Negation, Double Negation and Optimality Theory

Mark Newson

1 Optimality Theory

1.1 General Introduction

There are two important aspects to Optimality Theory (OT) relevant to this paper: the structure of the theory itself and the way the theory accounts for cross linguistic variation.

The structure of the theory is relatively simple and can be represented thus:

(1)

$$
\begin{array}{lllll}
 & & \lceil \longrightarrow & \mathrm{E} \to \star & \\
 & & | \longrightarrow & \mathrm{V} \to \star & \\
 & & | \longrightarrow & \mathrm{A} \to \star & \\
\text{input} \longrightarrow & \text{GEN} & \dashv \longrightarrow & \mathrm{L} \to \star & \\
 & & | \longrightarrow & \mathrm{U} \to \star & \\
 & & | \longrightarrow & \mathrm{A} \longrightarrow & \Rightarrow \\
 & & \lfloor \longrightarrow & \mathrm{T} \to \star & \\
 & & & \vdots\, \mathrm{E} \,\vdots &
\end{array}
$$

An input of some sort (in phonology it is usually taken to be a lexical representation) is passed through the GEN component ('generate'), which affects the input in various ways to produce a potentially infinite number of outputs, known as **candidates.** The candidates are then passed through the **evaluation,** where at least one will be chosen as the optimal output (denoted $\Rightarrow$) and hence will be the grammatical surface form associated with the input: all non-optimal outputs are ungrammatical.

The evaluation is made up of a number of universal output constraints. Importantly, the constraints are not inviolable: in fact, they are often conflicting, so that in order to satisfy one, a candidate will have to violate another. The optimal output is the one that is the most harmonious with the set of constraints in a sense to be defined.

Another important factor concerning the constraints is that they are ranked. The candidate set is presented to each constraint in turn, starting with the most highly ranked. Candidates drop out of the candidate set if they violate the constraint and there is at least one surviving candidate. If there are no surviving candidates, then nothing is eliminated from the candidate set and the next highly ranked constraint is considered. The process continues until there is only one surviving candidate, the optimal one, or there are no more constraints, in which case any surviving candidate will be optimal.

The following abstract example serves to highlight the points made above:

(2)

	C1	C2	C3	C4
1	$\surd$	$\star$!	($\surd$)	($\surd$)
⇒ 2	$\surd$	$\surd$	$\star$	$\surd$
3	$\star$!	($\surd$)	($\star$)	($\star$)
4	$\surd$	$\surd$	$\star$	$\star$!

In (2), the candidate set, denoted by the numbers 1–4, are evaluated against constraints C1–C4. Note that a candidate is 'killed off' (represented by !) when it violates a constraint that other candidates do not violate (see columns C1, C2 and C4). Once an output is killed off, it cannot be considered optimal, no matter how many lower ranked constraints it conforms to. Thus, the parenthesised marks in the table represent irrelevant information. Also, if none of the surviving candidates conform to a constraint then none are killed off (see column C3). In this way it is possible for the optimal output to be grammatical even if it violates any number of constraints.

Cross-linguistic variation is handled in the evaluation: simply, languages differ in how they rank constraints. This will then determine different outputs as being optimal in different languages. So, to take the above example again, reordering the constraints, we get the following result:

(3)

	C4	C3	C2	C1
⇒ 1	$\surd$	$\surd$	$\star$	$\surd$
2	$\surd$	$\star$!	($\surd$)	($\surd$)
3	$\star$!	($\star$)	($\surd$)	($\star$)
4	$\star$!	($\star$)	($\surd$)	($\surd$)

With this ranking of constraints 1 is optimal, whereas previously the optimal candidate was 2.

1.2 Optimality Theory and Syntax

With a few notable exceptions (Grimshaw 1993 and Legendre *et al.* 1993), OT has been taken as a theory of phonology. Although there is nothing in OT itself that would suggest that it cannot 'do syntax', the framework is too general to suggest specific ways in which to implement this. The standard response to this situation (in as much as there can be said to be 'a standard' established in the handful of papers written on the subject) is to import assumptions from other syntactic frameworks into OT. For example, Grimshaw (1993) adopts a basic GB framework using OT mechanisms simply to add functional projections to pre-formed 'thematic' structures and to filter out undesirable structures thus formed. It would seem, therefore, that there is no one way to do syntax in OT, but as many different ways as there exist generative theories. Yet it remains to see whether there exist conflicts between OT and other syntactic frameworks. Moreover, through importing other syntactic frameworks into OT we leave unexplored the full potential of what OT itself has to offer as a theory of syntax. At present, therefore, it might be better to try to develop an independent framework which is, where possible, motivated by OT considerations alone.

In this section, we sketch the outlines of a syntactic theory which is primarily motivated by OT itself. Of course, there are many details left to be worked out and here we will concentrate only on issues which bear directly on the topic of this paper and leave other issues for further research.

As outlined above, the OT framework consists of an input, the generation of a potentially infinite number of outputs and an evaluation of the output candidates against a bank of constraints. To do syntax in OT, we therefore need to state what we are to take as the input, what constitutes the generative procedure and what constraints are involved. The first two of these concern issues of a general nature and the last is more specific to a given analysis within the framework. However, there are general considerations concerning the evaluation which can be discussed separately from specific proposals concerning particular constraints. In this section we will confine our discussion to general issues and return to a specific proposal in section **3.**

1.2.1 The Input

The question of what constitutes the generative procedure partly depends on what we take the input to be. If we assume that the input is a pre-formed structure (*à la* Grimshaw 1993), then GEN will operate on this to form other structures. In this view, GEN is like a set of transformations: a function from structures to structures. However, this requires us to add mechanisms to the framework which are responsible for forming the input structures. A more minimal assumption is that GEN performs all structural operations and therefore that the input is not a pre-formed structure. An alternative view is that the input consists of a set of lexical items out of which candidate structures are formed by GEN. This is more in line with standard OT assumptions where the input is considered to be a form taken from the lexicon. It is also very similar to the suggestion made by Chomsky (1994) in the minimalist framework, where the input to the structure building process is a set of lexical items. We will pursue this idea here.

Let us suppose that the input is a set consisting of lexical items relevant for building a structure. One problem we face immediately is how to ensure that the input will contain exactly the right elements from which to build a grammatical structure. Note that according to OT principles, each input will be associated with at least one optimal and therefore grammatical output. Because of this, we cannot assume that lexical selection is random and that the grammar will generate grammatical structures only for those inputs which contain appropriate elements: under OT assumptions every randomly selected set of lexical items should be associated with at least one grammatical structure, obviously not a desirable result. Therefore, we need a mechanism for ensuring that the right elements get inserted into the input from the outset. To a large degree, the information needed to do this is present in the selectional properties of lexical elements: if the process of inserting elements from the lexicon into the input were done in accordance with the selectional properties of the inserted elements, then the only legitimate inputs would be those from which it is possible to form complete structures. For example, if a transitive verb is inserted into the input its selectional properties determine that it requires a nominal complement. Therefore a possible head of such a complement (*e.g.* a noun or determiner) should also be inserted into the input.[1] It is easy to see how, by this process, we can ensure that only legitimate sets of lexical items will be presented to GEN as inputs.

This at least provides us with a set of inputs for forming structures comprising of subcategorised elements. This process will have to be extended to take into account the inclusion of non-subcategorised elements such as adjuncts, but presumably this can be done based on the restrictive properties of the adjuncts themselves. I will leave this issue for further research.

1.2.2 GEN

Having determined the nature of the input, we must now consider the role of GEN in the grammar. In OT, GEN is usually regarded as an unrestricted mechanism which can perform virtually any operation on the input to form the candidate set. There is one important restriction on GEN however, which is that each output candidate must contain the input. This is a necessary requirement for the theory to enable output constraints to be more than mere output filters: if the input is recoverable in each output then constraints can make reference to processes carried out by GEN rather than just to the outcome of these processes. For the purposes of syntax, this requires that input lexical information be unaltered and complete in each output. It is not hard to see how this, along with the notion of a legitimate input discussed above, captures the effects of the Projection Principle of Chomsky (1981): each input, being legitimate, will contain all the lexical elements needed to form a complete structure and therefore the candidate set will contain structures in which the selectional requirements of the elements included in the input will be met in some way. Of course, the candidate set may contain structures in which the selectional requirements are not met, due to the combining of the elements in an incompatible way. But presumably these structures will be less than optimal, there being at least one structure in which the selectional requirements are appropriately met, and will therefore be deemed ungrammatical.

1.2.2.1 The Structure of GEN

A second issue concerns what GEN does to the input to turn it into candidate structures. The interesting question is whether GEN is capable of performing any conceivable operation on the input, or whether GEN is itself structured. The first would appear to be more in line with OT principles as GEN is supposed to be unconstrained. However, it is not entirely clear that this is necessarily so. Obviously GEN is restricted to carrying out relevant linguistic processes on any given input and we would not want to entertain the possibility that it could perform irrelevant operations (*e.g.* colour the input pink or add the day of the week to every second consonant) which would then be filtered out by some constraint. This would be unwelcome for a variety of reasons. First, it would swell the number of necessary constraints to an unmanageable size; second, it would make necessary the assumption that some constraints are inviolable in any language (there are no possible human languages with pink sentences) and thirdly it is plainly inconceivable that languages could work in this way.

But if we accept that GEN is restricted to the use of linguistically relevant operations, it becomes an issue what we are to take as a linguistically relevant operation. For example, is the concatenation of more than two elements at a time to form a structure to be considered a possible relevant operation, or is GEN restricted to making only binary concatenations? This is evidently an empirical question: if no language makes use of structures with more than two branches from any node, then it would seem that GEN is restricted to binary concatenations in building structures from inputs. The alternative of allowing GEN to create tertiary branching trees and then filter these out with a constraint is as pointless and as problematic as allowing GEN to perform linguistically irrelevant operations. Thus it would seem more plausible to assume that GEN is structured in that it is made up of a specific set of relevant operations.

While it is true that stating that GEN is structured entails that it is limited, in that there are operations that it cannot perform, this is not the same as saying that GEN is constrained. Within its structure GEN performs its operations without constraint: *i.e.* there are no rules to tell GEN which operations to apply when and how often. As to the question of what the structure of GEN is? This is an empirical question that we cannot hope to answer in a few short paragraphs and again this is a topic I leave for further research, though the operations discussed by Chomsky (1994) in his 'bare phrase structure' framework would be a good place to start from. In what follows, we assume that, above and beyond its basic structure building functions, GEN is at least capable of inserting non-input elements and moving elements within a structure.

1.2.2.2 Output

A third issue concerns the output of GEN. Much work in generative grammar has concentrated on the issue of how many levels of structural representation are needed. The same question arises with respect to OT: does GEN produce candidate outputs that are single structural representations which feed directly into the interpretative components, or does it produce sets of structural representations, the members of which interface with different components? To these questions, OT itself has nothing to say, though clearly one can structure GEN in different ways to produce as many levels of representations as are required. Again, it is an empirical question as to what levels of representation are needed, though as has always been known it is a far from straightforward one.

A particular issue which concerns this paper is whether it is possible to represent logical relations such as scope in the same structures that feed the phonological component: *i.e.* whether it is necessary to separate LF and PF. In this paper we pursue an approach which assumes that this is not the case. It may well be that investigation of other phenomena will suggest the opposite. I will remain tentative on the issue and once again put off a more definite conclusion awaiting further research.

1.2.3 The Evaluation

OT constraints are output constraints which means that all they see is the results of the processes that produced the candidates and not the processes themselves. Thus we cannot felicitously talk about constraining GEN, only of filtering its produce. However, it would be very useful if we could, in effect, constrain the processes that constitute GEN for the reason that we only want to sanction the application of a given process for exactly the number of times it is necessary to make a candidate optimal. For example, if the application of a movement benefited a given candidate in that it allowed it to overcome a highly ranked constraint, we would not want to also consider as optimal candidates that contained unnecessary movements. This can be simply achieved by constraining movement: movement will then be sanctioned only when it is necessary to overcome a constraint which is more highly ranked than the movement constraint and any further movements will be less optimal. This is very close to Chomsky's notion of economy in language (Chomsky 1986, 1989).

Effective constraints on processes through output constraints would be possible if the process were somehow visible in the output. This may entail output structures containing more information than has previously been assumed, but as these structures are abstract in nature, this is not a problem in principle. In some cases, we may not even have to assume any more than has already been assumed. For example, in the case of movement, movements are visible in the outputs through the traces they leave behind. Thus, we can effectively constraint movement by placing an output constraint on the appearance of traces.

The ranking of constraints which are associated with different phenomena is inconsequential as constraints that are not associated with the same phenomena are not in conflict with each other, and it is only conflicting constraints for which different rankings produce different optimality results. To demonstrate this, consider two non-conflicting constraints. As these do not conflict there are four possibilities for any output candidate: it may conform to both constraints, conform to none of them or conform to one but not the other. Obviously, the candidate which will be determined as optimal is the one which conforms to both constraints, no matter what the ranking of the constraints is. This is demonstrated in the tables in (4) and (5):

(4)

	C1	C2
output 1	$\surd$	$\star$!
output 2	$\star$!	($\surd$)
output 3	$\star$!	($\star$)
⇒ output 4	$\surd$	$\surd$

(5)

	C2	C1
output 1	$\star$!	($\surd$)
output 2	$\surd$	$\star$!
output 3	$\star$!	($\star$)
⇒ output 4	$\surd$	$\surd$

Now consider the case of two constraints which conflict in the sense that conformity to one entails the violation of the other. Here there are only two possibilities for any given output: either it conforms to the first constraint but not the second, or it conforms to the second but not the first. In this situation, the ranking of the constraints does make a difference in that the higher ranking constraint must be adhered to, as demonstrated in (6) and (7):

(6)

	C1	C2
⇒ output 1	$\surd$	($\star$)
output 2	$\star$!	($\surd$)

(7)

	C2	C1
output 1	$\star$!	($\surd$)
⇒ output 2	$\surd$	($\star$)

Obviously, it will only be constraints that are associated with the same phenomena that conflict: it is unlikely that a constraint that forces heads of phrases to be overt will conflict with one which states that an anaphor must have a close antecedent, unless, of course, overt heads play a role in determining the notion of 'closeness' relevant for anaphor binding. Thus, we can ignore the respective ranking of non-conflicting constraints.[2] In general, however, for ease of the representation of the relationships between constraints, we may assume a non-consequential ranking between two non-conflicting constraints, though as nothing depends on this ranking there can be no argument as to what the ranking should be.

A second issue concerning ranking is whether all (conflicting) constraints are strictly ranked. One of the assumptions we will make in this paper is that strict ranking of all constraints is not necessary and in this way we propose to capture optionality in a given phenomenon. Optionality is particularly interesting from an OT perspective as it represents the situation where the evaluation allows more than one optimal candidate. This, in itself, is the cause of a potential problem, which we will return to shortly. However, we must first attend to the matter of optionality itself. Often, we find that languages differ with respect to a given phenomenon and that this situation can be handled by the supposition of two conflicting constraints. Thus, the difference can be accounted for simply on the ranking of these constraints. However, it is not unusual to also find languages which allow both phenomena optionally. The question is that if we account for these phenomena in terms of two conflicting constraints, how can any language demonstrate both sets of phenomena at the same time: if one constraint outranks the other then one set of phenomena should be ruled out as non optimal. The solution to this problem lies in the possibility of not ranking the relevant constraints with respect to each other. If both constraints are of equal rank, and moreover they conflict, then every possible candidate will violate one or the other. In this situation, given that all candidates violate one constraint of the relevant rank, it will be impossible to determine optimality of any candidate and hence the whole set will be passed on for further evaluation. In other words, the non-ranking of conflicting constraints is a way of neutralising the effects of both constraints and thus allowing both sets of phenomena to be optimal. (8), (9) and (10) give a graphic demonstration of this argument. Braces denote non-ranking:

(8)

	Cx	Cy
$\Rightarrow$ 1	$\surd$	$(\star)$
2	$\star!$	$(\surd)$

(9)

	Cy	Cx
1	$\star!$	$(\surd)$
$\Rightarrow$ 2	$\surd$	$(\star)$

(10)

	{Cx	Cy}
$\Rightarrow$ 1	{$\surd$	$\star$}
$\Rightarrow$ 2	{$\star$	$\surd$}

The important point to make here is that in (10) both outputs violate only one constraint of the same rank position and hence both are as optimal as each other with respect to these two constraints.

Finally, we consider the question of whether all constraints are of relevance to all languages. There is an underlying assumption in OT that constraint ranking is the only source of linguistic variation. But while this may be the simplest view, it is not a necessary one and one can imagine the case that languages may differ in terms of which constraints they select as well as on how these are ranked. This is particularly relevant given the above assumptions about optionality. If we are to allow more than one optimal candidate, then all constraints must have potential effect. This is because as long as there remains more than one optimal candidate, the evaluation of these candidates must continue until either only one remains or the bank of constraints is exhausted. The situation in which there is more than one optimal candidate therefore means that each grammatical candidate must satisfy all constraints. What this means is that no constraint can be 'hidden' by placing it low in the ranking, which is one possible way of getting rid of inapplicable constraints in the OT framework. Thus, either all constraints are applicable to all languages, or we assume that languages can differ in terms of the constraints they select. It is, of course, an empirical question as to which view is correct and we shall see that there are arguments that the first should be rejected. We will return to this issue in section **3.**

1.3 Summary

To summarise, a syntactic theory which is, as far as possible, motivated by OT considerations takes its input from the lexicon, has an unconstrained but structured GEN which produces candidate outputs to be evaluated against a bank of ranked constraints. At least one of these candidates will be optimal with respect to the constraints and their ranking and this will be deemed the grammatical structure. The input will be a legitimate set of lexical items, where legitimacy is determined by the possible satisfaction of the selectional properties of each member of the set. GEN builds structures from the input using a set of linguistically relevant operations and produces an output set whose members are single structures. The evaluation may contain constraints that are not necessarily strictly ranked and hence there may be more than one structure which is defined as optimal, thus allowing for optionality. It may also be necessary to suppose that languages differ in the constraints they select as relevant.

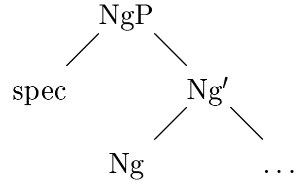
It remains now to discuss the kinds of constraints that are used to evaluate the candidate set. Given that this is where most of the syntax is done in OT, only a fragment of what is necessary can be offered. We will therefore make the practical assumption that the constraints that deal with all other phenomena, including that of basic clause structure, are in place and are not crucially involved in the analysis of the relevant phenomena. We first introduce the syntactic phenomena that we will be dealing with.

2 The Syntax of Negation

In this section we sketch the background assumptions on which we will base an OT analysis of negation phenomena. These are mainly derived from recent work in the principles and parameters approach, but as we shall see, the OT perspective offers a very elegant account of the data.

Like much recent work in negation (Pollock 1989; Laka 1989; Ouhalla 1990; Haegeman & Zanuttini 1991; Haegeman 1992), we shall assume that the seat of negation in a sentence is the Negative Phrase (NgP). Following Ouhalla (1990) we assume a universal structure for NgP which incorporates a head position and a specifier position which houses negative operators which mark the scope of the NgP, along the lines of (11) (order irrelevant):

(11)



This structure is not only relevant for simple sentential negation (12a), but also for cases involving negative quantifiers and adverbials (12b) and (12c):

(12) a. they did not see Bill
b. they saw no one
c. they never saw Bill

Taking *not* to be the head of NgP in English, we assume that it is accompanied by an abstract operator in (12a).[3] In the case of the negative quantifier we may also assume the involvement of a NgP, but in this case it contains an abstract element, which following Bródy (1994) we will refer to as a Scope Marker (SM), coindexed with the overt operator. In this paper we will not discuss negative adverbs, but we assume a similar treatment of them as for the negative quantifiers.[4] Thus, the corresponding structures for (12) are:

(13) a. they did $[_{\text{NgP}}$ Op not see Bill$]$
b. they saw $[_{\text{NgP}}$ SM_i e no one$_i]$
c. they $[_{\text{NgP}}$ SM_i e never$_i$ saw Bill$]$

We will give further discussion to different aspects of this analysis below.

There are at least four different but interacting phenomena that are relevant to a discussion concerning negation. One issue concerns the distribution of the NgP in a clause, languages differing in terms of where the NgP is placed and also how freely it is distributed. Thus, while some languages seem to have a fixed NgP position, others allow it to appear in a number of positions. Hungarian, for example, has a fixed NgP position. Note in the following sentences, the negative head *nem* is always above VP and can only be preceded by negative elements in its specifier position (negative operators), a subject or a topic:

(14) a. nem szereti János Marit
not loves J. M-ACC
'John doesn't love Mary'
b. János nem szereti Marit
c. János nem szeret senkit
J. not loves no one-ACC
'John doesn't love any one'
d János senkit nem szeret
e. *János mindenkit nem szeret
J. everyone-ACC not loves

f. mindenkit János nem szeret
everyone-ACC J. not loves
'it's not everyone that John loves'

Assuming that all pre-negative elements move to their respective positions, we can see that the NgP itself occupies a stable position about which the other elements move. English, on the other hand, allows more freedom in the placing of its negative head:

(15) a. John may not have been reading in the bath
b. John may have not been reading in the bath
c. John may have been not reading in the bath
d. I believed John not to be so stupid[5]

We will argue that English allows NgP in virtually all positions where it can have scope over some element, though there are restrictions on where the abstract negative operator can be placed.

A related issue concerns how many NgPs are allowed in a clause. English clauses permit more than one negative head and thus, assuming basic X-bar principles, there must be a corresponding number of NgPs:

(16) John may not have not been reading in the bath

Other languages, however, allow only one negative head:

(17) a. *Gianni non non telefona a sua madre
G. not not telephones to his mother
b. *nem nem látom a gyereket[6]
not not see-1SG the child-ACC

It is possible that it is because English allows a wider distribution for its NgP that it can have more than one of them in any clause, whereas those languages in which NgP is limited to a single position can only fill this position once. However, as we shall see, there are languages in which NgP seems to have limited distribution but which also allow more than one and thus there is more to consider concerning this phenomenon.

Interesting differences can be seen between languages in terms of how negative operators are related to NgP. We are assuming that SpecNgP provides the 'scope position' for all negative elements. It appears that in some languages, negative operators must move to this position to be given their scope. Haegeman & Zanuttini (1991) and Haegeman (1992) discuss the case of West Flemish (WF) in which negative operators with a non-echoic reading must move to SpecNgP:

(18) a. da ze nie ketent van eur werk en was[7]
that she *Op* contented with her work NG was
'that she was not pleased with her work'

b. da ze me niks ketent en was
that she with nothing contented NG was
'that she was not pleased with anything'

c. *da ze ketent me niks en was

(Haegeman 1992 : 4)

They account for this by assuming that negative operators and negative heads license each other in the spec–head configuration, a principle they call the Neg-criterion, and hence the negative operators must move to satisfy this licensing condition. The difference between a language like WF and one such as English, in which the negative operators do not appear to move, can be stated in terms of where the Neg-criterion is supposed to hold. If it applies at S-structure, then overt movement will be forced. However, if it is delayed until LF, then movement will be covert.

But there is another analysis for this phenomena made available by the assumption of SMs. If an operator has a SM in SpecNgP, there will be no need for it to move to this position: it will be assigned its scope via its SM. This would suggest that operators that move to SpecNgP themselves have no SM and therefore that languages which move their operators are unable to support SMs. We might therefore suggest that it is not negative heads or the operators themselves that need licensing, but SMs: the idea being that those languages which cannot support SMs are unable to license them. This seems natural as SMs, being empty categories, will need some grammatical condition under which their presence is made known. Similar considerations will hold for the abstract negative operator in SpecNgP.

Developing this idea, suppose that SMs are licensed by a local overt negative element. The relevant structural configuration seems to be adjacency, however, we will also see that one negative element can license a number of SMs adjoined to SpecNgP, not all of which can be adjacent to it, obviously. Thus let us define a condition of 'overt adjacency' where two elements will be overtly adjacent providing that no overt element intervenes between the two. Thus, a negative element can be overtly adjacent to any number of SMs, as any intervening SM being non-overt will not interfere with this relationship. The full licensing condition for SMs we assume is as in (19):

(19) α licenses β, β a SM, iff
   i. α is an overt negative element (or its trace)
   ii. α and β are overtly adjacent
   iii. β c-commands α[8]

Let us consider some examples of this licensing. The abstract negative operator in a simple negative sentence will presumably need licensing. This is the function of the negative head which therefore must accompany this operator:

(20) they did [$_{\text{NgP}}$ *Op* *(not) see Mary]

This captures an observation of Ouhalla's (1990) that NgP in general must have either its head or its specifier position filled with an overt element: if the operator is overt, there is no empty element to license and hence the head is not necessary, but if the operator is empty the head is necessary. However, it is possible to get a completely empty NgP, as long as the element in SpecNgP (a SM) is associated with an overt operator which is near enough to license it. With this in mind, consider the following example:

(21) they saw no one

The negative operator in (21) needs an NgP to assign it its scope, but obviously this NgP has an empty head. The question that arises immediately is: Where in the structure is the NgP? Given that the only potential licenser is the negative operator itself, its SM must be overtly adjacent to it. Thus, we claim, the appropriate structure is as in (22a) and not (22b):

(22) a. they saw [$_{\text{NgP}}$ $SM_i e$ no one$_i$]
  b. they [$_{\text{NgP}}$ $SM_i e$ saw no one$_i$]

A number of considerations support the assumption of (22a) over (22b). One is that (21) is syntactically positive, hence triggering phenomena associated with positive sentences such as negative tags and conjunction with *and so did*:

(23) a. they saw no one didn't they
  b. the girls saw no one and so did the boys[9]

Note that structures which have a NgP above the VP seem to be wholly negative:

(24) a. they did not see anyone did they (*didn't they)
  b. the girls didn't see the aardvark and neither did the boys (*and so did ...)

If (22b) were the correct structure for (21) we would have no account of the difference between this and syntactically negative sentences.

A second observation supporting (22a) concerns the scope properties of a negative quantifier in object position. The scope of the negative object is narrower than that of non-negative quantifiers in object position or a negative element in a higher position. Note that (25a) is unambiguous, unlike the sentences in (25b) and (25c):

(25) a. everyone saw no one
b. everyone saw someone
c. everyone didn't see the aardvark[10]

How are we to account for these facts?[11] We are assuming that the scope of a negative element is determined by the position of its SM. Let us assume that the same is true of all quantifiers (see Stowell & Beghelli 1994 for a similar idea). The fact that a non-negative object quantifier can have scope over the subject therefore indicates that its SM is high enough in the structure to allow this, as must be the SM for the negative element which is above the VP. Again, that the negative quantifier in object position does not behave in the same way, and in fact has obligatorily narrow scope, indicates that the NgP containing its SM is lower down in the structure. If the negative element just above the VP can have scope over the subject, again the indication is that the NgP of the object quantifier is below the VP.

Returning to the original issue, assuming structure (22a), we see that the SM is licensed by the negative operator itself. Two issues remain: why is the head of the NgP empty? and why can we not have a higher NgP containing the SM for the operator which is licensed by a negative head? In other words, why are the following ungrammatical:

(26) a. *they saw not no one
b. *they did not see no one (= they saw no one)[12]

We will see that a satisfactory account of these observations falls out from our analysis within the OT framework.

Languages also differ in the requirements placed on the negative head. In some, *e.g.* Hungarian, the negative head is always present whether it is needed to license a SM or not, while in others the head is only present when needed to license a SM. Thus, as we saw in (26a), in English, if the SM can be licensed by an overt negative operator, and therefore there is no need for the negative head, the head cannot appear. There are several variations on this theme. So, for example, while the English negative head only appears to support the abstract negative operator, the Italian negative head appears in negative sentences only when there is no pre-verbal negative element:

(27) a. Gianni *(non) telefona a sua madre
G. not telephones to his mother
'Gianni doesn't telephone his mother'

b. Gianni *(non) legge niente
G. not reads nothing
'Gianni doesn't read anything'

c. nessuno (*non) telefona
no one telephones
'no one calls'

d. a nessuno$_i$ Gianni (*non) dice niente t_i
to no one G. says nothing
'Gianni says nothing to anyone'

We leave accounting for this difference until the next section.

A final phenomenon we will discuss concerns the use of multiple negative elements in a single clause. Haegeman & Zanuttini (1991) and Haegeman (1992) provide an analysis of languages in which the use of multiple negative elements expresses a single negation, a phenomenon known as Negative Concord (NC), based on the assumption of NgP. The basic idea is that all negative operators will, at some point in the derivation, move to SpecNgP, as required by the Neg-criterion. Assuming multiple adjunction to SpecNgP, in this configuration there is an absorption of their negative features which results in the expression of just one negation. This is similar to the proposal by Higginbotham & May (1981) concerning the absorption of wh-features in multiple wh-questions. This easily translates into the current perspective. Either elements move to SpecNgP or they have SMs in this position. In either case, when there are multiple negative elements, SpecNgP will be multiply filled and in this configuration the absorption of negative features takes place. This is a fairly natural account, given current assumptions, as SpecNgP is assumed to be the position in which negative elements receive their interpretation: if there is only one such position, there will be only one such interpretation.

However, not all languages work in this way. In others, when more than one negative element is used each retains its negative features in the interpretation and hence there is a 'cancelling out' effect, as we get in logic. This situation is referred to as Double Negation (DN). This at first seems problematic for the simple account given above: why do some languages have negative absorption and other languages not? This is particularly poignant, given the observation that the NC situation seems natural if

SpecNgP is assumed to be associated with the interpretation of negative features.[13] The account that we have been developing here would lead one to expect that in the case of DN more than one SpecNgPs are involved, as each negative element would have to be associated with a different one thus blocking negative absorption. Developing this idea, we have already noted that some languages allow more than one NgP per clause and others do not. If our general speculations are correct, then we would expect that languages which allow only one NgP should be forced to allow only NC structures and those that allow more than one NgP can have DN readings. Indeed, this would appear to be true: Italian and Hungarian are both NC languages and as we have seen they only have one NgP per clause. English is a DN language and can have many NgPs per clause.

Thus, the proposal we make is simple and intuitive. Languages which allow only one NgP must interpret all negative elements in relation to one SpecNgP and when more than one negative element is associated with one SpecNgP there is an absorption of their negative features. Thus, such languages have only NC structures. Those languages which allow more than one NgP can provide a NgP for each negative element and therefore allowing each negative element to be associated with an individual SpecNgP. In this situation there can be no absorption of the negative features and in fact we predict that all such features will count in the interpretation and hence DN readings will arise. It remains to account for why languages which allow more than one NgP should have to have DN structures. We shall return to this issue in the following sections.

This proposal is given some support from WF. This language allows both NC and DN readings, but both are associated with different surface configurations. In NC structures, negative elements are moved to a position to the left of the negative operator generated in SpecNgP, but in DN structures, the negative elements move to a position to the right of this negative operator:

(28)	a.	da	ze	me	niemand	nie	ketent	en was
		that	she	with	no one	not	contented	NG was
		'that she was not pleased with anyone'						
	b.	da	ze	nie	me	niemand	ketent	en was
		that	she	not	with	no one	contented	NG was
		'that she was not pleased with no one'						

(Haegeman 1992)

We have already seen that WF does not allow SMs and hence in both of these examples the negative operator *me niemand* must move to SpecNgP for its interpretation. This is confirmed by the fact that the operator is displaced in both examples. But why there should be a difference in the interpretation given to these structures is at first mysterious. Suppose we assume that in (28a) the operator moves to adjoin to SpecNgP containing *nie.* From our point of view, this should give rise to a NC reading, which it does. This suggests that the operator is not adjoined to SpecNgP in (28b), rejecting the possibility of right adjunction if the language allows left adjunction. If the operator has moved to SpecNgP, but it has not moved to the SpecNgP containing *nie,* the conclusion is that it must have moved to a separate SpecNgP and hence we would expect a DN reading, which is what we have. Furthermore, Haegeman & Zanuttini (1990) state that in the case of the DN reading, the second negative element has narrower scope, confined to the VP. This cannot be accounted for if we assume that it has moved to the same SpecNgP that the operator in (28a) has moved to, but receives a natural account under the assumption that there is a second NgP lower that the first in (28b).

Finally, note that the WF case just discussed shows that whether a language allows one or more NgPs per clause is not necessarily related to restrictions that the language places on the distribution of its NgPs. WF has a very restricted distribution for its NgP, allowing this only above VP. However, we must also assume that it allows more than one NgP above VP to account for the DN structures. One can also imagine the opposite situation as possible: a language which has a freer distribution for its NgPs, but allows only one per clause. While I know of no language which is strictly like this, some non-standard varieties of English do allow NC readings, as well as DN ones, but are like Standard English in having a freely distributed NgP. Thus, in some structures there will only be one NgP providing the specifier position for a number of negative elements, but there is variation in where this NgP can be placed:

(29) a. he might not have seen nothing
'he might not have seen anything'
b. he might have not seen nothing
'he might have not seen anything'

In sum, we have seen that there are a number of dimensions along which languages differ with respect to negation phenomena. These seem to be highly interactive and choices made along one dimension may have consequences for phenomena dealt with by principles along another. For example, how freely NgP is distributed in a clause will have implications for the possibility of licensing SMs in SpecNgP: if the NgP must be removed from the operator, this element will not be able to license its own SM and hence a head must be inserted to fulfil the licensing condition. Furthermore, if a language moves its operators to SpecNgP instead of inserting SMs, then the licensing conditions will be voided and hence the language will behave differently with respect to when it inserts negative heads, for example. In the next section we will outline the OT principles we will be making use of to account for the negation phenomena that we have reviewed in this section.

3 An OT Analysis of Negation

Above we introduced several negation related phenomena and associated cross linguistic variation. We have discussed certain structural concepts which enable us to systematise the phenomena such as NgP, SMs, *etc.* However, these concepts themselves do not form a theory which explains the phenomena nor the variation. We must include these into a more general grammatical theory to provide real accounts of the observed phenomena, many of which are still mysterious even given the concepts introduced above. In this section then, we detail an OT account of negation which provides a very elegant explanation for much of what we have discussed above.

3.1 Gen and the Production of Structures

Before we begin our account of negation phenomena, there are a number of additional comments we should make about how the structures that are presented to the evaluation component of the grammar are formed, in the light of what we have seen above.

We have seen that in some languages, negative heads seem to be used as a 'last resort' to license SMs when there is no other way of doing so. In other languages negative heads are obligatory in all negative sentences. If we suppose that these heads are included as part of the input, then their necessity in certain structures becomes the responsibility of the processes that form legitimate inputs: GEN and the evaluation cannot force elements to be included in the input, but simply work on the inputs they are provided with. Thus, from this perspective, whatever property forces a negative head's presence must be lexical. It is difficult to think what the relevant lexical property should be: how is the lexicon to know when a given element is going to be needed when the conditions that determine this are structural?

Another possibility is that the presence of a negative head is determined by GEN: GEN has the option of inserting an overt head into a NgP or leaving it empty and the evaluation component will decide which choice is optimal.[14] This raises the question of what distinguishes simple negative from simple positive sentences, when in most languages this is indicated by the presence of a negative head. Rather than assume that negative and positive sentence are formed from the same inputs, which would make them opposing candidates, let us assume that it is the negative operator which is present in the input of simple negative structures. In some languages this is an overt element and in others it is abstract, but in all languages negative sentences have negative operators in their inputs. In this way, the treatment of a simple negative sentence is no different to that of one containing another negative operator, such as *no one* or *nothing.* We will also assume that GEN can insert SMs which are not in the input.

Now consider the question of what GEN actually does to form the structures that constitute the candidate set. We stated earlier that amongst the operations that GEN performs are processes like those assumed in Chomsky (1994). Thus, sanctioned by elements in the input, GEN will build structures to combine these elements, possibly moving them as part of this process. Above we mentioned the possibility of GEN inserting elements as well. Thus, certain elements and structures that contain them do not have to be sanctioned by the input. The evaluation will then determine what insertions of elements and structure produce optimal and non-optimal outputs.

Let us consider an example from English. Suppose we have the following input: {they, saw, no one}. The structure that is selected as optimal is:

(30) they $[_{\mathrm{VP}}$ saw $[_{\mathrm{NgP}}$ SM_{i} e no one$_{\mathrm{i}}]]$

This structure must be formed by building an empty NgP on top of the operator and then inserting a SM into the specifier of this structure. We can assume that these processes are unrestricted in that GEN is not obliged to perform exactly these operations for all candidates. However, (30) must be in the candidate set, as it is selected as optimal. Other possible candidates will be those formed by either not carrying out these processes or carrying them out a greater number of times, as well as making use of other operations such as movement. (31) represents some of these possibilities:

(31) a. they saw no one
b. they $[_{\mathrm{VP}}$ saw $[_{\mathrm{NgP}}$ SM_{i} e $[_{\mathrm{NgP}}$ SM_{i} e no one$_{\mathrm{i}}]]]$
c. they $[_{\mathrm{VP}}$ saw $[_{\mathrm{NgP}}$ no one$_{\mathrm{i}}$ e $t_{\mathrm{i}}]]$
d. they $[_{\mathrm{NgP}}$ SM_{i} e $[_{\mathrm{VP}}$ saw no one$_{\mathrm{i}}]]$
e. they $[_{\mathrm{NgP}}$ no one$_{\mathrm{i}}$ e $[_{\mathrm{VP}}$ saw $t_{\mathrm{i}}]]$

(31a) is the result of not inserting a NgP and therefore no SM either. (31b) is formed by inserting two NgPs each of which contains a SM. In (31c) GEN moves the negative operator to SpecNgP and in (31d) (31e) GEN inserts the NgP above the VP and either inserts a SM or moves the operator. There are numerous other possibilities, but these will suffice for expository purposes. Of course, some of these structures are relevant for other languages, but not English, and others are not relevant for any language at all. It is the role of the evaluation component to determine these issues.

An important issue is raised by the possible generation of (31a). Presumably no language accepts such structures, as the operator would not be assigned a scope and hence would be uninterpretable. Of course, we could propose a constraint to rule such structures out, and rank this constraint highly in all languages. But in effect this is to introduce an inviolable constraint into the evaluation, which is contrary to the spirit of OT. Supposing that the constraint is in principle violable, but just happens not to be in any language, is also unacceptable as it provides no explanation. However, consideration of the nature of the phenomena suggests a possible solution to the problem: (31a) is unacceptable because it is uninterpretable and this is a semantic fact not a grammatical one. It is not improbable that semantic constraints such as this are universal and hence lie outside usual OT considerations of violability and ranking. We can therefore ignore the possibility of such structures as they are dealt with by principles that lie outside the grammatical system.[15]

With these considerations in mind, we now turn our attention to the evaluation and the constraints that account for the phenomena we have reviewed.

3.2 Constraints on Negation

The crucial part of an OT analysis of any specific phenomena is the evaluation. In this section we introduce and motivate the constraints we propose to account for negation structures and their cross-linguistic differences.

We propose the following constraints to account for negation phenomena:

(32) INS (=Insert): Do not insert any element or structure

(33) MOVE: Do not move elements in a structure

(34) HEAD: All heads must be overt

(35) UNISPEC (=Unique Specifier): Any specifier position may only contain one element (no adjunction)

(36) LSM (=License SMs): All SMs must be licensed

Note that the first two constraints are constraints on the generative procedure rather on the output. As discussed above, however, it is possible to view them as output constraints if the operation of these procedures is visible in the output. Thus MOVE may be interpreted as a constraint on traces. INS is a little more problematic as there is no obvious correlate of a trace for the process of insertion. However, it is arguable that inserted elements must be marked in the output under the assumptions we have made so far. Recall that one basic constraint on GEN is that it must not alter input material and therefore the input may have to be marked as such to ensure that GEN does not accidentally affect it in non-legitimate ways. Inserted elements, then, will be picked out as unmarked in this respect and hence will be visible in the output. In what follows, I will continue to refer to these as constraints on processes because it is more convenient to do so. However, this must not disguise the fact that they are output constraints.

The first two constraints are obviously very general and as they constrain basic operations we expect them to be involved in many processes. The other constraints are not so general, but neither are they specific and thus may be motivated for a number of phenomena. The requirement that all phrases have heads, HEAD, may play a role in auxiliary inversion phenomena (see Grimshaw 1993) and in instances where overt complementisers are necessary, such as in sentential subjects and noun complement clauses.[16] The constraint on multiply filled specifiers, UNISPEC, may have a role in the analysis of multiple wh-question structures and, in general, any process that makes use of the specifier position, such as agreement and feature checking. Finally, negation is not the only phenomena to make use of SMs, any scope phenomena such as quantification, including wh-questions (see Bartos, this volume), may involve the use of SMs and hence the requirement that SMs be licensed may enter into a treatment of these.

Turning now to the question of how the constraints operate on negative structures, INS will militate against the insertion of any element or structure including negative heads, SMs and the NgPs that house these elements. This is in conflict with virtually all other constraints. We have said that, as a semantic fact, all negative operators will need some way to mark its scope. There are two ways of doing this: either a SM can be inserted for the operator, or the operator itself can move to a scope position (Spec-NgP). As both of these options violate one constraint or another, we can see that whether a language chooses to insert SMs or to move its operators will be determined by the relative ranking of INS or MOVE: if MOVE is ranked highest, insertion of SMs will be preferred, if INS is ranked higher operator movement will be favoured. Of course optionality of either moving the operator or inserting a SM will be produced by not ranking INS and MOVE with respect to each other, as discussed above.

INS also plays a crucial role in determining the presence or absence of negative heads. Of course, INS will generally rule against the insertion of negative heads. However, there are two cases where a head will be deemed necessary. First, in order to conform to HEAD an overt negative head must be inserted into NgP. Thus, the relative ranking of INS and HEAD will provide the distinction between those languages which always have negative heads and those where the negative head may be missing. However, in the latter case, the constraint that SMs must be licensed may enter the picture. If there are SMs, *i.e.* the language ranks MOVE above INS, and the distribution of NgP does not allow an operator to license its own SM, then a head will be necessarily inserted to conform to LSM. If there is no SM, or there is another way to license the SM without the insertion of a head, INS will force the head to be absent. In this way, we capture the cases where heads are obligatory and where heads are used only as a last resort. Again, optionality will be captured by not ranking these constraints.

Finally, INS also plays a role in determining the number of NgPs allowed in a clause. When there is more than one negative operator in the input, semantic considerations will force each to be assigned a scope, as discussed above. However, how the scope is assigned is partly determined by the structure of the sentence. If we insert a NgP for each operator, then we provide a unique position in which to assign the operator its scope, either by insertion of a SM or by operator movement. However, INS will work against the insertion of NgPs and thus will favour structures with as fewer NgPs as possible. If only one NgP were to be inserted, then the single specifier provided by that NgP would have to accommodate the scope assignments to each operator and hence will have to allow multiple adjunction to Spec-NgP. However, this is the situation that UNISPEC constrains, thus INS and UNISPEC are also in conflict. Once more, the relative ranking of conflicting constraints gives rise to the linguistic differences we have seen: if INS is ranked higher, the language will be restricted to a single NgP to provide the scope for any number of operators, if UNISPEC is ranked higher the language must insert a NgP for every operator. As there will be absorption of negative features in this case, this situation gives rise to NC phenomena. DN languages rank UNISPEC above INS, forcing a NgP for each operator and consequently preventing the absorption of their negative features. Languages which allow both structures do not rank INS and UNISPEC.

Thus, with reference to the above constraints, we are able to capture much of the syntactic phenomena we discussed in section two above. One phenomenon, however, goes unexplained: the distribution of NgP within the clause. We will have nothing to say about this in the present paper, assuming it to be properly dealt with by the theory of phrase structure rather than by principles concerning the syntax of negation *per se.* Of course, the differences that languages display with regard to the distribution of their NgPs interacts with other negation phenomena and thus we will not refrain from making use of the distributional facts to account for a given structure, but we will assume that the distributional facts themselves are determined by other constraints and lexical properties that we do not detail here.

To briefly sum up, we have the following interactions between the constraints:

i. INS and MOVE determine whether a language has SMs or operator movement
ii. INS, HEAD and LSM determine whether a language has obligatory negative heads or heads used as a last resort
iii. INS and UNISPEC determine whether a languages has NC or DN structures

Nothing of much consequence follows from the relative ranking of the non-conflicting constraints, such as MOVE and UNISPEC, as these do not force choices in a structure which are relevant to each other: the fact that MOVE forbids operator movement is inconsequential for whether there are more than one scope marking element in SpecNgP and the fact that UNISPEC prevents multiple elements in SpecNgP is unimportant for whether or not the language allows movement.

We now turn to the details of the analysis of negation phenomena in a number of languages, showing how linguistic differences arise through differences in the ranking of the constraints introduced above.

3.3 Analysis of the Languages

3.3.1 English

We start our analysis with English. The following data demonstrate that English is a DN language, which does not move its operators and hence has SMs. The negative head is used as a last resort licenser of SMs:

(37) a. I did [$_{\text{NgP}}$ *Op* not see Bill]
b. I saw [$_{\text{NgP}}$ *SM* *e* no one]
c. I did [$_{\text{NgP}}$ *Op* not see [$_{\text{NgP}}$ SM_{i} *e* no one$_{\text{i}}$]]
d. *I did [$_{\text{NgP}}$ *Op* *e* see Bill]
e. *I did [$_{\text{NgP}}$ no one$_{\text{i}}$ (not) see t_{i}]
f. *I saw [$_{\text{NgP}}$ SM_{i} not no one$_{\text{i}}$]
g. *I did [$_{\text{NgP}}$ SM_{i} *Op* not see no one$_{\text{i}}$]

Comparing (37a) and (37d), we note that the negative head is obligatory when the abstract negative operator is used. This is as expected, otherwise there would be no visible negative element. In this way, the operator is like a SM, in that it needs licensing by a local overt negative element. However, the head is impossible when it is not needed, compare (37b) and (37f). In the former, the position of the NgP enables the overt operator to license its own SM and hence the insertion of the head in (37f) produces an ungrammaticality. Displacement of the operator is not allowed, (37e), we therefore assume that insertion of SMs is obligatory in all cases.[17] When there is more than one negative operator, as in (37c), there must be more than one NgP, hence a DN interpretation is given to these structures. Finally, (37g) is ruled out as (Standard) English does not allow NC.

The fact that English prefers to insert SMs and NgPs rather than move its operators and fill its specifier positions with more than one element suggests that it ranks UniSpec and Move above Ins. The fact that the head is not obligatory and indeed cannot appear when there is no need for it suggests that Ins outranks Head and also that LSM outranks Ins. Thus we have UniSpec, Move and LSM above Ins and Head below it. As UniSpec, Move and LSM are non-conflicting, their ranking with respect to each other is unimportant. For a concrete proposal, let us assume the following ranking:

(38) UniSpec $\gg$ Move $\gg$ LSM $\gg$ Ins $\gg$ Head

We now demonstrate that this ranking does indeed account for negation phenomena in English. In a simple case of sentential negation such as in (39a), we can take the input to be as in (39b), complicating issues concerning other functional elements aside. (40) provides the analysis, in tabular form, that the ranking of constraints in (38) yields for this input:

(39) a. I did not see Bill
b. {I, *Op*, see, Bill}

(40)

	UniSpec	Move	LSM	Ins	Head
I [$_{\text{NgP}}$ Op *e* see Bill]	$\surd$	$\surd$	$\star$!		
$\Rightarrow$ I [$_{\text{NgP}}$ Op not see Bill]	$\surd$	$\surd$	$\surd$		

Note that we are assuming that the input contains the abstract negative operator, but not the negative head, as discussed above. As all input material must be present in each output, there is no need to consider the case where there is no NgP at all: this will have to be inserted to house the negative operator. We put aside the issue of *do*-support here. Presumably this is to be handled by consideration of other constraints which do not concern us here. Other options not represented in this table include the possibility of inserting superfluous NgPs and SMs, which of course will be ruled out by Ins: in each case the optimal structure will contain fewer violations of Ins at it represents the minimal number of insertions which make the structure compatible with the higher ranking constraints. Thus, the only real issue is whether or not to insert an overt head and, of course, LSM ensures that this is done.

Next, consider a case, similar to the above, in which there is a single overt operator, as in (41a):

(41) a. I saw no one
b. {I, see, no one}

(42)

	UniSpec	Move	LSM	Ins	Head
I [$_{\text{NgP}}$ SM$_i$ *e* saw no one$_i$]	√	√	⋆!		
I [$_{\text{NgP}}$ SM$_i$ not saw no one$_i$]	√	√	√	⋆⋆⋆!	
⇒ I saw [$_{\text{NgP}}$ SM$_i$ *e* no one$_i$]	√	√	√	⋆⋆	
I saw [$_{\text{NgP}}$ SM$_i$ not no one$_i$]	√	√	√	⋆⋆⋆!	
I [$_{\text{NgP}}$ no one$_i$ *e* saw t_i]	√	⋆!			
I [$_{\text{NgP}}$ no one$_i$ not saw t_i]	√	⋆!			

Here the issues are slightly more complex as there are more possibilities to decide on: do we insert a NgP close to the operator or further from it?;[18] do we insert an overt head or leave the position empty?; do we move the operator or insert a SM? Insertion of the NgP close to the operator ensures that operator can license its own SM and thus avoids the need to insert a licensing head. Hence Ins and LSM account for why the NgP is low in the clause and also why the head does not appear. Movement of the operator is of course ruled out by Move and hence the optimal structure is as expected. For simplicity, we have again omitted examples which a straightforwardly accounted for, such as the insertion of superfluous NgPs *etc.*

A similar case to the above, but with the negative operator in subject position, raises some interesting questions:

(43) a. no one saw Bill
b. {no one, see, Bill}

(44)

	UniSpec	Move	LSM	Ins	Head
⇒ [$_{\text{NgP}}$ SM$_i$ *e* no one$_i$] saw Bill	√	√	√	⋆⋆	√
⇒ [$_{\text{NgP}}$ SM$_i$ *e* no one$_i$ saw Bill]	√	√	√	⋆⋆	√

Issues such as whether to move the operator or insert an overt head are dealt with as above. The constraints we have proposed provide no way to distinguish between the case where the NgP is inserted directly above the operator in subject position or where the NgP dominates the whole clause.

Making the simple but plausible assumption that the scope of the operator, as provided by its SM, is determined under c-command, the difference between these options is that in the former the SM does not c-command anything other than the subject whereas in the second it c-commands the whole clause. The scope facts would seem to support the latter assumption, as a negative operator in subject position always has scope over any operator in object position:

(45) a. no one saw everyone $(\sim\exists x\ \forall y\ [x \text{ saw } y])$
b. no one saw any one $(\sim\exists x\ \exists y\ [x \text{ saw } y])$
c. no one saw no one $(\sim\exists x\ \sim\exists y\ [x \text{ saw } y])$

This suggests that we have not exhausted the list of constraints relevant to negation phenomena and that we need a way to rule out the possibility of placing the NgP directly above the subject operator. Just what the relevant constraint is, is at present mysterious, but we note that it has something to do with the distribution of NgP and once the principles that govern the distributions of phrases is better understood from an OT perspective, we may find a solution to this problem. We therefore put the issue aside pending further investigation.

Finally, we consider a case where there is more than one negative operator:

(46) a. I did not do nothing
b. {I, *Op*, do, nothing}

(47)

	UniSpec	Move	LSM	Ins	Head
I $[_{\text{NgP}}$ SM_i Op not did $\text{nothing}_\text{i}]$	⋆!				
I did $[_{\text{NgP}}$ Op SM_i *e* $\text{nothing}_\text{i}]$	⋆!				
⇒ I $[_{\text{NgP}}$ Op not did $[_{\text{NgP}}$ SM_i *e* $\text{nothing}_\text{i}]]$	√	√	√	⋆⋆ ⋆⋆⋆	
I $[_{\text{NgP}}$ Op not *e* $[_{\text{NgP}}$ SM_i *e* did $\text{nothing}_\text{i}]]$	√	√	⋆!		
I $[_{\text{NgP}}$ Op not did $[_{\text{NgP}}$ SM_i not $\text{nothing}_\text{i}]]$	√	√	√	⋆⋆⋆ ⋆⋆⋆!	
I $[_{\text{NgP}}$ Op not $[_{\text{NgP}}$ SM_i *e* did $\text{nothing}_\text{i}]]$	√	√	⋆!		
I $[_{\text{NgP}}$ Op not $[_{\text{NgP}}$ SM_i not did $\text{nothing}_\text{i}]]$	√	√	√	⋆⋆⋆ ⋆⋆⋆!	

In this example, we see how the combination of constraints conspire to select the correct structure as optimal. First, if each operator is to have a scope, there will have to be more than a single NgP as more than one element in a single specifier position is ruled out by UNISPEC. Inserting no licensing head for the abstract operator violates LSM and inserting a head for the SM of the overt operator produces an unnecessary violation of INS. Finally, inserting the overt operator's NgP too high will either produce an LSM violation if no head is inserted, or an unnecessary INS violation if the head is inserted. Either way, there is a more optimal candidate in which the NgP is inserted low enough down the tree to enable the operator to license its own SM.

Nothing we have said so far prevents the insertion of the abstract operator's NgP inside the VP, though as (48) shows, this ought to be ruled out:

(48) *I saw $[_{\text{NgP}}$ *Op* not $[_{\text{NgP}}$ SM_{i} *e* no $\text{one}_{\text{i}}]]$

However, note that although the general distribution of NgP is more or less free in English, this is not true of the abstract negative operator itself:

(49) a. *$[_{\text{NgP}}$ *Op* not I saw Bill]
b. *I saw $[_{\text{NgP}}$ *Op* not Bill]
c. *I saw Bill $[_{\text{NgP}}$ *Op* not]

We thus assume that it is a lexical property of this particular operator that prevents it from appearing in these positions and this property will be responsible for ruling out (48) as well.[19]

The particular ranking of the constraints that we have proposed straightforwardly accounts for the basic observations concerning negation in English. Looking back at the analyses given in the tables, note that each constraint has a role to play in determining the optimal candidate with the one exception of HEAD. This appears to be a superfluous constraint in English, at least concerning negation phenomena. However, being placed low in the hierarchy means that it never has an opportunity to have any effect: in each case the optimal candidate is selected by consideration of the higher ranked constraints. We will return to this observation below.

3.3.2 Italian

Italian offers a very interesting challenge for the present theory as it is similar to English in that its negative head seems to be used as a last resort, but a very different pattern of when the head is necessary emerges. Besides this difference, Italian is also a NC language and does not allow DN structures. Thus, consider the data in (50):

(50) a. Gianni *(non) telefona a sua madre
G. not telephones to his mother
'Gianni doesn't telephone his mother'

b. Gianni *(non) legge niente
G. not reads nothing
'Gianni doesn't read anything'

c. Gianni *(non) dice niente a nessuno
G. not says nothing to no one
'Gianni doesn't say anything to anyone

d. nessuno (*non) telefona
no one telephones
'No one calls'

e. *(non) telefona nessuno
not telephones no one
'No one calls'

f. a nessuno$_i$ Gianni (*non) telefona t_i
to no one G. telephones
'Gianni telephones no one'

g. a nessuno$_i$ Gianni (*non) dice niente t_i
to no one G. says nothing
'Gianni says nothing to anyone'

(Haegeman 1992)

As can be seen in (50a), Italian is like English in that in cases of simple sentential negation an abstract operator is used and the overt negative head is inserted to license this operator. Thus we can conclude that LSM outranks Ins in the Italian ranking. However, unlike English the head is also necessary to accompany a post-verbal overt negative operator, see (50b), (50c) and (50e). The reason for this is that the Italian NgP has a more restricted distribution than the English one. Just where the NgP is placed in the Italian sentence structure is a matter for debate. Zanuttini (1991) argues that NgP is the highest phrase in the clause and the negative head occupies its base position at the surface. Therefore all elements which precede it (pre-verbal subjects *etc.*) must occupy a topic position. Belletti (1994), on the other hand, argues that AgrP dominates NgP and that the negative head raises to adjoin to Agr. We shall see that Zanuttini's analysis is more compatible with our assumptions, hence we shall adopt it from now on. The important point at present is, however, that NgP has a restricted distribution well above the VP, and hence any post-verbal negative operator will not be able to license its own operator and an overt head is therefore necessary. This is compatible with the assumption that LSM outranks Ins.

However, the negative head is not necessary when there is at least one pre-verbal negative operator. We are assuming that the pre-verbal subject occupies a topic position above the NgP and the same, of course, is also assumed for a topicalised object.[20] Let us further assume that a post-verbal subject either remains inside the VP or at least does not move as high as the NgP.[21] Assuming a 'shortest distance' constraint on movement of the sort motivated by Rizzi (1990), it follows that any element that moves to the topic position passes through SpecNgP. Thus, the relevant structures are as follows:

(51) a. nessuno$_i$ [$_{NgP}$ t_i e telefona t_i]
  b. a nessuno$_i$ Gianni [$_{NgP}$ t_i e telefona t_i]

Given the assumptions we are making, it can be argued that these operators do not need SMs as they are assigned a scope via their trace in the scope position. Thus, as there are no SMs in these constructions, there is no need for the head.[22] As post-verbal operators do not move through SpecNgP, these must have a SM to assign them scope and hence a licensing head is necessary:

(52) a. [$_{NgP}$ SM_i non telefona nessuno$_i$]
  b. Gianni [$_{NgP}$ SM_i non telefona a nessuno$_i$]

To complete the picture we must also assume that the trace of a topicalised operator in SpecNgP can license SMs to account for the absence of the head in situations where there are both pre- and post-verbal operators:

(53) a nessuno$_i$ Gianni [t_i SM_j e dice niente$_j$ t_i][23]

The fact that the head is absent in cases where it is not necessary argues that Ins outranks Head. Thus in the relevant way, Italian is exactly the same as English with LSM outranking Ins and Ins outranking Head. The differing patterns of when a negative head is necessary in English and Italian therefore is due to the different distributions of NgP in these languages. Of course, as can be seen in (50c) and (50g), Italian is a NC language and hence differs from English in that Ins outranks UniSpec. Thus, we propose the following ranking for Italian:

(54) Move $\gg$ LSM $\gg$ Ins $\gg$ Head $\gg$ UniSpec

Again, this ranking contains a certain amount of redundancy as the relative ranking of Move and LSM and of Head and UniSpec have no consequence.

Some of the issues for Italian are exactly the same as for English, such as the necessity of the head with the abstract sentential operator. We will therefore not demonstrate the OT analysis of these. Here we restrict ourselves to more interesting cases.

Let us start by providing an analysis for the NC structure, as we have so far only given cases of DN structures. The example we consider concerns two post-verbal operators:

(55) a. Gianni non dice niente a nessuno
b. {Gianni, dice, niente, a, nessuno}

(56)

	Move	LSM	Ins	Head	UniSpec
$\Rightarrow$ Gianni $[_{\text{NgP}}$ SM$_{\text{i}}$ SM$_{\text{j}}$ non dice niente$_{\text{j}}$ a nessuno$_{\text{i}}$]	$\surd$	$\surd$	⋆⋆ ⋆⋆		
Gianni $[_{\text{NgP}}$SM$_{\text{i}}$ *e* $[_{\text{NgP}}$ SM$_{\text{j}}$ *e* dice niente$_{\text{j}}$ a nessuno$_{\text{i}}$]]	$\surd$	⋆!			
Gianni $[_{\text{NgP}}$ SM$_{\text{i}}$ non $[_{\text{NgP}}$ SM$_{\text{j}}$ *e* dice niente$_{\text{j}}$ a nessuno$_{\text{i}}$]]	$\surd$	⋆![24]			
Gianni $[_{\text{NgP}}$ SM$_{\text{i}}$ *e* $[_{\text{NgP}}$ SM$_{\text{j}}$ non dice niente$_{\text{j}}$ a nessuno$_{\text{i}}$]]	$\surd$	$\surd$	⋆⋆⋆ ⋆⋆!		
Gianni $[_{\text{NgP}}$ SM$_{\text{i}}$ non $[_{\text{NgP}}$SM$_{\text{j}}$ non dice niente$_{\text{j}}$ a nessuno$_{\text{i}}$]]	$\surd$	$\surd$	⋆⋆⋆ ⋆⋆⋆!		

The relevant observation is that any structure in which more than one NgP is inserted will either violate LSM, if no licensing head is also inserted, or will be at least one more violation of Ins than the optimal structure. Of course there can be no structures with fewer insertions as the SM is required for semantic reasons, the head is required to license it and NgP is required to house these elements.

Now let us consider the difference between pre- and post-verbal operators. We take the case of the subject, noting that the object will be given a similar analysis. A seemingly natural assumption is that structures with pre-verbal subjects are related to the same input as those with post-verbal subjects. Thus these structures represent options and should be equally optimal. We have argued that optionality is achieved by not ranking two conflicting constraints and we therefore suppose that whatever constraint forces the movement of the subject to the pre-verbal position must not be ranked with respect to Move. Violating Move, in this particular case, is harmless as conforming to it will violate a constraint of equal rank. However, this assumption is problematic, as is demonstrated by the following example:

(57) a. non telefona nessuno
b. nessuno$_{\text{i}}$ telefona t_{i}
c. {telefona, nessuno}

(58)

	Move	LSM	Ins	Head	UniSpec
$[_{NgP}$ SM$_i$ e telefona nessuno$_i]$	[$\surd$]	$\star$!			
$[_{NgP}$ SM$_i$ non telefona nessuno$_i]$	[$\surd$]	$\surd$	$\star\star\star$!		
$\Rightarrow$ nessuno$_i$ $[_{NgP}$ t_i e telefona $t_i]$	[$\star$]	$\surd$	$\star$		
nessuno$_i$ $[_{NgP}$ t_i non telefona $t_i]$	[$\star$]	$\surd$	$\star\star$!		

We want (57a) and (b) both to be optimal. However, when placed in competition with each other the pre-verbal subject wins out over the post-verbal one. For simplicity, let us assume that the post-verbal subject does not move out of the VP.[25] We are assuming that the difference between the two structures in terms of their conforming to Move is cancelled out by some other constraint relevant for subject raising in general. The post-verbal subject structure is ruled out if there is no overt negative head as the SM will be unlicensed otherwise and the pre-verbal subject structure does not require an inserted negative head as there is no SM. However, because there is no inserted SM or head when the subject moves through SpecNgP, this structure presents fewer violations of Ins and hence is optimal, contrary to fact.

The problem is given a satisfactory solution once we recognise the two presumed optional structures are in fact not options at all, but are related to two different inputs. Recall that we are assuming that pre-verbal elements move to a topic position and thus there is a topic structure present in these structures which is absent from those with only post-verbal elements. It would not be unreasonable to suggest that the topic structure is itself the result of some element in the input, a 'topic marker', and that this is the catalyst for the movement of the subject or object in these structures. Thus, we associate the inputs in (59a) and (59b) with the optimal structures in (60a) and (60b) respectively:

(59) a. {telefona, nessuno}
b. {TOP, telefona, nessuno}

(60) a. $[_{NgP}$ SM_i non telefona nessuno$_i]$
b. $[_{TopP}$ nessuno$_i$ $[_{NgP}$ t_i e telefona $t_i]]$

If these structures are related to different inputs, they are not in competition and therefore one cannot block the other. We can see, by splitting table (58) into two, the right structures are indeed deemed optimal under the assumptions we have made so far:

(61)

	MOVE	LSM	INS	HEAD	UNISPEC
[$_{NgP}$ SM$_i$ *e* telefona nessuno$_i$]	√	⋆!			
⇒ [$_{NgP}$*SM*$_i$ non telefona nessuno$_i$]	√	√			

(62)

	MOVE	LSM	INS	HEAD	UNISPEC
⇒ nessuno$_i$ TOP [$_{NgP}$ t_i *e* telefona t_i]	⋆	√	⋆		
nessuno$_i$ TOP [$_{NgP}$ t_i non telefona t_i]	⋆	√	⋆⋆!		

Note that this leads to the assumption that the constraint which forces the movement of the subject to the topic position outranks MOVE, thus movement is not an option in topicalised structures. This certainly would seem to be true in cases where the object is topicalised: if the object is not moved to the topic position it is not given a topic reading.

Italian negation phenomena, though complex at first sight, is capturable under the basic assumptions we have proposed with only a few additional, but justifiable, assumptions concerning the structure of the Italian clause. This adds further weight to our basic assumptions.

3.3.3 Hungarian

Hungarian is a NC language and therefore it ranks INS above UNISPEC. However, it differs from the languages we have discussed so far in two ways. First, the head of NgP is obligatory and hence HEAD is ranked above INS. Second, it optionally allows movement of its negative operators or the insertion of SMs for them, and hence MOVE and INS must be equally ranked. The data in (63) can therefore be accounted for by the ranking in (64):

(63) a. Gyula nem érti
G. not understands-def_obj
'Gyula doesn't understand it'

b. senki nem érti
no one not understands-def_obj
'no one understands it'

c. *senki érti

d. nem látott senki semmit
not saw no one nothing-ACC
'no one saw anything'

e. senki nem látott semmit
'no one saw anything'

f. senki semmit nem látott
'no one saw anything'

(64) HEAD $\gg$ {INS, MOVE} $\gg$ UNISPEC $\gg$ LSM

It is obvious how placing HEAD at the top of the ranking will ensure that the only optimal candidates will be those with an overt negative head. One consequence of this, however, it that any SM that appears in SpecNgP will always be licensed. We have placed LSM low down in the ranking, but it is clear that even if it were higher, it would always be satisfied and hence it effectively does no work in deciding the optimal candidate.

All that remains to be shown is how optionality of operator movement and SM insertion is achieved and how DN structures are ruled out. Consider the following example:

(65) a. senki nem látott semmit 'no one saw anything'
b. {senki, látott, semmit}

(66)

	HEAD	{INS	MOVE}	UNISPEC	LSM
$[_{\text{NgP}}$ SM$_{\text{i}}$ nem $[_{\text{NgP}}$ SM$_{\text{j}}$ nem látott senki$_{\text{j}}$ semmit$_{\text{i}}$]]	√	{$\star\star\star$ $\star\star\star$	√}!		
$[_{\text{NgP}}$ SM$_{\text{i}}$ nem $[_{\text{NgP}}$ senki$_{\text{j}}$ nem látott semmit$_{\text{i}}$ t_{j}]]	√	{$\star\star$ $\star\star\star$	$\star$}!		
$[_{\text{NgP}}$ senki$_{\text{i}}$ nem $[_{\text{NgP}}$ semmit$_{\text{j}}$ nem látott t_{j} t_{i}]]	√	{$\star\star$ $\star\star$	$\star\star$}!		
⇒ $[_{\text{NgP}}$ SM$_{\text{i}}$ SM$_{\text{j}}$ nem látott senki$_{\text{j}}$ semmit$_{\text{i}}$]	√	{$\star\star$ $\star\star$	√}	$\star$	√
⇒ $[_{\text{NgP}}$ senki$_{\text{i}}$ SM$_{\text{j}}$ nem látott semmit$_{\text{j}}$ t_{i}]	√	{$\star\star\star$	$\star$}	$\star$	√
⇒ $[_{\text{NgP}}$ senki$_{\text{i}}$ semmit$_{\text{j}}$ nem látott t_{j} t_{i}]	√	{$\star\star$	$\star\star$}	$\star$	√

The only candidates worth considering have overt negative heads otherwise they do not survive past the first constraint. In this example, the insertion of an extra NgP does not necessarily entail more violations of INS than any optimal candidate: the case where both operators are moved to separate SpecNgPs produces four INS violations as does the insertion of one NgP and two SMs in its specifier position. However, as the former also contains two MOVE violations whereas the latter does not violate MOVE at all, only the latter survives. While the insertion of two SMs into the same SpecNgP constitutes the greatest number of violations of INS of any of the optimal structures, note that any less insertions entails exactly the same number of MOVE violations in order to satisfy the operator's semantic requirements. Thus, taking the violations of INS and MOVE together, as long as there is only one NgP, there will always be exactly the same number of violations of these two constraints considered jointly. Of course, all surviving candidates after this violate UNISPEC, which therefore has no effect.

It is not only unnecessary, but also undesirable to treat these cases of operator movement as associated with different inputs, as we did with the case of pre-verbal arguments in Italian. This is because this is not a case of movement for any reason other than to provide the operator with its scope. The Italian example, though it made use of SpecNgP, ultimately involved movement to the topic position and it is clear that Italian operators do not move to get scope. Thus, movement to SpecNgP for its own ends is unmotivated in Italian. The movement of the operators in Hungarian is not a case of toplicalisation and hence this movement must be motivated as movement to provide the operator with its scope. This is supported by the fact that non-negative operators are not allowed to move to this position:

(67) *mindenki nem érti
'everyone not understands'

The fact that Hungarian neatly fits into the OT theory we have proposed without any outstanding problems is again strong support for this account. We now move on to consider the slightly more problematic case of West Flemish.

3.3.4 West Flemish

In the following analysis, we follow Haegeman & Zanuttini (1991) and Haegeman (1992) in assuming that NgP in WF dominates TP and is dominated by AgrP. However, we also allow a lower NgP which dominates at least VP. Whether this second NgP also dominates TP is unanswerable from the data and nothing in our analysis depends on this issue. We therefore minimally assume two NgP positions, both above the VP and below AgrP.

WF is more complex that the other languages we have so far examined as so many of the relevant considerations interact with each other. Being a V2 language, the word order inside embedded clauses is more restricted and thus to avoid unnecessary complication we will mainly restrict our attention to subordinate clauses. The following data demonstrates the main features of the WF negative constructions:

(68) a. da ze nie ketent (en-) was
that she *Op* pleased not was
'that she was not pleased'

b. da ze me niks ketent (en-) was
that she with nothing pleased not was
'that she was pleased with nothing'

c. over niemand (en-) klaapt ze
about no one not talks she
'she talks about no one'

d. da ze me niemand nie ketent (en-) was
that she with no one *Op* pleased not was
'that she wasn't pleased with anyone'

e. da ze nie me niemand ketent (en-) was
that she *Op* with no one pleased not was
'that she wasn't pleased with nothing'

Simple sentential negation is expressed in WF through the negative element *nie* and an optional clitic *ne-*. Following Haegeman & Zanuttini (1991), we assume that the clitic is the head of NgP and therefore that *nie* is a negative operator generated directly in SpecNgP. Thus, while the other languages we have examined make use of an abstract negative operator to express sentential negation, WF has an overt one. We will treat this similar to the abstract operators of other languages in assuming that this is the element that appears in the input.

Another point to note in the first example is the fact that the negative head is optional. As the negative operator is not empty, it will not need licensing and hence LSM is irrelevant for this example. The optional presence of the head must therefore be due to INS and HEAD not being ranked with respect to each other. The second example concerns a negative operator which is not generated inside NgP. However, this operator has moved out of its base generated position to the right of its head, into a pre-VP position, presumably SpecNgP. This would indicate that INS is ranked higher than MOVE. From these two examples, we can see that WF never has an empty element in SpecNgP and therefore any consideration of licensing such elements is irrelevant. We have already claimed that INS and HEAD are not ranked with respect to each other and given that the head is never needed to license any empty element, we predict that the head should be optional in all circumstances: this, on the whole, seems to be true.[27]

(68c) gives an example of a WF main clause in which the negative PP is moved to SpecCP. In many ways this situation is like the Italian pre-verbal argument case: the negative operator is moved to some position outside of NgP (SpecCP, in this case) moving through SpecNgP on its way. Hence it is assigned its scope via its trace in SpecNgP. Once again no SM is present, hence no licensing head is required and the head is optional.

The examples (68d) and (68e) demonstrate the NC and DN structures of WF, respectively. As discussed in section **2,** we assume a single NgP in (68d) and consequently that the operator is left adjoined to the SpecNgP containing *nie*. There are two NgPs in (68e) and both operators therefore occupy their own SpecNgP positions. As both structures are possible, we assume that INS and UNISPEC are ranked equal.

Thus, the ranking we propose for WF is:

(69) {INS, UNISPEC, HEAD} ≫ MOVE ≫ LSM

Simple sentential negation will be given a similar treatment in both subordinate and main clauses as V2 phenomena will not affect the negative operator generated inside NgP. We present here the case of the subordinate clause:

(70) a. da ze nie ketent (en-)was ‘that she wasn't pleased’
b. {. . . , da, ze, nie, ketent, was}

(71)

	{INS	UNISPEC	HEAD}	MOVE	LSM
⇒ da ze [$_{\mathrm{NgP}}$ nie t_i ketent t_i] was$_i$	{⋆	√	⋆}	√[28]	√
⇒ da ze [$_{\mathrm{NgP}}$ nie t_i ketent t_i] en-was$_i$	{⋆⋆	√	√}	√	√

As stated above, we assume that the operator *nie* is given in the input and hence that NgP must be inserted to accommodate it. The relevant choice is whether to insert the negative head or not. As there is no empty element to be licensed, LSM will treat both cases equally. However, one option violates INS and the other violates HEAD. As both of these are equally ranked there will be two violations of the constraints at this particular rank, including the one caused by the insertion of the NgP, whether or not an overt head is inserted. Thus, both structures are optimal.

With a single negative operator generated outside of NgP, the main issue is whether to insert a SM or move the operator to SpecNgP. Obviously the second option will be preferred as INS is ranked higher than MOVE. The optionality of the head is handled as above and we will not demonstrate this again.

(72) a. da ze me niks ketent en-was ‘that she was pleased with nothing’
b. {. . . , da, ze, ketent, me, niks, was}

(73)

	{INS	UNISPEC	HEAD}	MOVE	LSM
da ze [$_{\mathrm{NgP}}$ SM$_i$ t_j ketent me niks$_i$ t_j] en-was$_j$	{⋆⋆⋆	√	√}!		
⇒ da ze [$_{\mathrm{NgP}}$ me niks$_i$ t_j ketent t_i t_j] en-was$_j$	{⋆⋆	√	√}		

As the table shows, the optimality of the movement of the operator is decided on the violations of INS: inserting a SM adds an extra violation of this constraint and thus leads to a non-optimal structure.

Consider now a similar case in a main clause:

(74) a. over niemand (en-)klaapt ze 'she doesn't talk about anyone'
b. {ze, klaapt, over, niemand}

(75)

	{INS	UNISPEC	HEAD}	MOVE	LSM
⇒ $[_{\text{CP}}$ over niemand$_{\text{i}}$ (en-)klaapt$_{\text{j}}$ ze $[_{\text{NgP}}\ t_{\text{i}}\ t_{\text{j}}\ [_{\text{VP}}\ t_{\text{i}}\ t_{\text{j}}]]\ t_{\text{j}}]$	{($\star$)$\star$	$\surd$	$\surd$}	$\surd$	$\surd$
⇒ $[_{\text{CP}}$ ze (en-)klaapt$_{\text{j}}$ $[_{\text{NgP}}$ over niemand$_{\text{i}}$ $t_{\text{j}}\ [_{\text{VP}}\ t_{\text{i}}\ t_{\text{j}}]]\ t_{\text{j}}]$	{($\star$)$\star$	$\surd$	$\surd$}	$\surd$	$\surd$
ze en-klaapt$_{\text{j}}$ $[_{\text{NgP}}$ SM$_{\text{i}}$ t_{j} $[_{\text{VP}}$ over niemand$_{\text{i}}$ $t_{\text{j}}]]\ t_{\text{j}}$	{($\star$)$\star$ $\star$	$\surd$	$\surd$}!		

As with other cases of V2, the verb raises to C in main clauses and another element raises to SpecCP. In (74a) the negative operator is in SpecCP. Haegeman & Zanuttini (1991) assume that in this case the operator and the head license each other in the CP. However, from our perspective, we assume that as the negative operator moves to SpecCP it passes though SpecNgP, hence picking up its scope, and thus there is no SM to be licensed. Again, the movement is triggered by other considerations and is not relevant here, apart from the movement to SpecNgP which will be forced by INS. Of course the subject could have been the element to move to SpecCP, in which case the negative operator would move vacuously to SpecNgP. What is ruled out is the operator staying inside the VP.

Finally, we consider the case in which there are multiple negative elements. Here WF allows both NC and DN readings. However, both readings are related to sentences with different word orders. This is due to the fact that the NC reading has both operators in the specifier of a single NgP, where as the DN reading has one operator in the specifier of a lower NgP. The case we will consider involves the sentential negative operator *nie* and another operator:

(76) a. da ze me niemand nie ketent en-was
'that she wasn't pleased with anything'
b. da ze nie me niemand ketent en-was
'that she wasn't pleased with nothing'
c. {... da, ze, nie, ketent, me, niemand, was}

(77)

	{Ins	UniSpec	Head}	Move	LSM
⇒ da ze $[_{\mathrm{NgP}}$ me niemand$_{\mathrm{i}}$ nie t_{j} ketent t_{i} $t_{\mathrm{j}}]$ en-was$_{\mathrm{j}}$	{⋆⋆	⋆	√}	⋆	√
⇒ da ze $[_{\mathrm{NgP}}$ nie t_{j} $[_{\mathrm{NgP}}$ me niemand$_{\mathrm{i}}$ t_{j} ketent t_{i} $t_{\mathrm{j}}]$ en-was$_{\mathrm{j}}$	{⋆⋆⋆	√	√}	⋆	√

Most other possibilities have been considered above. One which has not is the possibility of inserting an overt head into the second NgP (presumably producing the verb form *en-en-was*). While this may be ruled out for morphological reasons, the same which rule out the form *wasn'tn't* in English, we can see that this candidate also yields one more violation of Ins and hence would not be optimal anyway.

3.3.5 British Non-Standard English

In this last section we turn to examine particular negation phenomena common in various non-standard varieties of English. There probably are a number of Non-Standard Englishes which differ in terms of how they encode negation, however, we will concentrate on one of these which is fairly common in Britain.[29] Let us refer to this as British Non-Standard English (BNSE).[30] The main difference between this variety and Standard English is that, like WF, it allows NC as well as DN constructions. It differs from WF in that the same surface form can often be given both interpretations. Consider the following data:

(78) a. I didn't see Bill[31]
b. I saw no one
c. I didn't see no one
'I did not see anyone/no one'
d. no one saw nothing
'no one saw anything/nothing'
e. he can't have not left

Of course, BNSE has much in common with Standard English: its NgP has a fairly unlimited distribution; the negative head is necessary only to licence empty elements in SpecNgP and it does not move its operators to give them scope. Thus, we expect the only difference between the ranking of constraints in these dialects will concern Ins and UniSpec. Given the ranking for Standard English in (79a), the ranking we propose for BNSE is (79b):

(79) a. UniSpec ≫ Move ≫ LSM ≫ Ins ≫ Head
b. Move ≫ LSM ≫ {Ins, UniSpec} ≫ Head

The case of simple sentential negation (78a) will be handled as for Standard English: the input contains an abstract operator which will require licensing by an inserted overt head as LSM outranks INS. Similarly for (78b), where the input contains an overt operator which licenses its own SM and hence the overt head is unnecessary. A question arises here as to whether the NC reading of (78c) can be related to the same input as for (78b): both constructions meaning the same thing. The theory we have developed however suggests that this is not so, as if both structures are in competition with each other, the one that does not insert an overt head should win out. Thus, given the input in (80), we compare the structure generated by inserting a high NgP containing a SM which is licensed by an overt head with that generated by inserting a low NgP where the SM is licensed by the operator.

(80) {I, saw, no one}

(81)

	MOVE	LSM	{INS	UNISPEC}	HEAD
I [$_{\text{NgP}}$ SM$_i$ not saw no one$_i$][32]	√	√	{⋆⋆⋆	√}!	
⇒ I saw [$_{\text{NgP}}$ SM$_i$ *e* no one$_i$]	√	√	{⋆⋆	√}	

As the second structure in (81) is optimal but yet the first is also grammatical, we can only assume that the two must be related to different inputs and hence are not in competition.

Of course, there is an input different from (80) that could be the basis of the NC structure in (81), *i.e.* the one associated with the DN reading of this sentence. Let us compare these then:

(82) {I, *Op*, saw, no one}

(83)

	MOVE	LSM	{INS	UNISPEC}	HEAD
⇒ I [$_{\text{NgP}}$ SM$_i$ Op not saw no one$_i$]	√	√	{⋆⋆⋆	⋆}	√
I [$_{\text{NgP}}$ Op not saw [$_{\text{NgP}}$ SM$_i$ *e* no one$_i$]]	√	√	{⋆⋆⋆⋆	√}	⋆!

Putting aside the problem that one of the grammatical structures is deemed non-optimal, we will return to this shortly, we can see that the structure under question is indeed an optimal one for the given input: as both the SM and the abstract operator occupy the same SpecNgP, they will be given the same scope and hence neg-factorisation will take place, giving a NC reading for the sentence *I didn't see no one.* The fact that this has exactly the same interpretation as a sentence formed from another input (*e.g. I saw no one*) is not problematic. What these structures have in common is that they are given the same interpretation, which is something which follows from the interpretative principles we are supposing. To insist that two structures which are given the same interpretation should be related to the same input is to ascribe to the input properties that we are not assuming for it: the input is simply a set of lexical items and is not itself a semantic object. Thus, whether an input gives rise to optimal structures which have different interpretations or whether two inputs have optimal structures which have the same interpretation is something which is determined by the system as a whole, not by the input alone.

Returning to the problem in (83), note that this is exactly the sort of problem that we envisaged in section **1.2.3.** The cause of the problem is the fact that we wish to define two structures as being optimal and as a result, all constraints lower down in the ranking have a potential effect. It is a fact that HEAD plays no role, other than the adverse one we are discussing, in determining the optimality of negative structures in BNSE. Indeed, as we have noted above, HEAD plays no role in determining the optimality of any structure in Standard English. The only ways to overcome this problem is either to ensure that there is only ever one optimal candidate associated with each input, or to assume that not all constraints are relevant for each language. The first option means that we must abandon our approach to optionality, which as we have seen has some very positive features.[33] The second option, however, requires a more complex account of linguistic variation: languages differ not only in terms of how they rank constraints, but also in terms of which constraints they select. The question is, are there considerations that would argue against this assumption besides that of complexity? If there are no such arguments, then the difficulties discussed in Note 33 amount to an argument in favour of adopting the more complex view of linguistic variation.[34] Thus, along these lines, we now claim that the ranking relevant for BNSE is (84):

(84) MOVE $\gg$ LSM $\gg$ {INS, UNISPEC}

There are possible other modifications to the rankings we have suggested above. However, as these modifications are inconsequential for the data we have examined, we will not consider them here.

Finally, we turn to example (78e), repeated below as (85):

(85) he can't have not left

This sentence obviously contains two NgPs as there are two negative heads and hence it is given a DN reading. As it is not the negative head which appears in the input, we assume that here the input must contain two abstract negative operators. However, a further problem arises as can be seen by considering table (87):

(86) {he, *Op*, *Op*, can, have, left}

(87)

	Move	LSM	{Ins	UniSpec}
he can $[_{\text{NgP}}$ Op not have $[_{\text{NgP}}$ Op not left]]	√	√	{⋆⋆⋆⋆	√}!
⇒ he can $[_{\text{NgP}}$ Op Op not have left]	√	√	{⋆⋆	⋆}

The DN structure is blocked by the more optimal NC structure and so we should not be able to have structures such as (85). This problem can be solved by noting that the abstract negative operator has a limited distribution in BNSE in exactly the same way that it has in Standard English. Thus, the sentences in (88) are also ungrammatical in BNSE:

(88) a. *$[_{\text{NgP}}$ *Op* not I saw Mary]
b. *I saw $[_{\text{NgP}}$ *Op* not Mary]

Thus, we may assume that the abstract negative operator cannot enter into a NC structure with another abstract negative operator as a lexical restriction and thus the 'optimal' structure in (87) is ruled out on other grounds, leaving the desired structure as optimal.

We can see, then, that although BNSE turns out to be one of the most problematic languages to analyse in terms of the theory we have proposed, a largely satisfactory analysis can be given if a number of not unsupported assumptions are also made.

4 Conclusion

In this paper, we have attempted simultaneously to build a general syntactic theory based on OT principles and to demonstrate this theory in a limited syntactic domain. It is worth reviewing the results of these separately.

As far as the broader theory is concerned, we have shown that interesting syntax can be done within the OT framework and without importing a substantial amount of mechanisms from other generative theories. Taking GEN to be the sole mechanism of structure building, we can simply take the input as being a set of items taken from the lexicon. From our discussion, it seems that the input is made up of lexical/thematic categories and certain operators which may or may not lack phonological content. Gen may introduce other elements, such as functional heads and other abstract operators. Furthermore, we also have to assume a notion of a legitimate input, where legitimacy is determined in terms of possible satisfaction of the lexical requirements of the elements contained in the input set. Gen, then, combines the elements of the input along the lines of certain well defined structural processes and while it is restricted to these processes, it can carry these out in an unconstrained way, thus producing an potentially infinite number of candidates.

The main innovation of this paper within the OT framework has been within the evaluation. We have made two main proposals. First, we claim constraints are not necessarily strictly ranked with respect to each other, and hence the system is capable of allowing languages to demonstrate two sets of phenomena which other languages demonstrate only one or the other. In other words, we can capture optionality fairly easily in these terms. Second, partially as a result of the fist, we claim that not all constraints are relevant to all languages and thus languages differ not only in how they rank constraints, but also in terms of which constraints they select. This gives a more complex theory of language variation, but the alternative theory is more complex in other ways: it must be guaranteed that each input is associated with a single optimal candidate and that optimality will be determined for each candidate by a certain point in the ranking so that irrelevant constraints can be 'hidden' below this point. Furthermore, optionality cannot be captured in such a system. We conclude that the more complex view of linguistic variation is to be preferred.

Turning now to the specific area of negation, we have demonstrated that with just five constraints a wide variety of phenomena and associated differences across a number of languages can be captured in an elegant and simple way which is completely harmonious with the OT theory we have developed. Obviously, there still remain problems that we have only partial or even no solutions to. This is to be expected, given the sate of the theory. These problems point the way for further research and demonstrate that the theory is capable of throwing up interesting questions about the nature of specific linguistic phenomena.

NOTES

[1] Of course, the question arises as to how the verb comes to be placed in the input. There are a number of possible approaches one could take depending on what view is taken about functional elements and whether these are inserted into the input or added by GEN. I leave this as an open issue.

[2] This is not to say that we will never have to state the relative ranking of non-conflicting constraints. Consider a situation in which three constraints, A, B and C, are such that A and B conflict with C, but not with each other. Suppose that in a given language A outranks C and C outranks B. In this case A outranks B, though nothing of consequence follows from this.

[3] The analysis of *not* as a head is not uncontended, see Rizzi (1990) and Newson (1991) for arguments that *not* may be a specifier element and Ouhalla (1990)

that it is a head. In the present paper, however, we will assume along with Ouhalla that *not* is a head.

[4] An alternative treatment of negative adverbials is to place them in the specifier of NgP itself as has been proposed for French and Italian negative adverbs (Pollock 1989; Belletti 1994). This might then be used to account for the limited distribution of the English negative adverb never. As we shall see, NgP in English has a fairly free distribution, but we might expect that it cannot be sentence final as it needs to have scope over at least part of the sentence. If *never* sits in SpecNgP, this would account for why, unlike other adverbs, it cannot be sentence final:
(i) I have never seen an aardvark
(ii)*I have seen an aardvark never

[5] It may be noted that some of these sentences involve consitituent negation rather than sentential negation. However, I am assuming that, in the relevant respects, constituent and sentential negation are the same in that they both make use of NgP. The difference between them can be seen precisely in terms of where the NgP is placed in the sentence: a higher NgP provides wider scope which is 'sentential' when it all the relevant elements within its scope and a lower NgP provides a narrower scope which is therefore confined to constituents of the sentence.

[6] Interestingly, some Hungarians would accept this sentence as grammatical, but with the second negative having scope only over the verb and hence the focus of the negative cannot include the object *a gyereket.* This differs from constituent negation (see Note 5) in that in the latter, the negation has scope over the entire VP, not just the verb. Thus, in the English equivalent of (35b) either the verb or the object can be the focus of the negation:
(i) I can't not 'see the child 'I can't hear her'
(ii) I can't not see the 'child 'its the adult that I can't see'
In Hungarian, only the first of these interpretations is available. This might be taken to indicate that this sort of negation is not 'constituent' negation but 'lexical' negation, in which case it might be argued not to involve NgP at all.

[7] West Flemish is like French in having an overt negative operator generated in SpecNgP as well as a head. Thus *nie* and *en-* are the equivalent of *pas* and *ne.*

[8] This definition is motivated by the following considerations. A negative head can license any number of SMs in its specifier position or adjoined to it, as long as no overt element intervenes. Furthermore, a negative operator can licence its own SM in SpecNgP as long as no overt material intervenes. Thus the following are legitimate licensing configurations:
(i) $[_{NgP}$ SM_1 SM_2 … SM_n NG …
(ii) $[_{NgP}$ SM_i e Op_i …
However, if a negative element could license a SM that it c-commanded, we would expect the following structure to be grammatical in English:
(iii) $[_{NgP}$ SM_i e no one_i $[_{NgP}$ Op e likes Bill]]
This would mean 'everyone likes Bill' due to the cancelling out properties of the double negation. Obviously, the sentence *no one likes Bill* cannot have this

meaning and hence we conclude that this structure is not possible. Thus, the negative operator can license its own SM, but not the abstract operator in the lower NgP which it is overtly adjacent to, but which it c-commands.

The following example shows that traces can license SMs:

(iv) they didn't$_i$ [$_{NgP}$ Op t_i like Bill]

The English contracted negative head is a clitic which moves with the finite verb to the inflections. The abstract operator will of course still need licensing, and hence we must assume that this is done via the trace of the moved head.

[9] Obviously the conditions under which a sentence is defined as syntactically 'negative' or 'positive' are complex. For example, (21) may also be co-ordinated with *and neither did,* which is a property of negative sentences:

(i) the girls saw no one and neither did the boys

(ii)*the girls saw an aardvark and neither did the boys

However, it does not trigger a positive tag (with the non-ironic interpretation), which is also a property of negative sentences:

(iii)*everyone saw no one did they

(iv) the girls didn't see the aardvark did they

Thus, while (21) has properties of positive sentences, it also possesses some negative properties. Importantly for our purposes, however, is the fact that this sentence *can* have positive properties, which sets it appart from other sentences which are wholly syntactically negative.

[10] It might be argued that this is not an ambiguous sentence, but merely vague as there are any number of conditions under which a negative proposition can be true. However, Kempson's (1977) VP deletion test for ambiguity would argue otherwise. Consider:

(i) Bill realised that everyone didn't see him and so did Fred

Concentrating on the meaning where the reconstructed *him* refers to *Fred,* this sentence is only two ways ambiguous:

(ii) Bill realised [$\forall x \ \sim[x$ saw Bill]] & Fred realised [$\forall x \ \sim[x$ saw Fred]]

(iii) Bill realised [$\sim\forall x$ [x saw Bill]] & Fred realised [$\sim\forall x$ [x saw Fred]]

This indicates that the scope difference in these interpretations is a true case of ambiguity, not just a vagary.

[11] Interestingly, the missing interpretation in (25a) present in (25b) with the object having wide scope is exactly the one we get if the negative object is fronted:

(i) no one did everyone see

Evidently here the NgP is higher and hence the negative operator has wide scope. It is interesting to note that the scope interpretation of negative operators in Hungarian does not alter, no matter where the operators are placed with respect to each other:

(ii) nem látott senkit senki

(iii) senki nem látott senkit

(iv) senkit senki nem látott

Each of these sentences has exactly the same interpretation: *everyone saw no one* (or equivalently, *no one saw anyone*). This can be accounted for under the assumptions we are making: the NgP in Hungarian is limited to a single position, no matter where the negative operators appear in the sentence. In English, however, the position of the NgP differs with respect to the positioning of negative operators. Thus, there is only one interpretation determined by the NgP in Hungarian, but variable interpretations are given in English as NgP has a variable distribution.

[12] Of course, this sentence is grammatical in Standard English, but only with the interpretation *they saw someone* and not the intended interpretation. It is, however, grammatical with the intended interpretation in certain non-standard varieties of English. We will examine one of these in the next section.

[13] Haegeman (1992) attempts to distinguish DN from NC in terms of the notion of 'strong' and 'weak' negative heads: the former being associated with NC and the latter with DN. A strong head is one that has its own negative features and hence enters into passive agreement with a specifier which also has its own negative features, or into active agreement with a specifier which does not. In the latter case the specifier recieves its negative features by dint of the fact that it agrees with the negative head. A weak negative head actively agrees with a negative specifier and receives its negative features from the element in the specifier position. Unfortunately, there are a number of problems facing this analysis. First, circularity is unavoidable: there is no independent way to tell whether a negative head is strong or weak other than whether the langauge has NC or DN. Second, if languages have either strong or weak negative heads, the prediction is that languages should either allow NC or DN structures, but some langauges allow both, *e.g.* West Flemish and Non Standard English. This leads to the contradictory conclusion that negative heads in these languages do and do not have their own negative features.

[14] This notion may be extended to all functional elements. However, rather than following up these possibilities, we will confine ourselves to the phenomena at hand and remain agnostic on this issue.

[15] The interaction between the grammar and the semantics obviously needs elaboration before we can declare these sorts of problems solved. If we assume, following standard generative practice, that the semantic component has an interpretative function, then such semantic constraints apply after the syntactic ones. This raises the possibility that the structures determined as optimal by the grammar may be unacceptable according to semantic principles. If the system is constructed so that each input has at least one optimal (and interpretable) structure, then the most optimal structure of the whole system may not be the most optimal of the syntactic evaluation component. In other words, when a syntactically optimal structure is rejected on semantic grounds, the next syntactically optimal candidate must be considered until one which satisfies semantic principles is found.

[16] The issues are far from clear here though, particularly in the light of a conclusion that we will reach later that HEAD is not a relevant constraint in all languages,

including English. If this is true then inversion and compulsory overt complementisers will have to be handled by some other means. This awaits further investigation.

[17] The one exception to this is in cases of negative fronting:

(i) no one$_i$ did everyone see t_i

The fact that the negative has wide scope in this construction would argue that it has moved to a scope position which is higher than that determined for the subject. As this seems to be exceptional, we might view it as a different case to those considered above. Specifically, it might be claimed that this construct is related to a different input to any of the clauses in (37): the input for (i) may contain an element which forces the movement and which is not present in the inputs for the sentences in (37). As (i) is obviously some form of a topicalised structure, it would not be outlandish to suggest that its input contained some abstract topic marker, which would obviously be absent for the other inputs.

[18] This is one case where our theory will have something to say about the distribution of NgP. This is because we are assuming that in English NgP has a fairly free distribution and thus whatever the structural principles that determine this will allow a NgP in either of the relevant positions. In this case, the constraints that we are concerned with are given the opportunity to decide on issues of optimality.

[19] Of course, (49c) will be ruled out by the fact that the negative operator has scope over nothing, which as we have previously said should not be allowed. However, the other examples in (49) are obviously not of this order and hence we maintain that the dstribution of the abstract operator is lexically limited.

[20] When there is both a pre-verbal subject and a topicalised object we assume that the object is adjoined to the specifier position occupied by the subject. This is compatible with our assumptions as, as we shall see, UniSpec is placed low in the ranking for Italian.

[21] Bonet (1991) proposes an analysis of Catalan in which the post-verbal subject occupies its base generated position inside the VP. The analysis may transfer to Italian.

[22] Haegeman (1992) reports that for some Italians an overt head in these structures is marginally possible, though intuitions do vary. One possibility might be that for some Italians the trace in SpecNgP is sufficiently like a SM, being empty, that it may trigger the presence of the head. The marginality and uncertainty of the data, however, make it difficult to state precisely what is going on here.

[23] To comply with the definition of licensing we gave earlier, we must assume that the SM c-comands the trace in this configuration. This is not problematic: as the trace is adjoined to SpecNgP it is not dominated by it and therefore both the SM and the trace share all dominating maximal projections.

[24] Here we assume that this structure violates LSM, despite the fact that there is no overt element intervening between the head and the 'unlicensed' SM, as we are assuming that a head only licenses the SMs that c-command it and not those

which do not. However, this assumption is not essential as the structure also violates Ins more times than the optimal candidate.

[25] Again this is harmless as presumably if the post-verbal subject does move out of the VP, this movement is forced by some constraint which presumably outranks Move and hence cancels the effect of this constraint.

[26] In fact, any of the operators can move to SpecNgP in any order and so there are six possible permutations of this sentence. For reasons of space, we restrict our attention to the three in (63).

[27] In fact, the negative head is not optional in all cases and there are some places where it cannot occur, in non-finite clauses, for example. However, here other considerations will rule against the appearance of the head. The head is a clitic, which therefore must attach to some element to satisfy its own morphological properties. Clitic attachment in WF appears to be done as the verb moves through NgP on its way to the inflectional nodes. In infinitival clauses there is no need for the verb to move to the inflections and hence it will not pass through the NgP. The head, if it is present, will therefore be unsupported and hence cause an ungrammaticality. Presumably, then, a more optimal structure is formed by leaving the head out and hence we account for why the negative head is obligatorily absent in infinitival clauses on independent considerations.

[28] Obviously Move is violated in these cases by the movement of the verb. However this is irrelevant to our concerns and thus we are only concentrating on negative operator movement. As no operator is moved we take Move to be satisfied. Of course, we assume a higher ranking constraint will force the verb movement in these cases.

[29] For the simple reason that I am a speaker of this dialect.

[30] I do not mean to imply by this that there is just one non-standard variety of English in Britain, but merely that from informal investigation, it seems that the relevant features of this dialect are quite wide spread in Britain and this contrasts with the situation in America, for example.

[31] I will stick to the contracted form of the negative head in order to avoid the incongruity caused by the use of the full form, which is normally associated with formal language use, in a non-standard variety. Of course, the full form is possible in BNSE, which must use it in cases where contraction is impossible or leads to an even more formal form:

(i) he might not come

Problems arise however in trying to determine what is grammatical in BNSE using more formal constructions. For example, in principle (ii) should have both NC and DN readings in BNSE, though because it would not normally be used outside of the standard variety, it is hard to determine how it is treated in the non-standard system:

(ii) I did not see no one

I assume that because (iii) and (iv) have both interpretations, then so does (ii),

though this is obviously a very atypical thing to say in this variety:
(iii) I didn't see no one
(iv) he may not have seen no one

[32] As with Standard English, here we ignore the process of *do*-insertion.

[33] Note that this in itself would not get rid of the problem as it is still possible for ties to occur even over a strictly ranked set of constraints. Further restrictions would have to be enforced to ensure that not only are inputs and optimal candidates in a one to one relationship, but also that optimality for the entire language be decided by a definite point in the ranking so that we could ensure that any constraint after that point would be inconsequential.

[34] One obvious consideration at this point is the learnability of the system: is it possible to distinguish between the case where one constraint is ranked lower than another and that in which the second constraint does not enter the system at all? One way to envisage the learning situation is that constraints are only included in the ranking if there is positive evidence that they have effect. Thus the situations we are envisaging are distinguishable as the first assumes that although the lower ranked constraint is often violated in order to conform to a higher ranked one, there will still be situations in which it has effect. In the case where the constraint never has effect, it will not be included in the evaluation. Obviously a fuller learning theory would have to be worked out to really test this situation. In particular, it must be demonstrated that for each constraint and combinations of constraint, there are positive data to allow the correct selection and ranking of constraints to be put into effect.

REFERENCES

Belletti, A. (1994) 'Verb positions: evidence from Italian', In Lightfoot, D. and N. Hornstein, eds. *Verb Movement.* Cambridge: Cambridge University Press. 19–40.

Bonet, E. (1991) 'Subjects in Catalan', In Cheng, L. L.-S. and H. Demirdash, eds. *Papers on Wh-movement: MIT Working Papers in Linguistics* **13.**

Bródy, M. (1994) 'Lexico-Logical Form—a Radically Minimalist Theory'. Ms., to appear.

Chomsky, N. (1981) *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. (1986) *Knowledge of Language: its Nature, Origin and Use.* New York: Praeger.

Chomsky, N. (1989) 'Some notes on economy of derivation and representation', In Laka, I. and A. Mahajan, eds. *Functional Heads and Clause Structure: MIT Working Papers in Linguistics* **10.** 43–74.

Chomsky, N. (1992) 'A minimalist program for linguistic theory'. *MIT Occasional Papers in Linguistics* **1.** Cambridge, MA: MIT.

Chomsky, N. (1994) 'Bare Phrase Structure'. *MIT Occasional Papers in Linguistics* **5.** Cambridge, MA: MIT.

Grimshaw, J. (1993) 'Minimal Projection Heads and Optimality'. Ms. Rutgers University.

Haegeman, L. and R. Zanuttini (1990) 'Negative heads and the Neg-criterion'. Paper presented at GLOW.

Haegeman, L. and R. Zanuttini (1991) 'Negative heads and negative concord'. Ms. Université de Genève.

Haegeman, L. (1992) 'Negative heads and negative operators, the Neg-Criterion'. Ms. Université de Genève.

Higginbotham, J. and R. May (1981) 'Questions, quantifiers and crossing'. *Linguistic Review* **1.** 41–79.

Kempson, R. (1977) *Semantic Theory*. Cambridge: Cambridge University Press.

Laka, I. (1989) 'Constraints on sentence negation'. In Laka, I. and A. Mahajan, eds. *Functional Heads and Clause Structure: MIT Working Papers in Linguistics* **10.** 199–216.

Legendre, G., W. Raymond and P. Smolensky (1993) 'An Optimality-Theoretic typology of case and grammatical voice systems'. In *Proceedings of the 19th Meeting of the Berkeley Linguistics Society*.

Newson, M. (1991) 'Negative phrases: further considerations'. Paper presented at LAGB meeting, York.

Ouhalla, J. (1990) 'Sentential negation, relativised minimality and the aspectual status of auxiliaries'. *Linguistic Review* **7.** 183–231.

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

Rizzi, L. (1990) *Relativized Minimality*. Cambridge, MA: The MIT Press.

Stowell, T. and F. Beghelli (1994) 'The direction of quantifier movement'. Paper presented at GLOW, Vienna.

Zanuttini, R. (1991) 'Syntactic properties of sentential negation: a comparative study of Romance languages'. PhD. dissertation, University of Pennsylvania.