxed{3}, w\right)$$
$$\wedge\ \texttt{replace}(\boxed{4}, \boxed{5}, w, \boxed{6})$$
$$\wedge\ \texttt{replace}(\boxed{2}, y, \boxed{3}, w)$$

In the recursive case, the TERM value is passed up, and the reduction step is taken care of by the relations `find-redex` and `replace`. y is a redex in AUX. w is a subformula of REC AUX such that w is obtained from y by β-reduction.

(75) The relation `find-redex`:
Base case:
$$\texttt{find-redex}(x, y, v, w) \stackrel{\forall}{\Longleftarrow}$$
$${}^{x}\begin{bmatrix} \textit{appl} \\ \text{FUNC} \ \textit{abstr} \end{bmatrix} \wedge\ x = y\ \wedge\ v = w$$
An example of a recursive case:
$$\texttt{find-redex}(x, y, v, w) \stackrel{\forall}{\Longleftarrow}$$
$${}^{x}\begin{bmatrix} \textit{abstr} \\ \text{TYPE} & \boxed{1} \\ \text{VAR} & \boxed{2} \\ \text{ARG} & \boxed{3} \end{bmatrix} \wedge\ {}^{v}\begin{bmatrix} \textit{abstr} \\ \text{TYPE} & \boxed{1} \\ \text{VAR} & \boxed{2} \\ \text{ARG} & \boxed{4} \end{bmatrix} \wedge\ \texttt{find-redex}(\boxed{3}, y, \boxed{4}, w)$$

x is the input formula, y is a redex in x, v is the reduced input formula, and w is the reduced version of y. `find-redex` is used to identify a redex subterm y of x. In the base case, x itself is a redex, and is identical to y. In the recursive cases, the recursively structured subterms of x are examined for redexes.

(76) The relation `replace`:
The base cases:
$$\texttt{replace}(x, y, v, w) \stackrel{\forall}{\Longleftarrow}$$
$$x = y\ \wedge\ v = w$$
$$\texttt{replace}(x, y, v, w) \stackrel{\forall}{\Longleftarrow}$$
$$\neg x = y\ \wedge\ {}^{x}\left[\textit{var} \vee \textit{const}\right] \wedge\ x = v$$

An example of a recursive case:

$$\texttt{replace}(x,y,v,w) \stackrel{\forall}{\Longleftarrow}$$
$$\neg x = y \wedge \ {}^{x}\!\begin{bmatrix} \textit{abstr} & \\ \text{TYPE} & \boxed{1} \\ \text{VAR} & \boxed{2} \\ \text{ARG} & \boxed{3} \end{bmatrix} \wedge \ {}^{v}\!\begin{bmatrix} \textit{abstr} & \\ \text{TYPE} & \boxed{1} \\ \text{VAR} & \boxed{2} \\ \text{ARG} & \boxed{4} \end{bmatrix} \wedge \texttt{replace}(\boxed{3}, y, \boxed{4}, w)$$

v is obtained by replacing y by w in x. The recursion goes over the term structure of x. There are two base cases: if x is identical to y, then the reduced version of y, namely w, is identical to v. If x is not identical to y, but an atomic term, i.e., either a variable or a constant, then x is identical to v. In this case, we were checking a subterm of x that did not contain y. In the recursive case, the structure of x is copied to v, and the relation `replace` is applied to all recursive subterms of x.

Note that our β-reduction differs from the standard definition in textbooks (c.f. Hindley and Seldon (1986)): We never rename variables. Except for the variable whose occurrences are replaced in the reduction, the variables in the redex and in the reduced term are identical. Variables that are free in a β-redex may, thus, become bound in the reduced term. It follows that our β-reduction does not necessarily preserve the meaning of a term. This property of β-reduction is, in fact, a prerequisite for our treatment of indefinites with the Existential Closure DR, (84). There are, of course, alternative technical solutions for implementing existential closure for indefinites that employ the standard β-reduction. One way would be to introduce existential closure at the phrase level instead of at the word level. As far as we can see, there is no empirical difference that could distinguish between these two implementations. We adopt the present solution because it is easier to formalize and leads to a simpler Semantics Principle.

Given the theory of the sort *β-reduction*, we only need to define one more auxiliary relation, `redex-free`, before we can state the linguistically interesting principles on logical forms. The definition of `redex-free` in (77) uses the relation `component`, (71).

(77) The relation `redex-free`:

$$\texttt{redex-free}(x) \stackrel{\forall}{\Longleftarrow}$$
$$\neg\, \exists y \left({}^{y}\!\begin{bmatrix} \textit{appl} & \\ \text{FUNC} & \textit{abstr} \end{bmatrix} \wedge \texttt{component}(y,x) \right)$$

According to (77), a meaningful expression x is redex free exactly if it does not contain an application whose functor is an abstraction.

The β-Reduction Principle, (78), requires that the LF value of a *content* object be always fully β-reduced:

(78) The β-Reduction Principle:

$$content \rightarrow \begin{bmatrix} \text{LOGICAL-FORM } \boxed{1} \\ \beta\text{-REDUCTION TERM } \boxed{1} \end{bmatrix} \wedge \texttt{redex-free}(\boxed{1})$$

For ease of exposition, we assume an overall binary branching syntactic structure when we state the Semantics Principle (SP). It is easy to see how it could be modified to account for other assumptions about the recursive syntactic structure of linguistic signs. For each type of phrase, the SP provides the possibility of functional application of one syntactic daughter to the other. In conjunction with the theory of *β-reduction*, (74), and the β-Reduction Principle, (78), the SP says that the LOGICAL-FORM of a phrase is the redex free term which is derived by β-reduction from the term obtained by functional application of the LOGICAL-FORM value of one syntactic daughter to the LOGICAL-FORM value of the other syntactic daughter of the phrase.

(79) The Semantics Principle:

$$phrase \rightarrow \begin{bmatrix} \text{SYNS L CONT } \beta\text{-REDUCT} \begin{bmatrix} \beta\textit{-reduction} \\ \text{AUX} \begin{bmatrix} \textit{appl} \\ \text{FUNC } \boxed{3} \\ \text{ARG } \boxed{4} \end{bmatrix} \end{bmatrix} \\ \text{H-DTR SYNS L CONT LF } \boxed{1} \\ \text{NH-DTR SYNS L CONT LF } \boxed{2} \end{bmatrix}$$
$$\wedge((\boxed{1} = \boxed{3} \wedge \boxed{2} = \boxed{4}) \vee (\boxed{1} = \boxed{4} \wedge \boxed{2} = \boxed{3}))$$

9.3. **The Principle of Full Interpretation (PFI).** Traditionally, the PFI imposes well-formedness conditions on the possible logical forms of certain distinguished expressions of a natural language. Our version of the PFI governs the logical form of unembedded signs (Richter 1997), i.e., expressions which carry illocutionary force. Matrix sentences are a prototypical instance of unembedded signs.

(80) PFI

In the logical form of an unembedded sign,

- every quantifier binds an occurrence of a variable,
- there are no free variables, and
- every logical operator ($\lambda, \forall, \exists$) binds a distinct variable.

For the formalization of (80), we need a relation, `free`, which relates variables to the complex expressions in which they occur free. In (81), x occurs free in y:

(81) The relation `free`:

$$\texttt{free}(x,y) \stackrel{\forall}{\Longleftarrow} x = y$$

$$\texttt{free}(x,y) \stackrel{\forall}{\Longleftarrow} {}^{y}\begin{bmatrix} \textit{appl} & \\ \text{FUNC} & \boxed{1} \\ \text{ARG} & \boxed{2} \end{bmatrix} \wedge \left(\texttt{free}(x, \boxed{1}) \vee \texttt{free}(x, \boxed{2})\right)$$

$$\texttt{free}(x,y) \stackrel{\forall}{\Longleftarrow} {}^{y}\begin{bmatrix} \text{ARG1} & \boxed{1} \\ \text{ARG2} & \boxed{2} \end{bmatrix} \wedge \left(\texttt{free}(x, \boxed{1}) \vee \texttt{free}(x, \boxed{2})\right)$$

$$\texttt{free}(x,y) \stackrel{\forall}{\Longleftarrow} {}^{y}\begin{bmatrix} \textit{neg} & \\ \text{ARG} & \boxed{1} \end{bmatrix} \wedge \texttt{free}(x, \boxed{1})$$

$$\texttt{free}(x,y) \stackrel{\forall}{\Longleftarrow} {}^{y}\begin{bmatrix} \textit{quant} & \\ \text{VAR} & \boxed{1} \\ \text{SCOPE} & \boxed{2} \end{bmatrix} \wedge \neg x = \boxed{1} \wedge \texttt{free}(x, \boxed{2})$$

$$\texttt{free}(x,y) \stackrel{\forall}{\Longleftarrow} {}^{y}\begin{bmatrix} \textit{abstr} & \\ \text{VAR} & \boxed{1} \\ \text{ARG} & \boxed{2} \end{bmatrix} \wedge \neg x = \boxed{1} \wedge \texttt{free}(x, \boxed{2})$$

We can now express the PFI without any further definitions:

(82) $\textit{unembedded-sign} \quad \rightarrow$

$$\begin{bmatrix} \text{SYNS L CONT } \boxed{1} \end{bmatrix}$$

$$\text{a. } \wedge \forall x \left(\begin{array}{l} \texttt{component}\left({}^{x}\begin{bmatrix}\textit{quant}\end{bmatrix}, \boxed{1}\right) \\ \rightarrow \exists y \exists \boxed{2} \left(\texttt{component}(y, \boxed{2}) \wedge {}^{x}\begin{bmatrix} \text{VAR} & y \\ \text{SCOPE} & \boxed{2} \end{bmatrix}\right) \end{array} \right)$$

$$\text{b. } \wedge \forall x \left(\texttt{component}\left({}^{x}\begin{bmatrix}\textit{var}\end{bmatrix}, \boxed{1}\right) \rightarrow \neg\, \texttt{free}(x, \boxed{1})\right)$$

$$\text{c. } \wedge \forall x \forall y\, \forall \boxed{3}\, \forall \boxed{4}$$

$$\left(\begin{array}{l} \left(\texttt{component}\left({}^{x}\begin{bmatrix}\text{VAR } \boxed{3}\end{bmatrix}, \boxed{1}\right) \wedge \texttt{component}\left({}^{y}\begin{bmatrix}\text{VAR } \boxed{4}\end{bmatrix}, \boxed{1}\right) \wedge \neg x = y\right) \\ \rightarrow \neg \boxed{3} = \boxed{4} \end{array} \right)$$

9.4. **Derivational rules.** With the term *derivational rules* we refer to object-level lexical rules in the sense of Meurers (1995). Derivational rules relate *word* objects to *word* objects as opposed to meta-level lexical rules, which relate descriptions of words to descriptions of words. Since the term "lexical rules" is often understood in the latter sense, we choose a different terminology in order to prevent confusion.

Type-shifting rules that account for scope ambiguities between quantified NPs as arguments of verbs are encoded in the derivational rule (83). The idea behind (83) is to turn a verb into a functor if it combines with a quantified NP. If (83) is applied twice, a wide scope reading for a quantified NP in object position is achieved. (83) is an example for how the type-shifting rules of Hendriks (1993) can be imported to HPSG.

(83) The Quantifier Raising DR (QR-DR)

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT} \quad \lambda x_1 \ldots \lambda x_i \ldots \lambda x_n . \phi \end{bmatrix} \Longrightarrow$$

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT} \quad \lambda x_1 \ldots \lambda X_i \ldots \lambda x_n . X_i(\lambda x_i . \phi) \end{bmatrix}$$

We introduce existential closure at the lexical level with another derivational rule. In combination with the PFI, (82), which forbids the presence of free variables in the logical forms of unembedded signs, this DR introduces an unselective existential quantifier which binds all variables which would otherwise remain free.

(84) The Existential Closure DR ($\exists$-DR)

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L} \begin{bmatrix} \text{CAT HEAD VFORM } \textit{finite} \end{bmatrix} \\ \text{CONT} \quad \lambda x_1 \ldots \lambda x_n . \phi \end{bmatrix} \Longrightarrow$$

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT } \lambda x_1 \ldots \lambda x_n . \exists y_1 \ldots \exists y_m \phi \end{bmatrix}$$

Note that this treatment of existential closure at the lexical level implies a definition of β-reduction which allows variables to enter into the scope of a quantifier in a reduction step.[58] This is exactly how we defined the theory of *β-reduction*, (74). More precisely, the definition of `replace`, (76), is responsible for the nonstandard behavior of β-reduction.

The Adjunct Introduction DR augments the semantic valency of a word by one argument. It does not change the word's syntactic valency.

(85) The Adjunct Introduction DR (AI-DR)

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT } \lambda x_1 \ldots \lambda x_i \lambda x_{i+1} \ldots \lambda x_n . \phi \end{bmatrix} \Longrightarrow$$

[58] See Richter and Sailer (1999) for an example.

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L CONT } \lambda x_1 \ldots \lambda x_i \lambda a \lambda x_{i+1} \ldots \lambda x_n . a(\phi) \end{bmatrix}$$

The *pas*-DR is a variant of the AI-DR. It augments both, the syntactic and the semantic valency.

(86)
$$\begin{bmatrix} \textit{word} \\ \text{SYNS L} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM } \textit{finite} \end{bmatrix} \\ \text{VAL} & \text{COMPS } \boxed{1} \\ \text{MARKING} & \textit{unmarked} \end{bmatrix} \\ \text{CONT } \lambda x_1 \ldots \lambda x_n . \phi \end{bmatrix} \end{bmatrix} \Longrightarrow$$

$$\begin{bmatrix} \textit{word} \\ \text{SYNS L} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{VAL} & \text{COMPS } \left\langle \begin{bmatrix} \textit{synsem} \\ \text{L CAT H MOD L CAT MARKING } \boxed{2} \end{bmatrix} \middle| \boxed{1} \right\rangle \\ \text{MARKING } \boxed{2} \end{bmatrix} \\ \text{CONT } \lambda a \lambda x_1 \ldots \lambda x_n . a(\phi) \end{bmatrix} \end{bmatrix}$$

9.5. **Formalization of the collocation theory.** In this section, a few relations will be presupposed without definition: `collocation-cc`(x) describes what constitutes a complete clause, x, in French (or, correspondingly, in any other language under consideration). In `dominate`(x, y), phrase x (reflexively) dominates sign y; and in `member`(x, y), x is a member of list y.

Because we lack a formal semantic analysis of the non-negative contexts that license the occurrence of French n_2-words, we restrict the formalization to cases of logical negation.

(87) Example of a lexical entry of an n_2-word:

$$\begin{bmatrix} \textit{word} \\ \text{PHON } \langle \textit{personne} \rangle \\ \text{SYNS L CONT } \lambda P.P(x) \\ \text{COLL } \left\langle \begin{bmatrix} cc \\ \text{LIC } \boxed{1} \; me \end{bmatrix} \right\rangle \end{bmatrix}$$

$$\wedge\, \exists v \exists w \left(\begin{array}{l} \texttt{component} \left({}^{v}\begin{bmatrix} \textit{neg} \\ \text{ARG } \boxed{2} \end{bmatrix}, \boxed{1} \right) \\ \wedge\, \texttt{component}(w, v) \;\wedge\; \neg\, w = \boxed{2} \;\wedge\; {}^{w}\begin{bmatrix} \textit{exists} \\ \text{VAR } x \end{bmatrix} \end{array} \right)$$

Note that, in (87), the description of the lambda term employs an abbreviatory notation instead of AVM syntax, namely the traditional syntax of lambda terms. Thus, the notation, x, for a variable in that lambda term is an abbreviation which must be distinguished from variables of the AVM description language, such as v and w. In the description above, x abbreviates a path from the *word* object being described to the variable object in the Ty2 expression that corresponds to x, and ${}^{w}\!\left[\textsc{var}\ x\right]$ is a meta-syntactic notation which indicates that the value of the path x and the value of the path wVAR are identical.

Beside n-words, the pre-verbal particle, *ne*, has been of special interest in our discussion of French negation. Following Recourcé (1995) we assume that *ne* is a marker in infinitival constructions and an affix if combined with finite verbs. In (88), the lexical entry of the marker *ne* is given. The marker *ne* itself is semantically empty ($\lambda p.p$), but it requires that the clause in which it appears contain a logical negation over its main event. For the cases in which *ne* occurs as affix of finite verbs, we assume that the affix derivational rule that licenses its appearance on a finite verb also introduces the collocation condition of the marker in (88) on the verb. That means the COLL attribute of finite verbs with the affix *ne* has a value as described for the marker *ne* in (88), and the relation `contains-event-negation` holds of its licensing content.

(88) The lexical entry of the marker *ne*:

$$\begin{bmatrix} \textit{word} & \\ \textsc{phon} & \langle \textit{ne} \rangle \\ \textsc{synss l} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{marker} \\ \textsc{spec}\ \ \mathrm{V}[\textit{inf}] \end{bmatrix} \\ \textsc{marking} & \textit{NE} \end{bmatrix} \\ \textsc{cont} & \lambda p.p \end{bmatrix} \\ \textsc{coll} & \left\langle \begin{bmatrix} \textit{cc} \\ \textsc{lic}\ \boxed{1}\ \textit{me} \end{bmatrix} \right\rangle \end{bmatrix}$$
$$\wedge\ \texttt{contains-event-negation}(\boxed{1})$$

The relation `contains-event-negation` holds of an *me* object iff the main event in that *me* is negated. In a fully developed semantic analysis, the logical forms would contain explicit event variables. Using these variables, it could be determined whether an event is in the scope of a negation or not. For the simpler analysis and the limited fragment of French of the present paper, it is still possible to express the necessary conditions, although in a more roundabout fashion. Consider the following three sentences and their logical forms:

(89) a. Quelqu'un n'a rien vu.
$\exists x[\mathsf{human}'(x) \wedge \neg\exists y[\mathsf{thing}'(y) \wedge \mathsf{see}'(x,y)]]$
b. Un auteur que Pierre ne connais pas, (*n') a écrit *Le premier homme.*
$\exists x[[\mathsf{author}'(x) \wedge \neg\mathsf{know}'(p,x)] \wedge \mathsf{write}'(x, \textit{Le premier homme})]$
c. Jean n'a rien dit à personne.
$\neg\exists x\exists y[\mathsf{human}'(y) \wedge [\mathsf{thing}'(x) \wedge \mathsf{say}'(j,x,y)]]$

The main event of the propositions expressed by the logical forms in (89) can be recovered by describing the possible paths to it on the basis of the hierarchical structure of the terms (or the relative embedding of the corresponding *me* objects). The main event is always expressed as a non-logical constant. That constant can be embedded recursively in the functor of a functional application, under a negation, under a quantifier, or in the second argument of a logical constant. This observation is employed in the relation `main-event` in (91). It holds of a term iff there is an event in that term. To ensure that that that event is in the scope of a negation, the relation `contains-event-negation` is defined to hold of its argument exactly if it contains a negation whose argument contains the event. The negation may, of course, be recursively embedded in the scope of a quantifier or in the second argument of a logical constant.

(90) The relation `contains-event-negation`:

$$\texttt{contains-event-negation}(x) \overset{\forall}{\Longleftarrow}$$
$$\,^{x}\!\begin{bmatrix} \textit{neg} \\ \text{ARG } \boxed{1} \end{bmatrix} \wedge \texttt{main-event}(\boxed{1})$$
$$\texttt{contains-event-negation}(x) \overset{\forall}{\Longleftarrow}$$
$$\,^{x}\!\begin{bmatrix} \textit{l-const} \\ \text{ARG2 } \boxed{1} \end{bmatrix} \wedge \texttt{contains-event-negation}(\boxed{1})$$
$$\texttt{contains-event-negation}(x) \overset{\forall}{\Longleftarrow}$$
$$\,^{x}\!\begin{bmatrix} \textit{quant} \\ \text{SCOPE } \boxed{1} \end{bmatrix} \wedge \texttt{contains-event-negation}(\boxed{1})$$

(91) The relation `main-event`:

$$\texttt{main-event}(x) \overset{\forall}{\Longleftarrow} \,^{x}\!\begin{bmatrix} \textit{const} \end{bmatrix}$$
$$\texttt{main-event}(x) \overset{\forall}{\Longleftarrow} \,^{x}\!\begin{bmatrix} \textit{appl} \\ \text{FUNC } \boxed{1} \end{bmatrix} \wedge \texttt{main-event}(\boxed{1})$$

$$\texttt{main-event}(x) \overset{\forall}{\Longleftarrow} {}^{x}\begin{bmatrix} \textit{neg} \\ \text{ARG } \boxed{1} \end{bmatrix} \wedge \texttt{main-event}(\boxed{1})$$
$$\texttt{main-event}(x) \overset{\forall}{\Longleftarrow} {}^{x}\begin{bmatrix} \textit{l-const} \\ \text{ARG2 } \boxed{1} \end{bmatrix} \wedge \texttt{main-event}(\boxed{1})$$
$$\texttt{main-event}(x) \overset{\forall}{\Longleftarrow} {}^{x}\begin{bmatrix} \textit{quant} \\ \text{SCOPE } \boxed{1} \end{bmatrix} \wedge \texttt{main-event}(\boxed{1})$$

It is illuminating to compare the sketch of the lexical entry of French *personne* in (87) to a sketch of the lexical entry of Polish *nikt*, as it is needed for the theory of negative concord in Polish which is developed in Richter and Sailer (1999). Note that, apart from the phonology, the descriptions of the two n-words only differ in the last line of their collocation restrictions. Polish *nikt* requires that the variable it introduces be bound by an existential quantifier in a sequence of existentials which is immediately preceded by a negation.

(92) The lexical entry of Polish *nikt*:

$$\begin{bmatrix} \textit{word} \\ \text{PHON } \langle \textit{nikt} \rangle \\ \text{SYS L CONT } \lambda P.P(x) \\ \text{COLL } \left\langle \begin{bmatrix} cc \\ \text{LIC } \boxed{1}\ me \end{bmatrix} \right\rangle \end{bmatrix}$$
$$\wedge\ \exists v\ \texttt{component}\left({}^{v}\begin{bmatrix} \textit{neg} \\ \text{ARG } \boxed{2} \end{bmatrix}, \boxed{1} \right)$$
$$\wedge\ \texttt{sequence-of-existentials-binding-var}(\boxed{2}, x)$$

(93) The relation `sequence-of-existentials-binding-var`:

$$\texttt{sequence-of-existentials-binding-var}(x, y) \overset{\forall}{\Longleftarrow} {}^{x}\begin{bmatrix} \textit{exists} \\ \text{VAR } y \end{bmatrix}$$
$$\texttt{sequence-of-existentials-binding-var}(x, y) \overset{\forall}{\Longleftarrow} {}^{x}\begin{bmatrix} \textit{exists} \\ \text{SCOPE } \boxed{1} \end{bmatrix} \wedge \texttt{sequence-of-existentials-binding-var}(\boxed{1}, y)$$

The definition of the CCCP, (40), is straightforward. The relation `collocation-cc` is abbreviated as `coll-cc`.

(94) The Clausal Content Collocation Principle (CCCP):

$$\forall b \left(\begin{array}{l} \left(\texttt{dominate}\left(\left[\textit{sign}\right], b\right) \wedge \texttt{coll-cc}(b)\right) \rightarrow \\ \forall w\, \forall x\, \forall\boxed{1}\,\forall\boxed{2}\,\forall\boxed{3} \\ \left(\begin{array}{l} \left(\begin{array}{l} \texttt{dominate}\left({}^{b}\left[\text{SYNS L CONT LF } \boxed{1}\right], {}^{w}\begin{bmatrix} \textit{word} \\ \text{COLL } \boxed{3}\end{bmatrix}\right) \\ \wedge\, \texttt{member}\left({}^{x}\begin{bmatrix}\textit{cc}\\ \text{LIC } \boxed{2}\end{bmatrix}, \boxed{3}\right) \\ \wedge\, \forall c \left(\begin{array}{l} (\neg c = b \wedge \texttt{coll-cc}(c) \wedge \texttt{dominate}(c,w)) \\ \rightarrow \texttt{dominate}(c,b)\end{array}\right) \end{array}\right) \\ \rightarrow \boxed{1} = \boxed{2} \end{array}\right) \end{array}\right)$$

b is the licensing complete clause, w is an n-word licensed by b. The fifth and sixth line of the CCCP say that there is no other intervening complete clause between the word, w, and the licensing complete clause, b.

References

Abeillé, A. and D. Godard (1997). The syntax of French negative adverbs. In D. Forget, P. Hirschbühler, F. Martineau, and M.-L. Rivero (Eds.), *Negation and Polarity*, pp. 1–27. Benjamins.

Acquaviva, P. (1997). *The Logical Form of Negation: A Study of Operator-Variable Structures in Syntax.* Garland.

Baker, C. L. (1991). The syntax of English *not*: The limits of core grammar. *Linguistic Inquiry 22*(3), 387–429.

Bittner, M. (1994). Cross-linguistic semantics. *Linguistics and Philosophy 17*, 53–198.

Błaszczak, J. (1998a). Some reflexions on the interaction betweem indefinites and the clause: The case of indefinite pronouns. In M. Giger and B. Wiemer (Eds.), *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)*, Volume 1, pp. 55–65. München. In press.

Błaszczak, J. (1998b). Towards a binding analysis of negative polarity items in Polish. Draft, based on a talk given at Formal Description of Slavic Languages II, Potsdam, 1997.

Brown, D. R. (1995). Term operators. In A. M. Bolkestein, C. de Groot, and J. L. Mackenzie (Eds.), *Predicates and Terms in Functional Grammar*, pp. 127–145. Foris, Dordrecht.

Copestake, D. Flickinger, R. Malouf, S. Riehemann, and I. Sag (1995). Translation using Mininimal Recursion Semantics. In *Proceedings of The Sixth International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-95)*, Leuven.

Corblin, F. (1994). Multiple negation processing. Research Paper HCRC/RP-62, Human Communication Research Center.

Corblin, F. (1995). Compositionality and complexity in multiple negation. *IGPL Bulletin 3*(3–2), 449–473.

Cristea, T. (1971). *La structure de la phrase négative en français contemporain*. Société Roumaine de Linguistique Romane, Bucarest.

Déprez, V. (1997a). A non-unified analysis of negative concord. In D. Forget, P. Hirschbühler, F. Martineau, and M.-L. Rivero (Eds.), *Negation and Polarity*, pp. 53–74. Benjamins.

Déprez, V. (1997b). Two types of negative concord. *Probus 9*, 103–143.

Dik, S. C. (1997). *The Theory of Functional Grammar, Part 2: Complex and Derived Structures*. Mouton de Gruyter. Edited by Kees Hengeveld.

Fodor, J. D. (1992). Islands, learnability and the lexicon. In H. Goodluck and M. Rochemont (Eds.), *Island Constraints: Theory, Acquisition and Processing*. Dordrecht: Kluwer.

Gaatone, D. (1971). *Étude descriptive du système de la négation en français contemporain*. Droz, Genève.

Giannakidou, A. (1997). *The Landscape of Polarity Items*. Ph. D. thesis, Rijksuniversiteit Groningen.

Giannakidou, A. (1998). *Polarity Sensitivity as Nonveridical Dependency*. John Benjamins. In preparation.

Grevisse, M. and A. Goosse (1993). *Le bon usage. Grammaire française* (13 ed.). Duculot.

Groenendijk, J. and M. Stokhof (1997). Questions. In J. van Benthem and A. ter Meulen (Eds.), *Handbook of Logic and Language*, pp. 1055–1124. Elsevier.

Haegeman, L. (1997). The syntax of n-words and the Neg criterion. In D. Forget, P. Hirschbühler, F. Martineau, and M.-L. Rivero (Eds.), *Negation and Polarity*, Current Issues in Linguistic Theory, pp. 115–137. Benjamins.

Haegeman, L. and R. Zanuttini (1991). Negative Heads and the Neg Criterion. *The Linguistsic Review 8*, 233–251.

Haegeman, L. and R. Zanuttini (1996). Negative concord in West Flemish. In A. Belletti and L. Rizzi (Eds.), *Parameters and Functional Heads: Essays in Comparative Syntax*, pp. 117–179. Oxford University Press.

Heim, I. and A. Kratzer (1997). *Semantics in Generative Grammar*. Blackwell.

Hendriks, H. (1993). *Studied Flexibility.* ILLC Dissertation Series 1995-5. Institute for Logic, Language and Computation, Amsterdam.

Hindley, J. R. and J. P. Seldon (1986). *Introduction to Combinators and the λ-Calculus.* Cambridge University Press.

Hinrichs, E., D. Meurers, F. Richter, M. Sailer, and H. Winhart (1997). Ein HPSG-Fragment des Deutschen, Teil 1: Theorie. Arbeitspapiere des SFB 340, Nr. 95, Universität Tübingen.

Hoeksema, J. (1996). Review of Progovac 1994. *Studies in Language 20*(1), 196–207.

Höhle, T. N. (1999). An architecture for phonology. In R. D. Borsley and A. Przepiórkowski (Eds.), *Slavic in HSPG*, pp. 61–90. CSLI Publications. In press.

Jones, M. A. (1996). *Foundations of French Syntax.* Cambridge University Press.

Kim, J.-B. (1996). *The Grammar of Negation: A Lexicalist Constraint-Based Perspective.* Ph. D. thesis, Stanford University.

Kim, J.-B. and I. Sag (1996). French and English negation: A lexicalist alternative to head movement. Unpublished manuscript.

King, P. J. (1989). *A logical Formalism for Head-Driven Phrase Structure Grammar.* Ph. D. thesis, University of Manchester.

King, P. J. (1994). An expanded logical formalism for head-driven phrase structure grammar. Arbeitspapiere des SFB 340, Nr. 59, Universität Tübingen.

King, P. J. (1999). Towards truth in HPSG. This volume.

Kupść, A. and A. Przepiórkowski (1998). Morphological aspects of verbal negation in Polish. Paper delivered at the Second International Conference on Formal Description of Slavic Languages, Potsdam, November 20–22, 1997. To appear in the proceedings.

Ladusaw, W. (1995). Configurational expression of negation. In J. van der Does and J. van Eijck (Eds.), *Quantifiers, Logic and Language*, pp. 203–223. CSLI, Stanford University.

Laka, I. (1990). *Negation in Syntax: On the Nature of Functional Categories and Projections.* Ph. D. thesis, MIT.

Maillet, A. (1990). *La Sangouine.* Leméac.

Meurers, W. D. (1995). Towards a semantics for lexical rules as used in HPSG. In *Proceedings of the Conference on Formal Grammar*, Barcelona. Also in: Proceedings of the ACQUILEX II Workshop on Lexical Rules, 1995, Cambridge, UK.

Muller, C. (1991). *La Négation en Français. Syntaxe, sémantique et éléments de comparaison avec les autres langues romanes.* Librairie Droz, Genève.

Muller, C. (1997). *De* partitif et la négation. In D. Forget, P. Hirschbühler, F. Martineau, and M.-L. Rivero (Eds.), *Negation and Polarity*, pp. 251–270. Benjamins.

Nerbonne, J. (1992). Constraint-based semantics. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the 8th Amsterdam Colloquium*, pp. 425 – 444. Institute for Logic, Language and Information.

Partee, B. (1987). Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh, and M. Stokhof (Eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, pp. 115–144. Foris, Dordrecht.

Pollard, C. and I. A. Sag (1994). *Head-Driven Phrase Structure Grammar.* University of Chicago Press.

Pollard, C. J. and E. J. Yoo (1998). A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics 34*, 415–445.

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

Progovac, L. (1988). *A Binding Approach to Polarity Sensitivity.* Ph. D. thesis, University of Southern California.

Progovac, L. (1994). *Negative and positive polarity.* Cambridge University Press.

Przepiórkowski, A. (1998). 'A Unified Theory of Scope' revisted. In *Proceedings of the FHCG-98, 14–16 August 1998, Saarbrücken*, pp. 185–195.

Przepiórkowski, A. (1999). Negative Polarity Questions and Italian Negative Concord. This volume.

Przepiórkowski, A. and A. Kupść (1997a). Negative Concord in Polish. Technical Report 828, Institute for Computer Science, Polish Academy of Sciences.

Przepiórkowski, A. and A. Kupść (1997b). Unbounded negative concord in Polish: A lexicalist HPSG approach. In J. Landsbergen, J. Odijk, K. van Deemter, and G. V. van Zanten (Eds.), *Computational Linguistics in the Netherlands 1996. Papers from the Seventh CLIN Meeting*, pp. 129–144.

Przepiórkowski, A. and A. Kupść (1997). Verbal negation and complex predicate formation in Polish. In *Proceedings of the 1997 Conference of the Texas Linguistics Society on the Syntax and Semantics of Predication*, Number 38 in Texas Linguistic Forum, pp. 247–261.

Przepiórkowski, A. and A. Kupść (1999). Eventuality negation and negative concord in Polish and Italian. In R. D. Borsley and A. Przepiórkowski (Eds.), *Slavic in HPSG*, pp. 211–246. CSLI Publications. In press.

Recourcé, G. (1995). *L'association négative en français: Étude linguistique de la particule* ne. Ph. D. thesis, Université de Paris 7.

Recourcé, G. (1996). Une double analyse de la particule *ne* en français. *Langages 30*(122), 62–78.

Richter, F. (1997). Die Satzstruktur des Deutschen und die Behandlung langer Abhängigkeiten in einer Linearisierungsgrammatik. Formale Grundlagen und Implementierung in einem HPSG-Fragment. In (Hinrichs, Meurers, Richter, Sailer, and Winhart 1997).

Richter, F. (1999). RSRL for HPSG. This volume.

Richter, F. and M. Sailer (1997). Underspecified semantics in HPSG. In H. Bunt, L. Kievit, R. Muskens, and M. Verlinden (Eds.), *Proceedings of the Second International Workshop on Computational Semantics, 8–10 January 1997, Tilburg*, pp. 234–246.

Richter, F. and M. Sailer (1999). LF Conditions on Expressions of Ty2: An HPSG Analysis of Negative Concord in Polish. In R. D. Borsley and A. Przepiórkowski (Eds.), *Slavic in HPSG*, pp. 247–282. CSLI Publications. In press.

Richter, F., M. Sailer, and G. Penn (1999). A formal interpretation of relations and quantification in HPSG. This volume.

Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy 12*, 683–721.

Rowlett, P. (1998). A non-overt negative operator in French. *PROBUS International Journal of Latin and Romance Linguistics 10*(2), 185–206.

Sag, I. A. (1997). English relative clause constructions. *Journal of Linguistics 33*(2), 431–484.

Sturm, J. (1981). *Morpho-syntaktische Untersuchungen zur* phrase négative *im gesprochenen Französisch. Die Negation mit und ohne NE.* Verlag Peter Lang.

van der Wouden, T. (1992). Beperkingen op het optreden van lexicale elementen. *De Nieuwe Taalgids 85*(6), 513–538.

van der Wouden, T. and F. Zwarts (1993). A semantic analysis of negative concord. In U. Lahiri and A. Z. Wyner (Eds.), *SALT III: Proceedings of the Third Conference on Semantics and Linguistic Theory*, pp. 202–219.

von Stechow, A. (1993). Die Aufgaben der Syntax. In J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (Eds.), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*, Volume 1, pp. 1–88. Walter de Gruyter.

Wilmet, M. (1997). *Grammaire critique du Français*. Duculot.

Zanuttini, R. (1991). *Syntactic Properties of Sentential Negation. A Comparative Study of Romance Languages*. Ph. D. thesis, University of Pennsylvania.

Frank Richter

Seminar für Sprachwissenschaft
SFB 340, B8
Kleine Wilhelmstr. 113
72074 Tübingen
Germany

E-mail address: `fr@sfs.nphil.uni-tuebingen.de`
Homepage: `http://www.sfs.nphil.uni-tuebingen.de/~fr/`

Manfred Sailer

Seminar für Sprachwissenschaft
Abt. Computerlinguistik
Kleine Wilhelmstr. 113
72074 Tübingen
Germany

E-mail address: `mf@sfs.nphil.uni-tuebingen.de`
Homepage: `http://www.sfs.nphil.uni-tuebingen.de/~mf/`