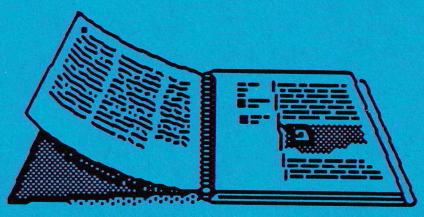
Forschungsschwerpunkt Allgemeine Sprachwissenschaft,
Sprachtypologie und Universalienforschung
der Förderungsgesellschaft Wissenschaftliche Neuvorhaben mbH

FAS Working Papers in Linguistics

Volume 2
September 1995



Edited by

Artemis Alexiadou
Nanna Fuhrhop
Paul Law
Sylvia Löhken

ISSN 0947-7055

Table of Contents

Optimal Questions*

Peter Ackema & Ad Neeleman
Utrecht University/OTS

1. *Constraint Interaction*

The traditional view of the organization of grammar is that, although facts are derived by a conspiracy of principles, the principles themselves do not compete. That is to say, no principle is violated in order to satisfy another principle. In a grammatical sentence, all principles are satisfied.

In the minimalist program (cf. Chomsky 1995) this picture is slightly different, due to the important role assigned to economy principles. Economy principles are not meant to evaluate a single derivation, but to choose from several possible derivations the one that best satisfies them. This means that with respect to economy conditions, though not with respect to other conditions, there is a difference with the traditional view in that economy conditions do not impose absolute requirements on a derivation or representation. This shifted perspective also implies that it is no problem if an economy condition conflicts with another constraint. In fact, all economy conditions disfavoring movement conflict with the conditions that trigger movement. For example, economy principles like Procrastinate and Shortest Steps conflict with the requirement to check strong features. This conflict is resolved by the assumptions that non-checking of a strong feature leads to a crashing derivation (i.e. 'check strong feature in overt syntax' is an absolute constraint on converging derivations) and that global economy takes into consideration converging derivations only. For example, in English it is necessary to place a WH in spec-CP to get a well-formed WH-question, despite the economy requirement to minimize overt movement. The fact that some constraints (the need to check strong features) must overrule others (Procrastinate, Shortest Steps) implies that these constraints are of different importance, which virtually means that they are ranked.

Constraint ranking is the key notion of optimality theory (Prince & Smolensky 1993). In the minimalist program 'ranking' is restricted to two types of constraints, absolute constraints and economy constraints, where in every language the absolute constraints outrank the economy constraints. Optimality theory proposes a more radical

* This paper grew out of discussions in the OTS Syntax Lunch, in particular with Tanya Reinhart, about the applicability of optimality theory in syntax. We would like to thank Frank Drijkoningen, Astrid Ferdinand, Jane Grimshaw, René Kager, Jacqueline van Kampen, Geraldine Legendre, Ludmila Menert, Alan Prince, Božena Rozwadowska, Eddy Ruys, Maaike Schoorlemmer, Paul Smolensky, and Fred Weerman for useful comments.

breach with the traditional view. There are no absolute constraints any longer (though see note 7). All constraints are ranked and evaluated collectively. Evaluation proceeds as follows. The structures to be compared are first evaluated with respect to the highest-ranked constraint. In the event of a tie, the remaining candidates are judged by the next highest constraint, and so on. Constraints therefore need not be surface true. A lower ranked constraint can be violated in an optimal form when this form scores better on higher ranked constraints than its competitors. (Note that even the highest-ranked constraint can be violated, namely in case there is no potential output that does not violate it).

Moreover, whereas in the minimalist program the ranking of economy constraints and other constraints is constant across languages, optimality theory seeks to explain language variation as a consequence of different rankings of constraints. Each possible ranking of the constraints must determine the grammar of a natural language (and each grammar of a natural language must be determined by some ranking of the constraints). This means that there is an interesting way of testing the validity of optimality theory. Given some well-motivated principles handling data in a specific language, reranking these constraints should result in a (partial) grammar of another language. If this indeed turns out to be the case, it is a strong argument for constraint ranking and reranking as proposed in optimality theory.

In this paper we will argue that language variation in the syntax of questions, in particular multiple questions, can indeed be accounted for by different rankings of a few well-motivated constraints. The paper is structured as follows. First, we will try to establish the constraints operative in the syntax of questions by considering English (section 2). We will then show how different rankings of these constraints account for the syntax of WH-questions in different languages. We will discuss Bulgarian, Czech, Chinese and French (section 3). The paper concludes with a brief summary of the proposed analysis (section 4).

2. *The Syntax of Questions*

Before we can test whether constraint reranking can account for the syntax of questions cross-linguistically, we must find out which constraints are operative in question formation. This is the topic of this section. In section 2.1, we will outline the theory of clause structure within which the analysis will be couched. In section 2.2, we will develop a minimal theory of question formation. From this theory two constraints can be derived, which in combination with the well-known Shortest Steps condition explain the syntax of questions in English (section 2.3).

2.1 *Functional Projections and Head Movement*

In the minimalist program, or more specifically in checking theory, WH-movement and the verb movement that accompanies it in some languages must be driven by the need to check WH-features. The difference between overt and covert WH-movement should be regulated by the strength of these features. In this paper we will take issue with this analysis, for the following reasons.

The theory of feature checking is only descriptively adequate, and not explanatorily, if no satisfactory definition of the strength of features can be given. Especially in the case of WH-features, this seems problematic. Compare for instance (1a) and (1b).

(1) a. Who(m)$_i$ did John talk to t_i
b. What$_i$ did John give t_i to who(m)
c. *John talked to who(m)

In a theory based on feature checking, one is forced to conclude on the basis of (1a) that *who(m)* has a strong WH-feature. If that were so, however, (1b) should crash. The hypothesis that *who(m)* can have either strong (in 1a) or weak (in 1b) WH-features is untenable. First, it is ad hoc since no morphological difference is detectable between strong and weak *who(m)*. Second, the ungrammaticality of (1c) (under a nonecho reading) is left unexplained, if a *who(m)* with weak features exists. Third, checking theory leads to a conceptually unattractive duplication of features in the syntactic representations. The WH-feature is not only present on the WH-expression (a necessity), but also in the functional position in which checking must take place.[1] Similar problems arise in connection with the accompanying verb movement. Given the principle of *Greed*, as assumed in checking theory, this movement must be caused by the need for the verb to check a strong WH-feature. The presence of inversion in (2a) versus the absence of inversion in (2b) then indicates that a verb like *will* has a variant with and without (strong) WH-features, even though no morphological evidence for this difference exists.

---

[1] This double representation of features might be used to solve the problem posed by (1) for checking theory. The solution consists of the assumption that the WH-feature of WH-words can be either strong or weak in English (as noted in the text this is an ad hoc assumption anyway, since there never is any morphological difference between the strong and the weak variant of the WH-word), while the WH-feature of C is always strong. This solution seems unattractive to us, since it assumes that the same feature can be strong or weak according to the position it happens to be generated in. This option is not otherwise attested, as far as we know. (In standard minimalism, features are assumed to be strong or weak, period). Note that it cannot be assumed that the WH-feature of C is a different feature from the WH-feature of the WH-expression, since in that case checking would not be possible.

(2) a. What$_i$ will$_j$ John [t_j see t_i]
b. John [will see a cow]

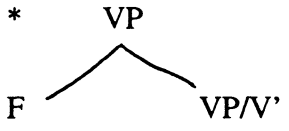
It seems, then, that in the domain of WH-movement, checking theory faces some serious difficulties, because ad hoc assumptions must be made about the featural content of lexical items.[2]

An alternative theory of question formation (that is, a theory of WH-movement and verb movement to functional positions) seems to be in order. We will now try to develop such a theory, based on the following two assumptions. First, functional structure is only generated when necessary. Second, generation of functional structure can be necessitated by the requirement that certain semantic functions of VP are syntactically marked. One of these functions would be that of 'question'. The remainder of this section deals with the first assumption. The second one is discussed in section 2.2.

Let us assume that, given the central role of economy, structure is only generated when required for wellformedness. Sentences minimally contain a proposition, which is syntactically encoded as a juncture of a subject and a verbal predicate (VP or V', depending on one's assumptions about the position of the subject). The question is under what circumstances further (functional) structure on top of this VP will be present.

Given a theory of phrase structure as proposed in Chomsky (1994), the insertion of a functional marker after the VP is generated leads to a functional projection headed by that marker. There is no other option, since further projection of the VP does not provide the functional marker with an appropriate position.

(3) a. 

b. * 

If we now take seriously the idea that projection depends on the presence of a lexical head, we are forced to the conclusion that the structure in (3a) is not base-generated if no functional marker is present. If only an inflected verb is taken from the lexicon, then only a VP will be base-generated.[3]

---

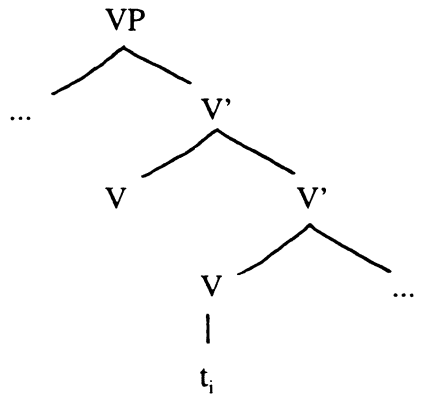
[2] In fact, there is an additional problem, pointed out by Reinhart (1993). According to her and others, WHs in situ do not move at all, not even at LF. If this is so (as we will assume for the largest part, see section 3), it is hard to see how their features could be checked.

[3] In accordance with the lexicalist perspective on inflection, we assume that verbs are inserted in fully inflected form. See Lapointe (1980), Lieber (1980), Jensen & Stong-Jensen (1984), Di Sciullo & Williams (1987), Chomsky (1993) and Ackema (1995) for various proposals. Moreover, the view expressed here presupposes that the VP-internal subject hypothesis is correct.

However, when necessary, functional structure may be derived in the absence of a functional marker in the numeration. In order to accomplish this we adopt a procedure proposed by Ackema et al. (1993). This procedure generates functional structure as a result of head movement. As opposed to Chomsky (1994), we would like to allow for the possibility of *self-attachment*. By this we mean that a head can move out of its base position and adjoin to a node of its own projection.

Suppose that, after VP has been generated, the verb moves out of its base position and adjoins to V'. The resulting structure would seem to be ungrammatical. It is standardly assumed that every head must be dominated by a maximal category and every (possibly segmented) maximal category must dominate a head. But in (4) one maximal projection dominates two heads.

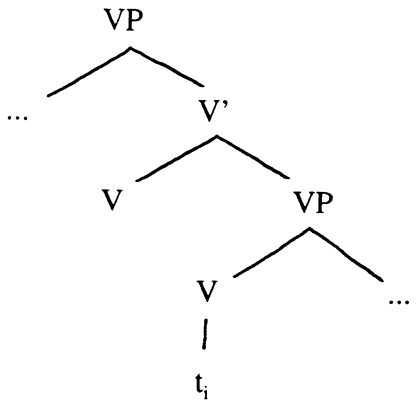
(4)



This problem is solved by Ackema et al. by assuming that X-bar theory has the following two properties. First, it applies at every stage of the derivation (cf. Van Riemsdijk 1989, Chomsky 1993). Second, the familiar X-bar levels are decomposed into the features [±M(aximal)] and [±P(rojected)] (cf. Muysken 1982), where at least the value of [±M] is defined contextually, namely as the highest projection of a head X. As a result, it follows that nodes which are nonmaximal at one point in the derivation may be maximal at another (and vice versa).

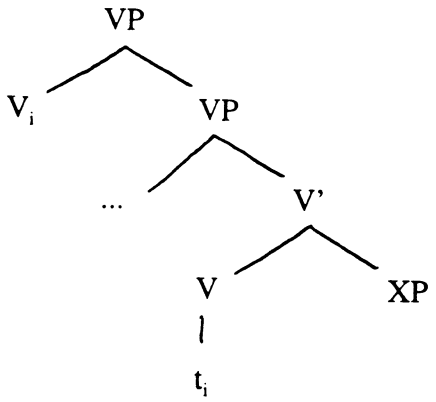
The structure in (4) is now rescued if the lower segment of V' is reanalyzed as VP, a reanalysis which takes place within the limits of the X-bar theory sketched above. The V'-node, which initially has the features [+P,-M], turns into a node with the features [+P,+M] at surface structure. In other words, this node is now conceived of as the highest projection of the verbal trace:

(5)



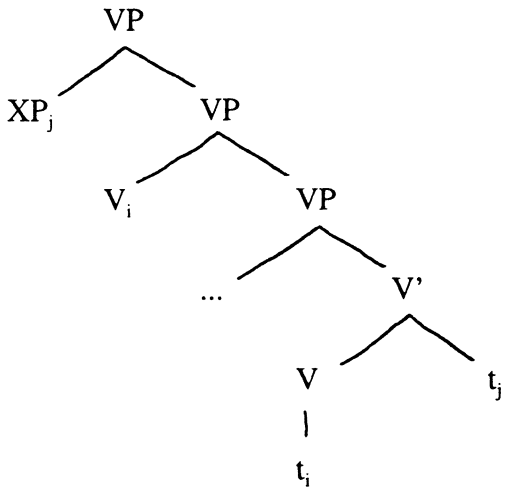
This is one way of establishing the structure in (3a) by movement. A second way, which will be the crucial one in this paper, also involves self-attachment. In this case, the verb adjoins not to V', but to VP:

(6)



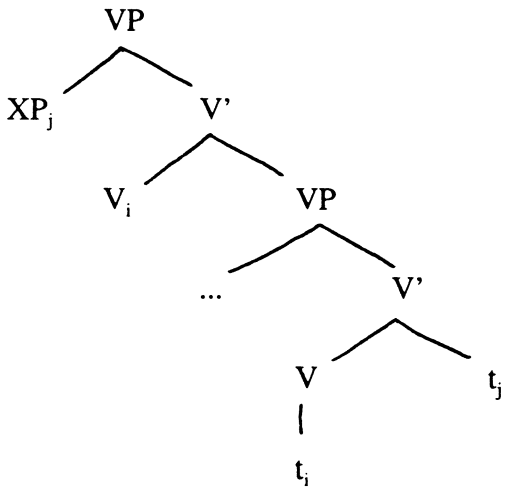
There is an important difference between self-attachment of the verb to V' and self-attachment to VP. In the former case, the specifier of the functional VP will contain the same element that the specifier of the lexical VP contained in the initial structure (without this element having actually been moved, compare (4) and (5)). Self-attachment to VP, however, creates the possibility of deriving a different specifier in the functional projection. This is accomplished as follows. First, a second element moves out of its base position and adjoins to VP. This could be the complement of V, as in (7), but in fact any element will do.

(7)



The element adjoined to VP can be turned into the specifier of the functional VP by reanalysis of the intermediate segment of VP. If this node's positive marking for maximality is changed into a negative marking, a structure results in which the moved V heads a functional VP with the moved XP in its specifier:[4]

(8)



An instantiation of (8) is question formation in (English) root sentences, as we will show below.

[4] It might be that the intermediate structures in (6) and (7) face some X-bar theoretical problems. The higher VP-nodes in these structures can either be analyzed as a projection of the moved verb, or as a projection of the lower VP (with the verb a [-P,+M] adjunct to this VP). If Chomsky (1994) is correct in claiming that syntactic representations may not have ambiguous structural properties, structures like (6) and (7) would be ruled out. This problem is solved by reanalysis as in (8), resulting in a structure that is not ambiguous. This means that, if Chomsky is correct, self-attachment of the verb to VP must be accompanied by the fronting of some constituent. See for discussion Koeneman (1995).

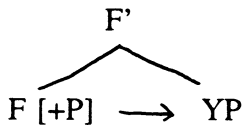
In sum, functional structure can be generated in two ways, namely by the insertion of a functional marker or by self-attachment. The latter option in turn consists of two subcases. If the head adjoins to an intermediate projection, reanalysis will result in a structure with a functional projection on top of VP that contains the same specifier as the VP contained in the base. If the head adjoins to its maximal projection, fronting of some element and reanalysis will result in a functional projection on top of VP with the fronted element in its specifier position.

2.2 *Q-Marking*

We have now seen how functional structure can be derived, but we have not yet considered the question why and when it is derived. We assume that one important reason to create functional structure is to derive a configuration in which certain properties of VP (i.e. the proposition) can be syntactically marked. VP can be marked for such notions as tense, modality and aspect, and, as we will argue in this section, for questionhood.

It is a pervasive property of natural language that heads may mark properties of their complements. A verb may, for instance, assign a thematic role and possibly case to its complement. It would therefore be a minimal extension of existing theories to assume that sentence-type marking in the sense just outlined involves a head and its complement. The general scheme for the marking of a property [+P] is the one in (9).

(9)



The marking of VP as a question is an instance of the general scheme in (9). In English, for instance, questions are clearly marked as such. A straightforward instance of Q(uestion)-marking is found in embedded yes/no questions. Here a base-generated functional head bearing the Q-feature takes the VP to be marked as its complement:

(10) John wonders [if [$_{VP}$ you have seen the soccer match]]
+Q ⤴

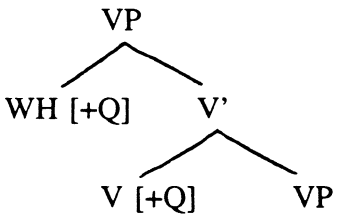
It is not obvious, however, that question marking always involves a head and its complement. In matrix WH-questions, for instance, the relevant feature seems to be present on the WH-expression, and not on the verb. Moreover, one may wonder whether Q-marking by a complementizer, as in (10), is an instance of the same process as Q-marking by a WH-expression, as in (11).

(11) [What$_j$ [have$_i$ [you t_i seen t_j]]]

To start with the first question, it is hardly a problem that the relevant feature is present in the specifier of the functional projection. By the general mechanism of specifier-head agreement, we may expect that a WH-specifier will transfer this feature to the related head, after which this head will be capable of marking its complement. This means, of course, that such a head must be present. In the absence of a complementizer in this root environment, self-attachment of the verb is necessary. If the verb did not move, there would be no head to mark the proposition as a question.

Assuming for the moment that the relevant feature indeed is the Q-feature we encountered earlier, (11) can be depicted as in (12), where, in line with assumptions made earlier, the projection hosting the WH-expression is labelled VP since it is headed by (the moved) V:

(12)



So, the derivation of a WH-question contains the following steps. First, the verb adjoins to VP (cf. 13b). Then, the WH-expression adjoins to VP (cf. 13c), and then VP-to-V' reanalysis takes places (cf. 13d).[5]

(13) a. [$_{VP}$ You have seen what]
b. [$_{VP}$ Have$_i$ [$_{VP}$ you t_i seen what]]
c. [$_{VP}$ What$_j$ [$_{VP}$ have$_i$ [$_{VP}$ you t_i seen t_j]]]
d. [$_{VP}$ What$_j$ [$_{V'}$ have$_i$ [$_{VP}$ you t_i seen t_j]]]

Let us now turn to the second question posed above: can the Q-feature of the interrogative complementizer and the WH-feature of a WH-expression be equated? The null hypothesis would be that both are identical. Although this might seem problematic, given that the Q-feature marks yes/no questions (cf. 10), while the WH-feature marks the questioning of a particular constituent (cf. 11), there are two observations that support unification. First, there are verbs like *wonder*, which select both *if* clauses and WH-questions (cf. 14), but there are no verbs that select one of these categories but not the other:

[5] This derivation does not comply with Greed, at least not in its strictest form. We therefore assume that this principle does not hold.

(14) John wonders [what [Mary will say]]

Second, in embedded WH-questions in Dutch, both a Q-complementizer and a WH-expression may be present (cf. 15). If the features of these elements are incompatible, one would expect the derivation to crash.

(15) Jan vraagt zich af [wat [of [Marie zal zeggen]]]
John wonders REFL PART *what if Mary will say*

If there is a uniform category 'question', the different interpretations of WH-questions and *if* questions must be derived in some other way. In fact, this can easily be achieved. WH-questions typically contain a variable accessible to the Q-marking head.[6] We assume that as a result of this the feature Q is interpreted as concerning this element. In yes/no questions there is no gap, and therefore the feature Q can only be interpreted as concerning the whole expression (i.e. its truth value).

The hypothesis that there is a general category 'question' and that English requires syntactic marking of this category faces two problems. First, in embedded WH-questions, there is a Q-bearing element, but there does not seem to be a head (cf. 16a). Second, in matrix yes/no questions there is a head, but there does not seem to be a Q-bearing element (cf. 16b). By earlier assumptions we are forced to say that (16a) contains a null head and that (16b) contains a null Q-operator. We believe that this is indeed what is going on.

---

[6] By 'accessible', we mean a variable that is not bound within the c-command domain of the Q-marking head. So, in (ia) a WH-question interpretation is possible, but in (ib) and (ic) it is not.

(i) a. WH_i C[+Q] [$_{VP}$ ... t_i ...]
b. C[+Q] ... WH_i ... [$_{VP}$ ... t_i ...]
c. C[+Q] [$_{VP}$ ... WH ...]

In (ia) the variable is accessible to the Q-marking head, since it is unbound within the c-command domain of this head. In (ib) there is a variable in VP, but it is not accessible to the Q-marking head, as it is already bound by its WH-antecedent. In (ic) there is no variable at all (at surface structure). As pointed out by Eddy Ruys (p.c.), this explains the illformedness of the following example:

(ii) ??Wie_i vraag je [t_i of [Jan slaat t_i]]
who ask you if/whether John hits

This example cannot be ruled out on a par with WH-islands, since *of* 'if' is the head, not the specifier of CP. The example is not so much ungrammatical, however, but uninterpretable. The trace in the embedded clause is accessible to C, hence the embedded clause should be interpreted as a WH-question. At the same time, the matrix clause must be interpreted as a WH-question as well. Since there is only one WH-operator present, this is impossible.

(16) a. John wonders [what ___ Mary will say]
b. ___ have you seen the soccer match

Concerning (16a), we have already seen that in certain languages the Q-marking head can be lexicalized. A Dutch example was given in (15). Apparently, embedded WH-questions do contain a Q-marking head. It is unclear, however, what the status of this head is in the theory of phrase structure we have adopted. Recall that functional structure could only be generated by insertion of lexical markers or self-attachment. Neither of these possibilities seems to apply in (16a).

This problem can be solved in several ways. The solution that seems most promising at the moment would be to say that a [+Q] complementizer is in fact inserted in embedded WH-clauses in English as well, but that, as a consequence of phonological constraints on the lexicalization of 'COMP', it is deleted (i.e. not pronounced) when the syntactic representation is mapped into a phonological one (cf. Pesetsky 1994). The relevant phonological constraint is usually referred to as the doubly filled COMP filter.

One might object to such an analysis by saying that *if* (as opposed to *that*) may usually not be omitted. This, however, can be seen as a result of the condition of recoverability. As is well-known, deletion is only possible if the features of the deleted material are recoverable from the overtly expressed parts of the sentence. This has the consequence that the complementizer in an embedded yes/no question like (10) may not be deleted: the Q-feature of C is not recoverable. In (16a), however, the Q-feature is encoded on the WH-expression as well as on the complementizer, and hence this property of the complementizer is recoverable after deletion:

(17) a. John wonders if Mary loves him --->
a'. *John wonders __ Mary loves him
b. John wonders what if Mary will say --->
b'. John wonders what __ Mary will say

Concerning (16b), it can be observed that, like the null head in (16a), the null Q-operator is phonetically realized in certain languages. Some examples are given below.

(18) a. **An** bpósfaidh tú mé (Irish)
Q *will-marry you me*
b. **Is** idda hmad s tmazirt (Berber)
Q *went Ahmed to country*
c. **Tsi** hot er geleient dos bux (Yiddish)
Q *has he read the book*

In addition to this, it might be that X-bar theory necessitates the presence of such an operator. Consider how a matrix yes/no question is derived. The sentence starts out as a VP (cf. 19a). Then the verb undergoes self-attachment, as in (19b). However, (19b) as it stands may be ungrammatical, because it is structurally ambiguous (see note 4). It can either be analyzed as an adjunction of a (minimal maximal) element to VP or a case of the shifted verb taking a VP-complement. Ambiguity of this type is resolved by VP-to-V' reanalysis if a phrasal element is adjoined to the top of the structure, as outlined in section 2.1. In (16b), this would be the (base-generated) null Q-operator:

(19) a. $[_{VP}$ You have seen the soccer match]
b. $[_{VP}$ Have$_i$ $[_{VP}$ you t_i seen the soccer match]]
c. $[_{VP}$ Q $[_{VP}$ have$_i$ $[_{VP}$ you t_i seen the soccer match]]
d. $[_{VP}$ Q $[_{V'}$ have$_i$ $[_{VP}$ you t_i seen the soccer match]]

Just as in the case of *if* deletion, one could argue that English does have a Q-operator in its lexicon, but that this element is not spelled out when the syntax is mapped onto a PF representation. This might follow from the constraints proposed by Pesetsky (1994). In particular, Left Edge (F), the constraint that requires left alignment of a functional projection with its head, would favor deletion of material preceding the finite verb in (19), up to recoverability. The Q-operator is in fact recoverable after its Q-feature has been copied onto the verb.

We will not pursue these issues here, assuming that the solutions suggested above, or something like them, are correct. In conclusion, Q-marking can be taken to be an instantiation of the general process of sentence-type marking, as outlined in section 2.1. We will now consider what consequences this has for the topic central in this paper: the syntax of multiple questions.

2.3 *Multiple Questions*

In this section we will begin our argument to the effect that the syntax of multiple questions is determined by constraint interaction in an optimality-theoretic way. First, we will formulate two constraints that follow from the theory of Q-marking introduced above. We will then show that interaction of these constraints with the well-known Shortest Steps constraint can account for the syntax of multiple questions in English.

The first constraint that follows from section 2.2 is obvious: a question must be marked as such, where marking requires VP to be the complement of the Q-marker. We will refer to this constraint as *Q-Marking*.

(20) *Q-Marking*
A question must be overtly Q-marked

We should point out here that Q-Marking (and the other two constraints we will introduce in this section) is a constraint holding of surface structures. That is to say, it requires that questions are visibly marked, thus forcing, as we will see below, overt movement of both one WH-expression and the verb (in main clauses). One can think of constraints of this type in two ways: either as S-structure constraints (re-introducing this level of representation) or as PF constraints (stretching the subject matter of this level of representation). We will adopt the former option here.

Q-Marking is a constraint about the element to be marked. From the theory, a plausible constraint about the marking element also follows. Due to the general notion of economy, it is desirable not to insert features into a tree without these features having a suitable function. If this is so, insertion of Q-bearing expressions requires usage of their Q-feature. We have already seen that 'Q' is a property added to propositions. Hence, the Q-feature must have a proposition in its scope. Since propositions are expressed by VPs (given the VP-internal subject hypothesis), the Q-feature must take scope over VP. The constraint we would like to propose requires that this scope be marked overtly.

(21) *Q-Scope*
[+Q] elements must c-command VP at surface structure

It will be clear that there is some overlap in the empirical effects of Q-Marking and Q-Scope. Both principles may trigger WH-movement. However, as we will show at length below, there are constructions in which one is crucially satisfied, while the other is not. For example, movement of a WH-expression without accompanying head movement will result in a structure that violates Q-Marking, but satisfies Q-Scope (for this WH-expression).

The third and last constraint that will be relevant needs no further introduction:

(22) *Shortest Steps*
Minimize the distance between chain links

Since zero is the minimal distance imaginable, (22) subsumes the 'Move-only-when-necessary' constraint that turns up in various forms in various works. It is an open issue how distance in chains should be measured. The most straightforward interpretation seems to be that each node in the path between two chain links results in a violation of Shortest Steps (cf. Collins 1994 for a similar approach). Length of chains is defined in (23).

(23) *Length*

The length of a chain is the (total) cardinality of the path(s) connecting the head of the chain and the foot of the chain, such that there are no paths connecting the head and the foot with a lower (total) cardinality.

We assume that, in calculating the cardinality of a path, more segments of a single category count only once. So, a path <X, Y, Y, Z> counts as equally long as a path <X, Y, Z>. The notion of 'path' and the relation of 'connectedness' (which we assume to be symmetrical and transitive) are defined as in (24) and (25).

(24) *Path*

A path is an ordered set of nodes $<N_1\ N_2\ N_3\ ...\ N_M>$ such that, if N_N and N_{N-1} are contained in this set, N_N immediately dominates N_{N-1}.

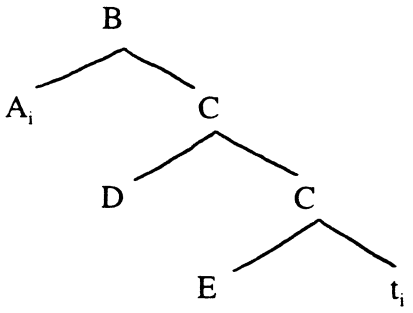
(25) *Connectedness*

a. A node N is connected to a path P if and only if there is a node in P that immediately dominates N.

b. A path P is connected to a path Q if and only if there is a node N that is contained in both P and Q.

Consider, for instance, the case of simple upward movement depicted in (26) (which might be a case of NP raising in a VP containing an adjunct).

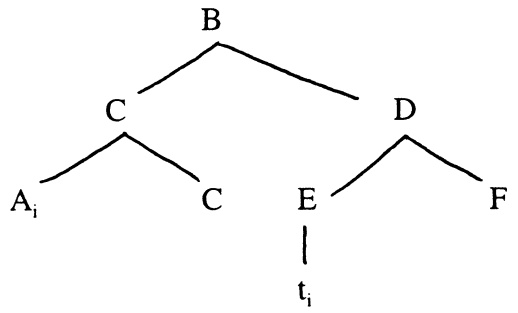
(26)



There is a single path connecting A and its trace, namely <B, C, C>. This set of nodes is a wellformed path because B immediately dominates the higher segment of C, while this segment immediately dominates the lower segment of C. The antecedent A is connected to this path since it is immediately dominated by B; the trace t is connected to this path since it is immediately dominated by (the lower segment of) C. The length of the chain between A and t is two, because the two segments of the single category C count only once. Hence, this movement results in two Shortest Steps violations.

Next, consider the slightly more complicated case in (27) (which might depict head-to-head adjunction).

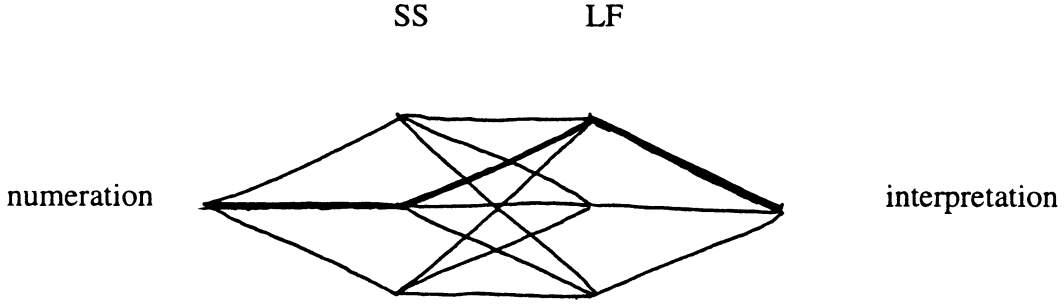
(27)



In this structure, there is no single path connecting A and t, because there is no set of nodes between A and t such that each of these nodes immediately dominates the next. Instead, A and t are connected by two paths: P_1 <B, C> and P_2 <B, D, E>. A is connected to P_1, t is connected to P_2, and P_1 and P_2 are connected because they both contain B. Hence, the length of the chain between A and t is five.

After having introduced the relevant constraints, we need to define the candidate set. Following work by Grimshaw (1993) and Reinhart (1993), we assume that sentences belong to the same candidate set if they are projected from the same set of lexical items (the same numeration) and target the same semantic representation. The intuition is that the syntax is a device mapping the lexicon to a semantic representation (and of course to a phonological representation). This mapping can take place in various ways, yielding various derivations and thus various candidates.[7] The actual derivation is the one that is optimal with respect to a number of ranked (S-structure and LF) constraints. This is schematized in (28), where the optimal route is bold-faced.

(28)



This view has several consequences. Trivially, two sentences that are projected from different lexical items are not part of the same candidate set. Two sentences that target different interpretations are not in competition either. Finally, sentences that are uninterpretable cannot block derivations that do lead to an interpretation, because such sentences belong to a different candidate set (if they belong to a candidate set at all).

[7] The device producing the various possible derivations (what is called GEN in optimality theory) might have certain intrinsic properties which (since they are part of the generator) cannot be violated, like the wellformedness conditions imposed by X-bar theory.

Consider now how the constraints interact in a nonmultiple question like *what have you seen*. In English, any question is Q-marked, so Q-Marking must be relatively high in the hierarchy. In particular, it must be ranked above Shortest Steps, or we would not expect WH-movement. This is shown in the tableau in (29), where Q-Scope is ordered lowest for reasons to be discussed below.[8]

(29) *English simple questions*

	Q-Marking	Shortest Steps	Q-Scope
☞ What have [you t seen t]		*******	
[You have seen what]	*!		*
What [you have seen t]	*!	**	
Have [you t seen what]	*!	***	*

The sentence *what have you seen* violates Shortest Steps seven times: three times because the number of nodes separating *have* and its trace is three (the functional V', VP and V'), and four times because the path connecting *what* to its trace consists of <VP, V', VP, V'> (cf. 13d). However, Q-Marking is satisfied. Due to spec-head agreement in the derived functional projection, *have* acquires a Q-feature with which it marks its VP-complement. Crucially, all other candidates violate Q-marking. For the second one, this is immediately

---

[8] A complication seems to arise in embedded clauses. Consider the following examples:

(i) a. *I wonder if John loves who
b. *I wonder whether John loves who
c. I wonder who ***if*** John loves

The problem is that, because Q-marking is possible without WH-movement in the above examples, one would expect the WH to remain in situ, given that Shortest Steps outranks Q-Scope in English. This problem is only apparent, however. As noted earlier, Q-marking a VP without there being an accessible variable contained in this VP leads to a yes/no question interpretation. We assume that the interpretation of the Q-feature in CP is determined when Q-marking takes place, i.e. at surface structure in English. A consequence of this is that (ia) and (ib) are not in the same candidate set as (ic), and hence they cannot block (ic). The same reasoning applies to a main clause structure like (ii):

(ii) *Q have$_i$ [you t$_i$ seen what]

Consequently, examples like (ia,b) and (ii) can only be interpreted as yes/no questions, but in that case the WH in situ must be interpreted as an echo-WH. (In this reading, the sentences are indeed possible).

clear, since nothing has been moved. In the third one, *what* has been adjoined to VP (thus complying with Q-Scope). An adjunction structure, however, is not a proper structure for marking, only a head-complement structure is, and no head has been fronted. In the last candidate, a head has been fronted, but this does not inherit a Q-feature, due to lack of movement of *what*. Note that it follows that there will be no language in which verb movement is triggered in WH-sentences (as it is in English), but in which the WH-element itself remains in situ (as in Chinese). Since moving the head without moving the WH-element does not lead to any improvement with respect to either Q-Marking or Q-Scope, Shortest Steps will rule out this possibility regardless of the ranking of these constraints. As far as we know, this prediction is correct.

Let us now turn to multiple questions. The high ranking of Q-Marking again ensures that the head and at least one WH-phrase must move in order to create the proper Q-marking environment. The question now is what will happen to the other WH-phrase(s). Here, the ranking between Shortest Steps and Q-Scope becomes relevant. If Shortest Steps is ranked higher than Q-Scope, as in the tableau in (30) below, it follows that the other WH-phrases remain in situ. Under the reverse order, one would expect all WH-elements to move out of VP.

(30) *English multiple questions*

	Q-Marking	Shortest Steps	Q-Scope
☞ Who has [t t seen what]		******	*
What has [who t seen t]		*******!	*
Who what has [t t seen t]		*******!***[9]	
Has [who t seen what]	*!	***	**

Although Shortest Steps is violable in English (the actual outputs in both (29) and (30) violate it), it has its effects. It does not only account for the fact that all WH-elements but one remain in situ, but also for the fact that the one WH-element that *is* moved (to comply with Q-Marking) is the one which makes the shortest possible move, i.e. the subject in (30) and the direct object in *what did Mary give to whom* (cf. Barrs & Lasnik 1986). In general, superiority effects follow from the global evaluation of Shortest Steps, as argued

[9] The way in which this number of violations is derived will be clearer in section 3.1, when Bulgarian multiple questions are discussed. It should be clear, however, that moving *what* is worse than leaving it in situ as far as Shortest Steps is concerned.

by Golan (1993) and Reinhart (1993) (a very similar analysis was already proposed by De Haan (1979:157 ff.), based on his Minimal String Principle).

Concerning the WH in situ, we assume, following Reinhart (1993) and others, that this element does not move at LF either. It is interpreted via a semantic procedure we need not go into here. One of the advantages of this assumption is that the null hypothesis can be adopted that all movement is subject to Subjacency, not just overt movement. The well-known fact that WH-in-situ can apparently violate Subjacency follows, since it is not interpreted by movement. (We will come back to this issue in section 3.3 on Chinese/Japanese).

So, constraint ranking is crucial for English. As noted in section 1, however, the most important type of evidence for constraint ranking and global evaluation lies in cross-linguistic variation. This variation should be explicable just by reranking the constraints. We will now try to establish whether the typology of multiple question formation can be accounted for in this way.

3. *Reranking*

There are six possible rankings of the three constraints we have proposed. It will turn out that in two instances two orderings have the same effects. This means that in practice four different types of question formation are to be expected. One of these was discussed in section 2.3, the English type. The remaining three are also attested, as we will show in sections 3.1 - 3.3. In section 3.4, we will discuss a possibility not yet mentioned, which is that two constraints are not ranked with respect to each other. This will account for optional WH-movement in French.

3.1 *Bulgarian*

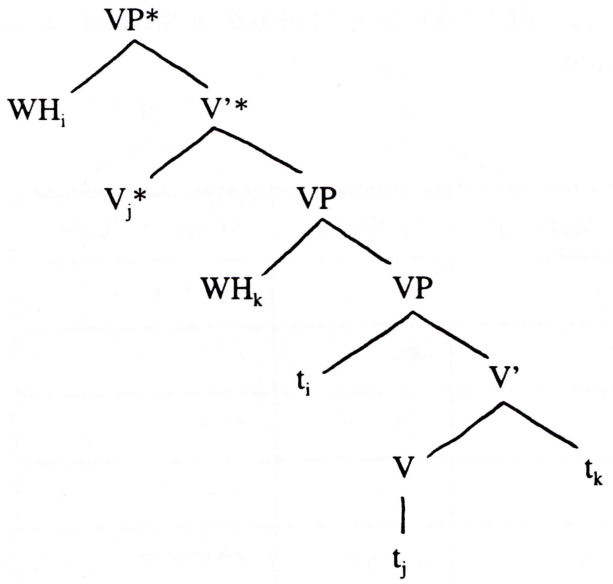
Suppose the constraints proposed in section 2 are ranked as in (31).

(31) Q-Marking >> Q-Scope >> Shortest Steps

The high ranking of Q-Marking ensures that, just like in English, the proper structure for Q-marking must be derived. That is, it ensures head movement plus movement of at least one WH-phrase to the specifier of the newly created VP (CP in traditional theories; for ease of reference we will henceforth refer to functional VPs as VP*). Compared to English, however, the ranking of Q-Scope and Shortest Steps has been reversed. This means that not just one, but all WH-phrases must move out of VP. The shortest possible

move is adjunction to VP, but this is ruled out for the following reason. The structure that would result if the other WHs would adjoin to VP is the one in (32).

(32)

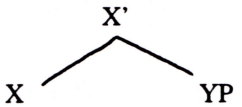


In this structure, Q-Marking is violated after all, because of a Relativized Minimality effect. Relations between an element of type A and another element are generally blocked by intervening elements of type A. This phenomenon is well-known from the literature on extraction, but it can also be observed in other domains. For example, a verb may select for the case of its complement, but not if another case-assigning head intervenes, such as a preposition. If Relativized Minimality holds of syntactic relations in general, Q-marking of the VP by the derived functional head in (32) is blocked by the Q-bearing WH_k, which is a potential Q-marker. Since WH_k is not in the proper configuration for marking, there is no Q-marking possible at all in (32).[10]

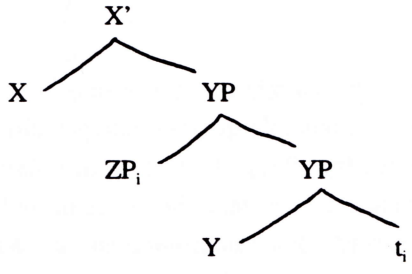
---

[10] Probably, even if the intervening element would not be a potential Q-marker, marking by V* would be blocked anyway in a structure like (32). This is because presumably adjunction to direct complements (that are to be marked) is generally ruled out. Adjunction to complement clauses, for instance, is ungrammatical (cf. Chomsky 1986, Grimshaw 1993, and others). In general, marking is possible in a configuration like (ia), but not in a configuration like (ib)

(i) a.



b.



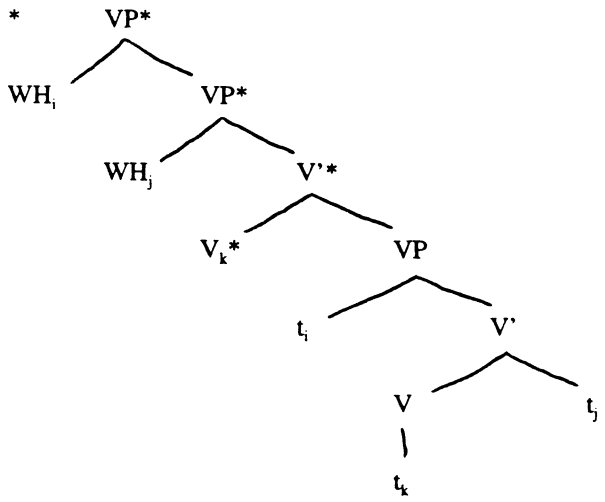
As a consequence, WHs may not adjoin to VP if Q-marking is necessary. If they must be moved out of VP, they must move to a position above V*. Hence, not just one, but all WHs must move to spec-VP*.[11] In short, the effect of the ordering in (31) is that all WH-expressions cluster together in spec-VP*. This is formalized in the tableau in (33) (where the brackets indicate VP-boundaries).

(33) *Bulgarian multiple questions*

	Q-Marking	Q-Scope	Shortest Steps
WH V [t t WH]		*!	******
[WH V WH]	*!	**	
V [WH t WH]	*!	**	***
[WH [WH [t V t]]]	*!		***
WH V [WH [t t t]]	*!		********
☞ WH WH V [t t t]			**********

[11] In principle, there still seems to be another possibility, namely moving one WH to spec-VP* to ensure that there will be Q-marking, and adjoining all other WHs to VP*, as in (i).

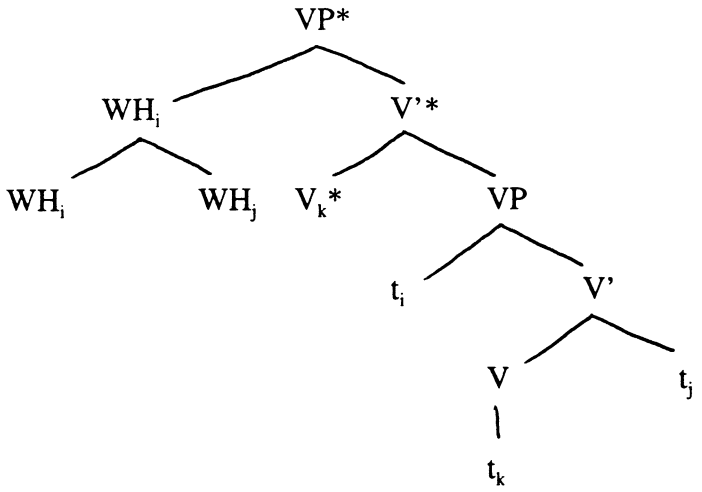
(i)



In order to obtain a multiple question interpretation, absorption must take place (cf. May 1977, 1985, Higginbotham & May 1981). We assume that absorption between moved WHs is possible under the standard condition of mutual m-command. Since the moved WHs in (i) do not mutually m-command each other (WH_j is dominated by VP*, which does not dominate WH_i), absorption is impossible. This structure is therefore uninterpretable, so it falls outside the relevant candidate set (see section 2.3).

It is predicted, therefore, that there are languages in which multiple questions are structured as in (34).

(34)



This prediction is correct. Languages in question are, for example, Bulgarian and Romanian, as shown by Rudin (1988). Rudin cites several pieces of evidence showing that all WH-expressions form an indivisible constituent in spec-CP (here: spec-VP*). Focusing on Bulgarian, it can be observed that no WH-expression remains in situ (all examples from Bulgarian are taken over from Rudin 1988):[12]

(35) Koj kogo vižda
who whom sees

Moreover, the WH-cluster cannot be interrupted. For instance, second place clitics in Bulgarian must come at the end of the entire WH-sequence, indicating that this sequence occupies a single syntactic position. This is shown for the pronominal clitic *ti* and the auxiliary clitic *e* in (36).

(36) a. Koj kakvo ti e kazal
who what you has told
'who told you what'
b. *Koj ti e kakvo kazal
who you has what told

Similarly, adverbials may not split up the fronted WH-sequence, as shown by the examples in (37).

[12] For reasons of exposition, we will restrict our attention to cases with two WH-phrases (subject and object) here, but the results can immediately be extended to cases with more than two WH-phrases.

(37) a. Zavisi ot tova, koj kogo prŭv e udaril
depends on this who whom first has hit
'It depends on who hit whom first'
b. *Zavisi ot tova, koj prŭv kogo e udaril
depends on this who first whom has hit

Another piece of evidence cited by Rudin comes from WH-words in free relatives (assuming that the syntax of relatives is parasitic on the syntax of questions, at least in these languages). Relative WH-words must be marked with a definitizing *-to*. It is possible that all WH-expressions are affixed with *-to* separately, but it suffices to add one *-to* at the end of the complete WH-sequence to yield the whole sequence relative. Adding *-to* to the first WH-constituent only is ungrammatical, however. As Rudin notes, this suggests that *-to* is suffixed to one WH-constituent, which consists of all fronted WH-phrases:

(38) a. Kojto kakvoto iska ...
who-DEF what-DEF wants
'whoever wants whatever'
b. Koj kakvoto iska ...
who what-DEF wants
c. *Kojto kakvo iska ...
who-DEF what wants

Another important ingredient of the structure in (34) is that there is obligatory verb movement in order to derive the VP* needed for Q-marking. This means that, if the subject is not a WH-element itself, we predict there to be obligatory inversion in WH-questions. This is correct. As noted by Kraskow (1992), languages of the Bulgarian-type do indeed display obligatory inversion in questions:

(39) a. Kakvo kupuva Ivan
what buys John
'What does John buy'
b. *Kakvo Ivan kupuva
what John buys

It seems, then, that the syntax of Bulgarian WH-questions indeed follows from the reranking in (31) of the constraints proposed in section 2.3.

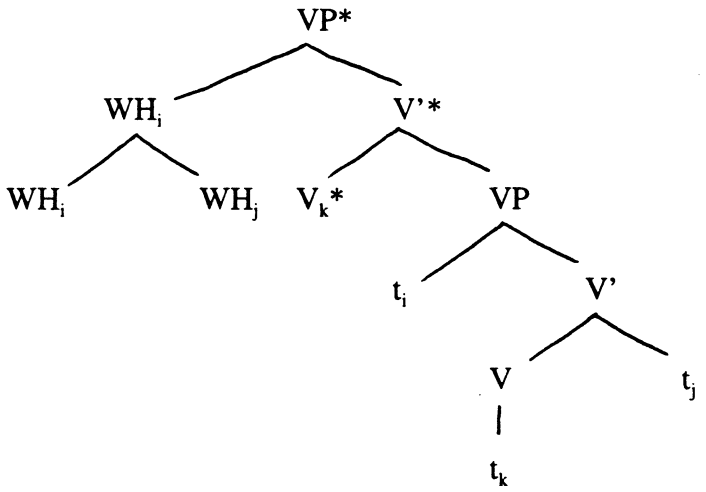
A further striking characteristic of the WH-sequence in Bulgarian resembles the superiority effects that can be observed in English. While in multiple WH-questions in English a WH-subject will be the element that is overtly moved, in Bulgarian a WH-subject must precede all other fronted WH-elements (cf. 40).

(40) a. Koj kogo vižda
who whom sees
b. *Kogo koj vižda
whom who sees

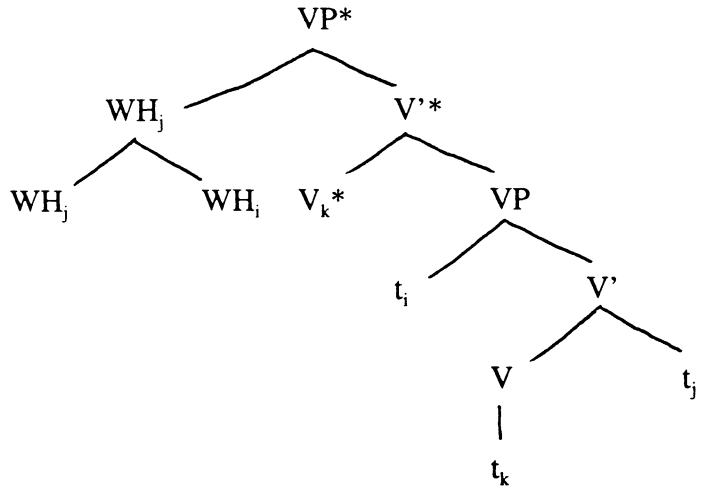
We will now argue that this is an indirect effect of the lowest ranked constraint in Bulgarian, Shortest Steps. This is an illustration, therefore, of the fact that low-ranked constraints can still have their effects when the different candidates are equal with respect to all the higher constraints. Note that (40) indicates that the relation between subject and object in the WH-cluster is asymmetric. Either the subject must be adjoined to the object, in which case (40) shows that WH-phrases must be left-adjoined, or the object must be adjoined to the subject, in which case (40) shows that WH-phrases must be right-adjoined. It cannot be the case, however, that it does not matter which WH-phrase adjoins to which, since free word order would be predicted then, even when assuming a uniform direction of adjunction. According to Rudin, the cluster in spec-VP* is headed by the subject, with the other WH-phrases right-adjoined to it. We will now show that this indeed follows from Shortest Steps in our analysis.

Suppose that both WH-phrases move to spec-VP* independently. No matter whether the subject is moved first, and the object adjoined to it (cf. 41a), or the other way around (41b), Shortest Steps will be violated equally, namely twelve times.

(41) a.



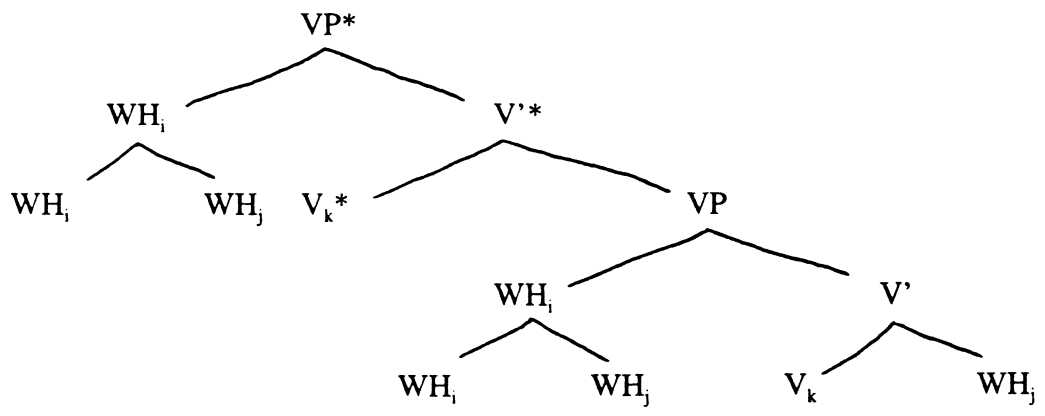
b.



The paths that connect antecedents and traces in (41a) are <V'*, VP, V'> for the moved verb, <VP*, V'*, VP> for the moved subject, and the two paths <VP*, WH$_i$> and <VP*, V'*, VP, V'> for the moved object. This amounts to twelve violations of Shortest Steps. In (41b) the paths are <V'*, VP, V'> for the verb, <VP*, V'*, VP, V'> for the object, and <VP*, WH$_j$> plus <VP*, V'*, VP> for the subject, which also results in twelve Shortest Steps violations. If these were the only options, we would therefore predict optionality as to which WH-phrase adjoins to which, while we just noted that this cannot be correct.

Fortunately, there is another option to derive the correct structure for a multiple WH-question in Bulgarian, which has the object adjoined to the subject and which turns out to be scoring better on Shortest Steps. This option consists of first adjoining the WH-object to the WH-subject and then moving this cluster as a whole to spec-VP*. The point is that moving the two WH-phrases together, as one constituent, is cheaper than moving them apart over the same distance. If the object can be taken for a ride by the subject for part of the distance it must move, this will therefore be cheaper. See the tree in (42), where for clarity we have spelled out the internal structure of the traces (we assume the copy theory of movement, cf. Chomsky 1993).

(42)



The paths we now obtain are the following: <V'*, VP, V'> for the moved verb again, <VP, WH_i> and <VP, V'> for the moved object, and <VP*, V'*, VP> for the moved WH-cluster. This results in only ten violations of Shortest Steps, and therefore is better than either of the trees in (41). Note that this derivation crucially depends on adjunction of the object to the subject, thus explaining the asymmetry in the WH-cluster.[13] The correct word order then follows from uniform right-adjunction of WH-phrases, as proposed by Rudin.[14]

To conclude the section, it should be pointed out that the constraint ranking in (31) is not the only one which results in the Bulgarian-type of multiple WH-questions. Reversing Q-Scope and Q-Marking has no effects, as long as Shortest Steps is ranked lowest. Consider the tableau in (43). Compared to (33), the column for Q-Scope and Q-Marking are reversed.

(43) *Bulgarian multiple questions (alternative ranking)*

	Q-Scope	Q-Marking	Shortest Steps
WH V [t t WH]	*!		******
[WH V WH]	*!*	*	
V [WH t WH]	*!*	*	***
[WH [WH [t V t]]]		*!	***
WH V [WH [t t t]		*!	********
☞ WH WH V [t t t]			**********

Q-Scope dictates that all WHs must move out of VP. The derivation that complies best with Shortest Steps would involve adjunction of the WHs to VP, without head movement. But since Q-Marking outranks Shortest Steps, the syntactic configuration for Q-marking must be derived. This can only be achieved by head movement plus WH-movement of at least one WH to the spec of the derived VP*. For reasons outlined above, the other WHs must then move to spec-VP* as well (if they have to move out of VP in the first place, as

[13] Note that first adjoining the subject to the object and then moving the whole cluster to spec-VP* is ruled out by the ban on downward movement, even though it also results in ten Shortest Steps violations at S-structure.

[14] There are in fact some cases where left-adjunction is possible as well, i.e. cases where the order object-WH before subject-WH is possible, dependent on pragmatic factors. See Billings & Rudin (1995) for discussion of this issue within an OT framework.

is the case because of the high ranking of Q-Scope), or Q-marking would be blocked after all. Hence, the same pattern reappears.

From the six possible rankings of our three constraints, three have now been discussed, leading to two different patterns of WH-question formation. The ranking Q-Marking >> Shortest Steps >> Q-Scope leads to movement of the verb and one WH, as in English. The rankings Q-Marking >> Q-Scope >> Shortest Steps and Q-Scope >> Q-Marking >> Shortest Steps trigger clustering of WHs in spec-VP*/CP and supporting verb movement in main clauses, as in Bulgarian. In the next section we will discuss a third pattern, derived from yet another ranking.

3.2 *Czech*

Suppose the constraints proposed in section 2.3 are ranked as in (44).

(44) Q-Scope >> Shortest Steps >> Q-Marking

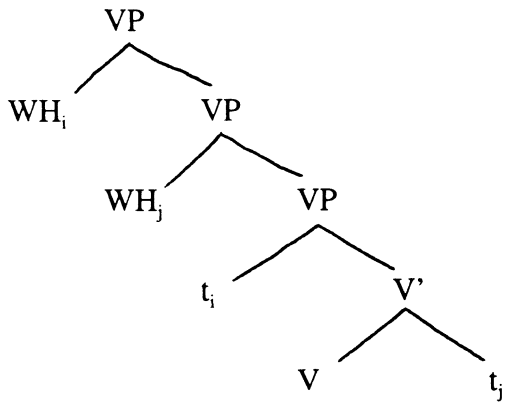
The high ranking of Q-Scope dictates that all WH-phrases be moved out of VP, like in Bulgarian. Unlike in Bulgarian, Shortest Steps outranks Q-Marking, so that it is more important that movement be minimized than it is to derive the proper Q-marking structure. As noted in section 3.1, the derivation that complies with the demand that all WHs take scope over VP which is best with respect to Shortest Steps is one in which all WHs are adjoined to VP and in which there is no head movement (and hence no derived VP* that would be necessary for Q-marking). This is illustrated by the tableau in (45).

(45) *Czech multiple questions*

	Q-Scope	Shortest Steps	Q-Marking
WH V [t t WH]	*!	******	
[WH V WH]	*!*		*
V [WH t WH]	*!*	***	*
☞ [WH [WH [t V t]]]		***	*
WH V [WH [t t t]]		****!****	*
WH WH V [t t t]		****!******	

The optimal candidate is structured as in (46).

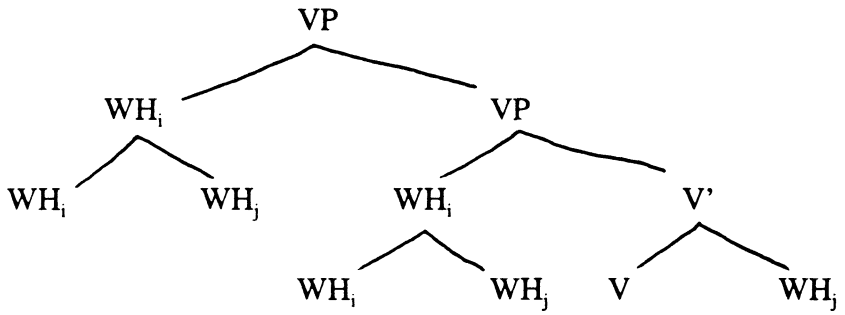
(46)



This representation violates Shortest Steps only three times. The path between the moved subject-WH and its trace consists of VP-segments only. As proposed in section 2.3, crossing segments of the same category does not lead to more than one violation. The path between the object-WH and its trace consists of two VP-segments and V', and hence induces two violations.

In section 3.1 it turned out that it was cheaper with respect to Shortest Steps to first adjoin the object-WH to the subject-WH, and then move this cluster out of VP. One might expect that a derivation like this is cheaper as well in the grammar under discussion, where WH-movement is adjunction to VP. It turns out, however, that in this case clustering the WHs and then moving them together is more costly. Such a derivation would result in the following tree (where the traces are spelled out).

(47)



The path for the moved WH-cluster is <VP, VP>, while the paths relevant to the moved object are <VP, WH$_i$> and <VP, V'>. This amounts to five Shortest Steps violations, two more than in the optimal candidate in (46).[15]

[15] Clustering is also suboptimal where the indirect object and the direct object of a double object construction are concerned. We assume the structure in (i) for these constructions. Adjoining DO- and IO-WHs separately to VP gives five Shortest Steps violations. First adjoining the DO-WH to the IO-WH, and then adjoining this cluster to VP gives six Shortest Steps violations.

It is predicted therefore that the ranking in (44) defines a grammar in which all WHs are fronted, but in which they do not form a cluster. There are indeed languages of this type. Examples are Czech and Polish. Focusing on Czech, all WH-expressions must be fronted in this language, just like in Bulgarian.

(48) Kdo koho viděl
who whom saw
'who saw whom'

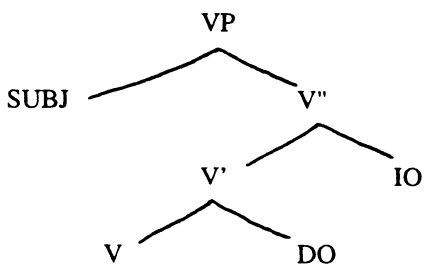
Unlike in Bulgarian, however, elements can appear between the fronted WH-expressions. Consider for example second place clitics, which in Czech must always follow the first major constituent of the clause, as shown by Toman (1986) and Rudin (1988). These authors further show that these clitics always follow the first WH-word in multiple WH-constructions. This indicates that the WH-expressions form separate constituents, not one:

(49) a. Kdo ho kde viděl je nejasné
who him_{CL} *where saw is unclear*
'it is unclear who saw him where'
b. *Kdo kde ho viděl je nejasné
who where him_{CL} *saw is unclear*

Moreover, parentheticals can appear anywhere in the WH-sequence in Czech. This is shown below for an example containing three WH-expressions (Rudin's 49):

(50) a. Kdo, podle tebe, co komu dal
who according to you what to whom gave
'who, according to you, gave what to whom'

(i)



This analysis of double object constructions comes close to the traditional analyses in Blom & Daalder (1977) and De Haan (1979) (where our V' is VP, our V" is PredP, and our VP is S). We assume (i) to be valid for English as well as a basic structure. The surface structure is derived by rightward movement of the DO, which takes place for reasons outlined in Neeleman (1994).

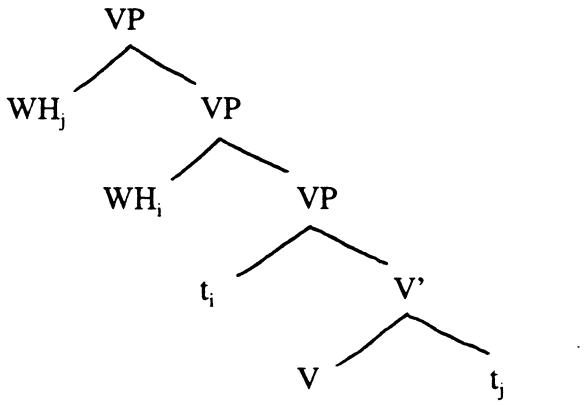
b. Kdo co, podle tebe, komu dal
who what according to you to whom gave

The same phenomenon occurs with adverbials. (All Czech data below were provided by Ludmila Menert, p.c.).

(51) a. Kdo rychle co komu dal
who quickly what to whom gave
'who quickly gave what to whom'
b. Kdo co rychle komu dal
who what quickly to whom gave

From the proposed ranking for Czech it does not only follow that the WH-words do not form a cluster, but also that no superiority effects should exist (unlike in Bulgarian again). Whereas in the Bulgarian WH-cluster the subject-WH precedes the other WHs (cf section 3.1), it is predicted that in a grammar defined by (44) the order between the fronted WHs is free. This is because the number of Shortest Steps violations will always be the same when all WH-words adjoin separately to VP, regardless of the order in which they are moved. Compare, for instance the tree in (52) with the one in (46). In (46) the object was adjoined first, after which the subject was adjoined. In (52) the order of adjunction is reversed.

(52)



Because segments count only once, there is no difference in the number of Shortest Steps violations of the paths connecting the subject and the object to their traces. The paths are <VP, VP> for the subject and <VP, VP, VP, V'> for the object, again giving three violations in total. We thus predict that in Czech-type languages there is free word order between the WHs (modulo possible stylistic preferences). In particular, it should be the case that the subject-WH does not have to precede the other WHs. This prediction is borne out. Alongside (48), (53) exists.

(53) Koho kdo viděl
whom who saw
'who saw whom'

Given the trees in (46/52) a further prediction follows. As opposed to the Bulgarian-type, there should be no verb movement in the Czech-type languages (since the proper structure for Q-marking need not be derived). It is predicted therefore that there should be no obligatory inversion in questions with a non-WH-subject. As noted by Kraskow (1992), this correlation indeed exists. Compare (54) with (39).[16,17]

(54) a. Co kupuje Ivan
what buys John
'What does John buy'
b. Co Ivan kupuje
what John buys

Although our analysis of Czech multiple WH-questions correctly predicts that the fronted WHs do not form a cluster, that they are ordered freely, and that there is no verb movement, there is a final point to consider. This concerns the position of the first WH-element. Until now, we have assumed that, like the other WHs, it is adjoined to VP. According to Rudin (1988), however, the first WH-expression is moved to spec-CP. This can, in fact, not be verified for main clauses, because C (V* in our terms) is not present.

---

[16] Note that inversion is possible, but this need not be due to verb movement, since Czech has an Italian-type of free subject inversion.

(i) a. Včera Jan koupil knihu
Yesterday John has-bought book
'John has bought a book yesterday'
b. Včera koupil knihu Jan
Yesterday has-bought book John

The same appears to be the case for the other languages of this type, like Serbo-Croatian (Maaike Schoorlemmer, p.c.).

[17] It also follows that there will be no languages in which one WH-expression is fronted without supporting verb movement while the other WHs remain in situ:

(i) *WH_i [$_{VP}$ ... t_i ... WH ...]

The reason for this is that, in the absence of verb movement, WH-movement must be triggered by Q-Scope, and not by Q-Marking (since this constraint is violated in (i)). This means that Q-Scope would also trigger movement of the other WHs. As far as we know, this prediction is correct.

Obviously, there is no complementizer in main clauses, and as we have seen above there is also no verb movement. Therefore, given the Shortest Steps requirement, we take it that the initial WH, like the others, adjoins to VP in main clauses.

In embedded clauses the situation is different. Rudin (1988) claims that the doubly-filled-COMP filter is present in languages like Czech and Polish. However, we are informed that in colloquial registers it is possible to combine an overt complementizer with a fronted WH-expression in Czech (though not in Polish):

(55) Chtěl bych vědět ...
I would like to know
a. *že co Marie četla
that what Mary read
a'. co že Marie četla
what that Mary read
b. *že komu co Marie dala
that who what Mary gave
b'. komu že co Marie dala
who that what Mary gave
b". *komu co že Marie dala
who what that Mary gave

These facts show that in embedded clauses the situation is as Rudin assumes: one WH-phrase has moved to spec-CP, while the others are again adjoined to VP (IP in Rudin's system). Apparently, in embedded clauses there is an extra factor forcing a longer movement of one of the WHs, in spite of Shortest Steps. This factor is tied to the presence of the complementizer *že*:

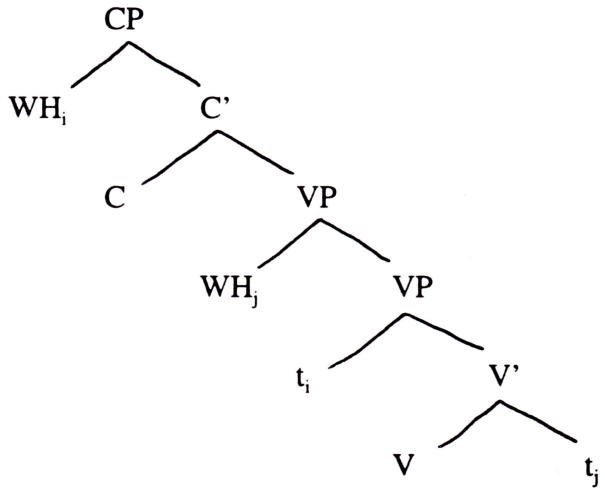
(56) Myslím/vím/říkám *(že) Janek přijde
I think/know/say that John comes

This complementizer marks its complement as a declarative. This implies that WH-expressions, at least at the level of interpretation, should take scope over this complementizer, or the clause would be uninterpretable.[18] However, as argued by Van

[18] In Czech embedded WH-questions, the complementizer that can appear is the declarative one, never the interrogative one. However, movement of at least one WH to spec-CP would be triggered just as well in the presence of an interrogative complementizer. This is because, as noted in section 2.2, a clause introduced by an interrogative complementizer and containing WH-phrases below this complementizer cannot be interpreted as a WH-question, only as an embedded yes-no question with an echo reading for the WH(s) (see also notes 6 and 8).

Riemsdijk (1978), a WH-expression that is A'-moved at surface structure cannot be moved at LF: its syntactic position is fixed. This implies that if all WHs would adjoin to VP in embedded questions (as they do in main clauses in Czech), a structure would result that is uninterpretable. So, in derivations leading to the desired interpretation the WHs must take scope over the complementizer already in overt syntax. In order to achieve this, it suffices to move one WH-expression to a position c-commanding the complementizer at surface structure, i.e. to spec-CP. This becomes apparent if we consider the resulting structure in (57). (The construction contains a CP as opposed to a VP*, since we are dealing with an embedded question. Hence, the functional projection is derived by insertion of a complementizer, rather than by movement of the verb).

(57)

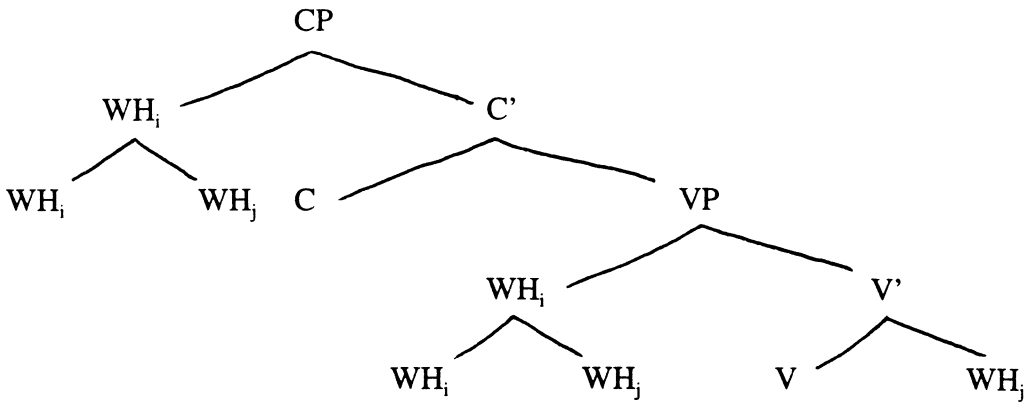


Absorption between the WHs must take place in order to obtain a multiple question interpretation. As noted in footnote 11, the condition on absorption between moved WHs is mutual m-command. The WHs do mutually m-command each other in (57), so absorption is possible. As a result, both WHs will have the same scope, namely that of WH_i (since, as noted, they must take scope over the complementizer). So, in order for interpretation to be successful one WH must move to spec-CP in order to provide an absorption site for the other WHs.

The other WHs will not move to spec-CP but adjoin to VP, because of Shortest Steps. In (57), Shortest Steps is violated five times (the path <CP, C' VP, VP> for the subject giving three violations, the path <VP, VP, V'> for the object giving two violations). The alternative derivation would involve formation of a WH-cluster within VP plus movement of this cluster to spec-CP (à la Bulgarian), as in (58), with traces spelled out. Here, Shortest Steps is violated seven times (the paths <VP, WH_i> plus <VP, V'> for

the object resulting in four violations, the path <CP, C', VP> for the WH-cluster resulting in three violations).[19]

(58)



Note that in the optimal structure in (57), as opposed to the suboptimal structure in (58), Q-marking is blocked by adjunction of a WH to VP. However, Q-Marking is ordered below Shortest Steps in Czech, and is therefore overruled. Hence the contrast with Bulgarian, in which (58) is the optimal structure. In this language, Q-Marking is ranked higher than Shortest Steps (see section 3.1 for discussion).

Finally, it is predicted that, like in main sentences, no superiority effects will obtain in Czech embedded clauses, not even with respect to the movement of one WH-expression to spec-CP. This is because all WHs must be moved out of VP anyway. Then, for Shortest Steps it does not matter which WH-phrase is selected to move further to spec-CP. This prediction is correct: both (59a) and (59b) are possible.

---

[19] It would seem to be predicted that when the distance to be covered by the WH-expressions in a multiple question increases, clustering will be favored. This is not the case, however. The structure of multiple WH-questions in Czech remains the same when the WHs are moved out of an embedded clause to a position c-commanding the matrix-clause (insofar as this is possible in the first place, cf. Rudin 1988:455). In (i), for instance, no clustering seems to take place:

(i) Co, podle tebe, komu Petr řekl [že Jan dal]
what, according to you, who Peter said that John gave
'What, according to you, did Peter say that John gave to whom'

This problem can be solved if we follow the suggestion of Fox (1995) that evaluation proceeds cyclically. That is to say, the ranked constraints first apply to the embedded CP, then to the combination of the embedded clause and the matrix clause. In the embedded CP, Shortest Steps favors separate movement of the two WH-expressions, as explained in the main text. This means that clustering can only take place when the larger cycle is taken into account, i.e. when the two WHs have already been adjoined to the embedded VP. However, then it is no longer possible, because it would have to take place within the embedded clause (the initial landing site of the WHs), which would be a violation of strict cyclicity.

(59) Chtěl bych vědět ...
I would like to know
a. kdo (že) co četl
who that what read
b. co (že) kdo četl
what that who read

In conclusion, like the first three, the fourth possible ranking of the constraints we have proposed results in an actually occurring type of question formation as well. In the next section we will show that the remaining two possible rankings lead to one more type, which is attested as well.

3.3 *Chinese and Japanese*

Suppose the constraints proposed in section 2.3 are ranked as in either (60a) or (60b), the two remaining possibilities.

(60) a. Shortest Steps >> Q-Marking >> Q-Scope
b. Shortest Steps >> Q-Scope >> Q-Marking

The high ranking of Shortest Steps ensures that it is more important not to move than it is to satisfy Q-Marking or Q-Scope. The result is that all WH-expressions will have to remain in situ in grammars defined by (60) (see the tableau in (61), in which the ranking between Q-Marking and Q-Scope is irrelevant).

(61) *Chinese multiple questions*

	Shortest Steps	Q-Scope	Q-Marking
WH V [t t WH]	*!*****	*	
☞ [WH V WH]		**	*
V [WH t WH]	*!**	**	*
[WH [WH [t V t]]]	*!**		*
WH V [WH [t t t]]	*!*******		*
WH WH V [t t t]	*!*********		

As is well-known, languages that adhere to the tableau in (61) exist. According to Huang (1982), Chinese is an example. In root questions as well as embedded questions, WH-expressions remain in their base position:

(62) a. Ni xihuan shei
you like who
'who do you like'

b. Wo xiang-zhidao Lisi mai-le sheme
I wonder Lisi bought what
'I wonder what Lisi bought'

At this point we need to be more precise about the interpretation of WH-expressions in languages like Chinese and Japanese. We have assumed earlier, following Reinhart (1993), that in English WHs in situ are not assigned scope by LF movement, but by an interpretational operation. One of Reinhart's arguments for this was that the interpretation of WH-in-situ is not subject to Subjacency. Although we cannot go into the details of the interpretational operation here, it is important to notice that in essence it makes use of the fact that a Q-operator obligatorily occupies spec-VP*/spec-CP in English WH-questions. Unmoved WH-phrases are parasitic on this Q-operator in a fashion reminiscent of unselective binding, so that for our present purposes the LF of a question like (63a) can be represented as (63b):[20]

(63) a. [Which persons]$_i$ did you ask [whether John read what]

b. Which$_{<i,j>}$ did you ask [t_i persons] [whether John read what$_j$]

We would like to assume that the Q-operator on which WH-phrases in situ depend for their interpretation must be introduced by lexical material. That is to say, it is not an abstract operator that can be freely inserted (as in English yes/no questions), but it is introduced by a moved Q-bearing element (i.e. *which persons* in the example above). Returning now to languages like Chinese and Japanese, this assumption has the effect that one WH-expression must move by covert movement to a position c-commanding the proposition. After the movement of this one WH, possible other WH-expressions can be interpreted in situ through the semantic operation proposed by Reinhart. So, an S-structure like (64a) must be mapped into a representation like (64b) at LF:[21]

---

[20] Unlike absorption of moved WHs, this interpretational mechanism obviously does not require mutual m-command of the WH in situ and the WH in spec-VP*/CP, which determines the scope of the question.

[21] Note that LF movement is possible in Chinese and Japanese, but not in Czech embedded questions (see section 3.2), because, as already noted in the section on Czech, the fact that WHs move overtly in this language blocks further movement at LF.

(64) a. $[_{VP} \ldots WH \ldots WH \ldots]$

b. $[_{VP} WH_{<i,j>} [_{VP} \ldots t_i \ldots WH_j \ldots]$

So, the prediction is that in languages with no overt movement one WH must move covertly, while the rest remains in situ.[22] Data presented by Watanabe (1992) indicate that this may well be the case. Watanabe shows that Subjacency effects can be observed in Japanese as well, despite the fact that all WHs are in situ at S-structure. Consider, for instance, the contrast in (65).

(65) a. ??John-wa [Mary-ga nani-o katta ka dooka] Tom-ni tazuneta no?
John-TOP *Mary*-NOM *what*-ACC *bought whether Tom*-DAT *asked* Q
'What did John ask Tom whether Mary bought'

b. John-wa [Mary-ga nani-o katta ka dooka] dare-ni tazuneta no?
John-TOP *Mary*-NOM *what*-ACC *bought whether who*-DAT *asked* Q
'Who did John ask whether Mary bought what'

Since the scope of the question is the matrix clause in these examples, a WH-expression must be moved to the top of the construction. The null hypothesis seems to be that Subjacency is a general constraint on movement, whether overt or covert.[23] If so, the ungrammaticality of (65a) can be explained if *what*, being the only Q-bearing element in the construction, must indeed be (covertly) moved. The grammaticality of (65b) can be explained as well: it follows from the presence of a WH-phrase in the matrix clause. This phrase can introduce the relevant Q-operator, so that *what* can remain in situ at LF and be interpreted in the way described by Reinhart.

---

[22] On apparent ECP effects with adjunct WHs in situ (or the lack thereof in certain languages) see Reinhart (1993), Aoun & Li (1993) and Cole & Hermon (1994).

[23] This is not what Watanabe concludes from this evidence. He assumes, following Huang (1982), that overt and covert movement fundamentally differ in that only the former is subject to Subjacency. This means that the Subjacency effect in (65a) gives evidence for S-structure movement. Watanabe then argues that at S-structure a null operator moves, while what we conceive of as the WH-expression remains in situ throughout the derivation (cf. also Aoun & Li 1993 on Chinese). Watanabe rejects analyses of the type assumed in the text (i.e. LF movement of one WH but not the others, with Subjacency holding for covert movement as well), on the basis of evidence suggesting that the spec-CP of embedded questions is already filled at S-structure. The evidence is based on the possible occurrence of Subjacency effects with certain overt movements in Japanese. This evidence is rather inconclusive, however. Watanabe notes (i) that Subjacency effects are lacking with the most straightforward type of overt movement in Japanese, namely scrambling, and (ii) that an alternative explanation can be given for most of the supposed S-structure Subjacency effects. But even if Watanabe is correct, this is not dramatic for our analysis. It would simply mean that Japanese is of the English type. We would only have a problem if there are no 'real' WH-in-situ languages at all.

To summarize, we have shown that the six possible rankings of Q-Marking, Q-Scope and Shortest Steps define four different patterns of question formation. All four patterns do actually occur. We take this as convincing evidence for the existence of these constraints, and for their operating in an optimality-theoretic fashion.

3.4 *French*

In French root questions a phenomenon can be observed that we have not discussed yet and which might seem problematic at first sight: optionality. It is possible that one WH moves to spec-VP* while the rest stays in situ, but it is also possible that all WHs remain in situ.[24]

(66) a. Qu' as-tu donné à qui
what have you given to whom
b. Tu as donné quoi à qui
you have given what to whom

One way in which optionality can arise in optimality theory is when there happens to be more than one optimal candidate with respect to a given ranking; it is possible that

---

[24] If the subject is not a clitic, but a full NP like *Jean*, simple inversion as in (66a) is impossible (cf. ia). Rather, there either must be stylistic inversion as in (ib), or complex inversion (again involving a clitic) as in (ic).

(i) a. *Qui a Jean vu
who has John seen
b. Qui a vu Jean
who has seen John
c. Qui Jean a-t-il vu
who John has he seen

Crucially, these examples do not incriminate an analysis in which the WH-movement in French is obligatorily accompanied by verb movement. On the contrary, this assumption makes possible an explanation of the fact that stylistic inversion and complex inversion do not occur in non-WH-sentences. (ia) shows that simple inversion in French is impossible when the subject is not a clitic. What is going on in (ib) and (ic), then, is that the subject is generated in a position such that it does not block verb movement (although this cannot be the sole trigger for stylistic inversion, since it is optionally available in embedded sentences as well). In (ib) it is in a right-peripheral position within VP (cf. Sportiche 1988, Friedemann 1991, and others); in (ic) it is degraded to an adjunct (adjoined to V*', cf. Rizzi & Roberts 1989) which is linked to the subject clitic. At this point we do not have an analysis for why full subjects block simple inversion (cf. Rizzi & Roberts 1989, Hulk 1993, for discussion).

candidates are different but nonetheless score equally on all constraints. However, given the discussion in section 3.1-3.3, this cannot be the case in (66). We discussed all possible rankings of the three relevant constraints and in all cases only one type of question formation turned out to be optimal. What must be the case, then, is that (66) is caused by the fact that in French two of the constraints are *unranked* with respect to each other. This leads to optionality as well. When two constraints A and B are unranked with respect to each other, then the candidate that is optimal when A outranks B and the candidate that is optimal when B outranks A are rated equally high (cf. Kager 1994). If these are different candidates, there will be a choice between the two. More specifically, the French paradigm can be explained by the constraint ranking in (67) (where '<>' indicates equal ranking).

(67) Shortest Steps <> Q-Marking >> Q-Scope

Under the interpretation of (67) in which Shortest Steps outranks Q-Marking, the optimal candidate will be the 'Chinese' one, as the tableau in (68) shows (compare section 3.3).

(68) *French WH in situ*

	Shortest Steps	Q-Marking	Q-Scope
WH V [t t WH]	*!*****		*
☞ [WH V WH]		*	**
V [WH t WH]	*!**	*	**
[WH [WH [t V t]]]	*!**	*	
WH V [WH [t t t]]	*!*******	*	
WH WH V [t t t]	*!*********		

Under the interpretation of (67) in which Q-Marking outranks Shortest Steps, the 'English' candidate is optimal, as can be observed in the tableau in (69) (compare section 2.3).

(69) *French WH-movement*

	Q-Marking	Shortest Steps	Q-Scope
☞ WH V [t t WH]		******	*
[WH V WH]	*!		**
V [WH t WH]	*!	***	**
[WH [WH [t V t]]]	*!	***	
WH V [WH [t t t]]	*!	********	
WH WH V [t t t]		*******!***	

Note that our analysis explains that, although both WH-movement and supporting V-movement are optional in French, they cannot apply independently. Both (70a) and (70b) are ungrammatical.

(70) a. *Quoi tu as donné
what you have given
b. *As-tu donné quoi
have you given what

The ungrammaticality of (70b) is an instantiation on a more general phenomenon. As we explained earlier, verb movement without movement of a WH will never lead to a more optimal candidate (see section 2.3). The ungrammaticality of (70a) also follows. The absence of inversion shows that the WH is in the 'Czech' position, adjoined to VP, but this structure is not optimal with respect to either of the two rankings of French.

Moreover, it follows that, if WH-movement takes place in French, superiority effects obtain. Under the 'English' ranking, which is the one that triggers WH-movement, Shortest Steps has the effect that the WH closest to spec-VP* must be moved:

(71) a. Qui a vu quoi
who has seen what
b. *Qu' a vu qui
what has seen who

An apparent problem for the analysis proposed here is that in embedded questions the optionality of WH-movement disappears. Here, there must be WH-movement of one WH-expression to spec-CP:

(72) a. *Je me demande (que) tu as vu qui
I wonder (that) you have seen who
b. Je me demande qui (~~que~~) tu as vu
I wonder who you have seen

Following Lasnik & Saito (1992), we would like to suggest that a simple explanation for the contrast between French main and embedded questions can be given once the notion of selection is taken into account. The difference between main and embedded questions is, of course, that the latter are selected by a matrix verb. A verb like *se demander* in (72), for example, selects a complement carrying a Q-feature. This type of selection is basically semantic in nature. Suppose, however, that there is a constraint which states that semantic selection must be reflected in syntax:

(73) *Selection*
Selectional requirements must be satisfied at S-structure

In the case of embedded questions, Selection will be satisfied when the highest projection of the embedded sentence carries the Q-feature. In (72b) this is the case, because the WH-phrase occupies spec-CP. By spec-head agreement and percolation the Q-feature will be present on CP. In (72a), however, the Q-feature will not reach the top of the projection, since it is present on the complement of the verb.

We must now ask ourselves what the status of Selection is within the optimality theoretic theory of question formation outlined above. We would like to propose that it is not evaluated on a par with the other constraints, for reasons to be explained below. Instead, two sets of constraints must be distinguished. One set contains the constraints proposed in section 2.3, the other contains Selection and possibly other constraints. The output of the evaluation of the first set then functions as input for the evaluation of the second set. So, if the first set of constraints allows for more than one optimal structure, as in the case of French questions, the second set of constraints may pick one of these as the ultimately optimal candidate. Thus the optionality disappears. This is what happens in French embedded questions. Like in main questions the first set of constraints defines two optimal candidates. The constraint Selection, which plays no role in main questions, then filters out the one with WH-in-situ. In strict WH-in-situ languages like Chinese, the first set of constraints defines only one optimal candidate, due to the fact that Shortest Steps is not equally ranked with respect to Q-Marking, but outranks it. Hence, there is only one candidate to be considered by Selection. Given that this is so, it will be the optimal candidate, despite the fact that it violates this constraint.

One may wonder whether the difference that in French Selection is not violated while in Chinese it is, cannot be explained as well under the assumption that Selection is evaluated on a par with the other constraints. This is indeed possible. In this view,

Selection would be the highest ranked constraint in French, while in Chinese it would be ranked somewhere below Shortest Steps. As is easy to see, the facts then follow. However, if language variation is explained by different rankings of Selection with respect to the other constraints, unattested language types are predicted. Consider, in particular, the ranking, in (74).

(74) Selection >> Shortest Steps >> Q-Marking >> Q-Scope

A language defined by (74) would obligatorily leave all WHs in situ in main clauses, while it would obligatorily move one WH to spec-CP in embedded clauses. The reason for this is obvious: Selection plays no role in main clauses, so that movement is suppressed by Shortest Steps; it does play a role in embedded clauses, so that WH-movement is triggered. To the best of our knowledge, no such language exists. We therefore feel that separate evaluation is empirically justified.

We think that there is some conceptual justification as well: there seems to be a parallel in phonology, as pointed out to us by René Kager (p.c.). In phonology, morphemes can impose selectional requirements on the prosodic shape of the base they attach to. For instance, if a language allows three prosodic patterns, a morpheme may require that its base has one of these patterns. However, in general a morpheme cannot require the base to have a pattern that is not part of the prosodic inventory of the language. That is to say, selectional requirements typically do not overrule the constraints determining the general prosodic make-up of the language, a phenomenon known as structure preservation (cf. Kiparsky 1985). This might be explained if these requirements operate on the output of the other constraints, like we propose for syntactic selectional requirements.

In sum, the optionality of WH-movement in French can be successfully analyzed in terms of equal ranking of Shortest Steps and Q-Marking. However, the possibility of equal ranking at first sight seems to reintroduce the problem of language typology. Recall that all possible rankings of Shortest Steps, Q-Marking and Q-Scope define an existing pattern of WH-question formation. As we saw for French, equal ranking has the effect that two of these patterns occur in the same language. If equal ranking is an option readily available in UG, one might expect many such co-occurrences of patterns of question formation. One might even expect a language in which all four patterns discussed in the previous sections are attested, namely when all constraints are ranked equally. In fact, however, optionality as in French seems to be rare and, to the best of our knowledge, languages allowing all four patterns do not exist. Apparently, what must be the case is that equal ranking is a marked option. This is in fact a plausible assumption, for reasons having to do with language acquisition.

Tesar & Smolensky (1993) address the question of how grammars are acquired in optimality theory. They propose a learning algorithm of which we can only give a rough

sketch here. Upon encountering a datum, the child concludes that this datum is the optimal form in the candidate set (in syntax: all outputs that target the same semantics and that are projected from the same set of lexical items). Crucially, the child then hypothesizes that all other outputs in the candidate set are less optimal. This hypothesis allows it to rank the constraints. Constraints that are violated by the encountered datum (i.e. the optimal output), but not by the other candidates, cannot be ranked among the highest ranked constraints: they must be outranked by constraints that are violated by the other candidates but not by the encountered datum. Consider for instance English WH-questions. Suppose a child encounters a sentence like *what will John read.* It will then hypothesize that other potential outputs of the underlying form (*John will read what*) are less optimal. Those potential other outputs include **what John will read*, in which a WH-expression has been fronted but the verb remains in situ, and **John will read what*, in which both elements remain in situ. Since the optimal candidate violates Shortest Steps to a greater extent than either of the two alternative outputs, this constraint must be ranked low. Since the two potential outputs, as opposed to the actual output, violate Q-Marking, this constraint must be ranked high. The child thus arrives at the correct partial ordering: Q-Marking >> Shortest Steps. Further data are required to order Q-Scope with respect to these two constraints.

The problem of acquiring a grammar with equal ranking should now be obvious. The learning strategy is based on the hypothesis that all potential outputs other than the actually attested form are suboptimal. Hence, a grammar will be acquired that rules out these other candidates. This means that if there is actually a second optimal candidate, due to equal ranking, part of the already acquired ranking must be unlearned. It seems plausible that this requires exposure to robust evidence, otherwise the child will ignore the evidence for a second optimal form like it ignores occasional mistakes in the input. This explains the relatively rare occurrence of equal ranking.

If this explanation is correct, we predict that, in the acquisition of WH-questions, French children will first go through a stage of optionality (the constraints are not yet ranked), then go through a stage in which the optionality displayed by the adult grammar is lacking (the constraints are ranked), and then learn the adult optionality (Shortest Steps and Q-Marking are 'disranked' again). This seems to be correct. Weissenborn (1993) reports that from 2;01;19 up to 2;03;21 WH-in-situ is as good as absent in Philippe (one example out of 114 WH-questions). From 2;06;13 onwards, WH-in-situ coexists with WH-movement (81 examples out of 199 WH-questions).[25]

In conclusion, the optionality in French question formation can be explained by assuming that Q-Marking and Shortest Steps are unranked with respect to each other. At

[25] A different matter is why French children (or Philippe at least) go through a stage of obligatory WH-movement instead of obligatory WH-in-situ. Our suggestion is that the evidence for WH-movement is more robust in the input because of its obligatory nature in embedded contexts.

the same time, we expect this type of optionality to be relatively rare, since it is difficult for the child to learn that two constraints are not ranked: the learning algorithm is designed to acquire total ranking.

4. *Summary*

Let us summarize the main proposal of the paper. We have argued that variation in the formation of WH-questions in such languages as English, Bulgarian, Czech, Chinese and French can be accounted for by a different ranking of three general constraints on movement and question marking, namely the following:

(75) a. *Shortest Steps*
Minimize the distance between chain links
b. *Q-Marking*
A question must be overtly Q-marked
c. *Q-Scope*
[+Q] elements must c-command VP at surface structure

Several properties of English-type questions follow from the constraint ranking in (76).

(76) Q-Marking >> Shortest Steps >> Q-Scope

The high ranking of Q-Marking ensures that the proper structure for Q-marking must be derived. Therefore, in English root WH-questions there must be verb movement plus movement of at least one WH to spec-VP*. (In embedded questions, we assume that a Q-marking complementizer is always present, though under certain cirumstances it is not spelled out; cf. Pesetsky 1994). In multiple questions the fate of WHs other than the one that moves to spec-VP*/CP is determined by the fact that Shortest Steps outranks Q-Scope: they remain in situ. Finally, Shortest Steps explains that the one WH that is moved is the one closest to spec-VP*/CP (superiority) (Golan 1993).

The properties of Bulgarian-type question formation follow from the constraint ranking in (77).

(77) Q-Marking >> Q-Scope >> Shortest Steps

Like in English, the proper Q-marking structure must be derived. So, one WH at least must be moved to spec-VP*. But since Q-Scope now outranks Shortest Steps, all WHs must be moved out of VP. We have argued that adjunction of the other WHs to VP blocks Q-marking, hence the ordering in (77) has the result that all WHs move to spec-VP*/CP.

The high ranking of Q-Marking also ensures that there must be verb movement. So, if the subject is not a WH itself, inversion occurs. Finally, we have shown that even though Shortest Steps is ranked lowest in (77), it still has its effects: it (indirectly) explains the effect reminiscent of superiority that within the WH-cluster the subject must come first. A grammar with identical properties is derived when the ranking of Q-Marking and Q-Scope is reversed in (77), because they are never crucially in conflict.

If the constraints are ranked as in (78), Czech-type questions are derived.

(78) Q-Scope >> Shortest Steps >> Q-Marking

The high ranking of Q-Scope ensures that all WHs must be moved out of VP. However, they do not move to spec-VP* in this case. Given that Shortest Steps outranks Q-Marking, it is better to adjoin all WHs to VP than it is to apply verb movement plus movement of all WHs to spec-VP*. The low ranking of Q-Marking implies that there will be no verb movement, and hence no inversion, in questions in these languages. Finally, we have argued that it follows from (78) that the order of the WHs is free in this case (there are no superiority effects).

The two rankings in which Shortest Steps outranks the other two constraints have the effect that all WHs must remain in situ, as in Chinese and Japanese. Summarizing, the six possible rankings of the three constraints result in four different patterns of question formation (in two cases two different rankings result in the same pattern). As we have seen, these patterns are all attested, while we do not know of other patterns.[26] Amongst other things, it is predicted that there are no languages in which WH-questions are formed by verb movement while all WHs remain in situ, and that there are no languages in which one WH is moved without accompanying verb movement while the other WHs remain in situ.

There is some further language variation: in French root questions, either all WHs remain in situ, or one WH moves to spec-VP*, with accompanying verb movement. This

---

[26] In some languages, like Italian and Irish, multiple questions seem to be impossible. This might be incorporated into the system by adding a constraint Parse (WH), which is violated if a WH-phrase in the input is not syntactically realized (cf. Legendre, Smolensky & Wilson 1995). Italian/Irish might then be accounted for by the ranking in (i) (in the languages discussed in the text Parse (WH) is ranked high).

(i) Q-Marking >> Q-Scope >> Shortest Steps >> Parse (WH)

The high ranking of Q-Marking ensures that there must be movement of one WH to spec-VP*/CP. The question is what happens to the other WHs in the input of a multiple question. Q-Scope for the other WHs can be satisfied in two ways: by moving them out of VP as well (see Bulgarian and Czech) or by not realizing them in syntax at all (in that case Q-Scope is satisfied vacuously). If Shortest Steps outranks Parse (WH), as in (i), the latter option must be chosen in the optimal candidate.

means it is possible to have two of the attested patterns occurring in a single language. This can be accounted for by assuming that two constraints can be unranked with respect to each other. Considerations of language acquisition suggest that this is a marked option.

If the analysis given here is correct, there is reason to believe that syntax and phonology are less different than is sometimes assumed.

References

Ackema, Peter: 1995, *Syntax below Zero*, PhD dissertation, Utrecht University.

Ackema, Peter, Ad Neeleman & Fred Weerman: 1993, 'Deriving Functional Projections', in A. Schafer (ed.), *Proceedings of NELS 23*, Amherst, GLSA, pp. 17-31.

Aoun, Joseph & Yen-hui Audrey Li: 1993, '*Wh*-elements in situ: Syntax or LF?', *Linguistic Inquiry* 24, 199-238.

Barrs, Andrew & Howard Lasnik: 1986, 'A Note on Anaphora and Double Objects', *Linguistic Inquiry* 17, 347-354.

Billings, Loren & Catherine Rudin: 1995, 'Optimality and Superiority: A new Approach to Overt Multiple-*wh* Ordering', Ms. Princeton/Rutgers and Wayne State College. To appear in *Formal Approaches to Slavic Linguistics* 3.

Blom, Alied & Saskia Daalder: 1977, *Syntaktische theorie en taalbeschrijving*, Muiderberg.

Chomsky, Noam: 1986, *Barriers*, Cambridge, MIT Press.

Chomsky, Noam: 1993, 'A Minimalist Program for Linguistic Theory', in K. Hale & S.J. Keyser (eds.), *The View from Building 20*, Cambridge, MIT Press, pp. 1-52.

Chomsky, Noam: 1994, 'Bare Phrase Structure', *MIT Occasional Papers in Linguistics* 5.

Chomsky, Noam: 1995, 'Categories and Transformations', Ms., MIT.

Cole, Peter & Gabriella Hermon: 1994, 'Is there LF *Wh*-Movement?', *Linguistic Inquiry* 25, 239-262.

Collins, Chris: 1994, 'Economy of Derivation and the Generalized Proper Binding Condition', *Linguistic Inquiry* 25, 45-61.

De Haan, Germen: 1979, *Conditions on Rules; The Proper Balance between Syntax and Semantics*, Dordrecht, Foris.

Di Sciullo, Anna-Maria & Edwin Williams: 1987, *On the Definition of Word*, Cambridge, MIT Press.

Fox, Danny: 1995, 'Economy, Scope and Semantic Interpretation', Talk given at the OTS Colloquium, Utrecht University.

Friedemann, Marc-Ariel: 1991, 'On the D-Structure Position of Subjects in French', in S. Barbiers, M. den Dikken & C. Levelt (eds.), *LCJL 3 Proceedings*, Leiden University, pp. 155-168.

Golan, Yael: 1993, 'Node Crossing Economy, Superiority and D-Linking', Ms., Tel Aviv University.

Grimshaw, Jane: 1993, 'Minimal Projection, Heads, and Optimality', Ms., Rutgers University, New Brunswick. To appear in *Linguistic Inquiry*.

Higginbotham, James & Robert May: 1981, 'Questions, Quantifiers and Crossing', *The Linguistic Review* 1, 41-80.

Huang, James: 1982, *Logical Relations in Chinese and the Theory of Grammar*, PhD dissertation, MIT.

Hulk, Aafke: 1993, 'Residual Verb-Second and the Licensing of Functional Features', *Probus* 5, 127-154.

Jensen, John & Margaret Stong-Jensen: 1984, 'Morphology is in the Lexicon!', *Linguistic Inquiry* 15, 474-498.

Kager, René: 1994, 'Ternary Rhythm in Alignment Theory', Ms., Utrecht University.

Kiparsky, Paul: 1985, 'Some Consequences of Lexical Phonology', *Phonology* 2, 85-138.

Koeneman, Olaf. 1995. 'Flexible Syntax and Sentence-Type Marking', MA Thesis, Utrecht University.

Kraskow, Tina: 1992, 'Multiple WH-Movement in Slavic', *The Slavic Syntax Newsletter* 2, 9-10.

Lapointe, Stephen: 1980, *A theory of grammatical agreement*, PhD dissertation, University of Massachusetts, Amherst.

Lasnik, Howard & Mamoru Saito: 1992, *Move α; Conditions on its Application and Output*, Cambridge, MIT Press.

Legendre, Géraldine, Paul Smolensky & Colin Wilson: 1995, 'When is Less More? Faithfulness and Minimal Links in *wh*-Chains', paper presented at the Workshop on Optimality in Syntax, MIT.

Lieber, Rochelle: 1980, *On the Organization of the Lexicon*, PhD dissertation, MIT.

May, Robert: 1977, *The Grammar of Quantification*, PhD dissertation, MIT.

May, Robert: 1985, *Logical Form; Its Structure and Derivation*, Cambridge, MIT Press.

Muysken, Pieter: 1982, 'Parametrizing the Notion Head', *Journal of Linguistic Research* 2, 57-75.

Neeleman, Ad: 1994, *Complex Predicates*, PhD dissertation, Utrecht University.

Pesetsky, David: 1994, 'Optimality Principles of Sentence Pronunciation', Ms., MIT.

Prince, Alan & Paul Smolensky: 1993, *Optimality Theory; Constraint Interaction in Generative Grammar*, Ms., Rutgers University/University of Colorado, Boulder.

Reinhart, Tanya: 1993, 'WH-in-Situ in the Framework of the Minimalist Program', Ms., Tel Aviv University/Utrecht University.

Riemsdijk, Henk van: 1978, 'On the Diagnosis of WH Movement', in S.J. Keyser (ed.), *Recent Transformational Studies in European Languages*, Cambridge, MIT Press.

Riemsdijk, Henk van: 1989, 'Movement and Regeneration', in P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht, Foris, pp. 105-136.

Rizzi, Luigi & Ian Roberts: 1989, 'Complex Inversion in French', *Probus* 1, 1-30.
Rudin, Catherine: 1988, 'On Multiple Questions and Multiple WH Fronting', *Natural Language and Linguistic Theory* 6, 445-501.
Sportiche, Dominique: 1988, 'A Theory of Floating Quantifiers and its Corollaries for Constituent Structure', *Linguistic Inquiry* 19, 425-449.
Tesar, Bruce & Paul Smolensky: 1993, 'The Learnability of Optimality Theory: An Algorithm and some Basic Complexity Results', Ms., University of Colorado, Boulder.
Toman, Jindrich: 1986, 'Cliticization from NPs in Czech and Comparable Phenomena in French and Italian', in H. Borer (ed.), *The Syntax of Pronominal Clitics*, New York, Academic Press, pp. 123-145.
Watanabe, Akira: 1992, 'Subjacency and S-Structure Movement of WH-in-Situ', *Journal of East Asian Linguistics* 1, 255-291.
Weissenborn, Jürgen: 1993, 'Matrix Infinitives and Economy in Language Development: Comments on Wexler', Talk given at the Conference on the Robustness of the Language Faculty, Utrecht University.

DRAFT, May 1995

Utrecht University
Research Institute for Language and Speech (OTS)
Trans 10
3512 JK Utrecht
The Netherlands

ackema@let.ruu.nl / neeleman@let.ruu.nl

Optimality theory and Greek syntax[1]

Gaberell Drachman
University of Salzburg

1. Background

This paper claims that the ranking of constraints, a fundamental tenet of Optimality Theory (OT), is a central part of Modern Greek syntax, and hence potentially of universal grammar.

As against Connectionism, OT respects symbolic representations and does not attempt to replace them or merely implement them. Thus on the one hand many old syntactic friends will be found in the present account, though perhaps transformed, e.g., the lexical categories, the constituent types, and constraints like case, the Projection Principle and the Wh-criterion; and on the other, not numerical weighting but simply ranking of constraints will be decisive.

As against traditional principles and parameters (P&P) treatments of phonology, morphology and syntax (including minimalistic ones), OT accounts involve no ordered rules or other derivational-processes or steps. OT involves only input representations, together with ranked constraints that filter candidate output representations. As for Minimalism in syntax, Chomsky's attitude remains arch-conservative. While for Bromberger-Halle (1989) 'Phonology is different' in requiring ordered rules, for Chomsky (1994) syntax absolutely depends on derivations.[2]

As against P&P and Minimality, the source of syntactic variation or change lies for OT not in the language-specific (morphological) properties of functional heads, but simply in the language-specific ranking or re-ranking of constraints.

And lastly, as against P&P and Minimality theory, which require a supplementary acquisition algorithm, OT is claimed to constitute its own ready-made (serial or parallel) acquisition algorithm: for Tesar (1993), a constraint hierarchy is essentially learnable.

2. Syntax

2.1. History

Bromberger and Halle (1989) proposes to answer a challenge from syntax. 'Questions about the ordering of transformations and about intermediate representations have all but disappeared from syntax..... This raises the question whether phonology should not undergo a similar development'. They add, however, that 'the facts ..are of a very different nature, and .. therefore there is no reason to assume a priori that they must be covered by formally similar theories'. And Chomsky (1994) remains somewhat underwhelmed by the suggestion, as already indicated.

But now that phonology and morphology have increasingly abandoned rules and derivations in favour of an OT approach, the converse question arises, do we really have to show why syntax is different? And in fact we will assume for present purposes that syntax is NOT different, and thus take the counter assumptions seriously, viz., (1) that there are no derivations -- there are only output representations of inputs -- and (2) that all constraints are violable, and there are no holy cows, not even the Wh-criterion, much less subjacency.

2.2. Constraints in syntax and OT

We survey first some general principles for clause structure

2.2.1.Ready-made constraints

Here are some old friends -- viz., Chomskian minimality constraints ready-made for taking over as OT constraints:

Move only if necessary, move minimally, Full-Interpretation,
ECP, Subjacency, Case Filter, Wh-criterion,
Last resort, No vacuous Quantifier/Operator-interpretation

2.2.2. Grouped

And here they are as in Grimshaw (1994); grouped, though not ranked.
Some refer to Specifiers, as in

(1) Op-in-Spec, part of the Wh-criterion;
(2) Spec-Phi (Spec of a head with phi-Features) must be filled at SS, equivalent to the Case Filter *NP [-Case]; and
(3) Subject, or the EPP - the highest A-specifier in an extended projection must be filled.

And some refer to Heads, e.g.,

(1) Ob-Hd - a projection obligatorily has a head;
(2) Hd-left - a head is leftward in its projection; and
(3) *Lex > Func - no lexical head in an F projection.

Others refer to government, e.g.,
T-(Lex) government: T is either governed, or lexically-governed.

And yet others include the well known triad

(1) Projection Principle. No movement to head of, and no adjunction to a selected clause;
(2) STAY ~ *trace = move only if necessary; and
(3) Full Interpretation (FI)

2.2.3. Further cases

From my own work I add

(1) Minimal structure is to be assigned for each construction token in its own right (Drachman (1989)) -- though contrast Sportiche (1993); and
(2) Corollary to (1): No generalising assumptions such as 'all complements are CP' are allowed. The size of a (Grimshaw-extended) projection is variable.

And in addition I shall employ

(3) Align -- borrowed from OT-morphology; and
(4) A split form of the Wh-criterion (adapted from Grimshaw, as the constraint True-Topic - Subjects may be truly Topics, with A' movement.

2.3. Clause division

2.3.1. The division

The functional division of clauses into Theta (VP, or Argument structure), morphological checking, and Typing/Operator components discussed in Drachman (1994) may now be handled partly under a general 'Align' constraint (checking, Typing, or Scope), perhaps parallel to Merge in Chomsky. We have the principle:

Left align Functions, from the right.

This gives us: (1) Align Head, Left of its complement; and (2) Spec, Left of Head-complement, as in:

a) in VP (Theta complex);
b) Align Verb-Morphology (Checking module) Left of Theta-complex;
c) Align Type (Decl, Interrog): Complementiser Left of checking module;
d) Align Scope: Align Operators (Topic, Focus, Wh) Left of Type.[3]

2.3.2. Consequences of the tripartite clause division
With leftward verb-head, a surface configuration XP-Verb is interpreted as resulting from left movement by XP over the verb; it must be motivated, by feature-checking or attraction by an operator. Subject raising in Greek can hardly serve nominative-assignment (Nom) , since the possibility of VSO shows that Nom-checking may (and therefore must) be delayed until LF under Procrastinate. Then, supposing a Topic operator is present, subject movement is obligatorily out of MP -- though via Spec-MP, to preclude violation of the EPP.[4] If only fully M-checked (e.g., Case-checked) elements can Wh-move, checking must intervene before Op-movement; this leaves NP-movement to operate before checking, for instance to guarantee Nom for raised passive objects. And finally, Op-movement always involves scope, so it may only be leftward up the tree. We thus at least partly derive the mutual ordering of the projection modules.

3. Case Studies

There follow now some tentative case studies, the aim of which is to show not merely that current theories are translatable into OT terminology, but also that OT brings a fresh and productive perspective to certain problems in Greek syntax.

3.1. Wh-movement
The output of so-called Wh-movement responds in Grimshaw (1994) to (among others) the three ranked U-constraints

Operator-in-Spec >> Minimal clause >> Move minimally, and/or *trace [5]

3.1.2. Wh-movement for Greek and English
Since Wh-movement is obligatory in the syntax for Greek as it is for English, we assume that Op-in-Spec is an undominated constraint in both languages. But of course a Wh-element cannot be moved to a Specifier already occupied by another constituent (or the trace of one). Compare English with Greek in this respect:

Wh goes to ⇒	Spec-CP	Spec-IP	Spec-VP
IN ⇓			
English:	Obligatory	NO: NP-Subj	NO: Deep-Subj
Greek:	Last resort	Oblig when possible	NO: Deep-Subj

If we assume 'Subject in VP' is a universal, since all the arguments of a verb must be in its (immediate) projection, it follows that Spec-VP is never available as an Op-landing site. But Spec-IP and Spec-CP are still candidates. What chooses between them?

3.1.3. English
English has Nominative-assignment overtly in syntax, in Spec-IP. So the first free Spec (recall, *trace) for English is indeed Spec-CP. Then for English, satisfying Op-in-Spec forces a violation of Minimal clause by projection-extension to CP.

3.1.4. Greek

Greek has SOV and VSO (not to mention VOS) orderings on the surface. As mentioned above, we assume all these orderings to implicate Subject-in-VP and verb-raising, so that VSO is the SS order with the minimum, i.e., only obligatory movement. But we also assumed Nom-checking movement in Greek may, and by Procrastinate, must be delayed until LF. Crucially, then, the VSO option leaves Spec-IP (or whatever it is in Greek ; say, Spec-MP, as I and others have claimed) free.

Drachman (1989) claimed, against conventional wisdom and the Uniformity Mafia, that Wh-movement may, and therefore must be to Spec-MP, at least for root sentences in Greek. Restated in the present OT context, Greek does not need to (and so may not) violate Minimal structure; MP is enough, e.g., for sentences like (1-2)

(1) *pyon filise i-Maria*? 'Whom did Mary kiss?'
(2) *pyos filise tin-Maria*? 'Who kissed Mary?'

where any further movement would be vacuous.

3.1.5. The Wh-criterion: a radical cure

But of course this proposal faces the principled and serious objection that it violates the Wh-criterion. I now offer a radical cure for this apparently fatal woe. Recall the OT tenet; that variation (whether between individuals or languages) is to be attributed to the 're-ranking of universal (but universally violable) constraints', so that a given constraint either emerges to greater (or even despotic) dominance, or sinks into submissive oblivion. What constraints could be relevant here? Well, obviously the Wh-criterion vs. Minimal-structure.

This idea might be carried out as follows. The Wh-criterion is first split into its two parts, viz., Op-in-Spec, and Filled head, as in Grimshaw (1994). Now look at the cases.

3.1.5.1. For English. In languages like English, there is a conflict-laden constellation (of course topped by the undominated Case-Filter), viz.,

Op-in-Spec (Wh-element in Spec) >>Filled head (by inversion) >>Minimal Structure.

As was said above, Spec-MP cannot be used as an escape hatch, since it is occupied by the Subject (moved for EPP and obligatory SS Case); and CP is created because of the overriding need for a Spec position to host the Operator. Minimal structure is violated, and *trace too, in favour of Op-in-Spec.

3.1.5.2. For Greek. But now suppose that for Greek Minimal Structure is the undominated constraint, thus;

Minimal structure >> Op-in-Spec/Filled Head (a tie)

Then not only is structure conserved, but we also avoid violating *trace for verb-movement; the trade-off this time is violation of (part of) the old Wh-criterion, the now-dominated constraint, Filled Head. No Greek speaker need lay down his life for this!

Compare also the idea of Subject Wh in-situ for English, characterised as avoiding vacuous movement in Chomsky (1986: 48-54), but perhaps now interpretable as the emergence of the default dominance of Minimal clause and *trace.

3.1.6. Graded responses

Of course, this Wh-in-MP version accounts for the blocking when we try to operate Wh-movement and simultaneously front a Subject, Object or Adverb. This was the original

problem data in Drachman (1989). But crucial to the present argumentation are the data on graded responses: Min-clause, Op-in-Spec, and Filled Head are again implicated.
Take now the graded triple (3-5)

(3) **pyon ton-filise i-Maria*? 'Whom did Mary kiss-him?'
(4) *?pyon i-Maria filise* 'Whom did Mary kiss?'
(5) *pyon filise i-Maria* ditto

Now 3) **Pyon ton-filise* is clearly catastrophic, and never subject to variable judgements. And conversely speakers agree that (5) *pyon filise i-Maria* is perfectly well formed. But why is 4) **?pyon i-M filise* bad yet not catastrophic for many speakers?
The gut reaction in this version is that (4) shows 'dialect variation'. But exactly what is it that is varying between the putative 'dialect' grammars? In the present framework we will not allow individual degrees of rigour for different parameters (cf. postulating that Subjacency is only a 'weak' constraint) or even the counting of violation-stars under OT. Rather, we interpret 'weak' as 'down-ranked' , and we thus re-rank our two constraints as the marked, last resort case for Greek[6]. This creates precisely that CP forbidden by an undominated Minimal structure in 1) , with Greek half-way to the English ranking, as in

CP[pyon MP[i-M filise .....

Op-in-Spec >> Minimal Structure >> Filled Head

with Part II of the old Wh-criterion, viz., Filled Head, downranked to inactivity; and the proper output is guaranteed. Clearly, the alternative in (5) above, repeated here as (6)

(6) *Pyon filise i-Maria* 'Whom did Maria kiss?'

with Op-in-Spec, Min-clause and Filled Head all respected, wins hands down -- so that our reversals constitute marked situations. As before, Fill Head is satisfied by a full verb, just as in the German in (7)

(7) *Wen kuesste die Maria*? 'Whom did Maria kiss?

3.2. Support from binding
Under Binding, the Minimal-clause constraint interacts with A/A'-movement. Support the kind of analysis just given comes from 'weak cross-over' facts. Take the example, used in another context in Horrocks 1994), in (8-9)

(8) **ton$_i$-agapay i-mitera tu-Petru$_i$* loves-him the mother-of-Peter
(9) ok/* *i-mitera tu-Petru$_i$ ton$_i$-agapay [t]$_i$* 'the mother-of-Peter loves-him'

(8) is of course a catastrophic sentence in the intended interpretation, since the pronoun c-commands the coindexed NP. What requires explanation is the variable status of (9), rejected by some speakers (X) but acceptable to others (Y). I will assume that there are two constraints involved, viz., Min-clause, and True-Topic, where True Topic implies that Subjects are 'really' topics and thus in an extra-clausal A' position in Greek.

3.2.1. Take first speakers X
These speakers have the constraint ranking

3.3.2. Other non-selected adverbs

Taking the enquiry to other non-selected adverbs. Suppose that such adverbs can indeed be licensed, roughly under 'sister to V-head' conditions, as in Drachman-Klidi (1992). If licensed, they behave like arguments, in that they obligatorily extract to quasi-Operator position, i.e., Spec-MP: and this extraction should and does block in the diagnostic configuration of a fronted Subject (the 'pos' intended here is of course the Manner adverb, and not 'how come?'), as in the pretty bad (18) above. We come back to the question, why this is not a catastrophic sentence, in a moment.

For the moment, compare the analysis of (18)

CP[pos [MP[o-Petros ilthe t

This violates Min-clause & Fill-Head. Now consider the alternative in (19) below, with its CP-analysis, including inversion:

(19) *Pos ilthe o-Petros* ?
CP[pos C[ilthe MP[o-Petros [t

This analysis would (ignoring *trace) violate only Minimal clause, but that is the undominated constraint for Greek, as we have suggested. So the alternative analysis, respecting Min-clause (as we showed, available for Greek though not for English) is preferred. On one interpretation, it does not even violate part two of the Wh-criterion, since Head is indeed filled:

MP [Pos M' [ilthe VP[o-Petros t

Now suppose there is an intervening constitutent, so that the adverb cannot be a sister to the verb? Then, the adverb is simply not licensed in VP. But adverbs can still extract, at least for some informants. We thus suppose that in such cases an adverb is inserted in-situ in the last resort Spec-CP, the position originally reserved for inherent Operators - those with in-situ adverbs, with no movement. This gives us (20-21)

(20) CP[Pos [MP[o-Petros filise tin Maria?
Pos o-Petros filise tin Maria?
'How did Peter kiss Mary?' (e.g., with passion!)

(21) *Pos o-Petros ilthe stin Athina*?
'How did Peter come to Athens?' (e.g., by train)

Of course, such constructions are far from immaculate; they violate at least Minimal clause, Fill-Head, and Stay (twice), and everyone prefers the minimal structure I suggested, as in

MP[Pos Agr[ilthe VP[o-Petros stin Athina

which of course violates none of our constraints except Fill Head (and trivially, *trace for verb-raising).

To conclude here, we revert to the construction in (18) above, repeated here as (22)

(22) ?*Pos o-Petros ilthe?*

This is for some informants not thoroughly bad -- again suggesting that some speakers use the marked possibility, Spec-CP of last resort, of course thereby violate at least the dominant Minimal clause constraint.

3.4. Negation: constraint conflicts

3.4.1. Neg-Scope

It is well known that the morphological imperative cannot be negated in MGk. The contrast between morphological and periphrastic imperatives is clearly seen in (23-4) vs. (25):

(23) *grapse-to*! 'Write it!' and
(24) **min grapse-to*! 'Don't write it!'
(25) *(na) min to-grapsis!*

Now as compared with Argument and Operator projections, checking projections have a questionable status. And we might infer that they should provoke no blocking effects. Then, instead of looking for ways to block V-movement in Neg Imperative (e.g., by assuming Neg is a head and thus blocks V-movement), we will suppose that there is no such blocking.
To support this, however, we will not take the problematic step of denying that Neg is a head. Rather, we will change the focus of attention. Suppose now, with Platzack-Rosengren (1994), we assume (1) that Imperative involves an abstract Imperative-Pronoun (Imp-pro), positioned in D-zero of a projection under Spec-VP, and distinct from the optional vocative addressee pronominal; and (2) that there is a condition such that Neg may NOT have scope over this abstract Imperative-pronoun. Thus for present purposes, under OT, we thus invoke the constraint

*Neg scope over Imp-pron

At first sight this condition seems too strong, for (ceteris paribus) it should block Neg Imperative everywhere. But now suppose Neg could cliticise to some X, thus blocking c-command of Imp-pro by Neg. For English, X might well be the inserted 'do' in a phrase like 'don't go!' For Greek, which has no equivalent to 'do', recall the earlier claim in Drachman (1994b), viz., that Neg-Imp somehow requires Modal 'na'; and we can now see why this is so. The modal particle 'na' functions as a host to the clitic Neg , just like 'do' in English, and with the same desired result.

So we now have a potential c-command-circumvention strategy of cliticisation as in

na-MIN + Imperative

However, in Greek, once na is introduced, its own selection properties are of course activated. Thus the further (and dominant) constraint Finite-Agreement is activated; the resultant clash between the [+finite] selected by 'na' and the [-finite] inherent feature of the imperative verb rejects the construction Na+M-imperative in favour of Na+subjunctive.
Thus it follows that, under OT, it is the tension between two constraints that produces the ban on Neg-Morphological-Imperative. The satisfaction of the anti-command constraint demands Neg-cliticisation; this in turn violates Full Interpretation in English and Greek. We seek to express prohibition, but cannot: and Last Resort adds 'do' in English and subjunctive in Greek.

3.4.2. OT and historical change

Now consider the classical Greek (AGk) situation as in (26-7):

(26) *kai me: vradine* 'don't tarry!'
(27) *me:de epimne:sthe:s eti Troias* 'neither think about Troy any longer!'

Clearly, AGk indeed allowed not only Aorist Subjunctive Neg Imperative, but also Morphological-Neg-Imperative constructions. In our OT model, AGk must have allowed low

ranking of the scope requirement. The outcome is that no cliticisation-insertion would be needed, so Finite Agreement is not violated; instead, the (low-ranked) Scope constraint is itself violated. Formally, the crucial constraints are ranked as follows:

Modern Greek: Finite-Agr >> Neg-Scope >> Full Interpretation
Ancient Greek: Finite-Agr >> Full Interpretation >> Neg-Scope

3.5. On emergence

We claimed that all constraints are Universal, and must be present as potential in all languages, just as the Distinctive Features in phonology are. For a given language, however, some never appear, so that they must be presumed so subordinate (low-ranked) that they never get a chance to operate. That this is not a vacuous (not to say, absurd) assumption is shown by what is called emergence, where in some context a very subordinate constraint is in fact suddenly up-ranked (cf. McCarthy-Prince (1994)). Consider briefly the following candidate cases in Greek and English.

3.5.1. Overt and null subjects

Suppose with Grimshaw (1994) we interpret the Extended Projection Principle (Nom in syntax/LF) as the constraint 'Subject', and the pro-Drop parameter as 'Free Pronoun'.

Then English has: Subject (NOM in syntax) >> Free Pronoun
But Greek has: Free Pronoun >> Subject

Yet an overt Subject (clitic) pronoun does in fact turn up in one small corner of Greek. Consider the sudden appearance of subject clitics in (28-9):

(28) Deictic: *na o-Petros! ~ na-tos!*
(29) Loc/Interrog: *pu ine aftos? ~ pundos?*[9]

The accident with *na-tos* is the existence of the uninflectable deictic verb *na*, which rather supports the notion that at least some post-clitics are phonological rather than syntactic clitics. The case of *pu-ndos* (by contraction from pu ine tos) is different.[10]
Conversely, pro-drop does emerge in English, by upranking, although only where discourse conditions allow Free Pronoun to also dominate Full Interpretation, thus permitting information-loss. Thus contrast the Declarative with the Interrogative prosodic patterns in (30-31)

(30) Declarative: (*I*) *Lost the car keys!* (interpreted as 1sg.)
(31) Interrog: (*YOU) Lost the car keys?* (interpreted as 2sg.)

The alternation of Subject dominant with Free Pronoun dominant of course makes the difference between English and Greek. Cf. Drachman (1975), Grimshaw (1994).

3.5.2. Emergence as a last resort

Here, recall the Clitic-projection in Greek, proposed in Drachman (1994). This projection is required (e.g.) to explain the Tobler-Mussafia effect, viz.,. that a clitic cannot stand clause-initially. This effect is normally inert in standard Greek, where the pronominal clitics are arguably associated with Agrs instead. But the emergence of a nominative-assigning clitic projection may be involved as a last resort to save a construction from Crash by *Case (details in Drachman ibid). Thus, in the construction in (32)

(32) *Pyon i-Maria*$_{Nom}$ *filise* 'Whom did Maria kiss?'

Nominative *i-Maria* can be in Spec-cliticP, a Nom position; but the corresponding Accusative could not, in the thus unacceptable (33)

(33) **Pyos tin-Maria*$_{Acc}$*filise* 'Who kissed Maria?'

4. Coda

4.1. On the sources of variations

Under P&P the variation was given by the values of parameters. Under Minimality, on the other hand, variation lies in the (morphological) properties of F-heads in the lexicon. Sportiche (1993) reduces this to Morpho-Phonemic properties, a theme recurring in Cardinaletti-Starke's (1994) Strength-Deficiency scale. In addition, is there convergence, at least of theories; e.g., does the growing importance of 'last-Resort' violation of hitherto absolute principles bring Minimality a step nearer to OT?

Sportiche (ibid) talks of the 'ranking of strength among these principles', viz

1) Lexical properties, such as that Q is a bound morpheme, are inviolable;
2) ECP is inviolable;
3) Paradigmatic Uniformity -- may be violated to avoid violations of 1) or 2); and
4) Greed may be violated to avoid violating 1), 2), or 3).

But it may be only a lexical coincidence or slip of the pen when Chomsky writes (1994:48) of 'the class of derivations that have to be considered in determining optimality': for indeed we recall the earlier passage (pg. 5 ibid) declaring that 'its derivation must (also) be optimal, satisfying certain natural economy conditions, e.g., conditions of locality of movement. Less economical computations are 'blocked' even if they converge'.

And in fact, Chomsky's aims (pg. 5 ibid) do not concern the properties of the computational system expressed in terms of output conditions -- whether through filters in Chomsky and Lasnik (1977), chain-formation algorithms in Rizzi (1986), or phonology in terms of OT as in Prince & Smolensky (1993).[11]

In turn, Optimality Theory lays the whole weight of variation on the ordering of universal constraints (many of which, as we saw, are all but identical with Minimality principles) -- only that the constraints are all violable. Questions of opacity as such have, on the other hand, hardly been dealt with in OT so far.

4.2. On Economy

We agree with those who, like Cardinaletti and Starke (1994) hold that all economy principles are of the 'minimise alpha' format. But the incorporation of the condition 'up to crash' makes such principles transderivational; 'you know you must stop if you know that going further will trigger ungrammaticality' (footnote p. 38).

The Optimality approach circumvents and trivialises this problem at once. Suppose the output candidate must satisfy 'weak pronominal'. Then if you must violate this, do so minimally: it follows that that output with the weakest pronominal that does not crash is optimal.

Notes

[1] Shortened from the paper read at the Workshop in Greek Syntax, FAS Berlin, Dec. 1994

[2] First, to the opacity problem. In phonology it is clear that, as compared to e.g., Korean Umlaut (Hume 1990), German Umlaut is opaque wrt the original phonetic front-vowel trigger. Yet it proved possible (Fery 1994) to predict a significant productive set of the German

Umlaut cases, viz., those in *-chen*, for both positive and negative sub-cases; and this by moving the synchronic motivation away from derivational history, in fact to the level of output prosodic structure. A parallel stratal treatment might be feasible in the case of syntactic traces, at LF.

Second, to the absoluteness of constraints. Over the years a certain falling off from the grace of absoluteness in classical 'constraints' is observable. Setting aside long-standing distinctions such as weak vs. strong Islands, weak vs. strong Crossover, and the characterisation of Subjacency as a weak constraint as compared (e.g.) to the ECP, -- and noting in passing that even weak ECP violations are countenanced (Chomsky Barriers) -- consider also a) Pollock/Chomsky weak (e.g., Theta-opaque) vs. strong (Theta-transparent) affixes; b) the deployment of Last Resort to ensure convergence in Chomsky 1994, and c) Cardinaletti-Starke (1994) on three degrees of element-strength, correlating with distance of movement.

Nevertheless, Chomsky (1994) shows no interest in characterisations of the properties of the computation system in terms of output conditions. On the contrary, he attributes greater prominence than ever to the derivational approach (ibid, pp.6-7), emphasising its inevitably step by step nature in the face of opacities between input and output strings, in syntax as in phonology.

For syntax, Chomsky notes crucially the opacity created by deletion and even replacement of traces. For example, despite the constraint on vacuous subject-movement in syntax, cases of ECP like the following arise:

*how do you wonder who fixed the car?

where in the syntax, 'how' moves to the inner Spec-CP, then to outer Spec-CP, leaving behind how-t'. At LF, however, 'who' moves to inner Spec-CP, deleting (or now co-occurring with) how-t'. Thus the original trace of 'how' is no longer p-governed, violating the ECP

[3] It is a question whether we also need Align-I: Agr to left of VP (State or Event) - making an I-relation. We will necessarily revert to align (but elsewhere, for lack of space) to deal with directionality in clitic placement.

[4] But there is a problem here. If ECM cases require independently motivated movement for EPP to Spec-Agr_SP before Object raising to matrix, as Iatridou (1994) claims, then EPP applies to simple clauses too. But then, what allows VSO structures to survive Spell-Out? We have assumed NP-Subject can remain in VP, so are these cases to be reinterpreted as cases of NP-subject to Spec-MP (motivated by EPP), followed by Verb-to-Comp? The problems are: 1) what would motivate Verb-to-Comp here? and 2) Negative always precedes the verb (main or auxiliary), for both SVO and VSO configurations; we would have to assume that Neg always syntactically cliticises to its verb.

[5] And note the tension whereby certain fully acceptable constructions will still necessarily contain (lower-ranked) violations: thus, eg.,

a) Wh-in-situ violates Op-in-Spec, but
b) Wh-in-Spec of course violates *trace, and may violate minimal clause as well.

So far as I know, LF movement is so far hardly taken into account in OT.

[6] Last resort is something of a problem, because subject-fronting is not the only way to achieve prominence for that NP.

[7] Compare also Tsimpli 1990, which takes Topics in Greek to be base-generated in CP-adjunction -- though that too is of course an A' position. So for Tsimpli there are only X-type speakers.

[8] Cf. pos2 and the 'how-come?' reading in
Pos o-Petros ilthe 'How did Peter come?'

where pos2 is probably an inherent Wh-word, like yati.

[9] The missing Qns (how/when/ why is he?) would be semantically anomalous.

[10] What about Who/what (is) he/it? These would have Agr-adjustment. And there is the problem of where *na-tos* get its case.

[11] Support for this judgment comes from the fact that the derivational approach is assigned even greater prominence under Chomsky's Minimalism (1994:6-7), emphasising step by step derivation, and pointing to the opacity relations often obtaining between input and output strings in phonology as well as syntax. Successive raising is an example. 'Thus...head-movement meets locality conditions, but several such operations may leave a head separated from its trace by an intervening head, as when N incorporates to V leaving the trace t_N, and the $[_V$V-N] complex then raises to I leaving the trace t_V, so that the chain (N, t_N) at the output level violates the locality condition satisfied by each individual step'. In Chomsky's view a fully derivational approach captures opacity in both syntax and phonology, 'and indeed suggests they should be pervasive, as seems to be the case'.

Chomsky claims we might indeed formulate the desired result in terms of outputs: since the trace is plausibly a copy, we could invoke a record of the original (and purely local) raising, within the intermediate trace. But he holds this is the wrong move, since the relevant chains at LF are (N, t_N) and (V, t_V), and in these the locality relation eliminated by successive raising is not represented. And he concludes that 'the computational system C_{HL} is strictly derivational' and that 'the only output conditions are the bare output conditions determined externally at the interface'.

Bibliography

Bromberger, S. (1986) What We Don't Know When We Don't Know Why, MS. MIT.

Bromberger, S., and M.Halle (1989) 'Why phonology is different', LI 20, 1. pp. 51-70

Cardinaletti, A., and M.Starke (1994) The typology of structural deficiency: on the three grammatical classes, MS. University of Venice and University of Geneva. pp. 1-19.

Chomsky, N. (1986) Barriers, MIT Press

Chomsky, N. (1992) 'A Minimalist Program for Linguistic Theory', MIT WPL. pp.1-71

Chomsky, N. 1994. Bare Phrase Structure., MS. MIT. pp. 1-48.

Chomsky, N., and H.Lasnik (1977) 'Filters and Control', Ling. Inq. 8.3. pp. 425-504.

Drachman, G. (1975) 'The syntax of casual speech', G.Drachman (ed.), Salzburger Beitraege zur Linguistik 1, Akten der 1. Salzburger Fruehlingstagung fuer Linguistik, Verlag G. Narr, Tuebingen. pp. 273-95.

Drachman, G. (1989) 'Raising in a Language without Infinitives', Paper at All-Austria Ling.Meeting, Vienna, Working Papers in Greek Grammar, University of Salzburg. pp.1-15

Drachman, G. and C. Klidi (1993) 'The extended Minimal Structure Hypothesis', Studies in Greek Linguistics, Proceedings of the 13th,. Annual Meeting of the Department of Linguistics, 1992. 371-90. University of Thessaloniki.

Drachman, G. (1994) 'Verb Movement and Minimal Clauses', I.Philippaki-Warburton et al. (eds.) Themes in Greek Linguistics, Papers from the First International Conference on Greek Linguistics, Reading, September 1993, J. Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 45-52.

Drachman, G. (1994a) 'Some properties of clitics', Paper at Inaugural Conference on
Universal Grammar and Typological Variation, FAS Berlin. To appear.

Drachman, G. (1995) 'Projections in Greek: there's no road back to Eden.' Studies in Greek Linguistics, Proceedings of the 15th. Annual Meeting of the Department of Linguistics, 1994. University of Thessaloniki.

Fery, C. (1994) German Umlaut in Optimality Theory, Paper at Inaugural Conference on Universal Grammar and Typological Variation, FAS Berlin. To appear.

Grimshaw, J. (1994) Projections, Heads, and Optimality, July MS. Rutgers University.

Horrocks, G. (1994) 'Subjects and Configurationality: Modern Greek clause structure', Journal of Linguistics 30.

Iatridou, S. (1994) 'To have and Have not: Participial Agreement and the lack thereof', Paper at the Workshop in Greek Syntax. FAS Berlin. To appear.

McCarthy, J.,and A. Prince (1994). The emergence of the unmarked Optimality in Prosodic Morphology, MS. University of Mass. and Rutgers University.

Platzack, C., and I. Rosengren (1994) 'On the Subject of Imperatives: a Minimalist Account of the Imperative Pronoun and Negated Imperatives', Sprache und Pragmatik, Arbeitsberichte 34. Lund. pp. 26-67.

Prince, A., and P. Smolensky (1993) Optimality Theory: Constraint Interaction in Generative Grammar. MS. Rutgers University and University of Colorado at Boulder, pp. 1-23

Sportiche, D. (1993) Sketch of a Reductionist Approach to Syntactic Variation and Dependencies. MS. UCLA.

Tesar, B. (1993) Optimality Theory and Genetic Algorithms. MS.

Tsimpli, I. (1990) 'The clause structure and word order of Modern Greek', UCL Working Papers in Linguistics 2, pp. 226-258.

Fast Speech in a Monostratal Model of Postlexical Phonology*

Ursula Kleinhenz

1 Introduction

Fast speech has largely been ignored by theoretical linguists. The literature is scant. In this paper, however, I will argue that fast speech has some properties phonology has to account for: first, the differences between the formal style (which is the speech level that is taken as the basis of phonological descriptions) and the fast speech level are systematic. Second, the phonology has to deal with is the large amount of variation within fast speech.

The general idea I propose is that the differences in speech rate or style do not result from different rule orderings, but from different *domain* sizes, on which one single block of rules applies. By comparing fast speech phonology to the phonology of (simple) cliticization, I will show that postlexical phonology provides further evidence for the prosodic view of cliticization (see Booij (ms); Lahiri et al. (1990)).

This paper is organized as follows: first, I will give an overview of the data this paper will deal with. Thereafter, I will apply the standard models of the phonology-syntax interface to the German fast speech data and will demonstrate their shortcomings w.r.t. fast speech. Then I will apply the model I will use instead. Finally, I will point out some difficulties that might arise under an Optimality-Theoretical (OT)-analysis concerning constraints on non-surface forms.

1.1 The data

The data I will present in this paper are of two different kinds: The first type is the "standard speech level". It is easy to gather data of this type, since this is the level phonologists usually talk about. Sources on standard speech come from recent phonologial descriptions of German, such as Hall (1992), or the judgments of native speakers.

---

* I would like to thank T.A. Hall and Paul Law for their helpful comments.

The fast speech data I have taken from my own tape-recordings of speakers of both standard German and a variety spoken in the Rhineland, which has additional rules showing the domain boundaries I will deal with.

Also, in these varieties it is easier to distinguish between a single speaker's standard and fast speech utterances. As I observed, speakers were aware of some of their dialectal characterics, which they suppressed when speaking slowly. One was the *g*-spirantization and other was the allomorphy of the neuter pronoun *das* 'that'. One and the same speaker switched from varieties such as those in (1.a) to those in (1.b), when using fast speech.

(1) Alternations between standard speech and fast speech

	(a) standard speech	(b) fast speech
geschichten 'stories'	[geʃɪçtən]	[jeʃɪçtən]
das/ dat 'this' - alternation	[das]	[dat]

The forms in (1.b) cooccur with other properties of fast speech, such as assimilation rules. In (2), some examples of data, this paper will deal with, are given.

(2) Nasal assimilation

(a) Standard speech

i	ii	iii	iv
in Köln	kommt man	kommt Markus	Rennbahn
[ɪnkœln] [ɪŋkœln]	[kɔmtman] [kɔmpman]	[kɔmtmɐkʊos] *[kɔmpmɐkʊos]	[ʀɛnba:n]*[ʀɛmba:n]
in Cologne	come3SG one	come3SG Markus	race+course
in Cologne	*does one come*	*does Markus come?*	*racecourse*

(b) Fast Speech

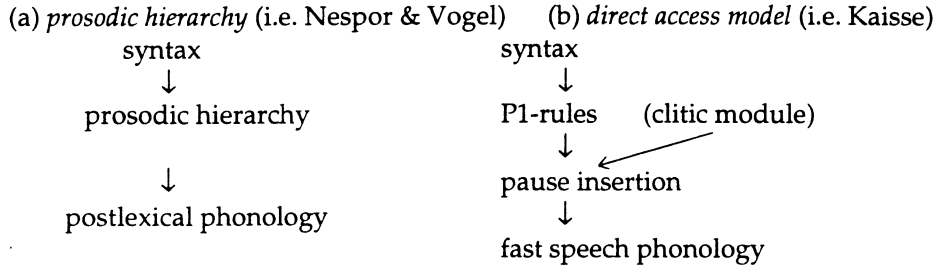
in Köln	kommt man	kommt Markus	Rennbahn
[ɪŋkœln]	[kɔmpman]	[kɔmpmɐkʊos]	[ʀɛmba:n]
in Cologne	come3SG one	come3SG Markus	race+course
in Cologne	*does one come*	*does Markus come?*	*racecourse*

One purpose of this paper is to explain the difference between the standard speech forms in (2.a) and the fast speech forms in (2.b). The forms in (2.i) and (2.ii) are function words. They can always be subject to postlexical assimilation rules. Contrary to that, the forms in (2.iii) and (2.iv) are lexical units, members of compounds, etc. In this paper I will argue that they form a boundary to phonology in standard speech, but not in fast speech. I will introduce a model of postlexical phonology that is based on different input environments for standard speech and fast speech.

1.2 The interface between phonology and syntax

Two models of postlexical phonology coexist for about ten years. One was proposed by Nespor & Vogel (1986), who assume a prosodic hierarchy providing rule domains, and the other was first proposed by Kaisse (1985), which assumes two different kinds of postlexical rules (P1 and P2-rules), among which the P1-type has access to some syntactic information.

(3) Two models of the phonology-syntax interface



The model in (3.b) has to assume two different levels of postlexical phonology in order to provide a separate level of rules that may refer to syntax. One reason for this is that postlexical phonology has to account for the different speech rates. P1-rules alone would only create a fairly elaborated speech level. The P2 rules (also called "fast speech rules") then can refer to nonsyntactic information, because this is overwritten by the intervening pause insertion rule.

The prosodic hierarchy model assumes that phonology does not refer directly to syntax. Instead, an algorithm derives a prosodic hierarchy, considering the syntax whose constituents are "restructured" for the purposes of phonology. This prosodic hierarchy is shown under (4).

(4) prosodic hierarchy (above the word) (Nespor & Vogel 1986)

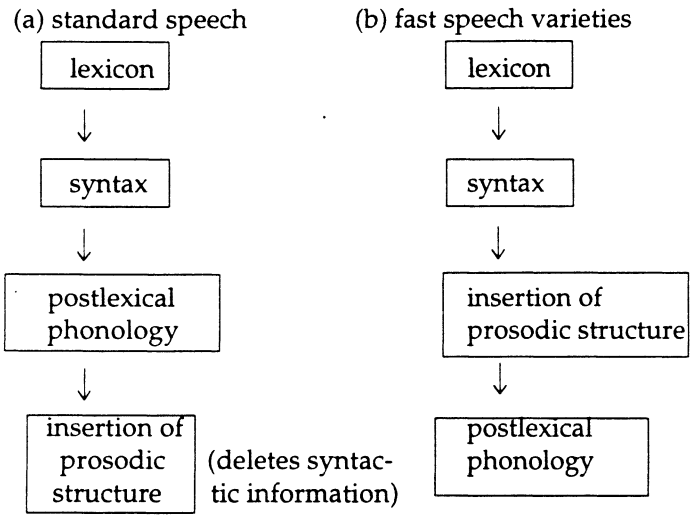
utterance (U)
intonation phrase (I)
phonological phrase (PhP)
clitic group (CG)
phonological word (Pwd)

These two models are not completely incompatible: Some of the properties of the direct syntax model are reflected in the prosodic hierarchy, i.e. the clitic group. Fast speech phonology, however, cannot be derived directly within this model, since the

large rule domains of fast speech cannot be derived by a prosodic algorithm (see section 2.3).

In this paper I will propose a model which combines some assumptions of these two models with some new observations from fast speech, in order to derive postlexical alternations in speech and style without requiring two or more levels.

(5) A monostratal model of postlexical phonology



The crucial idea behind this model is that variation in style and speed is not the result of the application of various rule blocks (as in the direct access model) but a result of the same rules applying on different domains. There are no type 1 and type 2 rules, but "standard" and "fast speech" domains, depending on the timing of the insertion of the prosodic structure. This algorithm marks boundaries prosodically. It can either apply after the postlexical phonology, creating large domains by inserting boundaries only after stressed constituents, or else it applies directly on the syntactic output, thereby considering major syntactic boundaries and transforming them into prosodic ones. In this case, the postlexical phonology can only apply on relatively small domains, creating a standard speech style.

As in the direct-syntax approach, this model provides access to syntax. But this model is more constrained: the syntactic boundaries are only mapped into the phonological representation if necessary, namely in slow speech, where phonological domains correspond to syntactic ones. In fast speech rate, syntax is only indirectly

involved. Instead, bigger units are the domains of phonology, which I here call "superwords". [1]

The latter part of this proposal is in a way contained in papers such as the ones by Hayes & Lahiri (1989) or Kenesei & Vogel (1993), who all propose restructuring algorithms in the prosodic phonology framework.

The most important aspect in (5) is that in infinite number of speech rates and varieties can be created, depending on the following two factors: first, the *timing of pause insertion*.[2] Prosodic boundaries are inserted either before or after the postlexical phonological rules apply.

Note that the motivation behind this is a functional one: elaborated speech has a different function than fast speech: speakers want to make things as clear as possible. This is the level of disambiguation forms and else. Fast speech, however, is the level of economy.

The prediction I make is that only at this formal level of speech speakers need to refer to syntax. The consequence is that only at this level, grammatically conditioned phonological processes are possible, i.e. cliticization. At fast speech, clitic pronouns behave the same way as other forms in the same (unstressed) environment. The remnants of grammatical information are contained in the domain boundaries: they mark the most prominent elements of discourse.

1.3 Characteristics of Fast Speech

I already mentioned that fast speech has largely been ignored in theoretical phonology. One of the reasons for this is that fast speech ignores or overrides the very basic assumptions of phonology and its interfaces, such as the prosodic hierarchy, sentence boundaries, or the coincidence between phonological and morphological boundaries.

However, I will argue that fast speech is part of the grammar in that it is a level with its own systematic characteristics.

---

[1] The term "superword" is more appropriate than the phonological phrase, because rules, which typically refer to the prosodic word apply to these entities as well.

[2] I have little to say about the insertion of prosodic boundaries. According to my observations, a prosodic boundary can (or must) be inserted after a focused constituent (obligatorily), after ellipsis, after coordination reduction, after a topicalized constituent, after enumerations, and so on.

2 Fast Speech is not Anarchy: Examples from Dialects

I discuss the data below in order to support the idea that there is a systematic distinction between two very basic patterns of speech, the formal one, i.e. the speech level to which the phonological descriptions of German refer and the fast speech level, which has attributes that make them an autonomous part of the speech system. Speakers kind of choose between the two levels by providing the postlexical phonology with different kinds of domains.

2.1 Prerequisites: Domains in Standard Speech

A derivation of the type (5.a) accounts for all phenomena in phonology which depend on syntax. In order to show which domains in phonology correspond to syntax and which do not, we first have to define the boundaries of phonological domains. Concerning Standard German, it is well known that there are at least two rules which indicate whether segments are separated by a syllable boundary: final devoicing and the insertion of a glottal stop.[3] In an example such as *Tag* 'day' the underlying *g* is realized as *[k]*, since is syllable-final, whereas in the plural form *Ta.ge* 'days' it remains *[g]*, since it is resyllabified to the onset of the following syllable. Postlexical resyllabification will be the most important indication that no boundaries intervene between the constituents involved, since a boundary would block resyllabification.

German, in contrast to other languages, has few rules of the syntax-sensitive type and they all have to do with cliticization. It has been doubted whether there are clitics in German at all (see Cardinaletti & Starke 1995). Showing that the so-called "simple" clitics[4] can be used in order to illustrate the difference between the two basic levels of speech, we can at the same time show that there is grammatically conditioned phonological reduction.

---

[3] The latter rule is optional: it can only indicate that a syllable boundary is indeed there. If no glottal stop is inserted in a vowel-initial context, especially if this vowel is unstressed, this does not automaltically indicate cliticization.

[4] According to Zwicky's (1977) definition, only simple clitics can be related to their full forms by phonological rules.

2.2 *Cliticization*

In this paper I adopt the prosodic view of cliticization i.e. that a clitic is incorporated prosodically into the host word (see Booij (ms)). Since cliticization is not the main topic of this paper, I will ignore further details about proclitics and enclitics. In each case, the result of cliticization is that the clitic and its host form a single phonological word, which can be tested by applying the well-known criteria of phonological words to these forms: according to Booij, phonological words are the domain of syllabification, i.e., clitics have to resyllabify, if the environment requires so.

A simple example is shown in (6), where the devoicing of the stem-final */d/* is blocked, because the consonant was resyllabified. If there were a syllable-boundary after the stem-final *d* in *find-*, it would have become voiceless (see 2.1).

(6) Resyllabification of a clitic

[dat fɪn.dɪçəʊ'rɪçtɪç]
das **finde ich** auch richtig
this find I also ok
I think it's ok this way

Assuming the model in (5), one would expect that, in standard speech, phonology can treat units differently according to their syntactic properties, while in fast speech it cannot, because the phonological rules apply after the syntax.

Cliticization can be traced back to systematic processes which also apply in fast speech level. The difference is that, in fast speech, only phonological and prosodic information can constrain their application. I will try to show that the difference between speech styles really is a difference between domains of rule application. This difference is induced by a different timing of prosody. Prosody is either applied in the syntax, deleting the syntactic boundaries (thereby making them inaccessible for the phonology) or it is applied after postlexical phonology. In the latter case, phonology is constrained by syntax, which results in a formal speech style. In (2) and (7), this difference is illustrated, using nasal assimilation as an example. Nasal assimilation is another rule which is constrained by boundaries.

A phenomenon typical of the Rhineland area (and some other German dialects in the more southern parts) is the lenition of obstruents in intervocalic position. But as one can see from (7), voicing does not apply across the board to all intervocalic obstruents. Strikingly, it applies to the obstruents in (7.a) but not to the ones in (7.b),

although generally, any phonological interaction would be more likely in environments such as in (7.b) compared to (7.a). In (7.a), the VCV-context is interrupted by a morphological word boundary whereas in (7.b), the intervocalic obstruent is word-internal in one case and between stem and suffix in the other case.

(7) Intervocalic lenition across morphological boundaries

(a) [da'blɪgɪç] nicht durch
da blick ich nicht durch
there see I not through
I don't get this
[das mʊzɐ] noch machen
das muß er noch machen
this must he still do
he still has to do this

(b) Peter *[pedɐ]
Peter
arbeit+en *[ʔɐbaɪdən]
work+3SG

Intervocalic lenition seems to refer only to the boundary between clitic and host word.

2.2.1 Evidence for the Clitic Group?

In this section I will discuss some data from fast speech, which seem to suggest the need for the category clitic group as a rule domain. However, I will show that these data can be explained neither with the clitic group, nor with a prosodic incorporation rule (see Booij ms.) alone. There is evidence for an enriched prosodic representation, such as the prosodic subcategorization (Inkelas 1989), because rules such as in (7) refer to this representation.

Phonological generalizations such as the one in (7), referring explicitly to the boundary between a clitic and its host word, but nowhere else, seem to be evidence for the so-called clitic group (see (4)), which was introduced by Nespor & Vogel (1986: 154). The clitic group consists of a nonclitic phonological word plus adjacent clitic Pwds, depending on their directionality. Still, a look the data in (8) shows that the clitic group is not the environment of this rule, neither does it serve as its domain, nor as its domain boundary.

(8) Further intervocalic environments[5]

(a) [[arbeitet]$_{PW}$ [er]$_{PW}$]$_{CG}$ → arbeite[d]er *but not* *arbei[d]e[d]er
works he works he works he
does he work?

(b) [[eine]$_{CG}$ [Tat]$_{CG}$ [ohne]$_{CG}$ → *eine ta[d]ohne
*eine [d]a[d]ohne
an action without

In (8), I made up some contexts, in order to look at the clitic group as a rule domain. Both examples in (8) are clitic groups containing contexts, which should trigger lenition: intervocalic obstruents within a clitic group ('domain span') as well as at the edge of a clitic group ('domain limit'). In both cases, lenition does not apply. Therefore, we can conclude that lenition marks the boundary between host and clitic. This, however, cannot be expressed in terms of the clitic group.

The only way to describe the lenition rule is to refer to the special kind of boundary involved, in other words, to give up the notion of the clitic group and refer to the prosodic properties of function words instead. This can best be captured in terms of the prosodic subcategorization framework (see Inkelas (1989)) (see also (20)).

(9) Subcategorization frame of an enclitic element (Inkelas 1989: 81)

[[ ]$_{PW}$ __]$_{PW}$

In this model, the idiosyncratic properties of clitics result from their lexical subcategorization frame which requires them to take a host word in order to be prosodically licensed. If we incorporate the idea of prosodic subcategorization into the model of postlexical phonology described here, it would be part of the lexical information, which is only accessible to phonology at the standard speech level. In other words, only in standard speech, clitics are expected to behave differently from other unstressed forms. The lenition rule can thus be stated as follows:

(10) Intervocalic lenition

$$\begin{bmatrix} +\text{obstr} \\ -\text{voice} \end{bmatrix} \rightarrow [+\text{voice}] \; / \; [[[-cons]\_]_{PW} [-\text{cons}]]_{PW}$$

[5] These restrictions refer to all speech levels.

According to (10), intervocalic obstruents con only be voiced if a prosodically weak form follows.[6] This can be seen in (11): only the boldprinted obstruents can be devoiced.

(11) Application of lenition according to (10)

da ['blɪg]$_{PW}$ ɪç]$_{PW}$ nicht durch
da blick ich nicht durch
there see I not through
I don't get this
das [mʊz]$_{PW}$ ɐ]$_{PW}$ noch machen
das muß er noch machen
this must he still do
he still has to do this
[Peter]$_{PW}$
Peter
[arbeit+en]$_{PW}$
work+3SG
[[arbeitet]$_{PW}$ [er]$_{PW}$
works he
does he work?
[[eine]$_{PW}$ [Tat]$_{PW}$ [ohne]$_{PW}$
an action without

2.3 *Domains at Fast Speech*

In this section I will apply the proposals I have made so far to fast speech data.

I will show that fast speech phonology has some properties which make it differ from the standard speech level in a systematic way. Therefore, phonology has to account for this.

Besides, I will show that the treatment of fast speech does not imply a new separate level of grammar. Instead, fast speech is the result of an application of the average postlexical phonology. The difference is that the rules apply in different domains.

Standard speech phonology applies on grammatical domains, its rules primarily have the function of supporting the grammatical boundaries (i.e. disambiguation etc.)

---

[6] The environment [[-cons] $_{PW}$ __ [-cons] $_{PW}$ cannot be tested, since there are no function words beginning with a voiceless obstruent in German.

Fast speech phonology applies in domains which are basically conditioned by the *direction* of speech processing. This means, as much material as possible is moved to the preceding phonological domain.

As a consequence, the diagnostics for fast speech versus standard speech is the directionality of rules. While in standard speech, rules which apply across word boundaries, such as assimilation, may either apply progressively or regressively, depending on the grammatical relationship of the words involved. At fast speech, however, the very same rules apply across all word boundaries, ignoring the grammatical restrictions at slower speech (see (7) to (11))

(12) voicing of obstruents in fast speech
(a)
[dat.zʊgər'i:ɐdətɔx]
das suggeriert es doch
this suggests it well
this does it well suggest
(b)
in EInem Geschoß ist [nɔɣ'aʊʃɛlʊŋ]
in einem Geschoß ist noch Ausstellung
at one floor is also exhibition
at one floor, there is also an exhibition

Another question of interest is why the second t in (12.a) doesn't voice. We will return to that question later (see example (24)).

In (12.b), we can see that the voicing rule is not restricted to plosives: the final *x* in *noch* 'still' becomes *ɣ*, the voiced velar fricative, which does not exist in standard German.

2.3.1 Resyllabification and Domain Boundaries

Ignoring syntactic information as well as categorial information is a well-known property of fast speech. Nespor & Vogel have proposed the Utterance (*U*)-restructuring algorithm to account for phonological rules across large domains

(13) *U*-Restructuring (Nespor & Vogel 1986: 244)

> Adjacent Us may be joined into a single *U* when the basic pragmatic and phonological conditions are met and when there exists a syntactic relation (ellipsis, anaphora) and/ or a positive semantic relation (*and, therefore, because*) between the *U*s in question.

The problem about this algorithm is that it is not restrictive enough to account for the standard speech data which typically respect these boundaries, but at the same time too restrictive to explain what really happens in fast speech. Below, I have listed some cases, where the phonological utterance just not seem to be relevant at all. The domain of obstruent voicing is the word. The cases listed in (14) require a domain even larger than an utterance would be.

(14) Obstruent voicing

diese linie. die ['ge:dybrɪgəns] hier durch diese türme da
geht übrigens
this line it goes by the way here through these towers there
you may have noticed that this line crosses the towers over there

(15) Reyllabification across major syntactic breaks

(a) embedded clauses
das letzte woranɐ sich **[ɐʔinɐdɐ]** is annɐ deutzer brücke
das letzte woran er sich erinnert: er ist an der deutzer brücke
the last thing to what REFL recalls, he is at the bridge of Deutz (place name)
the last thing he remembers was that he was at the bridge of Deutz (when it happened)

(b) relative clauses
ich **fintas** richt nich nach firsisch
ich finde, das riecht nicht nach pfirsisch
I thinks this smells not like peach
to me, this doesn't smell like peach

(c) main clause boundaries
is das ok wenn der hausmeister zwischen dem müll und dem wc **sitzich** mein jetz ma so
is that ok if the caretaker between the garbage and the toilet sits? I mean just...
do you think it is ok that the cartaker's office is located between the garbage and the toilet? I just wanted to know...

(d)
da is die ganze verwaltung [drɪnənzovaɪtɐ]
drin und so weiter
there is the whole administration there inand so on
the whole administration is in there, and so on.

3 Postlexical rules

In this section I will discuss the postlexical rules which derive the fast speech forms. According to the model in (5), the same set of rules should account for all postlexical rules, the different speech rates being a result of different rule domains.

Generally, the postlexical rules apply whenever their description is met. This means that they are blocked by boundaries of any kind. In the next sections, this will be demonstrated.

(16) resyllabification

C.V → .CV

Resyllabification has different outputs at different speech levels, because it only applies to adjacent segments. If a C.V-sequence is separated either by a syntactic or by a prosodic boundary, the consonant cannot be resyllabified.

3.1 Variation

At fast speech, we find lots of varying forms, such as the ones in (17.b).

(17)

(a) standard speech *sind* *aber* 'but they are'	['zɪnt.a:bɐ]	*['zɪn.da:bɐ]
be3SG but		
(b) fast speech	['zɪnt.a:bɐ] / ['zɪn.da:bɐ]	

The varations are due to two factors. First, one and the same utterance can be split up into different domains, depending on speech and style. Second, most postlexical rules are not ordered, so that their application may result in different varieties.

Below, I will introduce the most important rules, which interact in the postlexical component.

(18) Final devoicing (FD) (Hall 1992: 52)

[-son] → [-voice] / __]$_\sigma$

(19) r-vocalization (r-Voc.) (Hall 1992: 57)

$$\begin{bmatrix} +son \\ +cons \\ +cont \end{bmatrix} \rightarrow [\text{-cons}] \; / \; \overset{C}{\underset{}{|}} \; \underset{—}{} \; X$$

optional after short vowel

(19) vocalizes an r in a syllable-coda, obligatorily after a long vowel, optionally after a short one. For example, *fährt er* 'does he drive' is realized as [fe:ɐtɐ].

(20) Obstruent voicing (OV) ("feeding"-relation to (19))

$[\text{-son}] \rightarrow [\text{+sth}] / [[\text{-cons}] \_]_{PW} V]_{PW}$

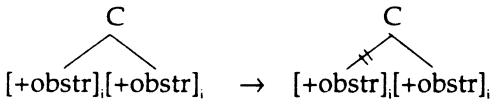
For example, *fährt er* 'does he drive?' reduces to fäh[ɐtɐ] (according to (19)) and then to fäh[ɐdɐ], according to OV. In standard German, this rule only occurs at the boundary between clitic and host word. This can only be expressed in terms of its prosodic subcategorization frame. This formal description rules out all other environments and ensures that the rule only applies in this environment and not on other forms (such as those in (8))

(21) voicing assimilation (VA)

$[\text{-son}] \rightarrow [\text{-sth}] / [\text{-sth}] \_\_$

I.e., fährt der 'does this one drive' assimilates to *fährt[t]er*. Subsequently, the geminate gets reduced by (22): *fähr[t.t]er* → *fähr[t]er*.

(22) Degemination



4 Comparing Standard Speech and Fast Speech

Below, I will demonstrate how these rules may account for the variations at the different speech levels. In the end of the section, I will discuss briefly some cases, which would be a problem for an analysis in an Optimality-Theoretical framework.[7] The postlexical rules apply in different ways, depending on the input they get. At a formal speech level, the input corresponds to the major syntactic breaks and does also consider information contained in the lexical subcategorization of the forms.

In fast speech, the syntactic and the categorial information are overwritten by the insertion of prosodic boundaries. These boundaries create domains in which the postlexical rules apply. The size of these domains again varies according to speech rate. The postlexical rules cannot cross boundaries.

[7] For overviews on OT see the other papers in this volume.

4.1 Examples

The following examples show some of the problematic cases, variation in fast speech as well as counterfeeding relations among rules. In a *d.d* - sequence as in (23), the first (syllable-final) consonant voices optionally, depending on the interaction of degemination and final devoicing.

(23) fast speech variation

	(a) *und.dann* 'and then'		(b) *und.dann* 'and then'
FD	unt.dan	Deg.	un.dan
Deg.	---	FD	---
vc.ass.	unt.tan		
Deg.	[un.tan]		[un.dan]

(23) shows how the nonordered application of rules accounts for the different realizations of the sequence und dann. In (23.a), final devoicng bleeds degmination, whereas in (23.b), it is the other way round. In (24), I contrast two cases, intervocalic *t* in *fährt er* 'does he drive' and the environment *t.d* in *fährt der* does this one drive?', in order to show the interdependencies of rules.

Note that the form *[feɐdɐ] in (24.b) is not a possible realization of *fährt er*.

(24) extrinsic rule ordering

(a) *fährt der*	(b) *fährt der* /fɛ:Rt.deR/		(c) *fährt er*	(d) /fe:Rt eR/			
VA.	fɛ:Rt.teR	*VA*	fɛ:Rt.teR	*VA*	---	r-voc	feɐtɐ
OV	---	*Degem.*	fɛ:R.teR	*OV*	---	OV	feɐdɐ
Degem.	fɛ:R.teR	*r-Voc.*	fɛ:ɐ.tɐ	*r-Voc*	feɐtɐ		
r-Voc.	fɛ:ɐ.tɐ	OV	* fɛ:ɐ.dɐ				

In (24.c) and (24.d) the rule-based explanation for the alternating forms [feɐdɐ] and [feɐtɐ] is given: r-vocalization feeds OV, since it creates the intervocalic environment required. These alternations could easily be explained in an OT-framework as unordered constraints. Forms such as in (24.a) and (24.b), however, would be a problem in an OT-based analysis (see below).

4.2 Problems with an OT-analysis

(24.a) and (24.b) are examples of a problem an OT-analysis would have to deal with. On the surface, the form *[feɐdɐ] should be in free variation with [feɐtɐ]. Since

the constraints blocking the deletion of one of the input consonants are violated anyway, both forms should be equally optimal (which can be seen from the surface forms in (24.c) and (24.d)).

The reason why the form *[feɐdɐ] is still ungrammatical is, I suppose, due to a constraint, which operates holds the derivation on the non-surface form [feɐt.tɐ], the *Linking Constraint*.

(25) Linking Constraint (Hayes 1986)

Association lines are interpreted as exhaustive

In order to explain why data such as the ones in (24)(b) do not exist, we must say that the linking constraint blocks them at one intermediate level of the derivation. After voicing assimilation, the two consonants in consideration form a linked structure. Thereafter, they cannot be input to rules treating them as single consonants, such as obstruent voicing.

The possible alternative explanations, such as extrinsic rule ordering, could not be explained in terms of OT, either. I conclude that, at the fast speech level, we have to deal with a certain amount of derivation.

References

Booij, G. ms. Cliticization as Prosodic Intregration: the Case of Dutch. To appear in: *The Linguistic* Review 1996.

Cardinaletti, Anna & Michal Starke. 1994. The Typology of Structural Deficiency. On the Three Grammatical Classes. FAS Working Papers In Linguistics 1, 1-56.

Hall, T.A. 1992. Syllable Structure and Syllable-Related Processes in German. Tübingen: Niemeyer.

Hayes, B. 1986. Inalterability in CV Phonology. Language 62: 321-350.

Kaisse, E. 1985. *Connected Speech. The Interaction of Phonology and Syntax*. London: Academic Press.

Kaisse, E. 1990. *Toward a Typology of Postlexical Rules*. In: Sharon Inkelas & Draga Zec (ed.): The Phonology-Syntax Connection. Chicago: The University of Chicago Press. 127-145.

Lahiri, A., A. Jongman & J. Sereno (1990). The Pronominal Clitic [dər] in Dutch: a Theoretical and Experimental Approach. *Yearbook of Morphology* 1993, 79-154.

Nespor, M. & I. Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Vogel, I. 1990. *The Clitic Group in Prosodic Phonology*. In: J. Mascaró & M. Nespor (ed.): Grammar in Progress. GLOW Essays for Henk van Riemsdijk. Dordrecht: Foris.

Zwicky, A. 1977. *On Clitics*. Bloomington, Indiana: Indiany University Linguistics Club.

Vowel weakening and vowel deletion in German*
A constrained-based analysis

Sylvia C. Löhken
FAS, Berlin

0. Introduction

During the Middle Ages, full vowels in German were weakened to schwa in most unstressed positions. In many cases, these vowels were subsequently deleted in a syncope process. An example is given in (1).

(1) OHG[1] *ˊêristo* > MHG *ˆêr(ə)stə*[2] > NHG *érstə*

I will show that both sound changes (i.e. the weakening of full vowels to schwa and the subsequent deletion of schwa) were caused by an interaction between prosodic and segmental criteria. Within the constraint-based approach of Optimality Theory (McCarthy/Prince 1993a, Prince/Smolensky 1993), these interactions can be described in a precise way. This description is the first aim of this study.

The second goal relates to the theoretical framework. Like most recent phonological theories, Optimality Theory was developed based on synchronic data only, and hence case studies in this framework relating to sound change are very rare (e.g. Zubritskaya 1994). However, it can be shown that the description of diachronic developments sheds a new light on several aspects of the theory itself. First, it will be shown that constraint reranking can serve as a descriptive base for variation and sound change data. Second, sound change also causes an alteration in the input that are not motivated grammatically but rather economically and thus relate to Lexicon Optimization (Prince/Smolensky 1993:192-196). Faithfulness relations between inputs and optimal candidates change because of the reranking of constraints. In this context, sound change can be interpreted as a lexicon optimization process in which inputs are changed according to modifications in faithfulness relations.

In the first part of this paper, the data will be presented. The following section is an introduction into the basic assumptions of Optimality Theory. Section 3 contains an analysis referring to segmental and syllable-related constraints whereas section 4 focuses on the prosodic context in which weakening and deletion processes occurred. In the last section I will sketch the interaction of segmental and prosodic conditions that favoured the given sound change phenomena.

---

* I would like to thank Wiebke Brockhaus, Nanna Fuhrhop, T. Alan Hall, Ursula Kleinhenz, Mark Verhijde and Wolfgang Wurzel for helpful comments and discussions.

[1] The following abbreviations will be used: OHG (Old High German), MHG (Middle High German), NHG (New High German).

[2] The parentheses around schwa indicate that the deletion of this vowel was optional in this stage.

1. Weakening and deletion of vowels in OHG and MHG

In OHG and MHG, vowels in unstressed syllables were very regularly weakened to schwa. Weakening did not take place in some derivational suffixes if these suffixes could bear secondary accent.[3] The words in (2) show that the morphological domains were not crucial to both weakening and syncope phenomena: Both processes took place in prefixes (2a) as well as in suffixes (2b) and stems (2c). As in many OHG and MHG texts, the circumflex is used to mark vowel length.

(2)

a.	**OHG**	**MHG**	**NHG**	
	gil´îhnissa	*gəl´îchnissə*	*Gléichnis*	'parable'
	ginâda	*g(ə)n´âdə*	*Gnádə*	'mercy'
b.	**OHG**	**MHG**	**NHG**	
	góugalâri	*góukəlærə*	*Gáuklər*	'clown'
	óffanunga	*óffənungə*	*´Öffnung*	'opening'
	áhsalôm	*áhsəl(ə)n*	*Áchsəln*	'shoulders', dat.pl.
c.	**OHG**	**MHG**	**NHG**	
	´êristo	*´êr(ə)stə*	*érstə*	'first'
	giméinida	*gəméinde*	*Gəméində*	'community'
	h´ulisa	*h´ülsə*	*H´ülsə*	'husk'
d.	**OHG**	**MHG**	**NHG**	
	hémidi	*hém(ə)də*	*Hemd*	'shirt'
	héngist	*héng(ə)st*	*Hengst*	'stallion'
	hérbist	*hérb(ə)st*	*Herbst*	'autumn'
	mérzo	*mérz(ə)*	*März*	'March'
	óbaʒ	*ób(ə)ʒ*	*Obst*	'fruit'

The words in (2d) are not subject to the following analyses. Relating to prosodic structure, they are special because the NHG words are monosyllabic after vowel deletion has taken place whereas syncope caused disyllabic foot structures in all the other words in (2). In addition to prosodic and segmental conditions, the data in (2d) also require reference to morphological markedness conditions which are not considered in this study.[4]

Vowel weakening in unstressed syllables was favoured by several conditions. As the examples in (2) show, main stress in all three stages is usually on the first stem syllable. There are only some very rare exceptions, some of which are listed in (3).

(3)	**OHG**	**MHG**	**NHG**	
	fórhana	*fórhel*	*Forélləə*	'trout'
	hóluntar	*hólundər*	*Holúndər*	'elder'
	hórnaʒ	*hórniʒ*	*Horníssə*	'hornet'
	lébêntîg	*lébəndic*	*lebéndig*	'alive''

[3] Such suffixes are *-ig*, *-in*, *-lich*, *-ung*, etc.

[4] See Löhken (to appear) for an analysis of the data in (2d).

In MHG, main stress was still on the first stem syllable (Paul 1969:24). Subsequently it shifted to the second stem syllable. Apart form these exceptions, however, the position of main stress in native German words is very regular.[5] This regularity contributed to the weakening of unstressed syllables.

Secondly, German is an accent-counting language, meaning that the distance between stressed syllables tends to be equal. In accent-counting languages, vowel weakening is particularly frequent (Wurzel 1994:50ff.). However, not all weakened vowels were deleted. Factors favouring syncope are listed in (4).

(4) **Factors favouring syncope in MHG**
- three or more syllables per prosodic word
- vowel position left or right from *n* or *l*

The data in (2) mostly meet these factors. Note that these are only general tendencies but not necessary conditions for vowel deletion. Especially the consonantal environment has to be interpreted only as a statistical but not as a decisive factor. However, both factors show that prosodic as well as segmental conditions should be considered with respect to the given data.

The rules in (5) formalize the weakening and syncope processes that took place in German. Since schwa does not have any terminal articulator features (Kenstowicz 1994a:159), weakening means that full vowels in unstressed syllables lost their place specification (5a). Since this affected all kinds of vowels, reference to place in general is sufficient and need not be specified. Syncope (2b) is a process in which a whole segment is removed from the syllable structure. In the present cases, the word had to be resyllabified because a nucleus was deleted.

(5) a. weakening b. syncope

(5) illustrates that weakening and syncope are subject to segmental and syllable-related conditions. These conditions can be formulated in terms of constraints. The next section introduces the theoretical framework.

2. Optimality Theory

During the last few years, phonological theory has changed its focus. Instead of derivational approaches and rule-governed processes, representations have become the basic instrument of structural descriptions. One of these non-derivational models is Optimality Theory.

[5] See Eis (1950:20f.), Paul (1969:25f.), and von Kienle (1969:16f.) for differences between nominal and verbal compounds since the Germanic period.

Within this theory, the evaluation of surface structures is essential. A morphological chain consisting of one or more lexical entries serves as an input. This chain is assigned phonological structure (such as syllables, feet, etc.). This assignment is formalized by the function GEN (6a). Via different ways of structure assignment, GEN generates a set of candidates.

(6) a. GEN (input) → $\{cand_1, cand_2, cand_3, ..., cand_n\}$
b. EVAL ($\{cand_1, cand_2, ..., cand_n\}$) = output

The evaluation function EVAL accounts for the selection of the candidate with the optimal structure. Evaluation parameters are universally valid constraints. The functions GEN, EVAL, and the constraint inventory form the three modules of a grammar. Within a language, constraints are ranked in a specific way. Due to this ranking, the degree of well-formedness of single candidates can be measured. Constraints refer to all areas of phonological structure, such as syllable structure, foot structure, correspondence between affixes and syllables, epenthesis, deletion, etc.

The importance of a constraint in a language depends on whether the constraint is active in this language, and - if this is so -, how it is ranked in relation to other constraints. Constraints can be violated. The higher a constraint is within the ranking, the more selective is a violation. A violation of a lower ranked constraint can be accepted to avoid a violation of more important constraints.

For example, the constraint ONSET ("Syllables have onsets") has a high ranking position in German. Therefore, epenthesis of additional segments can take place to avoid a violation against ONSET, although epenthesis causes a violation of the constraint FILL ("Syllable structures are filled with underlying segmental material"): FILL is dominated by ONSET (7a). The tableau in (7b) shows how the optimal candidate is evaluated.

(7)
a. ONSET >> FILL
b.

Input: *teatər*	ONSET	FILL
te.a.tər	*!	
☞ *te.ʔa.tər*		*

Explanatory note: The dot (.) indicates a syllable boundary. The candidates which are generated by GEN are listed underneath the input. Potentially, the number of candidates is infinite. It is for the sake of transparency that the list contains only two candidates. The hand (☞) points to the optimal candidate. The asterix (*) indicates a constraint violation. The exclamation mark is placed at the right hand side of a violation that is crucial for the ruling out of a candidate. Shaded fields are not relevant for the selection.

In the course of GEN, candidates can be assigned both too much or not enough phonological structure. In the first case ('overparsing'), the segmental chain is parsed into more prosodic constituents than necessary, or additional segments are inserted and thus cause FILL violations. Underrepresentation of phonological structure ('underparsing') occurs if segments or prosodic constituents (8) are not parsed into constituents of the next level of the prosodic hierarchy.

(8) **Prosodic Hierarchy** (Nespor/Vogel 1986)

prosodic word
foot
syllable
subsyllabic constituents

Unparsed features are excluded from the phonetic surface of the respective segment. Unparsed segments belong to the candidate but are similarly not phonetically realized. The PARSE-constraints that will be needed in the following analysis are listed in (9).

(9) a. **PARSE-FEATURE** Features are parsed into segments.
specification: **PARSE-PLACE** Place features are parsed into segments.
b. **PARSE-SEGMENT** Segments are parsed into syllables.
c. **PARSE-SYLLABLE** Syllables are parsed into feet.

PARSE-PLACE is one of the possible specifications of PARSE-FEATURE.[6] As I show below in section 3, PARSE-PLACE is the constraint that accounts for vowel weakening. Both vowel weakening and vowel deletion violate one of the PARSE constraints: In the first case, PARSE-PLACE is violated (10a), whereas vowel deletion can be described as a violation of PARSE-SEGMENT. In the latter case, a whole vowel is not parsed into syllable structure (10b).

(10) a. OHG *gina:da* > MHG *gəna:də*
b. MHG *gəna:də* > NHG *gna:də* 'mercy'

A violation of PARSE-SYLLABLE (9c) occurs if a syllable is not licensed by a metrical foot. In contrast to PARSE-PLACE and PARSE-SEGMENT, PARSE-SYLLABLE does not influence the phonetic realization of the unparsed syllable. German prosody allows unparsed syllables in certain contexts. Vowel weakening could arise either in non-heads or in unparsed syllables.

3. An analysis relating to segments and syllable structure

A first analysis of the data given in (2) can be presented by focussing on constraints that belong to the PARSE family. Since all phonological features are supposed to be contained in the input, the non-specification of a feature implies a PARSE violation.[7]

---

[6] Other possible specifications are PARSE-[voice], PARSE-[nasal], etc.

[7] However, the assumption that features are specified in the input does not cause a change in the candidate evaluation. If certain features had to be assigned to candidates (e.g. default values in underspecification theory), this would cause violations of another FAITHFULNESS constraint, namely FILL-F ("Features are underlyingly

(11) $*\sigma_w$
|
V
|
PLACE

Vowels in unstressed syllables do not have a place specification.

The PARSE constraint in (11) (Kiparsky 1994) relates to both prosodic and segmental structure. In many languages, vowels in non-prominent metrical positions have less phonetic specification, whereas vowels in prominent syllables contain relatively more features and/or a higher degree of scalar values, such as sonority (Trubetzkoy 1939, Kiparsky 1994).

Within the framework of OT, sound change can be described by a reranking of constraints: Due to changes in the hierarchy, the optimal candidate in ranking 1 can differ from the one in ranking 2 (12).

(12) Ranking 1 (t_1) ---> $cand_{opt} = x$
Ranking 2 (t_2) ---> $cand_{opt} = y$

However, this concept is idealized: In the course of language change, there are also variation phenomena which have to be accounted for. Candidate x and candidate y may both be acceptable variants during a certain period. Since Ranking 1 and Ranking 2 are two different grammars, variation can be accounted for as grammar competition.[8]

I will show that the variation phase can also be described as a reranking that has not yet been fixed because the dominance relations are not yet established. In that respect sound change is not reduced to a switch from ranking 1 to ranking 2, but is rather seen as a gradient shift of one or more constraints within a given hierarchy. During this development, dominance relations weaken, are neutralized and rearranged, and eventually new dominance relations emerge. The analysis will show that this concept can account for the empirical facts in a more adequate way.

The example in (1) has the input *êrist+o* in OHG. The tableau in (13) shows the evaluation of the optimal candidate. It contains only candidates that can be evaluated with the constraints given in (9) and (11). The prosodic structure is not considered yet so that syllable and feet edges are assumed without evaluation.[9] In section 4, prosodic conditions will be accounted for separately.

---

specified"). Aside from the content of underlyingly featural specification, the optimal candidate remains the same (cf. Itô/Mester/Padgett 1994).

[8] In historical linguistics, the notion of competition between parts of the grammar is well-known. For instance, Zubritskaya (1994:335) relates the notion of grammar competition in terms of markedness to the wave model which was developed by Bailey (1973).

[9] Round brackets mark edges of metrical feet. Square brackets enclose unparsed segments or syllables.

(13) **OHG**

Input: *êrist+o*	PARSE-SEG	PARSE-PL	*V(σ_w) | PL	PARSE-SYLL
1. ☞ *(ˊê.ri)<sto>*			**	*
2. *(ˊê.rə)<stə>*		*!*		*
3. *(ˊêr<i>sto)*	*!		*	

In OHG, PARSE-PLACE dominated the constraint militating against place features in the nuclei of unstressed syllables. Therefore, place features were assigned in the optimal candidate. PARSE-SYLLABLE is violated by two candidates. However, this constraint is ranked so low that it does not have any influence on the selection of the optimal candidate. PARSE-SEGMENT dominates all the other constraints: Syncope (as in candidate 3) was prevented whereas it could well occur at later stages after vowel weakening had taken place. In addition, a chronological ordering has to be assumed because PARSE-PLACE had to get a lower position within the ranking before PARSE-SEGMENT could also be violated and thus cause syncope.

The weakening of [o] to [ə] in MHG is evidence that *-sto* was left unparsed and not parsed into a degenerate foot which would have had secondary stress (cf. section 4): Such stress would have prevented the later weakening process since vowels in stressed syllables are supposed to keep their place specification via PARSE-PLACE.

The data in (2) show that in MHG the ranking had changed: The assignment of place features was now disfavoured in the nuclei of unstressed syllables so that the candidates which contained weakened vowels were better than the candidate with full place specifications in all vowels. This development can be accounted for by a reranking within the hierarchy: The third constraint, *V(σ_w), gained importance and hence changed places with PARSE-PLACE. The tableau in (14) shows the new evaluation.

(14) **MHG** (preliminary)

Input: *êrist+o*	PARSE-SEG	*V(σ_w) | PL	PARSE-PL	PARSE-SYLL
1. *(ˊê.ri)<sto>*		*!*		*
2. ☞ *(ˊê.rə)<stə>*			**	*
3. *(ˊêr<ə>stə)*	*!		**	

However, the ranking in (14) is still incomplete because PARSE-SEGMENT also changed its position: In MHG, syncope became so common that PARSE-SEGMENT, which is violated by

vowel deletion, obviously lost influence in the evaluation process and therefore lost its position in the ranking. The example in (1) shows that the variant *erste* instead of *erəstə* was already possible. Candidate 3, which contained an unparsed segment, met the constraint PARSE-SYLLABLE because the remaining syllables could be parsed into one binary foot. Since both variants were of equal importance in MHG, it is empirically inadequate to assume a ranking between PARSE-SYLLABLE and PARSE-SEGMENT. Since there is no further evidence for direct conflicts between both constraints, they can be equally ranked. The absence of ranking is indicated by the dotted line in (15).

(15) **MHG**

Input: *êrəst+ə*	*V(σ_w) | PL	PARSE-PL	PARSE-SYLL	PARSE-SEG
1. *('ê.ri)<sto>*	*!*		*	
2. ☞ *('ê.rə)<stə>*			*	
3. ☞ *(ê'r<ə>stə)*				*

The MHG tableau in (15) contains a different input in which the place features in unstressed vowels are no longer specified. The new input accounts for the fact that vowel weakening in unstressed positions turned out to be a systematic phenomenon in MHG. However, it is essential that the different input does not have any influence on the candidate selection. Even if the input contained place specifications in all vowels, candidate 3 would win.[10] However, in (15) PARSE-PLACE is not violated by the candidates 2 and 3 any more.

The quality of the input therefore does not depend on grammatical criteria but rather on learnability aspects: The less difference there is between input and output, the less information has to be learnt during the acquisition process and the more harmonic is the input-output-relation (Lexicon Optimization; cf. Prince/Smolensky 1993:192ff., Itô/Mester/Padgett 1994:20ff.). Implications for the description of sound change will be discussed in section 4. The tableau in (16) shows the evaluation in NHG.[11]

[10] Candidate 1 would be ruled out because of the dominating constraint. Candidates 2 and 3 would both violate PARSE-PLACE. The result of EVAL would be the same. Note that candidate 1 also violates FILL because non-underlying place specifications are added. Since FILL would not influence the candidate selection, it is not introduced here.

[11] In the next section it will become clear that the input changed not only in MHG (*êrəstə*) but also in NHG (*erstə*). For the sake of transparency I do not introduce constraints which will be needed in a prosodic context, hence the MHG input is still preserved in the segment-related analysis. The most harmonic input will be introduced in section 4.

(16) **NHG**

Input: *êrəst+ə*	*V(σ_w) | PL	PARSE-PL	PARSE-SYLL	PARSE-SEG
1. *(ˊê.ri)<sto>*	*!*		*	
2. *(ˊê.rə)<stə>*			*!	
3. ☞ *(êˊr<ə>stə)*				*

The difference between the MHG and the NHG tableau consists in a ranking relation that emerged between PARSE-SEGMENT and PARSE-SYLLABLE. Candidate 2 and candidate 3 are not longer of equal importance: The variant with an unparsed vowel is the only one that is accepted in NHG whereas candidate 2, containing an unparsed syllable, is not well-formed. Therefore, PARSE-SYLLABLE now dominates PARSE-SEGMENT.

4. The relevance of stress-related constraints in weakening and deletion processes

Up to this point, the prosodic structure was not part of the evaluation. However, the rhythmic structure played an important role in both weakening and deletion processes so that they should be considered in the candidate selection.

Accent is defined as the relative prominence of a syllable within a prosodic word. Suffixes, such as *-in* and *-ung* in (17a-c), are not prosodic words and thus do not form a domain for the assignment of metrical feet.

(17) **a.** *[(Gˊöttin)$_\varphi$]$_\omega$* 'goddess'
[(Máhnung)$_\varphi$]$_\omega$ 'warning'

b. *[(Léhrə)$_\varphi$ (rìn)$_\varphi$]$_\omega$* 'teacher', f.
[(Wándə)$_\varphi$ (rùng)$_\varphi$]$_\omega$ 'hike'

c. *[(ˊÖffnung)$_\varphi$]$_\omega$* 'opening', sg.
[(ˊÖff)(nùng ən)$_\varphi$]$_\omega$ 'openings', pl.

The data in (17) illustrate that a suffix may be stressed or unstressed, depending on its metrical environment. (An exception are schwa syllables, which can never be stressed in any context.) For example, *-in* is unstressed in *Göttin* (17a), but bears secondary accent in *Lehrerin* in (17b). Stress can also vary within a paradigm as illustrated in (17c). Two suffixes can form a foot containing secondary accent if the first suffix does not contain schwa.

Constraints can refer to the edges of phonological and morphological categories.[12] In German, the first syllable of a native stem is always stressed.[13] Therefore, the left edge of a stem has to be simultaneously the left edge of a trochee or a degenerate foot (18).

(18) **ALIGN-L (Stem, L, Foot, L)**
Every stem begins at the left edge of a foot.

German has the following foot structures:

(19) **Inventory of foot structures in German**

a.	$(\acute{}x\ .)_{\varphi}$	*(H´än.də)*	'hands'
b.	$(\grave{}x\ .)_{\varphi}$	*(Wán.də)(**rùn.gən**)*	'hikes'
c.	$(\acute{}x)_{\varphi}$	*(Hand)*	'hand'
d.	$(\grave{}x)_{\varphi}$	*(Wán.də)(**rùng**)*	'hike'

I assume that the maximal foot is binary. Since stress in German native words is not quantity-sensitive,[14] the unmarked foot structure is a syllabic trochee, i.e. a weak syllable following a strong (i.e. accented) one so that the head is on the left as in (19a-b).[15] Degenerate feet as in (19cd) have the logically smallest possible size. In quantity-insensitive systems they consist of one syllable only (Hayes 1995:75) and are marked cases. Accented stems carry primary accent in derived words whereas accented suffixes can only be assigned secondary accent.[16]

The regularities of foot parsing in native German words are summarized in (20).

(20) **Metrical rules for German**

1. The unmarked foot is a syllabic trochee.
2. The first stem syllable is always stressed.
3. The accented stem prevails metrically over accented suffixes.

The constraints in (21) express these stress facts:[17]

(21) **FOOT-MAX**
Feet may be no larger than two syllables.
FOOT-MIN
Feet may be no smaller than two syllables.
RHYTHMTYPE=TROCHEE
Feet are trochees.
***HEADHEAD**
Heads of feet must not be adjacent (Kager 1994).

The maximal foot is binary (Hayes 1995). According to Everett (1994), the constraint FOOTBINARITY ("Feet are binary under moraic or syllabic analysis"), which was proposed

---

[12] McCarthy/Prince (1993b) show that such alignments should be part of universal grammar. This is expressed in the general constraint format of Generalized Alignment.
[13] Exceptions are listed in (3).
[14] See Giegerich (1985) for a survey on quantity-sensitive stress assignment in non-native words.
[15] See Hayes (1995) for typologically motivated generalizations between quantity-sensitivity and foot structure.
[16] The suffix *-eréi*, which carries primary accent, is not a counterexample because it was borrowed from French during the MHG period (Henzen 1957:185).
[17] Within the given context, main stress and secondary stress need not be distinguished. It is only important to locate any kind of stress that could prevent vowel weakening and deletion.

by McCarthy/Prince (1993a) and Prince/Smolensky (1993), should rather be split into two constraints, FOOT-MIN and FOOT-MAX. As Everett points out, FOOTBINARITY alone can not account for the difference between ternary and degenerate feet: Both would violate FOOTBINARITY although they usually are not equally evaluated in a given language: Degenerate feet are possible in many cases whereas ternary feet are prohibited. By assuming two separate constraints, this distinction can be derived from the ranking of FOOT-MIN, which is dominated by FOOT-MAX. An illustration is given in (22).

(22)

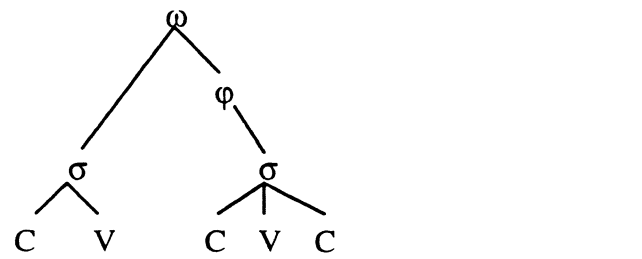
a. FOOT-MAX >> FOOT-MIN

b.

Input: *kolibri*	FOOT-MAX	FOOT-MIN
(kó.li.bri)	*!	
☞ *(kó.li)(brì)*		*

Weak syllables cannot be parsed in certain contexts and are therefore dominated directly by a prosodic word as in (23):

(23)



The prefix [gə] in (9) cannot be stressed and is left unparsed. Since the domain of footing is the prosodic word and the prefix is a prosodic domain on its own (prosodic word or appendix, see Booij 1985: 154), it cannot be incorporated into the foot at its right. The constraint PARSE-SYLLABLE, which was introduced in section 3, is violated in those cases:

In addition to the constraints listed above (including PARSE-SYLLABLE and PARSE-SEGMENT), the present analysis also requires the constraints in (24):

(24) a. **NOCODA**
Syllables don't have codas.
***H/ə**
The head of a foot does not contain schwa.

NOCODA is a well-known syllable structure constraint (see e.g. McCarthy/Prince 1993a: 10). The last constraint in (24) is based on Kenstowicz (1994b), who assumes a ranking expressing the degree of markedness of nuclei in stressed syllables. This markedness is inversely proportional to the degree of sonority of the respective vowel (cf. Prince/Smolensky 1993:127-167). In other words, the more unmarked a vowel is as a syllable peak, the more

unmarked it is as a nucleus within a head. The hierarchy underlying this hypothesis is presented in (25).[18]

(25) **Markedness of heads in metrical feet** (Kenstowicz 1994b)
*H/ə >> *H/i,u >> *H/e,o >> *H/a

*H/ə says that [ə] is unstressable. In German, this constraint is undominated because schwa is not stressed under any circumstances. By contrast, constraints such as *H/a are ranked very low. A syllable head must have a nucleus, and the vowel *a* is the most unmarked option. This constraint is only included here for reasons of completeness because all vowels should be specified regarding their degree of markedness in heads of metrical feet.

However, it does not make much sense to consider a constraint like *H/a to be relevant for the EVAL function: A violation is never relevant for the selection of a candidate. Therefore, only the constraint rejecting schwa in a stressed syllable (*H/ə) should be considered. The constraints required in the present analysis are listed in (26):

(26)
PARSE-SEGMENT; PARSE-SYLLABLE
ALIGN-L
FOOT-MAX; FOOT-MIN
RHTP=T
*HEADHEAD; *H/ə
NOCODA

For the sake of descriptive simplicity, I will assume syllable structures in the following tableaux. The candidates are compared with respect to their metrical structure only.
The tableau in (27) shows the OHG ranking.

(27) **OHG**

Input: *êrist+o*	ALIGN-L	FT-MAX	RhTp=T	*HDHD	PARSE-SEG	FT-MIN	PARSE-SYLL	NOCODA
1. *<ê>(rí.sto)*	*!						*	
2. ☞ *(ˊê.ri)<sto>*							*	
3. *(ˊê.ri)(stò)*			*!			*		
4. *(ˊê)(ri.stò)*			*!*			*		
5. *(ˊê)(rì.sto)*			*!	*		*		
6. *(ˊê)<ri>(stò)*			*!*			**	*	
7. *(ˊêr<i>sto)*					*!			*
8. *(ˊê.ri.sto)*		*!	*					

The structure of candidate 7 implies a structural reanalysis due to syncope: Since the second vowel is deleted, a binary foot can be assigned.

[18] Kenstowicz uses the term 'peak' instead of 'head'.

The constraint ranking in (27) is motivated as follows:

ALIGN-L and FOOT-MAX are both undominated, and there is also no evidence for an internal ranking between them. The rhythm constraint is ranked lower than FT-MAX. This can be justified by considering the scope of RhTp=T: Degenerate and ternary feet both violate the trochaic rhythm constraint. Since degenerate feet are not as bad as ternary feet, the rhythm constraint is ranked below FT-MAX. This ranking accounts for why candidate 8, for instance, is worse than candidate 3. Such a gradation could not be expressed by the violation of FT-MAX in 8 if FT-MAX were not ranked higher than the rhythm constraint.

The rhythm constraint dominates *HDHD because a violation of *HEADHEAD within a prosodic word only occurred if a trochaic foot could be assigned in return. In OHG, stress clash could occur after a long vowel in favour of a binary foot as the example in (28) shows:

(28) OHG *(âh)(tùn.ga)* 'persecution', nom.sing.

RhTp=T dominates PARSE-SEG because variants with vowel deletions in favour of trochaic rhythm can already be observed in OHG. See (29) for examples.

(29) **Syncope in OHG**

ab(a)	'from'
als(ə)	'when'
dritt(i)o	'third'
zunt(a)ra	'tinder'

However, this ranking was not yet fully established: Since the syncopized forms were only optional variants, the inverse ranking PARSE-SEG >> RhTp=T was also possible.

PARSE-SEG and PARSE-SYLL are dominated by *HDHD: Generally, the parsing of neither a syllable nor a segment occurs if a stress clash is the consequence. The only exception is a violation of PARSE-SEG if a trochee can emerge (see (29)). In this case, the violation of RhTp=T can be prevented whereas only the lower ranked PARSE constraint is violated.

PARSE-SYLL is dominated by PARSE-SEG although this ranking is not very stable because syncope also took place in OHG - in favour of the parsing of an additional syllable into a trochaic foot (29). On the other hand, syncope is prevented in cases like *éristo* because, all the other constraints being equally ranked, PARSE-SEGMENT would be violated. Therefore, candidate 2 is more optimal than candidate 7 in OHG.

*HDHD dominates FT-MIN because stress clash was avoided if only one syllable could be parsed into a degenerate foot. An example is shown in (30):

(30) a. *<gi>(n´âda)* 'mercy'
b. **(gi)(n´âda)*

The later weakening of *i* to schwa is evidence that the prefix *<gi>* was rather left unparsed and was not parsed into a degenerate foot carrying secondary stress instead: Such stress would

have prevented the weakening process. Thus, a violation of PARSE-SYLL (30a) was not as bad as a violation of FT-MIN (30b).

However, FT-MIN was dominated by PARSE-SEG in OHG: Degenerate feet were preserved although the non-parsing of a segment could have produced a trochee. An example is shown in (31). The syllable in bold type is degenerate. Note that there is also stress clash in OHG.

(31) OHG *(lé)(wìn.na)*[19] > MHG *(l´ö.win)* > NHG *(L´ö.win)* 'lioness'

Candidate 3 in (27), which has secondary stress, is not optimal. Evidence for the assumption that the final vowel was unstressed lies in the fact that this vowel was weakened to schwa. Secondary stress would have prevented this weakening (as in OHG *-ùnga* > MHG *-ùngə* > NHG *-(`)ung).*

Candidate 7 with unparsed <i> is the second best alternative following candidate 2. This is already an indication for later syncope.

The MHG tableau in (32) shows that only one change in ranking was necessary to evaluate candidate 7 as optimal.

(32) **MHG**

Input: *êrəst+ə*	*H/ə	ALIGN-L	FT-MAX	RhTp=T	*HDHD	FT-MIN	PARSE-SYLL	PARSE-SEG	NOCODA
1. *<ê>(r´ə.stə)*	*!	*					*		
2. ☞ *(´ê.rə)<stə>*							*!		
3. *(´ê.rə)(st`ə)*	*!			*		*			
4. *(´ê)(rə.st`ə)*	*!			**		*			
5. *(´ê)(r`ə.stə)*	*!			*	*	*			
6. *(´ê)<rə>(st`ə)*	*!			**		**	*		
7. ☞ *(´êr.<ə>stə)*								*	*
8. *(´ê.rə.st ə)*			*!	*					

In MHG, all unstressed syllables were weakened to schwa. Schwa syllables are without exception unstressable both in MHG and in NHG. Therefore *H/ə is undominated. Since there is no evidence for any ranking between *H/ə and ALIGN, both constraints are considered to be equally ranked.

[19] In the OHG form, there is a stress clash and a violatiion of RhTp=T. Both violations should have selected the form as non-optimal. However, the morphological structure is different from the structure of *êristo*, and other constraints (such as a constraint relating to the length of derivational suffixes) would also have to be considered.

The tableau shows that five of the eight candidates are ruled out because the schwa constraint is violated. This is evidence that the 'wrong' rhythm is prevented more effectively by the segmental weakening which had taken place at this time. Conversely, weakening could only occur in unstressed syllables, meaning that there is a mutual dependency between stress and segmental quality.

The modifications are motivated by two changes in the constraint hierarchy: Due to segmental weakening, *H/ə must be considered in the ranking. The second change consists in the lower ranking of PARSE-SEG. Since syncope is so very frequent in MHG, there is sufficient empirical evidence for this new constellation. In section 3 it was shown that there is no dominance relation between PARSE-SYLLABLE and PARSE-SEGMENT in MHG. Of course, the same relation is true within the context of a prosodically motivated ranking: Due to the non-ranking between both PARSE constraints, there were two optimal candidates.

In certain contexts vowel deletion is preferred to an unparsed syllable, especially if stress clash and/or a degenerate foot can be avoided. Therefore *HDHD and FT-MIN dominate PARSE-SEG in (32). Syncope in 'certain contexts' means that segmental adjacency conditions have to be fulfilled: If one consonant can't follow another because of sonority constraints, syncope is blocked. See (33) for such a case: In contrast to *ginâda* becoming *gnáːdə* (2a), syncope could not occur in the prefix because **gb-* was never a possible syllable onset in German.[20]

(33) OHG *gibˊûidi* > MHG *gəbˊûwədə* > NHG *Gəbˊäude* 'building'

Candidate 2 in the MHG tableau is the second best candidate. This is plausible because this was the optimal candidate in OHG. Conversely, the optimal candidate 7 in the MHG tableau was second best in OHG.

The NHG constraint rankings in (34) below do not differ from those in MHG. Some structures have become impossible because the morphological input changed on the base of Lexicon Optimization as it was already shown for MHG in section 3. There are only two syllables left, and so the number of candidates is limited. The optimal candidate has become even better: It only violates the NOCODA constraint (which plays a minor role compared to the metrical constraints) whereas there still was an additional violation of PARSE-SEG in MHG.

[20] This restriction can be formally expressed by a general constraint 'SONORITY'. Since in this context only syncopes that really occurred are considered, such a constraint would never be violated. Therefore, it is not necessary to integrate it into the rankings. However, it is obvious that SONORITY would be undominated because it must be met by acceptable candidates without any exception.

(34) **NHG**

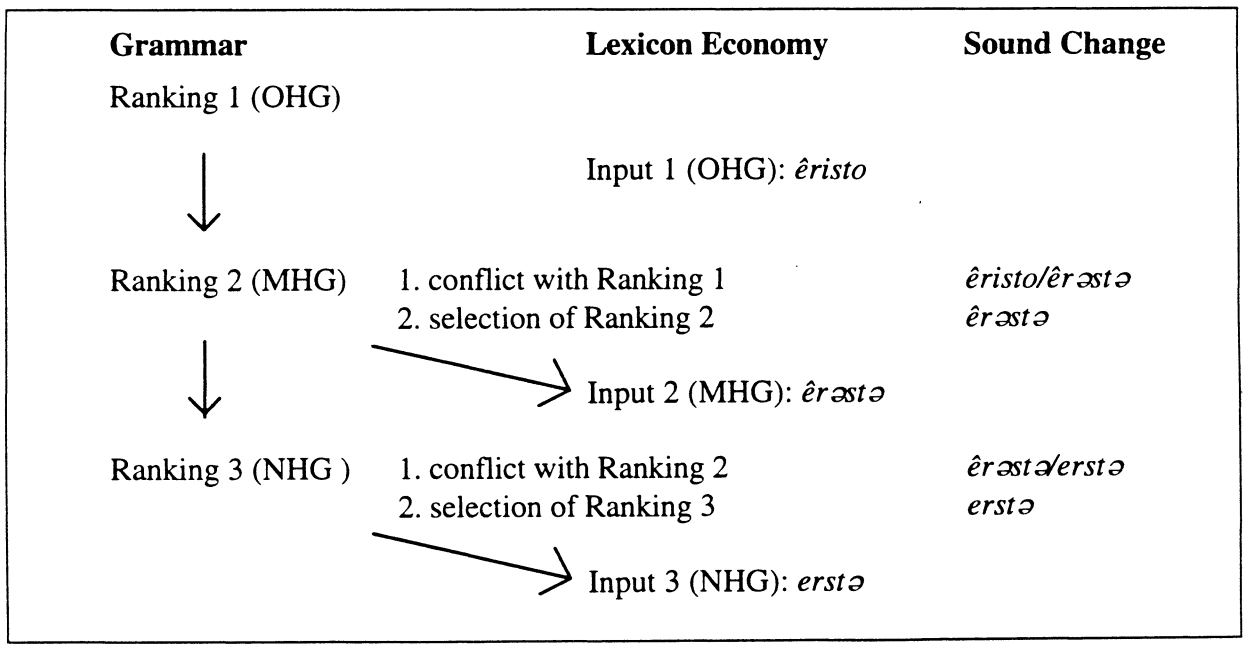
Input: *erst+ə*	*H/ə	ALIGN-L	FT-MAX	RhTp=T	*HD̈HD	*FT-MIN	PARSE-SYLL	PARSE-SEG	NOCODA
1. *<ér>(st`ə)*	*!	*	*	*		*	*		*
2. *(ér)<stə>*			*!	*		*	*		*
3. *(ér)(st`ə)*	*!		**	**	*	**			*
4. ☞ *(ér.stə)*									*

In section 3 it was argued that the most harmonic input in the MHG evaluation also differs from the lexical entry in OHG. Both weakening and deletion thus give evidence for the assumption that a reranking of constraints can subsequently cause a change in the lexicon. This means that sound change is related to two different domains that are connected in a temporal order. The first domain of change is the grammar: Since rankings can conflict, these rankings can lead to more than one optimal candidate at a given stage (such as *êrəstə* versus *êrstə* in MHG). They are the basis of variation. If one of the conflicting rankings gets predominant at a later stage, (which is called selection in historical linguistics), it is highly probable that the most harmonic input will not correspond to the input which was best in the previous stage.

In summary, sound change in its first phase is a turbulence within the grammar which is caused by parameters such as articulatory economy and other factors which are encoded in constraints. Due to this turbulence, there are conflicting rankings and therefore several optimal variants that are all well-formed. In a second phase, sound change affects the lexicon: One variant - and therefore one ranking - prevails, and the most harmonic input with respect to the new hierarchy is supposed to change.

The sequence of phases in a sound change process can be sketched as follows:

(35) **Reranking and change of the input**



5. Interaction of segmental and prosodic constraints

In the previous sections, segmental and prosodic constraints were treated separately. However, this should not obscure the fact that these two factors interacted. Certain constraints do not interact, whereas the constraints PARSE-SEGMENT and PARSE-SYLLABLE were active in both contexts. The changes within the ranking that made weakening and deletion processes possible are summarized in (36):

(36) a. **OHG**

PARSE-SEG >> PARSE-PL >> *V(σ_w) >> PARSE-SYLL
FT-MIN

b. **MHG**

*H/ə; *V(σ_w) >> PARSE-PL >> PARSE-SEG; PARSE-SYLL
FT-MIN

c. **NHG**

*H/ə; *V(σ_w) >> PARSE-PL >> PARSE-SYLL >> PARSE-SEG
FT-MIN

References

Bailey, Charles-James (1973). Variation and lingustic theory. Arlington: Center for Applied Linguistics.

Booij, Geert (1985). Coordination reduction in complex words: a case for prosodic phonology. In: van der Hulst, Harry/Smith, Norval (eds.)(1985). Advances in nonlinear phonology. Dordrecht: Foris, 143-159.

Braune, Wilhelm/Mitzka, Walter (1967). Althochdeutsche Grammatik. 12th edition. Tübingen: Niemeyer.

Broe, Michael (1993). Specification Theory. The treatmen: of redundancy in generative phonology. Ph.D. Dissertation, University of Edinburgh.

Dresher, B. Elan/Lahiri, Aditi (1991). The Germanic foot: metrical coherence in English. In: Linguistic Inquiry 22, 251-286.

Eis, Gerhard (1950). Historische Laut- und Formenlehre des Mittelhochdeutschen. Heidelberg: Winter.

Eisenberg, Peter (1991). Syllabische Struktur und Wortakzent: Prinzipien der Prosodik deutscher Wörter. In: Zeitschrift für Sprachwissenschaft 10, 37-64.

Everett, Dan (1994). Foot binarity. Ms., University of Pittsburgh.

Giegerich, Heinz (1985). Metrical phonology and phonological structure: German and English. Cambridge: Cambridge University Press.

Hall, Tracy Alan (1992). Syllable structure and syllable-related processes in German. Tübingen: Niemeyer.

Halle, Morris/Vergnaud, Jean Roger (1987). An essay on stress. Cambridge, MA: MIT Press.

Hayes, Bruce (1995). Metrical stress theory: principles and case studies. Chicago; London: The University of Chicago Press.

Itô, Junko/Mester, Armin/Padgett, Jaye (1994). NC: Licensing and underspecification in Optimality Theory. Ms. FTP.

Jacobs, Haike (1989). Nonlinear studies in the historical phonology of French. Ph.D. dissertation, Nijmegen University.

Kager, René (1993). Alternatives to the iambic-trochaic law. In: Natural Language and Linguistic Theory 11, 381-432.

Kager, René (1994). On eliminating directional foot-parsing. Paper read at the GLOW-Colloquium, Vienna, April 8, 1994.

Kenstowicz, Michael (1994a). Phonology in generative grammar. Oxford: Blackwell.

Kenstowicz, Michael (1994b). Sonority-driven stress. Ms., FTP.

von Kienle, Richard (1960). Historische Laut- und Formenlehre des Deutschen. Tübingen: Niemeyer.

Kiparsky, Paul (1994). Remarks on markedness. Paper read at the Trilateral Phonology Weekend, University of California at Santa Cruz, January 22, 1994.

Levin, Juliette (1988). Generating ternary feet. In: Texas Linguistic Forum 29, 97-113.

Löhken, Sylvia (to appear). The emergence of complex consonantal clusters in German: an Optimality-Theoretical approach. In: Nespor, Marina/Smith, Norval (eds.) Proceedings of the Second Phonology Conference of the Holland Intitute of Generative Linguistics. den Haag: Holland Academic Press.

McCarthy, John J./Prince, Alan S. (1993a). Prosodic Morphology I. Constraint interaction and satisfaction. Ms., University of Massachusetts/Rutgers University.

McCarthy, John J./Prince, Alan S. (1993b). Generalized alignment. In: Yearbook of Morphology (1993), 79-154.

Nespor, Marina/Vogel, Irene (1986). Prosodic phonology. Dordrecht: Foris (Studies in Generative Grammar; 28).

Paul, Hermann (1969). Mittelhochdeutsche Grammatik. 20th edition. Tübingen: Niemeyer.

Prince, Alan S./Smolensky, Paul (1993). Optimality Theory: constraint interaction in generative grammar. Technical Report #2, Rutgers University, Center for Cognitive Science.

Prokosch, Eduard (1939). A comparative Germanic grammar. Baltimore: Linguistic Society of America.

Rice, Curtis (1992). Binarity and ternarity in metrical theory: parametric extents. Ph.D. Dissertation, University of Texas at Austin.

Steriade, Donca (1995). Underspecification and markedness. In: Goldsmith, John (ed.) (1995). The handbook of phonological theory. Cambridge; Oxford: Blackwell, 114-174.

Trubetzkoy, Nikolaj S. (1939). Grundzüge der Phonologie. 7. Auflage 1989. Göttingen: Vandenhoeck & Ruprecht.

Wiese, Richard (1995). The phonology of German. Oxford: Clarendon Press.

Wilmanns, Wilhelm (1911). Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsch. Third edition. Strassburg: Trübner.

Wurzel, Wolfgang Ullrich (1994). Grammatisch initiierter Wandel. Ms. FAS, Berlin.

Zubritskaya, Katya (1994). Sound change in Optimality Theory and constraint tie. In: CLS 30, Vol. 2, 335-349.

Prominence hierarchy and phrase ordering-

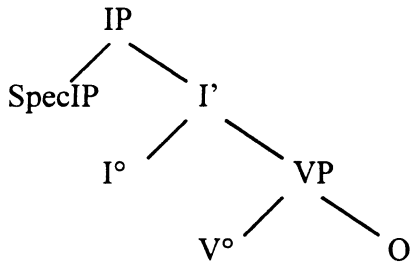
On why equidistance is not the right concept for explaining A-movement crossing and about how it can be accounted for in an alternative theory of relativized relativized minimality[1]

André Meinunger
andre@fas.ag-berlin.mpg.de

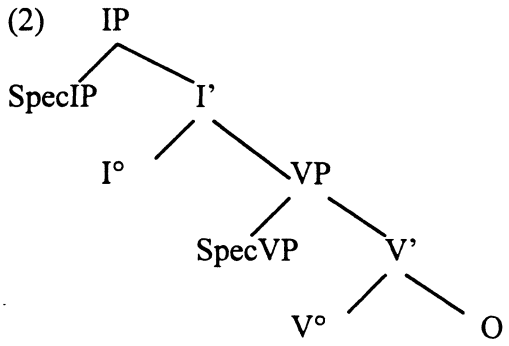
1. Introduction, where some new proposals within the theory of generative grammar are presented which will later be used as tools for the aim announced in the above title of the paper

In the mid eighties, it was proposed by several authors that the subject should be base generated VP internally and then start from this position, which is presumably [Spec,VP], and raise to the specifier position of some INFL projection. The most convincing theoretic argument for such an analysis is theta-theory, insofar as under the VP internal subject hypothesis (VPISH) the subject starts in a position that is within the projection of the verb from which it gets a theta-role, though the external one. The most compelling empirical argument comes from languages that seem to provide more than one, usually two, positions which the subject may occupy. These two positions then are the derived position, something like the traditional [Spec,IP] on the one hand, and the base position [Spec,VP] on the other. Thus, the version in (1) is replaced by the more flexible one in (2):

(1)



(2)



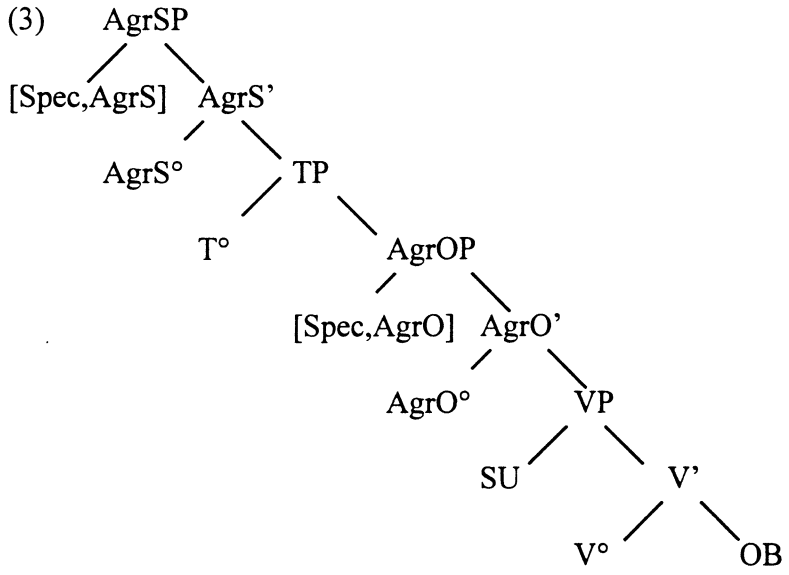
This proposal also had the desirable consequence of assigning unitary structures to both lexical and functional projections. Before the VPISH, there was no agreement whether VP should have a specifier position at all, nor were there reasonable proposals for what could be the specifier of VP.

The next standardization that is relevant for our purposes is the unitary treatment of Case assignment. To my knowledge, one of the first to propose that Case assignment to the object works parallel to Case assignment to the subject was Sportiche (1990) with his Strong

[1] I would like to thank several people. First, I thank Chris Wilder for checking my English and making insightful comments. Second, I am grateful to Enric Vallduví and Josep Quer for providing me relevant Catalan data. In further discussions I learnt from Artemis Alexiadou and Ilse Zimmermann.

Correlation Hypothesis (SCH). It says that structural Case is generally assigned in a Spec-head configuration of an argument NP with an agreement morpheme. Whereas formerly, nominative Case was assigned to the subject that was base generated under [Spec,INFL] (or had raised there in raising constructions) in a Spec head configuration with the inflectional element in INFL°, and accusative was assigned by the verb under c-command, structural Case is now uniformly assigned (or checked) in a spec head configuration between an Agr° head and an NP. Combining everything said so far, we get a tree that could be taken from Chomsky's minimalist paper:

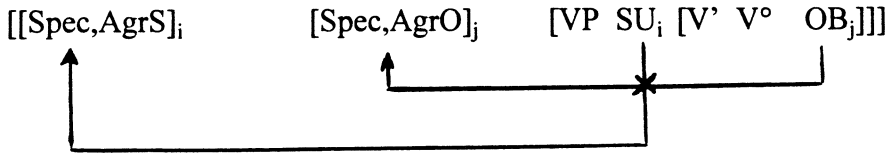
(3)



2. Equidistance and its problems

Now, the harmony and elegance of the tree in (3) is matched by the problem it poses. The minimalist paper develops the idea that the lexicon feeds the computational system with items which project according to X-bar theory. Generalized transformations (GT)[2] operate on them until a legitimate PF-object is created and Spell-out applies. After Spell-out, the work of GT continues until the final LF representation is achieved. The LF representation is supposed to be universal and thus languages should not (considerably) differ at that level. At LF, the subject of an (unmarked declarative) sentence has to stay in [Spec,AgrS], the object in [Spec,AgrO]. The problem is that, if both specifier positions are of the same type, presumably A-positions, we face a relativized minimality violation. Relativized minimality says that the closest potential governor blindly governs. That means, the object in [Spec,AgrO] interrupts the chain C= {[Spec,AgrS]$_i$, SU$_i$}:

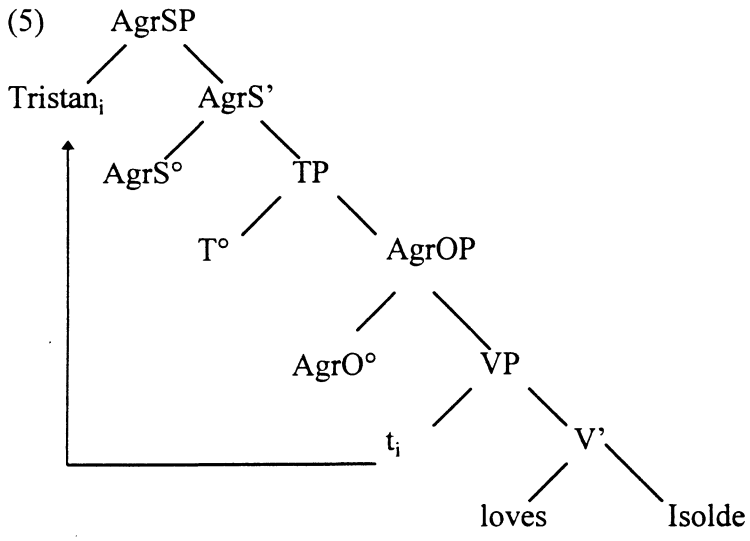
(4)



[2] In the most recent Chomsky paper ('Bare Phrase Structure') GT does not play a role any more. The relevant operation is now called 'merge'.

Chomsky develops a theory to avoid this problem. His solution is based on a derivational view of structure creation. To see how Chomsky's theory works, let's start with an unproblematic case. For English, Chomsky assumes that the only thing that moves in overt syntax is the subject. With the further assumption that specifier positions are not necessarily projected, but only created when needed[3], the VP internal subject may (and must) raise to [Spec,AgrS] without crossing any intervening element, and thus the Spell-out representation in (5) is well formed: which is the

(5)



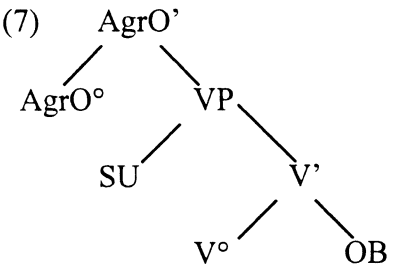
The RM violation only arises when also the object moves. Well, Chomsky's clever trick is the following. The verb has to move as well. At LF, at the latest, English verbs also have to have raised and adjoined to AgrS° (or possibly C°). The verb raises through head-to-head movement. That means the verb starts in its base position and adjoins to AgrO° (step (7) to (8)). This move creates the chain C = {V+AgrO°, t_{verb}} with the head in AgrO° and the foot as the trace heading the VP projection. This move enlarges the so-called minimal domain of the verb(al element). The minimal domain of V° is every node contained in VP, not including V itself, thus (SU, OB). When V° adjoins to AgrO°, the resulting head V+AgrO° has a minimal domain with one more member, namely [Spec, AgrO°]. Now, Chomsky introduces the following definition:

(6) If α, β are in the same minimal domain and c-command Γ, they are equidistant from Γ.

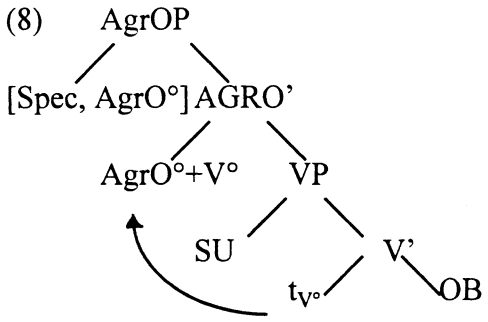
For our example in (8), that means that [Spec, AgrO°] and SU are equidistant from OB. Thus the object may leave its base position and raise to [Spec, AgrO°]. The subject trace does not count as an intervener as it is not closer, but exactly as close as [Spec, AgrO°] with respect to the object trace. Thus for this moment in the derivation, the out put does not violate RM (9).

[3] 'Needed' means 'present by virtue of being filled or targeted for movement within the derivation.'

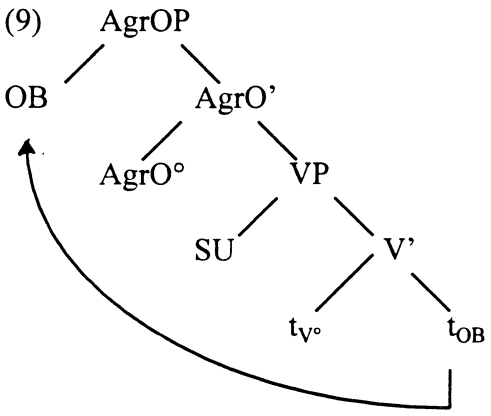
(7)



(8)



(9)



Consider that under such an analysis, overt movement of the object should be prohibited if no overt verb raising had taken place before. Only the moved verb makes the two relevant positions equidistant. This seems to be confirmed by the behavior of Scandinavian object shift. Since the Scandinavian languages are V2, the verb raises to the highest functional head (C°) in root clauses. Assuming HMC, this implies that the first steps in the derivation above must have been as described in the preceding paragraph. Thus in V2 sentences, object shift is allowed:

(10) Peter læste den_i ikke t_i.
Peter read-past it_i not t_i.
(Peter didn't read it.)

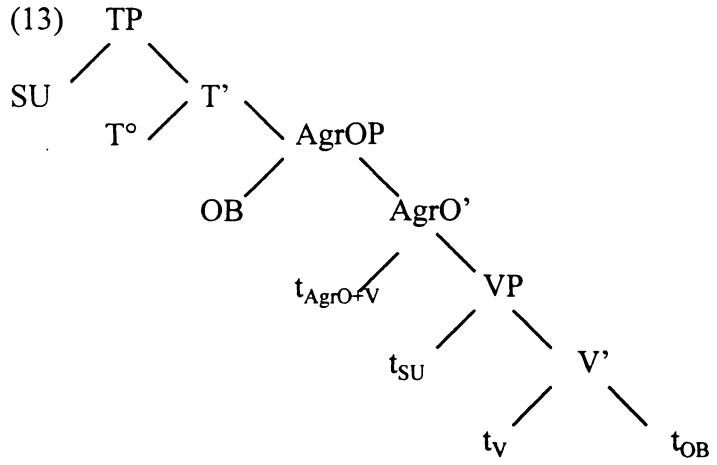
In embedded sentences the verb does not move. Its position relative to the negation adverb indicates that no verb movement has taken place (11). In that case then, object movement is also impossible (12).

(11) at Peter ikke læste den
that Peter not read it
(that Peter didn't read it)

(12) *at Peter den_i ikke læste t_i
that Peter it not read

The equidistance idea reminds of Baker's (1988) Government Transparency Corollary (GTC). However, as Jonas and Bobalijk (1993) observe, the equidistance concept is not transitive as GTC is. That means that further movement of the V°+AgrO° complex does not render more positions equidistant. In a system of the order of functional categories Chomsky assumes, TP is the next higher functional layer above AgrOP. Thus obeying HMC, the V°+AgrO° complex adjoins to T°. This step, however, does not render [Spec,TP] and [Spec, VP] equidistant. The head chain C = {[T[AgrO^V]] , [AgrO^V]} is different from the very bottom chain C' = {[AgrO^V] , V}. There is no chain for which more than two specifier positions are equidistant. What the further additional movement does is to render [Spec,TP] and [Spec, AgrO] equidistant. And this is what Chomsky needs. As soon as these positions both count as equidistant to VP and everything it contains, the subject is allowed to move from its original position without causing a RM violation, at least for the Scandinavian structure in (10).

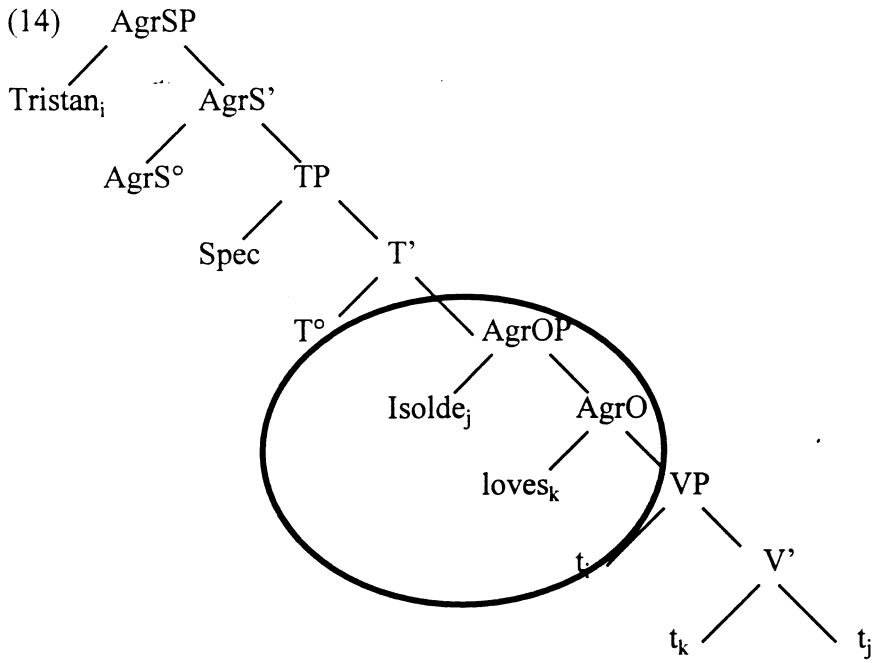
(13)



This way, the Danish sentence in (10) comes out as grammatical.

Now, if we don't ask further, we could be satisfied with what Chomsky has proposed. However, a closer look reveals some problems. Let's go back to the English example in (5). This is the spell-out or pre-spell-out structure. However, the derivation continues to obtain the final representation where every XP and X° is in its designated position, i.e. the verb under AgrS° (or even C°) and the object in [Spec,AgrO]. The derivation should go like this: V° raises to AgrO°. That move renders [Spec,AgrO] and [Spec,VP] equidistant. This provides the chance for the object to move to [Spec,AgrO], as described above, no RM violation is triggered since [Spec,VP] does not count as an intervener. However, at this point in the derivation we get a RM violation.

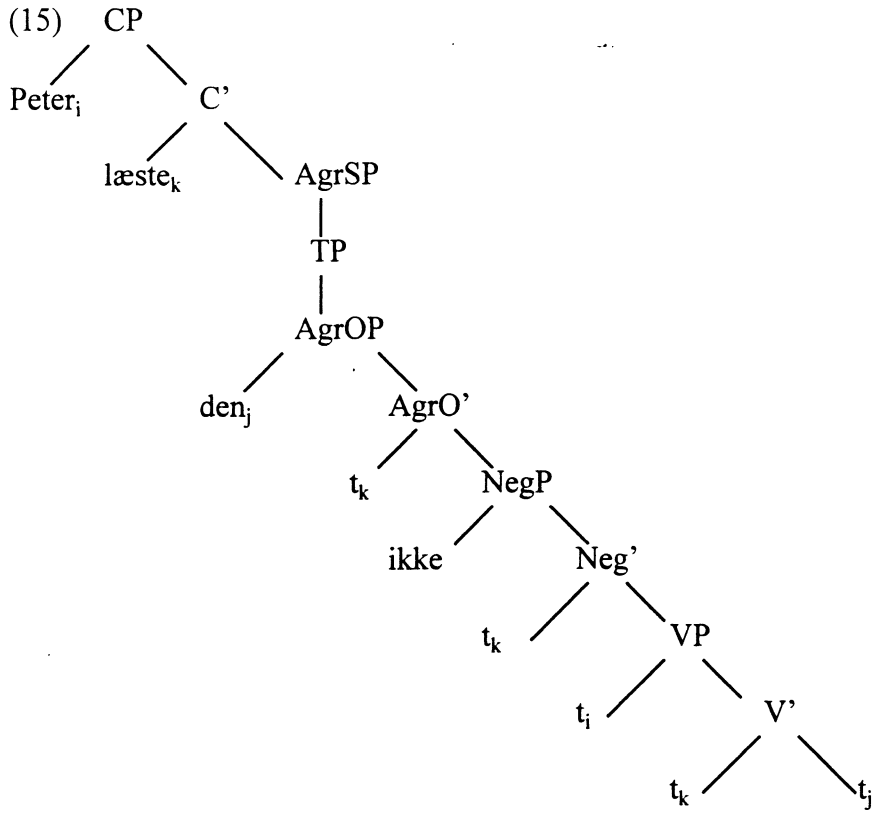
(14)



Only [Spec,AgrO] and [Spec,VP] are equidistant with respect to what VP contains. That means that now the object in [Spec,AgrO] is an intervener. It interrupts the chain C = {Tristan$_i$, t$_i$}. Since the subject is already in [Spec,AgrS] in English, the equidistance relation between the subject position [Spec,AgrT] and the object in [Spec,AgrO] can never be obtained. I don't see any reasonable solution of the problem. I just see stipulations. The other thing is that the equidistance story crucially depends on a purely derivational view of structure generation. Under a representational perspective or a combination of derivation and representation the theory does not work at all. I think we should eliminate the equidistance story and think of a better explanation. Nevertheless, before I will try to do that I will show some empirical problems with the equidistance explanation.

One point that is also important for Chomsky's idea is that for his explanation to work it is crucial that VP is selected by AgrO. Only this configuration permits such a local relationship of a subject and a object position that these are potentially equidistant. If something else than AgrO (immediately) selects VP, this head would be the first target of the verb. Then its specifier would be equidistant with the subject, however without any advantage for the object. Then, there would be no way for it to raise out of its base position. Is there such a configuration? The most recent treatments of negation in syntax propose that negation follows X-bar syntax and projects according to it. Thus, negation is a head with a complement and a specifier that shares the negative property of the head (see Haegeman). The structure of negation is universal, languages differ in how they make use of it. There are languages that have a negative head (Italian, Russian); in those languages this head behaves like a verbal affix and cliticizes onto the verb. Then there are languages that have a morphophonological spell-out of both the negative head and the specifier. French is such a case. And finally, there are languages that only use a negative adverb to be base generated in [Spec,Neg]. Such languages are represented by German and Scandinavian for example. Nevertheless, there is a phonologically empty head, that hosts and licenses the specifier position of *nicht, net, niet, ikke, ekki* and the like. If we incorporate negation into the syntactic tree for the representation of (10), we get the following tree:

(15)



There we cannot get a eqidistance creating structure where [Spec,AgrO] and [Spec,VP] have the same distance from the object.[4]

We face the same problem in the analysis for German. Nowadays there are two proposals for German sentence structure. The more traditional one, which I will adopt later, deals with head final structures. Except for order, which is {complement > head} for the verb and all functional heads (but C°), we get the same representation as in the Danish example in (15). The object has scrambled (shifted) and negation intervenes. This should cause the same RM violation as in Scandinavian.

(16) Peter las das Buch nicht.
Peter read the book not.
(Peter didn't read the book.)

If we choose the other version and analyze German as SVO language, as it is fashionable now, we get into even more trouble. The analysis of Dutch in Zwart (1993), following the main idea of Kayne (1993), proposes that the Germanic pattern is {head > complement} throughout. In such analyses the verb occupies two positions. Either it raises and adjoins to

[4] The problem becomes even more relevant if one adopts Kayne's (1993) or Hoekstra's (1991) X-bar theory. They propose that there is no adjunction anymore. A maximal phrase either has to be a complement or a specifier. Within this framework adverbials are licensed in specifier positions of functional heads (see also Alexiadou (1994)). Thus any adverb type that precedes the VP creates this equidistance blocking structure, like negation in example (15). Object shift always crosses those adverbs (formerly analyzed as VP adjoined) and should then introduce a RM violation. Interestingly, the evidence that is always given to show that object shift has taken place is adverb positioning.

C°, or it stays in its base position, i.e. exactly as in Scandinavian. The first case is triggered in V2 contexts, i.e. main clauses; the latter one in embedded sentences (Jan-Wouter Zwart p.c.). To account for the linear order of {object > verb} he is forced to say that, while the verb remains in situ, the object moves to [Spec,AgrO]. This is exactly what the equidistance story wants to rule out. Movement to [Spec,AgrO] is only possible when the verb has moved as well.

A more general problem is raised when we consider double object constructions with indirect objects. There is no consigns currently on whether dative should be analyzed as a structural Case. Nevertheless, it is claimed more and more often that dative should be regarded as such. The following arguments favor such an analysis. One important characteristic of oblique case is that it is lexical[5]. Dative, however, does not have a (completely) unpredictable occurrence. In the case of bitransitive verbs, a certain thematic structure automatically forces dative assignment. In that respect, dative patterns like nominative and accusative, the cases traditionally analyzed as structural ones. The second argument has to do with agreement. In the minimalist program and elsewhere, assignment or checking of structural Case is triggered under spec-head agreement (see above). Thus if we find morphological verbal agreement with dative objects, this should support the dative-as-structural-Case-hypothesis. There are indeed many languages whose morphology induces agreement morphemes for (nominative, accusative and) dative objects (Givón for Swahili (1976); Suñer for Spanish (1988)).The next point is, that dative is not more, or less, closely linked to any particular theta-role than nominative or accusative. Being associated with a theta-role is a characteristic of oblique cases, not one of structural ones[6]. The last, and maybe most compelling argument, is that (some) languages have a dative passive. Constructions where for certain reasons regular Case assignment is blocked and raising of the Case-less NP to some other Case position is forced to save grammaticality are an indicator that structural positions are involved. Lexical Case never allows for passive constructions; dative, like accusative, does. German illustrates that very nicely:

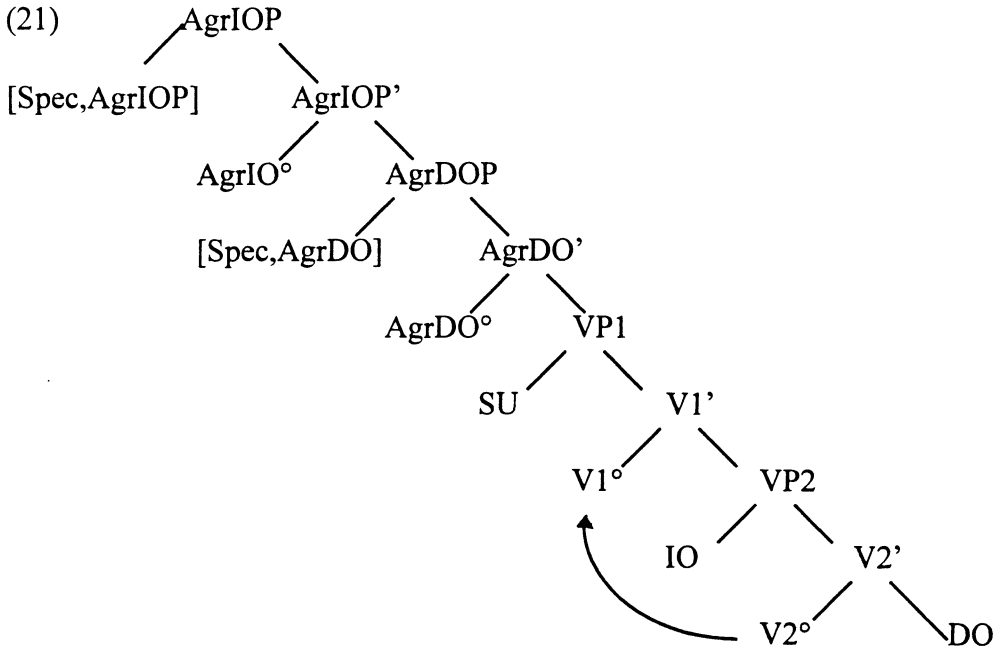
(17) Wotan verhieß Siegmund ein Schwert.
Wotan$_{nom}$ promised Siegmund$_{dat}$ a sword$_{acc}$
Wotan promised a sword to Siegmund.
(18) Siegmund bekam (von Wotan) ein Schwert verhießen.
Siegmund$_{nom}$ pass$_{aux}$ (by Wotan) a sword$_{acc}$ promised
Siegmund got promised a sword (by Wotan).
(19) Alberich stielt den Rheintöchtern das Gold.
Alberich$_{nom}$ steals the Rhinemaidens$_{dat}$ the gold$_{acc}$.
Alberich steels the gold from the Rhinemaidens.
(20) Die Rheintöchter bekommen (von Alberich) das Gold gestohlen.
The Rhinemaidens$_{nom}$ auxpass (by Alberich) the gold$_{acc}$ stolen
The Rhinemaidens were robbed of the gold (by Alberich).

If we now implement dative as structural Case into the minimalist framework we see again that Chomsky's trick is untenable. By combining X-bar theory, Larsonian structure and a

[5] Sometimes, oblique Case and lexical Case are even used as synonyms.
[6] Of course, there is some affiliation between dative case and goal. However, there is a comparable one between nominative and agent, an accusative and theme. Further theta-roles of dative bearing NPs are experiencer, bene-(male-)factive, agent (in causative constructions), and all the mysterious cases of the so-called free dative.

thematic hierarchy where goal is higher than theme (for justification thereof see below and references quoted there) we get a VP like that in (21). Furthermore, as cross-linguistic data show, the (relative) order of the arguments when outside the verb phrase parallels the one downstairs in the VP. Thus the lowest Agr head should be associated with the lowest object, i.e. AgrDO should select VP and should itself be selected by AgrIO. (We skip here possible intervening functional categories.)

(21)



V2° raises to V1° in order to link all arguments together. Within the equidistance theory this move would render SU and IO equidistant from DO. However, there is no position the object could move to. The next available one is its designated position [Spec,AgrDO], which, nevertheless, is to far away. The first step has already shown that the equidistance theory doesn't work here either. The problems multiply as the derivation continues.

3. chapter - within which Relativized Minimality is presented and criticized, and then an algebraic semantic account is adopted

In his book „Relativized Minimality", Rizzi (1990) shows that movement of any sort obeys the same constraint: movement to position X cannot cross a position of the same type. This theory is a representational one, in that an output structure is ruled out if there is an intervening element between the moved element and its trace, with both the moved and the intervening element being of the same type. This explains the unacceptability of the following sentences.

(22) *Why do you wonder [who left t]
(23) *John seems that it is unlikely [t to win]]

(22) is ruled out because *who* is in an A'-position and intervenes between *why* which also occupies an A'-position, and its trace, and thus blindly binds it. In (23) both *John* and *it* are in

A-positions. *it* is closer to the trace of *John*, binds it and thus causes the RM violation. To summarize: what is crucial for the further argumentation is that A'-movement and A-movement are restricted in the same way. For this reason, Rizzi gives a formulation that does not make reference to a special type of position.

Rizzi, Relativized Minimality (1990), page 7:

Relativized Minimality: X α-governs Y if there is no Z such that
(i) Z is a typical potential α-governor for Y
(ii) Z c-commands Y and does not c-command X

Rizzi was well aware of the fact that this formulation was too restrictive. In some cases, an element may intervene without inducing an ungrammatical structure. Compare (22) with (24):

(24) (?)Which paper do you wonder who reviewed t ?

Although *who* in an A'-position intervenes, *which paper* - also in an A'-position, but further away - remains capable of binding and thus identifying its trace. Rizzi stipulates that referential expressions are not subject to RM, they carry a referential index that renders them able to identify their trace from anywhere. For Rizzi, a referential index is linked to a referential theta-role. He modifies the classical argument/adjunct distinction and proposes that theta-roles like agent or patient make phrases referential whereas roles like manner do not. This way he explains the contrast between (25) and (26).

(25) Which linguist do you wonder whether I like t ?
(26) *How do you wonder whether Artemis behaves t ?

Thus, though the manner phrase is theta-marked by the embedded verb in (26), i.e. argumental, it cannot be extracted from a weak island since it lacks a referential theta-role.
However, Rizzi's RM is still too rigid to explain all data. Within his theory of referential indices, only arguments can bear a referential index, since only arguments are linked to certain thematic roles. Nevertheless, extraction of adjuncts out of weak islands is possible. Normally, adjuncts do not extract (27a), (28a), however, if the context allows for a discourse linked interpretation, even an adjunct can be extracted without causing (sharp) ungrammaticality (27b), (28b).

(27a) *Why$_i$ do you wonder [if they can fire you t$_i$]
(27b) For which of these reasons$_i$ do you wonder [if they can fire you t$_i$]

(28a) *How$_i$ were you not able to solve the problem t$_i$
(28b) (Our boss said that one could solve this problem with every computer here in this room. Now you are saying this is not true. So tell me:)
[With which of the computers here]$_i$ were you not able to solve the problem t$_i$?

On the other hand, if certain interpretations are forced, extraction of complements becomes ungrammatical:

(29) *How much wine$_i$ did you not poison t$_i$?

(30) *Who the hell$_i$ do you regret that our aunt saw t_i ?

These data show that Rizzi's proposal is not completely correct. (27b) and (28b) should be ungrammatical, as the extractees do not get assigned a (referential) theta-role by the verb. On the other hand, if bearing a „referential" theta-role like patient made a phrase referential, it is unclear what explains the binding failure of the extractees in (29) and (30).

One of the most promising theories that tries to explain extraction facts that has been recently elaborated is to be found in Szabolcsi and Zwarts (1991, 1993). Their idea is that phrases that (are supposed to) take scope are associated with Boolean operations. Then, when a wh-phrase (i.e. a potential scope taker) scopes over some intervening other scopal element, all relevant operations that are associated with the wh-phrase must also be associable with the intervening scopal element. If this condition is not met, the wh-phrase cannot scope over the intervener. That means, either that sentences become ungrammatical, or that only a subset of potentially possible scope readings is available. In order not to misinterpret the quoted authors, I cite their rule (57) from the 1993 paper

> Scope and Operations:
> Each scopal element SE is associated with certain operations (e.g., not with complements). For a wh-phrase to take wide scope over some SE means that the operations associated with SE need to be performed in the wh-phrase's denotation domain. If the wh-phrase denotes in a domain for which the requisite operation is not defined, it cannot scope over SE.

Let me explain how this works by giving some examples. Boolean operations are: taking complements, intersection and union. Now, let's apply this to the following questions.

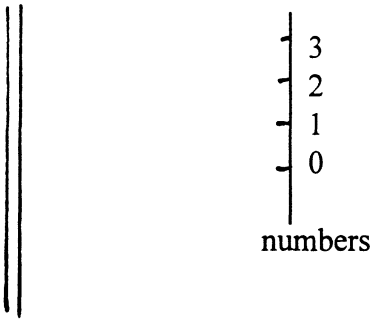
(31) Which European countries do you like ?
(32) Which European countries do you not like ?
(33) Which European countries does every American like ?

Szabolcsi and Zwarts make the reasonable assumption that the interpretation of questions ensures that an exhaustive list is determined by the answer. So to answer (31), one has to list all relevant, i.e. liked European countries. What taking complements means becomes clear when answering (32). All European countries form a set. The countries from the answer to (31) also form one. This is a subset of the total set. The remainder which is not in this subset forms the complement to the set of liked countries. This complement is the answer to (32). For (33), one has to look at each American and list the European countries (s)he likes. Then the lists are intersected. Intersection singles out the names of European countries that show up in every American's list, and this intersection is the answer to (33).

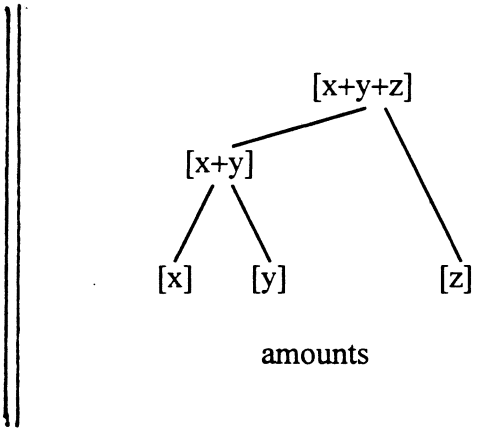
Reflections on the meaning (denotation) of potential extractees, Szabolcsi and Zwarts propose the following hierarchy:

(34)

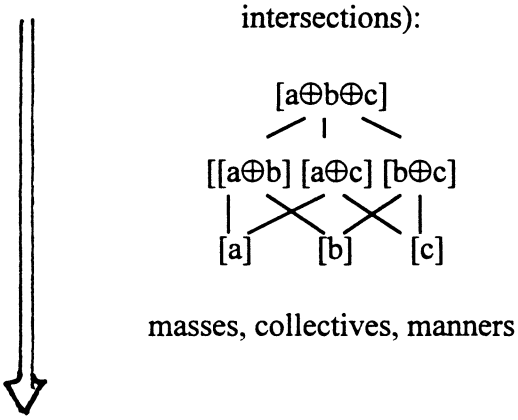
Chain (not closed under either unions or intersections or complements):



Join Semi-lattice (closed under unions, lacking closure under complements and intersections):



Free Join Semi-lattice (closed under unions, lacking closure under complements and intersections):



Set of unordered, discrete individuals (not restricted in the application of Boolean operations):

This hierarchy determines which scopers may scope over which others. Thus scope dependencies are relative. As the number of operations possible to be performed in a scopal element's domain decreases, the more difficult it becomes for that element to take scope. For an SE1 to take scope over some SE2, SE1 must at least allow for all the operations under which the domain of SE2 is closed as well. That means, the possible operations of an element with narrower scope must be a subset of the operations associated with the element that is supposed to take wider scope. The reason why discourse linked phrases are such good extractees is because they (usually) range over a domain of discrete individuals. However, non-discourse-linked phrases are also extractable from some weak islands, when they allow

for the necessary operations (for reasons of space I advise the reader to look at Szabolcsi and Zwarts (1991), (1993)). Now it is clear why (27a) and (28a) are grammatical, though they involve adjunct extractions. They range over concrete, salient reasons (27a); the computers from (28a) also form a set of unordered elements. The extracted object in (29) has an amount reading, and therefore denotes in a domain that forms a join semi-lattice, i.e. a partially ordered domain. Join semi-lattices are not closed under complements. This however is a condition for scope over negation. Thus the sentence is ruled out.

4. On hierarchies that are relevant for argument structure

In her book on argument structure, Grimshaw (1990) develops the idea of ordered argument structure. She argues that argument structure (AS) is not a collection of unordered thematic roles as had been assumed previously. She claims that AS is an ordered representation over which relations of prominence are defined. That means that the arguments of a verb (or of lexical categories in general) obey some principle that orders them, i.e. establishes a hierarchy, and that principle is prominence. Knowing that there are hypotheses of hierarchy that propose almost every permutation possible, she gives the following one with which I agree:

(Agent[7] (Experiencer (Goal / Source / Location (Theme))))

For her AS contains no information about particular theta roles, but only information about the relative prominence of the arguments. She explicitly states that she assumes the goal to be more prominent than the theme. This however is a point of debate. In the list of hierarchies Speas (1990) gives, only 3 ½ of 8 hierarchy proposals locate the goal argument higher then the theme. Since I agree with Grimshaw I first give her two main arguments and then add one by myself. Grimshaw refers to an earlier article by herself and Mester (1988) and brings evidence from a particular construction in Japanese. There is the light verb *suru* which does not have an argument structure. This verb however is accompanied by a direct object NP, and that NP in her example *shoomi* (= prove) brings arguments with it that integrate into the clause. The sentence is the grammatical only in case the goal argument precedes the theme.

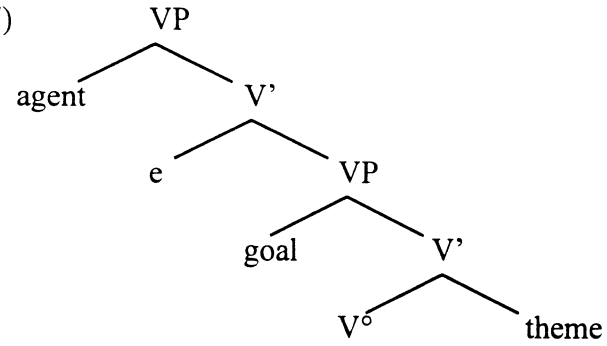
(35) Sono deeta-ga wareware-ni [[kare-no riron-ga machigatteirru-to]-no shoomei]-o shiteiru
that data-nom us -to he-gen theory-nom mistaken-be-C-gen prove-acc suru
'That data proves to us that this theory is mistaken'

(36) *Sono deeta-ga [kare-no riron-ga machigatteirru-to] [wareware-e-no shoomei]-o shiteiru
that data-nom he-gen theory-nom mistaken-be-C us -to- gen prove-acc suru

Her second argument comes from compounding in English. She considers bitransitive verbs like English *give*. A hierarchical structure like the one given above, put into a X-bar theoretic (Larsonian) representation, gives the following tree:

[7] Agent seems to be an exception in that it appears as highest thematic role in everybody's hierarchy I know of.

(37)



If incorporation (compounding) takes place, it is most likely the theme that incorporates into the verb. Incorporation crucially depends on c-command. In the tree given above only the theme may incorporate, the goal may not. When we look at compounds we get a confirmation for the correctness of the prominence hierarchy goal > theme:

(38) Gift-giving to children
(39) *Child-giving of gifts
(40) from: the giving of gifts to children

The argument that I wanted to add concerns the unmarked word order in German and Dutch. As I hinted above I assume a head final VP for both languages. I furthermore assume that existential indefinites don't move, i.e. they stay in their base position The linear ordering of arguments we get is subject (mostly agent) > indirect object (mostly goal) > direct object (mostly theme)

(41) weil zu dieser Zeit viele Lehrer einigen Schülern zu gute Noten gaben
since at that time many teachers$_{nom}$ some pupils$_{dat}$ too good grades$_{acc}$ gave
'since at that time many teacher gave too good grades to some pupils'

This data in my opinion strongly supports the Grimshaw hierarchy. In the same sense the proposed hierarchy is confirmed by languages that express dative positionally. According to Emonds (1993) there are four ways of identifying goals: (1) by productive morphological dative case, (2) by NPs which agree with an appropriately inflected verb, (3) by using an empty preposition like element that is best analyzed as case marker (K°), and (4) - and that's the crucial fact - by fixed word order and no case morphology. In that case precedence, which implies a structurally higher position, locates dative objects higher in the hierarchy than accusative ones. English illustrates that nicely:

(42) Sieglinde gave Hunding the sleeping draught.
The pope did not forgive Tannhäuser his sins.

Now I want to give my explanation for crossing - the actual topic of this paper. It has been observed that in languages with (more or less obligatory) movement of argumental phrases out of their base position to some higher one in the clause, there is a semantic effect. Scrambling is such a movement. Objects that scramble introduce some kind of anti-definiteness effect, i.e. scrambled objects (and subjects as well) only allow for a strong

reading. NPs with a weak determiner (indefinite article, adjective like quantifiers *many, some, few*; numerals and the like) are systematically ambiguous. They may either have a strong (presuppositional) or a cardinal reading. Thus sentences in (42), (43) reflecting base order are ambiguous[8], the scrambled versions (44) and (45) are not ambiguous any more, i.e. for *many linguists* a partitive reading is forced as the translation suggests.

(42) daß die Polizei gestern viele Spachwissenschaftler verhaftet hat (German)
(43) dat de politie gisteren veel taalkundigen opgepakt heeft (Dutch)
that the police yesterday many linguists arrested has
'that the police arrested many linguists yesterday'

(44) daß die Polizei viele Spachwissenschaftler gestern verhaftet hat (German)
(45) dat de politie veel taalkundigen gisteren opgepakt heeft (Dutch)
that the police many linguists yesterday arrested has
'that the police arrested many (of the) linguists yesterday

5. Tree splitting - attempts to account for argument positions depending on the reading that those trigger, at the end of this chapter the actual proposal will be made

In the past few years, several theories have been proposed that account for these data. I will briefly present the main idea of three of these. The first one is Diesing's Mapping Hypothesis (MH) (1988, 1992). Using quantifier structures in the style of I. Heim (1982) which split quantificational structure into a tripartitite one containing an unselelectively binding quantifier, a restrictive clause (RC) and a nuclear scope (NS), she proposes the following:

MH: Material from the VP is mapped into the NS
Material from the IP is mapped into the RC

Assuming that VP is the domain of existential closure the above readings are (almost) accounted for. If the indefinite NP moves outside VP into the RC domain, it gets a quantificational reading there.

The second proposal is by de Hoop (as well as the Dutch examples (43), (45) 1992). She proposes that there are two types of structural cases. One of them she calls Weak Case and reserves it for weakly quantified objects. This Case is assigned in situ. The other case she calls Strong Case. That one is assigned in a derived position and triggers a strong reading for its bearer. If the derived position is the target position of scrambling, the above facts are explained as well. She divides a sentence into (the set of real) arguments that are predicated of, on the one hand, and the predication itself on the other. The predication obeys a principle she calls Principle of Contrastiveness (POC), p. 166:

For all NPs Q of type <<e,t>,t> and predicates P :
Q(P) is only appropriate if $|P| \ni C$ & $|C| >$ or $= 2$

The complicated looking principle ensures that there must be alternatives to the predicate that could serve a possible assertion to the same Qs, i.e. true arguments.

[8] The ambiguity however is not a real free choice one. The weak reading is much more preferred in the non scrambled version.

In his dissertation, Vallduví (1992) proposes that a sentence can be divided into several parts that have different informational tasks. He suggests that sentences are structured according to how the speaker wants the addressee to retrieve the relevant information. This he calls information packaging. According to him, every sentence contains a focus. It furthermore may contain a ground. The ground is the forms what is supposed to be old information for the hearer and serves as an anchor in the preceding conversation. For reasons of cohesion the ground may be useful, it is, however, not necessary (for an opposite assumption see Jäger (1994)). The splitting is the following 'trinominal hierarchical articulation' (p. 46):

Sentence = {Focus, Ground}
Ground = {Link, Tail}

Links are what elsewhere is often called topic, in topic structures like

(46) Marc I will never forget.
Yesterday I went to a Wagner opera.

Tail elements are those phrases that do not constitute a topic in the sense that the sentence is about them, but are elements that are known to hearer and do not deliver new information.
In earlier work of mine (Meinunger (1994)) I use the above ideas and try to give a syntactically more refined analysis of what is going on with the movement out of VP. Under my analysis there, verbal arguments are projected according to Grimshaw's thematic hierarchy in the VP. In case some argument serves, as I and Jäger (1994) call it, a topic, it has to leave the VP and targets the specifier position of an agreement projection. Very informally, topic is to be understood a referential anchor about which something is asserted. It turns out that the domain of Diesing's RC, which is IP minus VP, de Hoop's set of true arguments and Vallduví's ground is the same and correspond to the topic part in my Console paper (1994) where I propose the following tree splitting:

[CP...[Agr... | [VP...]]]
Topic | Comment

The more refined syntactic analysis is what Adger (1993) calls the local versus global proposal. In the theory presented there (see Adger (1993) and Runner (1993) for very similar proposals) the specifier positions of agreement phrases host the topical NPs. Using evidence from morphological case realization, agreement data, clitic doubling, word order from typologically very different languages, I show that NPs get their thematic or rhematic reading in the relevant position[9], i.e.:

- direct objects that are part of the comment in the position that is closest to the verb or its trace (=complement position, i.e. sister of V°)
- indirect objects that are part of the comment in the specifier position of the lowest V°
- subjects in the highest specifier position of VP

[9] Actually I do not talk about the projection of indirect objects there. However the step from the proposal there to the integration of dative objects is straightforward.

Scambling in Dutch suggests that the same order is required for the hierarchical ordering of scrambled NPs, thus the order of agreement projection has to parallel the VP internal order[10]:

(47) a. dat Jan de mannen deze film met plezier toont
that Jan the men this movie with pleasure shows

b. *dat de mannen Jan deze film met plezier toont
that the men Jan this movie with pleasure shows

c. *dat deze film Jan de mannen met plezier toont
that this movie Jan the men with pleasure shows

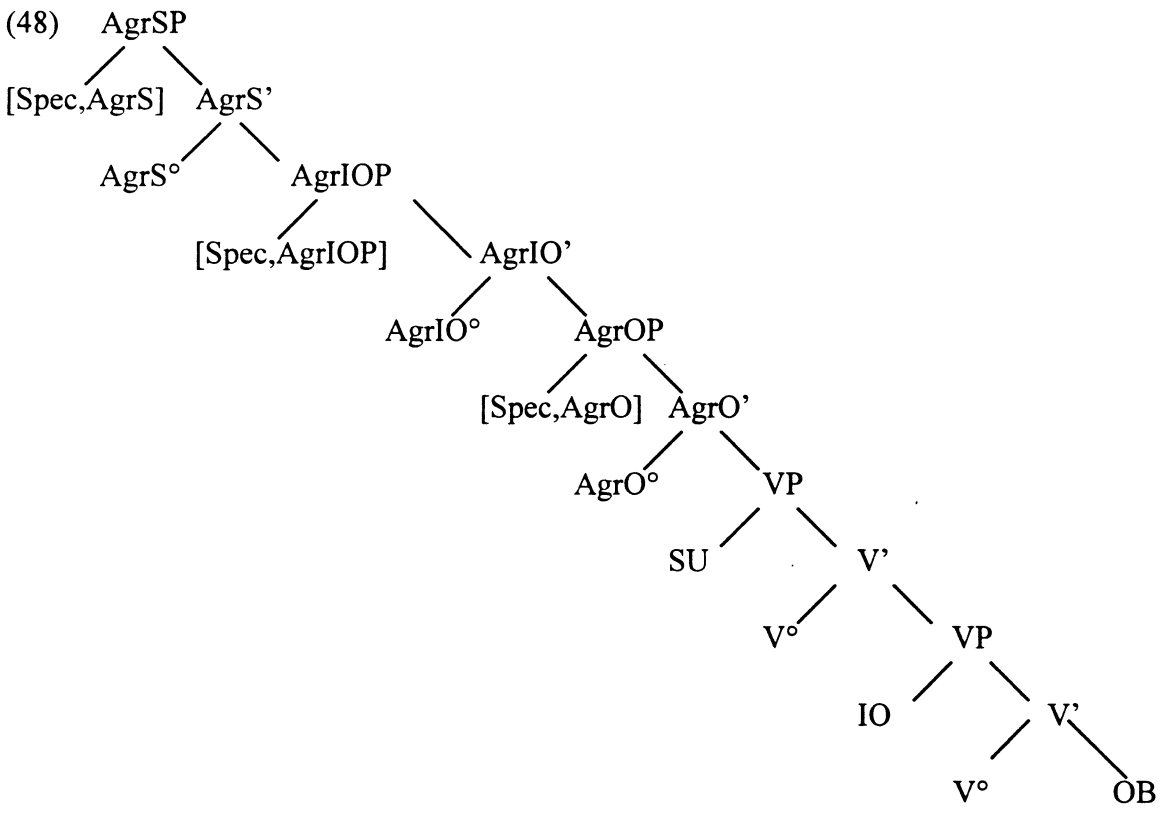
d. ??dat Jan deze film de mannen met plezier toont
that Jan this movie the men with pleasure hows

That suggests that AgrS is higher than AgrIO and that in turn is higher than AgrO. Thus the target position

- of thematic direct objects is [Spec,AgrO]
- of thematic indirect objects is [Spec,AgrIO] and
- of thematic subjects is [Spec,AgrS]

Accordingly we get the following tree:

(48)



[10] Examples taken from Neeleman (1994)

In order to get to their designated position all arguments cross each other several times. If we now implement Szabolcsi and Zwart's idea that crossing is possible if a certain hierarchy is preserved, we are able to account for the crossing mess. The hierarchy that must not be violated is the same that orders the arguments in their base: prominence. The intuitive idea behind prominence is salience, and that makes sense. By undergoing topic movement, i.e. movement to Spec,Agr positions, crossing is not only possible, but even forced. The only constraint is that is to be obeyed is hierarchy preserving of prominence.

In his article 'Topic, pronoun, and grammatical agreement' Givón (1976) also justifies the hierarchy which is assumed in this paper. He presents conceptual reflections which lead him to the conclusion that „there are grounds for believing that with respect to the topicality hierarchy, datives stand above accusatives. This is reflected in the higher percent of definites and humans for datives as compared to accusatives." And indeed, agents are always the highest arguments in a hierarchy. For agents as well, it is normally the case that they are human. Insofar the hierarchy of arguments is somehow a structuring of mankind's anthropocentric viewpoint. Prominence, in terms of communicative salience, thus orders arguments (see also Haftka (1980)). However, the scale of prominence is not always the same one, and sometimes there seem to be conflicts. One apparent conflict is shown in Grimshaw's book. There she gives the hierarchy given at page 12, here repeated as (49)

(49) (Agent (Experiencer (Goal / Source / Location (Theme))))

One class of experiencer verbs - the fear class (or Belletti and Rizzi's *temere* class) - is well-behaved. That means the experiencer becomes the subject of the sentence, the theme the object.

(50) Lohengrin fears Elsa's question.
Artemis likes Kayne's theory.

However, there is the class of ill-behaved verbs - the *frighten* class (Belletti and Rizzi's *preoccupare* class)

(51) Alberich frightens the Nibelungs.

Here the experiencer appears as a postverbal object, and the theme occupies the subject position. Grimshaw however presents a way out of the dilemma. Her proposal is that there is not only one scale of hierarchy but more, at least two. She shows that the ill behaved verbs have something to them which the other class lacks. There is a causative element involved such that (51) can be paraphrased by:

(52) Alberich causes the Nibelungs to experience fear.

Then she states that the causal structure of a predicate also defines a hierarchy, just as the thematic structure does, a hierarchy in which the cause argument is most prominent:

(53) (cause (....))

She claims that the causativity hierarchy overrides the other one(s) and imposes a structure where the causer is the most prominent argument. She furthermore attributes the hierarchy to event structure. This however is not that crucial to our point. What matters is the relative prominence that has to be preserved.

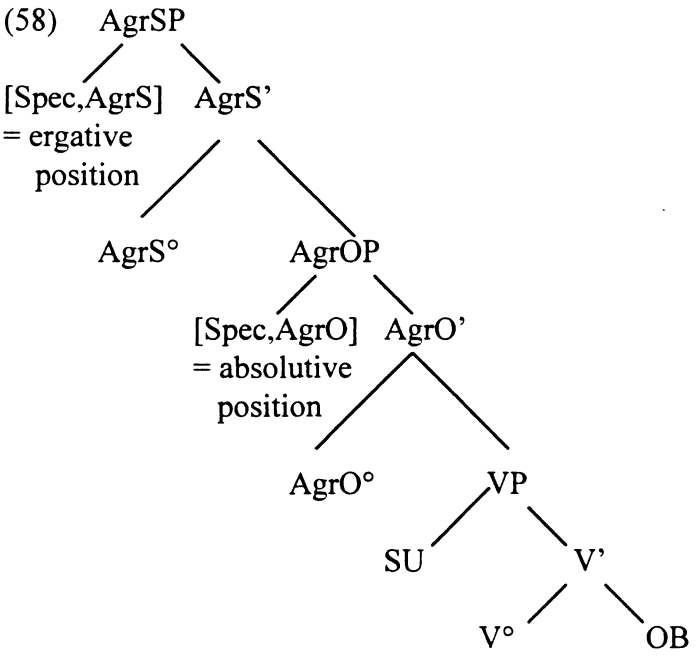
Additional support for my theory can be taken from ergative absolutive languages (e-a languages). Typologically those languages differ from nominative accusative languages (n-a languages) in that in the former, the same Case is assigned to subjects in intransitive sentences as is assigned to objects in transitive ones[11]. Nominative accusative languages normally uniformly assign nominative case to subjects (and accusative to objects).

(54) I kissed him. (I = nom, him = acc)
(55a) I came.
(55b) *Me came. (me = acc)

(56) Balan d^{y}ugumbil baŋgul yaraŋgu balgan (Dyrbal, examples taken
woman-abs man-erg hit from Comrie (1989))
'The man hit the woman.'
(57a) Bayi yara banin^{y}u.
man-abs came-here
'The man came here'
(57b) * Bayi yaraŋgu banin^{y}u.
man-erg

If we translate this case pattern into the minimalist framework, we get the following representation for e-a languages.

(58)



[11] This is a very rough sketch of the n-a / e-a difference. The actual data are much more complicated. The proposal however is not affected.

Now the difference between nominative (-accusative) and (ergative-) absolutive structures reduces to the target position of SU and OB. In transitive structures the thematically more prominent argument has to raise to the highest Spec,Agr position and the object raises to a lower one. This triggers SU movement to [Spec,AgrS] (the nominative checking position in n-a languages and the ergative checking position in e-a languages) and OB movement to [Spec,AgrO] (the accusative checking position in n-a languages and the absolutive checking position in e-a languages). These movement constraints are easily explained within our prominence preserving theory. In intransitive structures only one argument is present. In that case either landing site [Spec,AgrS] or [Spec,AgrO] could be targeted. No intervention or crossing structure arises. Since there is only one element, the hierarchy is trivially preserved. Languages may choose which Spec position a sole argument targets[12]. This is in the full spirit of Grimshaw's theory or the hypothesis presented here (prominence preserving): what matters is not fixed, rigid positions, but the relative prominence among arguments.

6. Prominence theory extended

6.1. The apparent problem with more flexible languages

Now I want to show that prominence theory extends even further. First I have to enlarge the data base and to eliminate some doubts which the new data might raise concerning the strict prominence theory. Using the Dutch examples in (47) above I tried to show that the basic order SU > IO > DO may not be changed. (47) was supposed to show that among all possible permutations, only the one in (47a) comes out as grammatical. This is indeed the case in Dutch, but not in German. In German it is possible for a direct object to move over an indirect one (59), and also for an object to cross an unmoved subject (60):

(59) weil Johannes das Buch einem Freund gegeben hat
since John the book$_{acc}$ a friend$_{dat}$ given has
'since John gave the book to a friend'

(60) (Mensch, die Gegend sieht aber eigenartig aus. Die ganzen Gebäude passen gar nicht zusammen. Wie kommt denn das ?)
Ja das ist so, weil hier jedes Haus ein anderer Architekt entworfen hat.
well that is so,since here every house$_{acc}$ a other architect$_{nom}$ designed has
'(Oh, this neighborhood looks strange. All the buildings here don't really fit to each other. How come ?)
Well, that's because every house here has been designed by an other architect.'

These are not only possible word order devices. They are even obligatory to express the intended meaning. If *einem Freund* is fronted in (59) then the NP loses its existential reading. The dative gets a specific reading, which may not be intended. In (60) the sentence even becomes ungrammatical, see (60'):

[12] That the choice is actually not free, but fully determined is shown in Mahajan's work on ergativity. His proposal however is independent of the prominence story. It explains why the languages on the one hand choose the n-a option, whereas the others use the e-a strategy. The idea is thus located on a different level and fully compatible with the prominence proposal.

(60') * weil hier ein anderer Architekt jedes Haus entworfen hat.[13]

How do we account for this data? As already stated above, Meinunger (1994) proposes the topic-comment mapping where topic NPs leave the VP and the other ones that belong to the comment stay in situ. Thus it follows that *das Buch* in (59) is discourse linked and acts as an element which is predicated of (exactly as *Johannes*) and therefore has to move to [Spec,AgrO]. The comment NP which happens to be the dative object remains in its base position. There it gets its purely existential rhematic reading. The same happens in (60). *'jedes Haus'* is a strongly quantified NP. Thus *Haus*, being presupposed and belonging to the restrictive clause, has to leave its base position. *'ein anderer Architekt'* is obviously not referential. It is contained in the comment and therefore stays inside VP in German. Lenerz (1977) also argues for the IO > DO order. His argument is that there are no restrictions for the IO > DO order. The other linearization is limited to certain circumstances. DO > IO is (only) possible, when IO is focused while DO is not. That is in full agreement with what the topic-comment mapping predicts.

The reason why German, as opposed to Dutch, allows for this kind of linear ordering, is probably because the morphology is rich enough to tell the thematic role. In that sense, Dutch is more 'configurational' than German. Only the position is able to identify the argumental status of a bare NP not inflected for Case. Anyway, what one can conclude from the German data is that topics move whereas comment elements stay. Topics are definitely more prominent than non-topics. As stated above, they act as anchor in the ongoing conversation. Thus they are prominent to the degree that they are used as point of departure for processing of new information. That means that we now have a case that somehow parallels Grimshaw's dilemma. We have two hierarchies that are ideally congruent, but not necessarily; thus we have two hierarchies that exist next to each other and are incompatible. Like in Grimshaw's case, one hierarchy wins the competition: here the topic prominence triumphs over the thematic prominence.

The preceding discussion delivers a picture where prominence relations govern the following hierarchies: argument structure is organized through prominence relations between thematic roles. This gives the instructions to create lexical projections. Clause structure is triggered by sentence functional perspective, or, in other words, as Vallduví would call it : information packaging. That means topics reflect their higher prominence with respect to comment parts by moving out of the VP. Their order with respect to one another outside VP is again determined by thematic hierarchy.

Now the story goes on. The last sentence of the preceding paragraph is actually not the complete truth. As the data presented till now (are supposed to) show, the order of arguments in the base position as well as in the derived one is IO > DO. That suggests that the order of topic NPs with respect to each other is fixed. This again, however, is true only for languages like Dutch. Languages that have the possibility to allow for crossing, i.e. fronting of a thematically deeper, but topic argument with respect to a thematically higher but comment element, also seem to allow for some freedom among scrambled elements. German again exemplifies that (negation shows that scrambling has taken place):

[13] After some time, the sentence does not sound bad any more. The meaning of the sentence, however, is completely different from (60).

(61) weil Lohengrin der geliebten Elsa seine Herkunft nicht preisgeben möchte
since Lohengrin the loved Elsa his descend not reveal wants
'since Lohengrin does not want to reveal his descent to his possible deliverer'

(62) weil Lohengrin seine Herkunft der geliebten Elsa nicht preisgeben möchte
since Lohengrin his descend the loved Elas not reveal wants

In a certain sense, Catalan shows similar behavior and can be analyzed in a similar manner. For the time being, the analysis I will suggest is not very fashionable, in face of Kayne's 'Antisymmetry of syntax' (1993). However; I will follow Vallduví in assuming that there is rightward movement in this language. And even more, I will propose a VP structure that in terms of X-bar structure completely mirrors the German VP (see (68)). Vallduví shows that in Catalan (semantic, informational) focus is also phonologically encoded, in that the deepest element in a structure carries the accent:

(63) L'amo odia el BRÒQUIL.
the boss hates the broccoli
'The boss hates BROCCOLI.'

This sentence is a canonical one, i.e. it is not restricted to a narrow focus reading on *bròquil* and allows for focus projection. Under my analysis *bròquil* necessarily belongs to the comment and should therefore not move. In constructions that differ from the one in (63) in terms of information packaging, things change. If it is clear between the communicants that there is some relation between the boss and broccoli, both arguments may function as topics. The relevant relation then constitutes the comment.

(64) L'amo l'ODIA, el bròquil.
the boss it +hates the broccoli
'The boss HATES broccoli.'

The following facts are in favor of a rightward movement analysis. First, Cinque (1993) argues that there are lots of languages that mark their deepest embedded element by assigning them neutral accent. In some sense, Vallduví argues for the same, i.e. he analyzes Catalan as such a language (without any reference to Cinque's work) . If this is adopted, the object in (64) cannot be the deepest element anymore since it the verb that carries phonological stress and delivers the relevant information of the sentence. Thus the object must have moved outside the c-command domain of the verb.[14]

The second argument is clitic doubling. Note that clitic doubling is ungrammatical in (63), but obligatory in (64). In Meinunger (1994), mainly following Suñer (1988), it has been argued that some clitics (especially those that occur in clitic doubling constructions) are best analyzed as agreement markers. Furthermore, it has been shown there that the presence of clitics triggers a topic reading of the doubled argument. If, in turn, clitic doubling is triggered by an argument in its relevant Spec,Agr position at something like S-structure, *el bròquil* in (64) should have moved there. This implies that [Spec,AgrO] is on the right. Such an analysis

[14] I have to confess that I am not particularly happy with the explanation. Actually I believe that there is some correspondence between stress assignment and the deepest embedded constituent's head. I think, however, that this is true for constituents only, i.e. the verb as X°-element has a syntactically fix position which should not vary whether the verb itself is stressed or not.

seems to be supported by constructions involving double object constructions. The neutral order in Catalan is verb > direct object > indirect object (/ directional argument), see (65). However, verb > indirect object (/ directional argument) > direct object is also possible. In that case then, the direct object must be clitic doubled, and additionally, the indirect one must constitute the focus, (66)[15] .

(65) No he donat encara les notes als ALUMnes.
not have1sg given yet the marks to-the students
'I haven't given the marks to the students yet.'

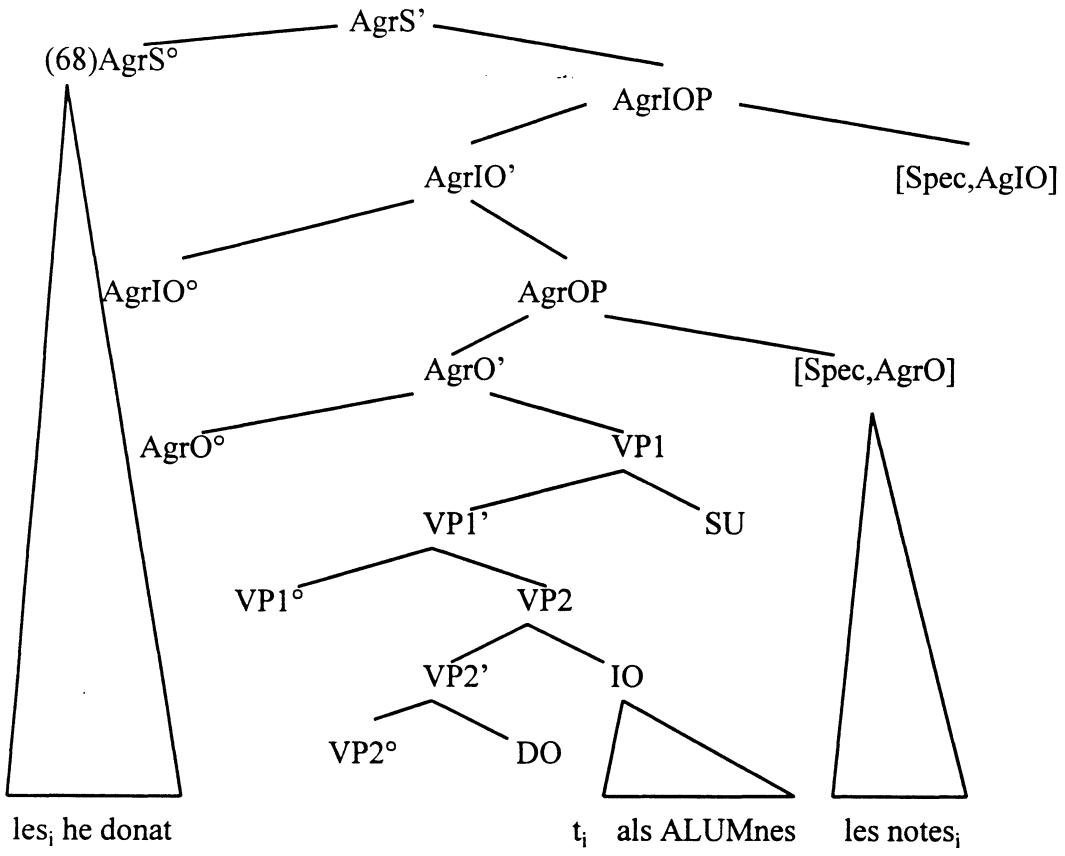
(66) No les_j he donat encara als ALUMnes les $notes_j$
not cl_{DO} $have_{1sg}$ given yet to the students the notes

Further evidence for the rightward movement analysis is provided by the placement of clause peripheral particles. These may not occur between the verb and its arguments unless they are defocused and clitic doubled, which we took both as indicator for movement (to the right).

(67a) Ficarem (*oi) el ganivet (*oi) al CALAIX, oi ?
put1pl prt the knife prt into-the drawer
'We'll put the knife in the drawer, right ?'
(67b) El_i ficarem t_i al CALAIX, **oi**, el $ganivet_i$, (oi) ?

Considering that Catalan overtly identifies the thematic role of arguments using (empty) prepositions, we get the mirror image of German.[16]

[15] Thanks to Josep Quer for the data.
[16] Interestingly, van Riemsdijk presented a theory the other day which reinforces the rightward hypothesis from a totally different perspective. In his talk 'Adjunktion und Adjazenz' held at FAS on June 11, he tried to show that verb cluster formation, i.e. the morphophonological conflation of verbal lexical and functional material depends on strict adjacency. The consequence of such a theory is that only languages which project complements and specifiers on the same (!) side of the head (heresy in Kaynians' eyes) should allow for inflected forms of verbs with bound inflectional morphemes. Thus all head final languages are good candidates, and they indeed behave as expected, normally. The other group would be languages that are completely head initial. The modern romance languages could be analyzed as such, and Catalan belongs to them. (Since in German as well as in Catalan, as representatives a larger group of more languages, the CP level is undoubtedly head medial, verbs that incorporate or adjoin to C° should be excluded, and interestingly, in matrix declarative sentences this step is forbidden. I don't know whether the theory holds. However, I find it interesting and promising, well maybe less interesting than Kayne's proposal, but more promising.)



Vallduví attributes that to information packaging. I also would like to argue that in German-like languages phrases are linearized according to their communicative contribution. This has been observed a long time ago. What we get is the German scrambling behavior, i.e. since in Catalan overt elements identify an argument's thematic role, the order of topic elements is rather free:

(69) Encara	no els$_i$ les$_j$ he donat t_i t_j,	les notes$_j$ als alumnes$_i$.
Encara	no els$_i$ les$_j$ he donat t_i t_j,	als alumnes$_i$ les notes$_j$.
Les notes$_j$ encara	no els$_i$ les$_j$ he donat t_i t_j,	als alumnes$_i$.
Als alumnes$_i$ encara	no els$_i$ les$_j$ he donat t_i t_j,	les notes$_j$.
Les notes$_j$ als alumnes$_i$ encara	no els$_i$ les$_j$ he donat t_i t_j.	
Als alumnes$_i$ les notes$_j$ encara	no els$_i$ les$_j$ he donat t_i t_j.	

The idea that linearization is triggered by communicative weight goes back to Behagel (1909) and his 'Gesetz der wachsenden Glieder'. An other pioneer is the Czech scholar Firbas. He proposes a theory of 'communicative dynamism' (mainly Firbas (1964)). This theory rejects a binary analysis of sentences that divides a sentence into a thematic and rhematic part[17]. His dynamism theory proposes that all phrases are part of a continuum that is a scale of communicative importance. For him there is an additional part to theme and rheme. He calls it transition. Its task is to mediate between the (proper) thematic and (proper) rhematic part of

[17] In the immense work on functional sentence perspective there are lots of proposals about how to split a sentence and then how to call the parts (theme-rheme, topic-comment, topic-focus, focus-open proposition, hearer old- hearer new...) For the details see Vallduví. What matters here is that this binary analysis is not adequate or at least not explanative enough.

the sentence. Thus the theme is constituted by the element(s) carrying the lowest degree of communicative dynamism within the sentence. Towards the end of the sentence the degree of information becomes higher, that means, the informative part is on the right. In-between, i.e. within the transition part, phrases are ordered according to their communicative weight, which I will call communicative prominence. Thus again prominence comes as an ordering principle. This ordering, however, should probably be analyzed as the product of A-bar-movement.

6.2. A place for Optimality Theory?

This is the right place to use another theoretical framework that has been introduced into syntax by the same linguist: Jane Grimshaw, namely optimality theory (Grimshaw (1993)). Roughly, optimality theory works like this: there are a number of constraints that evaluate some output of a grammatical process. These constraints state what output is good in a (grammatical sense). These constraints are autonomous, i.e. they do not depend on each another. Optimality theory says that these constraints are ranked. This means that that there is a language particular hierarchy of relevant constants. These may be obeyed or violated. The output of the relevant grammatical construct which violates the least comes out as grammatical. Thus where some language decides to give priority to one constraint over another, the output may violate the more deeply ranked one, but not the more highly ranked one. A different language may reverse the ranking and, consequently, the result must be different.

The ranking that I propose to account for the German type - Dutch type difference is not strictly parallel to the usual ranking. However, if we adopt the following constraints for word order, we have an explanation for the different behavior:

- ($\Theta\vartheta$) order constituents according to their thematic hierarchy
- ($\Pi\vartheta$) order constituents according to their weight with respect to communicative dynamism

Dutch prescribes ($\Theta\vartheta$) > ($\Pi\vartheta$). In German, the ranking is not rigid. It allows for either ordering, maybe with a preference for ($\Pi\vartheta$) > ($\Theta\vartheta$).
Thus, in a generalized fashion one could state, that one of the parametric differences between non-configurational languages and (more) configurational languages is the constraint ranking ($\Theta\vartheta$) > ($\Pi\vartheta$) vs. ($\Pi\vartheta$) > ($\Theta\vartheta$). The possibility for the latter ranking is probably dependent on the language having a rich enough case morphology.

7. Summary

It has been shown that many linguistic hierarchies are ordered by prominence in the sense of communicative salience. If thematic hierarchy is one of these and a requirement is imposed, namely that this hierarchy be preserved under case checking, we do not depend on the mysterious equidistance theory any longer.

References:

Adger, D. (1993) *Functional Heads and Interpretation*, Edinburgh diss.

Alexiadou, A. (1994) *On Adverbial Licensing*, FAS manuscript and talk, to appear in the proceedings of 15th Annual Meeting of the Linguistics Department of The University of Thesaloniki.

Baker, M. (1988) *Incorporation: A Theory of Grammatical Function Changing*. Chicago.

Behagel, O. (1909) *Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern.* Indogermanische Forschungen 25, 110-142.

Chomsky, N. (1992) *'A Minimalist Programm for Linguistic Theory'*, MIT Occational Papers in Linguistics.

Comrie, B. (1989) *Language Universals and Linguistic Typology*, 2nd edition.

Diesing, M. (1990) *Indefinites*, MIT Press, LI monographes.

Emonds, J. (1993) 'Projecting indirect objects' in *The Linguistic Review vol. 10*, No. 3, p. 211 - 265.

Firbas, J. (1964) *On Defining the Theme in Functional Sentence Analysis*. TLP 1.

Givón, T. (1976) 'Topic, Pronoun, and Grammatical Agreement.' in Li, Ch. & S. Thompson (eds.) *Subject and Topic*. Academic Press New York.

Grimshaw, J. (1990) *Argument Structure*. MIT LI monographes.

Grimshaw, J. (1993) *Minimal Projection, Heads, and Optimality*, manuscript, Rutgers University.

Haegeman, L. (1992) *Negative Heads and Negative Operators: The Neg Criterion*, manuscript, University of Geneva.

Haftka, B. (1980) Bewußtseinspräsenz und aktuelle Gliederung von Äußerungen, in *Linguistische Studien 68*, Akademie der Wissenschaften der DDR, Zentralinst. f. Sprachwissenschaft.

Hoekstra, E. (1993) *Licensing Conditions on Phrase Structure*, Groningen Diss.

de Hoop, H. *Case Configuration and NP Interpretation*, Groningen Diss.

Jäger, G. (1994) *Der Stadien-/Individuenkontrast als Diskursfunktion*, Reaserch proposal, MPG Berlin.

Jonas, D. & J.D. Bobaljik (1993) 'Specs for Subjects: The role of TP in Icelandic'. manuscript, to appear in *MIT Working Papers in Linguistics vol. 19. Papers on Case and Agreement.*

Kayne, R. (1993) *'The Antisymmetry of Syntax'*, CUNY manuscript.

Lenerz, J. (1977) *Zur Abfolge nominaler Satzglieder im Deutschen*, Studien zur Deutschen Grammatik.

Mahajan, A. (1993) The Ergativity Parameter: have-be alternation, Word order and Split Ergativity, to apper in the *Proceedings of NELS 1993.*

Meinunger, A. (1994) *'Case Configuration and Referentiality'* to appear in the Proceedings of CONSOLE II.

Neeleman, A. (1994) *Complex Predicates*, OTS Diss.

Partee, B et al. (199) *Mathematical Methods in Linguistics.*

Rizzi, L. (1990) *Relativized Minimality*. MIT LI monographes.

Speas, M.(1990) *Phrase Structure in Natural Language*. Studies in Natural Language and Linguistic Theory.

Suñer, M. (1988) 'The Role of Agreement in Clitic Doubled Constructions' in *Natural Language and Linguistic Theory 6*, 391-434.

Sportiche, D. (1990) *Movement, Agreement and Case*, UCLA manuscript.
Szabolcsi, A. & F. Zwarts (1991) *Unbounded Dependencies and Algebraic Semantics*. Course material. Summer school LLI 1991.
Szabolcsi, A. & F. Zwarts (1993) 'Weak Islands and an Algebraic Semantics for Scope Taking', in *Natural Language Semantics*, Vol. 1.
Vallduví, E. (1992) *The Informational Component*. Outstanding dissertations in Linguistics.

Potential verbs in German: the emergence of a productivity gap

Renate Raffelsiefen
Heinrich-Heine Universität, Düsseldorf

This paper explores the theoretical consequences of two parallel changes in the historical development of German.[1] As the result of a change in the phonological wellformedness conditions for verbs all dactylic feet were systematically reduced to trochaic feet by schwa deletion (e.g. MHG ât[ə]m[ə]n > NHG atm[ə]n 'to breathe'). Simultaneously words ending in a schwa syllable closed by non-liquid consonants were excluded from the domain of the highly productive morphological rule of verbalization.

I argue that the emergence of the morphological gap does not reflect a change in the subcategorization requirements of the verbal suffix, but rather is a direct consequence of the autonomous change in the prosodic wellformedness conditions for verbs. On this view the set of potential verbs are those phonologically wellformed nonce verbs which can be related to a base in a phonologically transparent manner. Assuming that the (surface oriented) phonological relatability-conditions remain constant, a gap in the domain of verbalization with precisely the phonological characteristics stated above is predicted to accompany the changes in prosodic structure which marked the transition from MHG to NHG. The theoretical significance of the data thus concern the morphology-phonology interface and in particular the notion of the input.

The paper is structured as follows. In section 1 I describe the prosodic wellformedness conditions for verbs in NHG informally. In section 2 I describe the prosodic wellformedness conditions in MHG in terms of constraints showing how a reranking of those constraints would yield the current patterns. The description is based on Vennemann's wellformedness conditions for syllable structure (cf. Vennemann 1982, 1988), which are formalized within Optimality Theory (cf. Prince and Smolensky 1993). In section 3 I describe the conditions for verb formation in German. In section 4 I discuss the question of why a gap in the domain of verbalization results from historical changes in the prosodic wellformedness conditions of verbs.

---

[1]Earlier versions of this paper (cf. Raffelsiefen 1994) were presented in Cologne, Potsdam, Wuppertal, Düsseldorf, and Berlin. I thank the audiences for many helpful comments, in particular I am grateful to Manfred Bierwisch, Peter Eisenberg, Stig Eliasson, Tracy Alan Hall, Joachim Jacobs, Paul Kiparsky, Martin Neef, Roland Noske, Heinz Vater, Theo Vennemann, Richard Wiese, and Wolfgang Ullrich Wurzel. Also I would like to thank Caroline Féry and Bruce Straub for discussion and Rüdiger Harnisch for a letter with detailed comments.

1. Wellformed verbs in NHG[2]

German infinitives are subject to two general conditions on phonological wellformedness. They always end in the alveolar nasal [n], and with the exception of the two high-frequency verbs tun 'to do' and sein 'to be',[3] they always end in exactly one schwa syllable.[4] Historically, this prosodic restriction is the result of a series of changes including 'schwa epenthesis', 'schwa deletion' and 'metathesis'.[5]

(1)	Middle High German:	New High German:	
	holn ~ hol[ə]n	hol[ə]n	'to fetch'
	ât[ə]m[ə]n	atm[ə]n	'to breathe'
	hag[ə]l[ə]n ~ hag[ə]ln ~ hagl[ə]n	hag[ə]ln	'to hail'

The historical changes illustrated in (1) not only 'conspired' to bring about a uniform prosodic shape, that is, the occurrence of exactly one final schwa syllable in Modern High German. In addition, there is no longer any variation in the position of the schwa. That is, for any given sequence of postvocalic consonants there is generally only one possible site for the schwa, regardless of the stress contour of the

---

[2]For a criticism of relevant descriptions in Lexical Phonology (cf. Giegerich 1987, Wiese 1986, 1988, Hall 1992), see Raffelsiefen 1995.

[3]The verb tun is often pronounced with a final schwa syllable in colloquial speech (i.e. tu[ə]n).

[4]Phonetically, the sequence schwa plus sonorant is often realized as a syllabic sonorant or as a vocalic r respectively, as shown in (ib):

(i)a.	red[ə]n	b.	red[n̩]	reden 'to talk'
	wick[ə]ln		wick[l̩]n	wickeln 'to wind'
	stolp[ə]rn		stolp[r̩]n	stolpern 'to stumble'
	geb[ə]n		geb[m̩]	geben 'to give'
	reg[ə]n		reg[ŋ̩]	regen 'to move'

The place of articulation of a syllabic nasal is always identical to that of the preceding consonant. The description presented here is based on the perhaps overarticulated variants in (ia). The variants in the b column, as well as other register-dependent variants, can be derived from the representations in column a.

[5]Orthographically the schwa is always represented by the grapheme <e>.

verb or its internal morphological structure.[6] For example, the position of the schwa in the verb hageln in (1) is mirrored by every other verb in which the last full vowel is followed by the consonants g, l, and n. Examples are given in (2):

(2) XV**gln**: XV**g**[ə]**ln**

nág[ə]**ln** 'to nail'	fräg[ə]**ln** 'to ask cunningly'
schúrig[ə]**ln** 'to bully'	schmúgg[ə]**ln** 'to smuggle'
móg[ə]**ln** 'to cheat'	máßrèg[ə]**ln** 'to reprimand'

Regarding the question of what determines the site of the schwa we find that almost all patterns follow from two principles none of which is specific to verbs. The first of these principles concerns the wellformedness conditions on sonority relations within the syllable stated in (3):

---

[6]Investigating the verbs listed in Muthmann's reverse dictionary we find a total of 323 distinct sequences of consonants following the last full vowel (affricates are considered bisegmental). For all but one of those sequences the position of the schwa is fixed. In the table in (i) the verbs are classified according to the sounds which follow the schwa; consonants preceding the schwa are represented by C_o:

(i)	Position of schwa:	Number of patterns:	Example patterns:	Example verbs:	gloss:
a.	XVC_o[ə]n	156	XV**rl**[ə]**n**	quirl[ə]n	'to whisk'
b.	XVC_o[ə]rn	88	XV**nt**[ə]**rn**	kent[ə]rn	'to capsize'
c.	XVC_o[ə]ln	74	XV**pst**[ə]**ln**	herbst[ə]ln	'to *autumn*'
d.	XVC_o[ə]rln	6	XV**s**[ə]**rln**	buss[ə]rln	'to kiss'

The verbs illustrated in (id) are not only rare, but are also considered dialectal by many. Nonetheless, whether or not hearers accept such verbs as part of the standard language they will clearly reject any alternative positions for the schwa. The examples in (id) thus support the claim that for any given sequence of postvocalic consonants there is only one possible site for the schwa. The only exception to this generalization concerns verbs in which the last full vowel is followed by the consonants r and n. For such verbs the schwa usually precedes the final n (cf. fahr[ə]n 'to drive', probier[ə]n 'to try') but in the two verbs näh[ə]rn 'to approach' and wieh[ə]rn 'to neigh' the schwa precedes the r. These two verbs (along with tun and sein, cf. p. 2) are the only counterexamples to the wellformedness conditions formulated here.

(3) SON
A sonorant in the syllable onset may only be followed by segments of higher sonority; a sonorant in the syllable coda may only be preceded by segments of higher sonority.

According to (3) for every sonorant in the syllable shell (i.e. head and coda) the sonority level must increase toward the nucleus. The constraint in (3) is consistent with both Vennemann's 'Head Law' and 'Coda Law', which say that syllable heads and codas are the more preferred the more sharply the sonority increases towards the nucleus (1988:13ff). The sonority hierarchy with reference to which the constraint in (3) is evaluated is given in (4):

(4) increasing sonority decreasing sonority
<--->

Vowel	r i̯ u̯ y̯	l	m n ŋ	Obstruent

The ranking among r, l, and the nasals in (4), is consistent with the 'discontinuous gradings' of sonority proposed by Sievers on auditory grounds (1901:198f).[7] According to the table in (4) sonorants must precede obstruents in coda position and the only permissible sonorant clusters are rN, lN, rl, and rlN. Any other combinations would constitute 'sonority violations' and are obligatorily 'broken up' by the schwa as shown in (5):

(5)a.	Schem[ə]l]$_N$ 'stool'	b. *Sche**ml**
	wied[ə]r]$_{ADV}$ 'again'	*wie**dr**
	At[ə]m]$_N$ 'breath'	***Atm**
	Ab[ə]nd]$_N$ 'evening'	*Ab**n**d
	alb[ə]rn]$_A$ 'silly'	*al**br**n
	Am[ə]n]$_{INTERJ}$ 'amen'	***Amn**
	hund[ə]rt]$_{NUM}$ 'hundred'	*hun**dr**t
	MacDon[ə]ld]$_{NAME}$	*Macdo**nl**d

Since the schwa forms an additional syllabic nucleus all words in (5a) satisfy the condition in (3). Without the schwa

---

[7]There seems to be general agreement among phonologists working on German that r is more sonorous than l which in turn is more sonorous than the nasals. The overall structures of the hierarchies proposed, however, differ considerably (cf. Vennemann (1982:284), Strauss (1982:97), Hall (1992:64)).

these words would violate (3) as a result of including the boldfaced clusters in coda position as is shown in (5b).

Consider next the ranking of glides in the sonority hierarchy given in (4). This classification is based on the distribution of schwa following diphthongs:

(6)a.	faul 'lazy'	b.	Mau[ə]r 'wall'
	Geheul 'howling'		teu[ə]r 'expensive'
	Pfeil 'arrow'		Fei[ə]r 'celebration'

On the assumptions that a) German diphthongs consist of a vowel in nucleus position followed by a glide, that is, a high vowel in coda position (i.e. [au], [ai], and [ɔy])[8], and b) that (3) holds, the data in (6a) indicate that glides are more sonorous than l. The fact that a diphthong is never followed directly by r in coda position indicates that glides are equally (or less) sonorous than r.[9] The sonority table in (4) is accordingly consistent with the data considered so far.

The generalization in (3) along with the table in (4) also rule out the occurrence of adjacent identical sonorants, which in fact are broken up by a schwa as well (cf. Pfarr[ə]r 'priest' and Lein[ə]n 'linen', etc.). Coda clusters involving obstruents as second members are not regularly broken up by a schwa regardless of the sonority relation within the cluster. This is the reason for restricting SON to sonorants as formulated in (3).

With reference to the sonority constraint in () one can state the generalizations determining the site of the schwa in NHG verbs concisely. First, when the consonants following the last full vowel include a single cluster in which sonority fails to decrease, that cluster is 'broken up'. Examples are given in (7):

---

[8]Sievers claims that only high vowels can appear in coda position which is probably related to the fact that sonority in vowels decreases with hight (1901:204). On the assumption that sonority relations are universally constant but that, in individual languages, speechsounds may range over adjacent slots of the sonority hierarchy it would follow that only the least sonorous vowels can also appear in the slot for (the most sonorous) consonants.

[9]For evidence that glides (i.e. high vowels) and the r occupy the same slot in the sonority hierarchy for Icelandic see Vennemann (1988:51f).

(7)

Consonantal Pattern	Sonority Violation:	Example:
XVgln	*gl	ha**g**[ə]**l**n 'to hail'
XVgrn	*gr	zö**g**[ə]**r**n 'to hesitate'
XVrntn	*tn	ern**t**[ə]**n** 'to harvest'
XVnzln	*zl	häns[ə]**l**n 'to tease'
XVsrln	*sr	bus**s**[ə]**r**ln 'to kiss'
XVurn	*ur	da**u**[ə]**r**n 'to last'
XVymdn	*dn	verleum**d**[ə]**n** 'to slander'
XVrpstln	*tl	herbs**t**[ə]**l**n 'to turn fall'
XVlpsn	*sn	rülp**s**[ə]**n** 'to burp'
XVnstrln	*tr	fens**t**[ə]**r**ln 'to visit a lover by climbing through his or her bedroom window'

Infinitives such as ha**gl**[ə]n, zö**gr**[ə]n, etc. are thus impossible in German, although they do not violate sonority.

In all NHG verbs involving two potential sonority violations the two violations always overlap in that they share a sonorant (i.e. the boldfaced sonorant in (8)). In such cases the schwa always breaks up the rightmost sonority violation:

(8)

Consonantal Pattern	Sonority Violations:	Example:
XVtmn	*t**m**, *m**n**	at**m**[ə]**n** 'to breathe'
XVgnn	*g**n**, *n**n**	reg**n**[ə]**n** 'to rain'
XVknn	*k**n**, *n**n**	trock**n**[ə]**n** 'to dry'
XVšlrn	*š**l**, *l**r**	tisch**l**[ə]**r**n 'to do woodwork'
XVkslrn	*s**l**, *l**r**	drechs**l**[ə]**r**n 'to work the lathe'
XVmpnrn	*p**n**, *n**r**	klemp**n**[ə]**r**n 'to do plumbing'
XVrtnrn	*t**n**, *n**r**	gärt**n**[ə]**r**n 'to garden'

Third, in the absence of potential sonority violations the schwa immediately precedes the final n as is shown in (9):

(9)

Consonantal Pattern	Sonority Violation	Example:
XVn	-	sä[ə]n 'to sow'
XVln	-	hol[ə]n 'to get'
XVrn	-	hör[ə]n 'to hear'
XV̂in	-	schnei[ə]n 'to snow'
XV̂un	-	hau[ə]n 'to slap'
XVrln	-	quirl[ə]n 'to whisk'
XV̂uln	-	faul[ə]n 'to rot'
XV̂yln	-	heul[ə]n 'to cry'

The conditions determining the site of the schwa in German verbs are summarized in (10):

(10)a. Given one potential sonority violations, the schwa breaks it up. Exceptions: none.
b. Given two 'overlapping' sonority violations, the schwa breaks up the rightmost one. Exceptions: none.
c. In the absence of sonority violation, the schwa appears before the last segment. Exceptions: tun, sein, nähern, wiehern.

The generalisations in (10) raise the question of why there are no examples with nonoverlapping sonority violations. This question concerns the notion of the input and will be addressed in section 4.

Note that the generalizations in (10) make no reference to the internal morphological structure of verbs. In particular, the site of the schwa in the verbs in (11) is not influenced by the phonological structure of their respective (etymological) base:

(11)

	Base:
regn[ə]n 'to rain'	< Reg[ə]n 'rain'
atm[ə]n 'to breathe'	< At[ə]m 'breath'
zent[ə]rn 'to center'	< Zentrum 'center'
knäul[ə]n 'to squeeze sth. into a ball'	< Knäu[ə]l 'ball'

Consider also the morphological rule of "l-Infixation" in verbs illustrated in (12). Generally the l directly precedes the final n as shown in (12a). However, if the segment preceding the final n in the base is more sonorous than l we find the pattern in (12b):

(12)		(etymological) base:
a.	äug[ə]**l**n 'to glance secretly'	< äug[ə]n 'to look searchingly'
	streich[ə]**l**n 'to caress'	< streich[ə]n 'to stroke'
	kusch[ə]**l**n 'to snuggle up'	< kusch[ə]n 'to knuckle under'
	dräng[ə]**l**n 'to jostle'	< dräng[ə]n 'to push; to press'
b.	krau**l**[ə]n 'to fondle'	< (†)krau[ə]n 'to scratch'
	grau**l**[ə]n 'to be scared'	< grau[ə]n 'to be terrified'
	wur**l**[ə]n 'to swarm'	< †wurr[ə]n 'to roar'
	†wir**l**[ə]n 'to surge confusedly'	< wirr[ə]n 'to surge confusedly'

In the examples in (12a) the position of the schwa is determined by rule (10a); in the examples in (12b) it is determined by rule (10c).

2. From MHG to NHG: a constraint-based analysis.

2.1. Basic assumptions

MHG schwas are largely the result of a process of vowel reduction in unstressed syllables which characterizes the transition from OHG to MHG:[10]

(13)	OHG	MHG	gloss:
	géban	geb[ə]n	'to give'
	hábe:n	hab[ə]n	'to have'
	hólo:n	hol[ə]n, holn	'to call'
	mángolo:n	mang[ə]l[ə]n, mangl[ə]n, mang[ə]ln	'to lack'
	á:tamo:n	a:t[ə]m[ə]n	'to breathe'

In MHG we find variation for some verbs (e.g. mang[ə]l[ə]n, mangl[ə]n, mang[ə]ln), but not for others (e.g. a:t[ə]m[ə]n). The same type of variation extends also to new verbs in MHG (hag[ə]l[ə]n ~ hagl[ə]n ~ hag[ə]ln, but wa:p[ə]n[ə]n ~ *wa:pn[ə]n) which shows that the distribution of schwa, including the variation, was rule-governed. The description of the relevant rules to be presented below is

[10]The change in (13) indicates a reranking of constraints to the effect that *SCHWA is dominated by some constraint on foot structure which will not be discussed here.

based on the assumption that there is a general constraint against schwa as shown in (14):[11]

(14) *SCHWA

The constraint in (14) implies that schwas are unstable unless their presence is needed in order to satisfy some equally high or higher ranking constraint. For example, the assumption that SON as stated in (3) dominates *SCHWA explains the stability of the schwa in geb[ə]n 'to give' as is shown in (15):

(15)			SON	*SCHWA
		gebn	*!	
	→	geb[ə]n		*

For now it will be assumed that inputs consist of morphological structures which include abstract stems. For example, the candidates in (15) are generated by some context-free epenthesis rule based on the input geb+n, where n is the infinitival suffix. The constraint-based evaluation of the candidates amounts to a phonological interpretation of the input in accordance with the standard view of the morphology-phonology interface in Generative Grammar.

The description in (15) refers exclusively to phonological constraints which concern the wellformedness of syllable structure. The constraint SON prohibits certain sequences of sounds in the syllable coda and onset, whereas *SCHWA refers to the structure of nuclei. As will be shown, the entire range of schwa patterns in MHG verbs can be described exhaustively in terms of an interaction of constraints all of which refer to syllable structure. In order to ensure that the constraints are independently motivated they are consistent with Vennemann's (1988) 'Preference Laws for Syllable Structure'.

2.2. The list of constraints

Apart from SON which constrains the wellformed sequence of speech sounds in the syllable, there are also preferences for noncomplex heads and codas. These constraints are consistent with Vennemann's Head Law and Coda Law, respectively:

(16)a. HEAD
A syllable head must be neither empty nor complex.

b. CODA
A syllable coda must not be complex.

---

[11]This constraint is also proposed by Mester and Ito (1994).

Evidence for restricting syllable codas with respect to complexity rather than requiring empty codas will be presented in the next section.

Apart from constraining the complexity of syllable heads and codas there is also a constraint against complex rhymes. That constraint refers not to speech sounds but rather to moras and is stated in (17) (Cf. Vennemann's "Weight Law" 1988:30ff[12]):

(17) WEIGHT
The maximal weight of a syllable is three moras.

Consider next the wellformedness condition for syllable structure referred to as 'Shell Law' in Vennemann (1988:11). This law says that identical speech sounds flanking the nucleus are disfavored, especially if they are not the only ones within their margin. The Shell Law is expressed in the following constraint:

(18) SHELL
*CC_iVC_j$$, where C_i and C_j are equally sonorous.

Finally there is a constraint which does not refer to intrasyllabic structure but rather to the wellformedness of syllable contacts. According to the "Syllable Contact Law" proposed by Vennemann (1982, 1988) the syllable contact A.B is the more preferred, the more sonorous the offset A and the less sonorous the onset B. The 'Contact Law' is adopted here in the following formulation:

(19) CONTACT
In a syllable contact A.B, A must be more sonorous than B.

The possible domains for syllabic wellformedness are hereby exhausted. As will be shown in the next section all constraints listed above play a role in MHG verb prosody.

2.3 Wellformed verbs in MHG

As was noted before MHG differs from NHG in that certain verbs show free variation in the position of the schwa. Variants such as those in (20) are often attested to even within a single document (cf. Moser & Stopp (1970:84ff):

---

[12]The 'Weight Law' says that "in stress accent languages an accented syllable is the more preferred, the closer its syllable weight is to two moras". Limiting the constraint 'WEIGHT' to two moras would only yield the correct results if the last consonant in a word was ignored.

(20) hageln ~ haglen ~ hagelen 'to hail'
handeln ~ handlen ~ handelen 'to handle'
wundern ~ wundren ~ wunderen 'to wonder'
toppeln ~ toplen ~ topelen 'to play at dice'
kîfeln ~ kiflen ~ kifelen 'to chew'
mangeln ~ manglen ~ mangelen 'to lack'
sateln ~ satlen ~ satelen 'to saddle'
roteln ~ rötlen ~ röttelen 'to play the *rotte*[13]'
rütteln ~ rütlen ~ rütelen 'to shake'

The cooccurrence of the variants in (20) is accounted for by the constraint ranking in (21).:

(21)			SON	*SCHWA	CONTACT	HEAD	CODA
		hagln	*!				
	→	ha.g[ə].l[ə]n		**			
	?→	hag.l[ə]n		*	*		
	?→	ha.gl[ə]n		*		*	
	→	ha.g[ə]ln		*			*

The "!" marks a fatal constraint violation. Unorderedness among constraints is represented by the dotted lines. Unorderedness among *SCHWA, CONTACT, HEAD, and CODA results in a tie among four candidates thereby expressing the variation observed in (20). Tshe question of whether the written form <haglen> represents hag.l[ə]n, ha.gl[ə]n, or maybe both variants, cannot easily be decided. We will return to this question below.

Characteristic for the verbs in (20) is the presence of a potential contact violation (i.e. <g,l>) followed by a less sonorous final segment (i.e. <l,**n**>). Verbs in which two consonants potentially forming a wellformed contact (i.e. <r,k>) are followed by a more sonorous final segment (i.e. <k,**n**>) yield a single optimal candidate as the example merken 'to watch' in (22) illustrates:

(22)			SON	*SCHWA	CONTACT	HEAD	CODA
a.		merkn	*!				
		me.r[ə]kn	*!				
		me.rk[ə]n	*!				
	→	mer.k[ə]n		*			
		me.r[ə].k[ə]n		**!			

[13]A "rotte" is a musical instrument.

Given a decrease in sonority (i.e. <r,k>) followed by an increase (i.e. <k,n>) we find that for every candidate which violates *SCHWA twice, there will always be a candidate which violates *SCHWA only once without incurring additional violations. Variants with two final schwa syllables such as mereken are accordingly always eliminated. Additional examples are given in (23):

(23)	wërfen	*wërefn, *wërefen	'to throw'
	trinken	*trinekn,*trineken	'to drink'
	hëlfen	*hëlefn, *hëlefen	'to help'
	warnen	*warenn, *warenen	'to warn'
	firmen	*firemn, *firemen	'to make firm'
	knarschen	*knareschn, *knareschen	'to grind one's teeth'
	smirken	*smirekn, *smireken	'to be rancid'
	slurken	*slurekn, *slureken	'to swallow'

The absence of ordering between the last four constraints in (22) does not clear the way to a random variety of forms but accounts just for the kind of variation which is attested to.[14]

Consider next the variants in (24):

(24)	kebsen ~ kebesen	*kebesn	'to commmit adultery'
	rîchsen ~ rîchesen	*rîchesn	'to govern'
	houpten ~ houbeten	*houbetn	'to decapitate'
	markten ~ marketen	*marekten, *marketn	'to trade'
	vögten ~ vögeten	*vögetn	'to protect'
	lëchzen ~ lëchezen	*lëchezn	'to dry out'
	rofzen ~ roffezen	*roffezn	'to belch'
	guckzen ~ guckezen	*guckezn	'to stare'

In order to account for the occurrence of the dactylic variants in (24) the sonority hierarchy in (4), in particular the ranking among obstruents, must be specified further. According to the table in (25) fricatives and affricates are more sonorous than stops, which is consistent with Sievers' description (cf. Sievers 1901:205).

---

[14]The dactylic forms in (i) are not accounted for by the tableau in (22). All counter-examples involve nasals followed by nonhomorganic obstruents.

(i)	vremeden ~ vremden	'to alienate'
	(er)lemeden ~ lemden	'to make lame'
	baneken ~ banken	'to romp about'
	ruomesen	'to brag'
	trumeten	'to trumpet'

The existence of such variants indicates perhaps that a constraint ruling out clusters of nasals and nonhomorganic obstruents ranks as high as *SCHWA in MHG.

(25)

Vowel	r	l	Nasal	Fric./Affr	Stop

Reference to the higher sonority of fricatives vis-a-vis stops is crucial in order to explain the occurrence of dactylic variants in (24), but not in (26):

(26)

zaspen	*zasepen, *zasepn	'to drag one's feet'
haften	*hafeten, *hafetn	'to stick'
vristen	*vriseten, *vrisetn	'to delay'
vrêsken	*vrêseken, *vrêsekn	'to find out'
lispen	*lisepen, *lisepn	'to lisp'
veiJten	*veiJeten, *veiJetn	'to fatten'

All potential syllable contacts in the verbs in (26) are wellformed which is not true for the verbs in (24). As a result, a trochaic candidate is invariably superior to dactylic candidates in (26), but not in (24) as is shown by the next tableau:

(27)			SON	*SCHWA	CONTACT	HEAD	CODA
a.		kebsn	*!				
		ke.b[ə]sn	*!				
	→	keb.s[ə]n		*	*		
	→	ke.bs[ə]n		*		*	
	→	ke.b[ə].s[ə]n		**			
		zaspn	*!				
		za.s[ə]pn	*!				
	→	zas.p[ə]n		*			
		za.sp[ə]n		*		*!	
		za.s[ə].p[ə]n		**!			

Consider next the variation, versus lack thereof, in verbs which do not potentially violate SON.

(28)

a.			b.		
varn ~ var[ə]n	'to go'		hôr[ə]n *hôrn	'to hear'	
holn ~ hol[ə]n	'to call'		vall[ə]n *valln	'to fall'	
wern ~ wer[ə]n	'to last'		teil[ə]n *teiln	'to share'	
steln ~ stel[ə]n	'to steal'		vüer[ə]n *vüern	'to lead'	

Verbs with a short vowel followed by a single (e.g. nongeminate) liquid typically show free variation between mono- and bisyllabic forms as is illustrated in (28a). By contrast, verbs with a long vowel or a geminate liquid are never monosyllabic. This observation indicates that WEIGHT ranks at least as high as *SCHWA:

(29)			SON	WEIGHT	*SCHWA	CONTACT	HEAD	CODA
	→	varn						*
	→	va.r[ə]n			*			
		va.ern			*			*
		hôrn		*				*
	→	hô.r[ə]n			*			
		hô.[ə]rn			*			*

The fact that varn alternates with varen supports the claim that *SCHWA and CODA are unordered (cf. the analysis of the variants in (21)).

The ranking in (29) predicts that in every verb which potentially violates WEIGHT but not SON the schwa precedes the wordfinal n. The lack of variation in the following verbs is thereby accounted for:

(30) bërlen *bërln, *bëreln, *bërelen 'to decorate'
murlen *murln, *mureln, *murelen 'to murmur'
turlen *turln, *tureln, *turelen 'to be dizzy'

The starred variants in (30) are ruled out because each of them involves more violations than the actual form:

(31)			SON	WEIGHT	*SCHWA	CONTACT	HEAD	CODA
		bërln		*				**!
		bë[ə]rln		*	*			**!
		bë.r[ə]ln			*			*!
	→	bër.l[ə]n			*			
		bë.r[ə].l[ə]n			**!			

The fact that bër.l[ə]n, rather than bë.r[ə]ln, is the actual form is the main motivation for defining CODA with respect to complexity rather than requiring syllable codas to be empty.

If neither WEIGHT nor SON are violated, monosyllabic variants are always possible. In fact, the constraint ranking in (31) predicts that there is one case in which the optimal candidate must be monosyllabic, that is, all verbs where the wordfinal n is the only postvocalic consonant:

(32)			SON	WEIGHT	*SCHWA	CONTACT	HEAD	CODA
	→	gân						
		gâ[ə]n			*!			

Other relevant examples are MHG stân but *stâ[ə]n 'to stand', sîn but *sî[ə]n 'to be'.

Apart from the case illustrated in (32) where verbs are obligatorily monosyllabic there are also cases where verbs are obligatorily dactylic. Consider the examples in (33):

(33)a.	tugenden	*tugneden, *tugnden, *tugendn	'to lend virtue to
	abenden	*abneden, *abnden, *abendn	'evening comes'
	jugenden	*jugneden, *jugnden, *jugendn	'to be youthful'
	zëhenden	*zëhneden, *zëhnden, *zëhendn	'to give a tenth'
	rêterschen	*rêtreschen, *rêtrschen, *rêterschn	'to puzzle'
	ritterschen	*rittreschen, *rittrschen, *ritterschn	'to militarize'
b.	hilderlen	*hildrelen, *hildrlen, *hilderln	'to nag'
	vingerlen	*vingrelen, *vingrlen, *vingerln	?
	lecherlen	*lechrelen, *lechrlen, *lecherln	'to smile'

The examples in (33) differ in that those in (33a) require two schwas in order to avoid any violations of SON whereas those in (33b) require one schwa for that purpose. The nonvarying dactylic forms in both cases are accounted for by the tableau in (34):

(34)			SON	WEIGHT	*SCHWA	CONTACT	HEAD	CODA
		tugndn	**!					
		tugn.d[ə]n	*!					
		tug.nd[ə]n	*!					
		tug.n[ə].d[ə]n			**	*!		
		tu.gn[ə].d[ə]n			**		*!	
	→	tu.g[ə]n.d[ə]n			**			
		hildrln	*!					
		hildr.l[ə]n	*!					
		hild.rl[ə]n	*!					
		hil.d[ə]rln		*	*			**!
		hil.dr[ə]ln			*		*	*!
	→	hil.d[ə]r.l[ə]n			**			

Consider finally the verbs in (35):

(35)	rëgenen	?rëgnen	*rëgenn	'to rain'
	âtemen	?âtmen	*âtemn	'to breathe'
	morgenen	?morgnen	*morgenn	't procrastinate'
	krademen	?kradmen	*krademn	'to make a noise'
	wolkenen	?wolknen	*wolkenn	'to be full of clouds'
	ëbenen	?ëbnen	*ëbenn	'to level'
	truckenen	?trucknen	*truckenn	'to dry'

The verbs in (35) differ from those in (20) in that the dactylic variants appear to be preferred in MHG although trochaic variants for some verbs are also attested to. Preference for dactylic forms is not expressed in the tableau in (36):

(36)			SON	WEIGHT	*SCHWA	CONTACT	HEAD	CODA
		rëgnn	**					
		rë.g[ə]nn	*					
	→	rëg.n[ə]n			*	*		
	→	rë.gn[ə]n			*		*	
	→	rë.g[ə].n[ə]n			**			

How do the verbs in (35) differ from those admitting both trochaic and dactylic variants considered earlier? In particular, on what basis could the trochaic candidates in (36) be ruled out?

Consider first the trochaic candidate rë.gn[ə]n. Unlike trochaic variants of the verbs in (20) and (24) that candidate can be eliminated by ranking SHELL al least as high as *SCHWA. However, in order to eliminate the other trochaic candidate, e.g. rëg.n[ə]n, CONTACT must rank higher than *SCHWA. Such a ranking would imply that variants such as haglen are admitted with the syllabification ha.glen only. The question of whether or not such a conclusion is desirable calls for further phonological investigations.[15] The tableau in (37) rules out the trochaic variants of the verbs in (35) whereas the tableau in (36) allows for them. Both descriptions account for the wellformedness of dactylic variants.

[15]Conceivably reference to processes of vowel lengthening in open syllable could help decide the issue.

(37)			SON	CONTACT	WEIGHT	SHELL	*SCHWA	HEAD	CODA
		rëgnn	**!						
		rë.g[ə]nn	*!						
		rëg.n[ə]n		*!					
		rë.gn[ə]n				*	*	*	
	→	rë.g[ə].n[ə]n					**		

It can be concluded then that although there is some uncertainty regarding the ranking among CONTACT and *SCHWA, the interaction between various syllable structure constraints and *SCHWA accurately describes the prosodic form of MHG verbs. In particular there is no need to refer to the internal morphological structure of verbs. Most importantly, the constraint-based description allows for a straightforward account of the seemingly complex changes of prosodic form characterizing the transition to NHG discussed in the next section.

2.4. The transition to NHG

As was noted above the changes from MHG to NHG in the prosodic form of infinitives "conspired" to yield trochaic forms only where for each verb the schwa is in a fixed position. Consider first the fact that among the three or four variants of the verb <u>hagelen</u> in MHG, only <u>ha.geln</u> is left in NHG. This observation suggests the following reranking of constraints:

(38)			SON	*SCHWA	CONTACT	HEAD	CODA
		<u>hagln</u>	*!				
		ha.g[ə].l[ə]n		**!			
		hag.l[ə]n		*	*!		
		ha.gl[ə]n		*		*!	
	→	ha.g[ə]ln		*			*

The illformedness of <u>hagelen</u> in NHG indicates that *SCHWA ranks higher than CONTACT, HEAD, and CODA.[16] The preference of <u>hageln</u> over *<u>haglen</u> shows that CODA ranks lower than CONTACT and HEAD. The analysis in (38) does not indicate a specific

---

[16]Except for CONTACT all constraints in (38) are identical to the corresponding constraints in MHG. Evidence for restricting CONTACT to sonorants in NHG comes from verbs in which a stop-fricative cluster is followed by a liquid-nasal cluster (cf. [vɛksəln] <u>wechseln</u> 'to change'). The reader may convince herself that the ungrammatical form *[vɛkslən] would emerge as optimal candidate if CONTACT would also apply to obstruent clusters.

order among CONTACT and HEAD. The existence of the variants in (39) indicates that these constraints are either unordered or that the order differs for different speakers:

(39) re.[g]nen ~ re[k].nen 'to rain'
or.[d]nen ~ or[t].nen 'to order'
e.[b]nen ~ e[p].nen 'to level'

The voicing alternation in (39) follows from the fact that only voiceless obstruents occur in syllable codas in German. A lack of order between CONTACT and HEAD would result in a tie among the two variants in (39), thereby accounting for the acceptability of both forms in standard German. What is no longer acceptable is the dactylic variant which is explained by the constraint ranking:

(40)			SON	*SCHWA	CONTACT	HEAD	CODA
		regnn	*!				
		re.g[ə]nn	*!				
	→	reg.n[ə]n		*	*		
	→	re.gn[ə]n		*		*	
		re.g[ə].n[ə]n		**!			

The ranking of the phonological constraints in (40) accounts for the generalisations in (10a,b).[17] There is no evidence that either WEIGHT nor SHELL play any roll in the prosodic form of NHG verbs.[18,19]

The fact that verbs are trochaic in NHG even if they do not potentially violate SON (e.g. sä[ə]n, hol[ə]n, etc.) indicates that *SCHWA is dominated by a constraint which requires verbs to end in a schwa syllable. That constraint presumably has the

---

[17]The same ranking also accounts for the distribution of the schwa in uninflected words (cf. the examples in (5)).

[18]The observation that WEIGHT no longer plays a role in NHG is in accordance with the general loss of quantity-sensitivity in German. The MHG length contrasts in consonants (i.e MHG ta[n:]e 'pine' vs. va[n]e 'flag') and arguably also in vowels (i.e. MHG m[a:]ge 'relative' vs. m[a]ge 'stomach') have disappeared in NHG. The phonological analysis of phonetic length contrasts in NHG vowels is a matter of debate (cf. Ramers 1988).

[19]While there is no direct evidence for SHELL, its existence requires perhaps a ranking among CONTACT and HEAD in order to account for the variant re.gn[ə]n. This is because re.gn[ə]n would eventually lose to the candidate reg.n[ə]n when evaluated with respect to SHELL, regardless of how low that constraint ranks. The order CONTACT >> HEAD yields the form re.gn[ə]n, whereas the order HEAD >> CONTACT yields reg.n[ə]n. Both variants and consequently both orders exist.

function to adjust the prosodic shape of such verbs to the shape of the majority of verbs where the trochaic form is determined by a potential SON violation (e.g. geb[ə]n, ruf[ə]n, etc.).[20] Because the constraint in question is irrelevant for the morphological issues under investigation, it will be ignored here.

To sum up, the variation in the prosodic patterns found in MHG verbs indicate a lack of order among various phonological constraints. The specific order among those constraints in MHG has resulted in trochaic forms only, where liquids systematically precede the final n in all verbs which involve a potential SON violation. The changes are summarized in (41):

(41) MHG: SON >> *SCHWA, CONTACT, HEAD, CODA, SHELL
NHG: SON >> *SCHWA >> CONTACT, HEAD >> CODA >> SHELL

The next section discusses the morphological changes which accompanied the phonological changes in (41).

3. Potential verbs

In German there is a highly productive morphological rule for forming new verbs. Apparently any uninflected word, regardless of its category or morphological complexity, can be verbalized by n-suffixation. Examples of verbs based on words ending in a schwa syllable are given in (42):

[20]Such a constraint dominates *SCHWA in all words with a sonorant suffix in NHG which shows that the prosodic form of those words is no longer determined by strictly phonological wellformedness conditions alone. The claim that the historical schwa insertion in words such as sä[ə]n 'to sow' is morphologically conditioned is also supported by considerations concerning word frequency. Words with a high frequency are first to undergo phonological rules but are last to undergo morphologically motivated change (cf. Philipps 1984). As was noted before, the only exception to the historical schwa insertion rule are the verbs sein 'to be' and tun 'to do', both of which are very frequent.

(42)

Category of base:		derived verb
adjective	sicher 'safe'	sichern
	dunkel 'dark'	verdunkeln[21]
comparative	milder 'milder'	mildern
	neuer 'newer'	erneuern
	weiter 'wider'	erweitern
	schlechter 'worse'	verschlechtern
	leichter 'easier'	erleichtern
noun	Zwiebel 'onion'	zwiebeln
	ferkel 'piglet'	ferkeln
plural noun	Eier 'eggs' (sg.:Ei)	eiern
	Löcher 'holes' (sg.:Loch)	löchern
	Geister 'ghosts' (sg.:Geist)	geistern
	Hämmer 'hammers' (sg.:Hammer)	hämmern

Evidence for the high degree of productivity of verbalization comes from the observation that native hearers of German are typically unsure of whether or not nonce verbs such as the italicized formations in (43) "exist":

(43)

amseln -> Amsel 'blackbird'	*faltern* -> Falter 'moth'
nesseln -> Nessel 'nettle'	*kadern* -> Kader 'cadre'
sesseln -> Sessel 'armchair'	*ebern* -> Eber 'boar'
rüsseln -> Rüssel 'trunk'	*messern* -> Messer 'knife'
disteln -> Distel 'thistle'	*katern* -> Kater 'tomcat'
wachteln -> Wachtel 'quail'	*kaisern* -> Kaiser 'emperor'
gürteln -> Gürtel 'belt'	*pfarrern* -> Pfarrer 'priest'
schnabeln -> Schnabel 'beak'	*bibern* -> Biber 'beaver'

Actual words such as Amsel or Falter imply apparently that amseln and faltern are potential verbs. The productivity of the rule also extends to loanwords. Again, most of the italicized verbs in (44) are not listed in dictionaries and yet they are acceptable:

[21]Verbalizations of adjectives or comparatives tend to include a prefix as well.

(44)				
	Butler	*butlern*	Level	*leveln*
	Label	*labeln*	Model	*modeln*
	Panel	*paneln*	Navel	*naveln*
	Container	*containern*	Oldtimer	*oldtimern*
	Cracker	*crackern*	Teenager	*teenagern*
	Poster	*postern*	Hamburger	*hamburgern*
	Gangster	*gangstern*	Power	*powern*
	Insider	*insidern*	Cover	*covern*
	Computer	*computern*	Joker	*jokern*
	Laser	*lasern*	Trawler	*trawlern*
	Charter	*chartern*	Compiler	*compilern*

In view of the high degree of productivity of verbalization in German it is striking that one class of words is excluded from the domain of the rule. The productivity gap concerns words which include a postvocalic schwa followed by a non-liquid consonant. The claim that such words do not give rise to potential verbs is illustrated in (45) with nouns ending in a nasal. The reader may convince herself that the italicized nonce formations listed in (45a,b) are the only possible derivations which satisfy the phonological wellformedness-conditions for verbs. That is, their final segment is n, they end in exactly one schwa-syllable where the position of the schwa always conforms to the rules in (10). Yet, they are clearly unacceptable verbalizations of the nouns in the lefthand column. Note that the nonce verbs in (45) cannot be ruled out on semantic grounds because for all corresponding nouns in English verbalisations are attested.

(45)		a.	b.
	Besen 'broom'	**besen*	**besnen*
	Daumen 'thumb'	**daumen*	**daumnen*
	Bogen 'bow; arch'	**bogen*	**bognen*
	Riemen 'strap'	**riemen*	**riemnen*
	Drachen 'kite'	**drachen*	**drachnen*
	Kragen 'collar'	**kragen*	**kragnen*
	Volumen 'volume'	**volumen*	**volumnen*
	Kissen 'pillow'	**kissen*	**kissnen*
	Schmarren 'trash'	**schmarren*	**schmarrnen*
	Magen 'stomach'	**magen*	**magnen*
	Hafen 'harbor'	**hafen*	**hafnen*
	Boden 'ground'	**boden*	**bodnen*
	Rachen 'throat'	**rachen*	**rachnen*
	Balken 'beam'	**balken*	**balknen*
	Garten 'garden'	**garten*	**gartnen*
	Fladen 'fritter'	**fladen*	**fladnen*
	Knochen 'bone'	**knochen*	**knochnen*
	Kuchen 'cake'	**kuchen*	**kuchnen*
	Busen 'bosom'	**busen*	**busnen*

How do the the unacceptable verbalizations in (45) differ from those considered so far? As for the nonce formations in

(45a) an obvious difference concerns the fact that they are formally identical to their base (cf. (46a)) whereas the formations in (43) and (44) include a segment which is not present in their base (cf (46b):

(46)a. [bézən] = [bézən] 'Besen]$_N$' b. [ámzəl**n**] ≠ [ámzəl] 'Amsel]$_N$'

The nonce formations in (45b), on the other hand, differ from the examples in (43) and (44) in that they lack phonological transparency with respect to their base. Phonological transparency is given in (47b) because the base is phonologically identical to a part of the derived verb, but not in (47a):

(47)a. [**béznə**n] - [**bézən**] 'Degen]$_N$' b. [**ámzəl**n] - [**ámzəl**] 'Amsel]$_N$'

The unacceptability of the verbalizations in (45a) and (45b) is accordingly due to distinct causes which is reflected in a difference in hearer judgments. The verbalizations in (45b) are consistently judged to be considerably worse than those in (45a), an observation, to which we will return below.[22]

Any verbalizations which involve neither of the two problems (e.g. phonological identity or lack of phonological transparency) would necessarily be phonologically illformed as is shown in (48). The formations in (48a) are unpronounceable in German because they include a geminate; those in (48b) are illformed because they end in more than one schwa syllable:

(48)a. *[bézənn] b. *[bézənən]

The illformedness of the verbalizations in (47) concerns accordingly their form in *relation* to the form of their base; the illformedness of the verbalizations in (48), on the other hand, concerns the form of the verb itself. The conditions delimiting the set of potential verbs can informally be stated as in (49):

(49) A word with the phonological (surface) form [X] can be verbalized iff either [Xn] or [Xən] satisfy the phonological wellformedness conditions for verbs.

The generalization in (49) leads us to expect that all words ending in a schwa syllable closed by a consonant less sonorous than l cannot give rise to potential verbs in NHG. If verbs were formed from such words by adding the sequence [ən] the resultant formation would end in two final schwa syllables and hence be illformed (cf (48b)). Adding just n to a word

---

[22]Neef (1994:178) points out that the verbalizations in (45a) are more acceptable with a prefix, which is consistent with the claim that their unacceptability is due to phonological homonphony.

ending in any consonant other than liquids would necessarily yield a sonority violation (cf. (48a). Avoiding the sonority violation by schwa epenthesis while at the same time preserving trochaic foot structure accommodates phonological wellformedness but has the consequence that the relation to the base no longer satisfies phonological transparency (cf. (47a). The dilemma described here affects all words ending in a schwa syllable closed by any consonant other than liquids which accounts for the observation that such words do not give rise to potential verbs.

The productivity dilemma encountered by words ending in a schwa syllable closed by a non-liquid can be further illustrated with the relatively few nouns listed in standard dictionaries in which a schwa is followed by a wordfinal obstruent. For such words there is only one possibility of forming a phonologically wellformed verb, which, however, is always entirely unacceptable as a derived form (the capital letters indicate the provenience of the nouns, 'Y': Yiddish, 'R': Rhenish, 'E': English):

(50)	Nippes 'knick-knack'	**nipsen*
	Kirmes 'kermis'	**kirmsen*
	Pommes 'French fries'	**pomsen*
	Kokolores 'rubbish; fuss'	**kokolorsen*
	Tinnef (Y) 'trash'	**tinfen*
	Schabbes (Y) 'sabbath'	**schabsen*
	Dokes (Y) 'bottom'	**doksen*
	Dalles (Y) 'poverty, money troubles'	**dalsen*
	Schammes (Y) 'shammes'	**schamsen*
	Zores (Y) 'anger'	**zorsen*
	Schmackes (R) 'zest, verve'	**schmacksen*
	Kappes (R) 'cabbage; rubbish'	**kapsen*
	Köbes (R) 'waiter (in a Cologne pub)'	**köbsen*
	Ticket (E)	**tickten*
	Racket (E)	**rackten*
	Krocket (E)	**krockten*
	Kricket (E)	**krickten*
	Velvet (E)	**velvten*

The starred verbalizations in (50) lack phonological transparency with respect to their base, a condition which could only be rectified at the expense of phonological illformedness (i.e. *nipp[ə]sn, *nipp[ə]s[ə]n). The dilemma illustrated in (45) and (50) is argued here to be at the root of the productivity gap which delimits the set of potential verbs. To be sure, there are *actual* verbs which are etymologically related to words ending in a schwa syllable closed by a nonliquid. First, Duden (1989) includes twenty-six pairs of etymologically related words, which seem to illustrate

precisely the pattern of the starred nonce verbs in (45a) and the respective nouns to their left.[23]

(51)	Graben]$_N$ 'ditch'	graben]$_V$ 'to dig'
	Husten]$_N$ 'cough'	husten]$_V$ 'to cough'
	Schnupfen]$_N$ 'cold'	schnupfen]$_V$ 'to take snuff'
	Rechen]$_N$ 'rake'	rechen]$_V$ 'to rake'
	Tropfen]$_N$ 'drop'	tropfen]$_V$ 'to drip'
	Schaden]$_N$ 'damage'	schaden]$_V$ 'to damage'
	Streifen]$_N$ 'strip'	streifen]$_V$ 'to brush (against)'
	Glauben]$_N$ 'belief'	glauben]$_V$ 'to believe'
	Zapfen]$_N$ 'plug; cone'	zapfen]$_V$ 'to tap (beer)'
	Nutzen]$_N$ 'use, benefit'	nutzen]$_V$ 'to be of use to'
	Fetzen]$_N$ 'shred'	fetzen]$_V$ 'to rip'
	Flicken]$_N$ 'patch'	flicken]$_V$ 'to patch'
	Funken]$_N$ 'spark'	funken]$_V$ 'to spark'
	Haken]$_N$ 'hook'	haken]$_V$ 'to hook'
	Schrecken]$_N$ 'fright, horror'	schrecken]$_V$ 'to scare'
	Flecken]$_N$ 'stain'	(be)flecken]$_V$ 'to stain'
	Brocken]$_N$ 'lump, chunk'	brocken]$_V$ 'to break (bread)'
	Rücken]$_N$ 'back'	rücken]$_V$ 'to move'
	Ballen]$_N$ 'bale'	ballen]$_V$ 'to clench (a fist)'
	Stopfen]$_N$ 'stopper; cork'	stopfen]$_V$ 'to stuff; to darn'
	Gefallen]$_N$ 'favor'	gefallen]$_V$ 'to please'
	Schatten]$_N$ 'shadow'	(be)schatten]$_V$ 'to shadow, to tail
	Rahmen]$_N$ 'frame'	rahmen]$_V$ 'to frame'
	Klumpen]$_N$ 'lump'	klumpen]$_V$ 'to go lumpy'
	Knoten]$_N$ 'knot'	knoten]$_V$ 'to knot'
	Fohlen]$_N$ 'foal'	fohlen]$_V$ 'to foal'

If phonological identity were the reason for why the formations in (45a) are unacceptable then what accounts for the existence of the verbs in (51)? There is evidence that none of the verbs in (51) were derived from the nouns in the lefthand column in their present phonological form. In particular, the final n in those nouns is innovative as the corresponding MHG words show:[24]

---

[23]The nouns listed in (51) do not include gerunds, which are always neuter (cf. das Graben 'the digging' vs. der Graben 'the ditch').

[24]I do not claim that the verbs in (52) were necessarily derived from the nouns to their left in MHG. In fact, in many cases both forms already existed in Old High German where some of the nouns were possibly derived from the corresponding verbs. At least the strong verbs (graben, braten, laden, gevallen) are clearly not historically derived from nouns. The table in (52) merely shows that there exists a stage at which the nouns did not meet the description characterizing the gap (e.g. words wnding in a schwa syllable closed by a nonliquid) and most of the verbs are already attested.

(52)

graben]$_V$ ~ grabe]$_N$	haken]$_V$ ~ hâke]$_N$
huosten]$_V$ ~ huoste]$_N$	schrecken]$_V$ ~ schrecke]$_N$
snupfen]$_V$ ~ snupfe]$_N$	vlecken]$_V$ ~ fleck(e)]$_N$
rechen]$_V$ ~ reche]$_N$	brokken]$_V$ ~ brocke]$_N$
tropfen]$_V$ ~ tropfe]$_N$	rücken]$_V$ ~ rück(e)]$_N$
schaden]$_V$ ~ schade]$_N$	ballen]$_V$ ~ balle]$_N$
streifen]$_V$ ~ strife]$_N$	stopfen]$_V$ ~ stopf]$_N$
gelouben]$_V$ ~ g(e)laube]$_N$	gevallen]$_V$ ~ geval]$_N$
zapfen]$_V$ ~ zapfe]$_N$	(be)schatewen]$_V$ ~ schate(we)]$_N$
nutzen]$_V$ ~ nutze]$_N$	(?ramen]$_V$) ~ rame]$_N$
vetzen]$_V$ ~ vetze]$_N$	(?klumpen]$_V$) ~ klumpe]$_N$
vlicken]$_V$ ~ flick]$_N$	(?knoten]$_V$) ~ knote]$_N$
vunken]$_V$ ~ vunke]$_N$	(?volen]$_V$) ~ vol(e)]$_N$

The process by which nouns like grabe developed a final -n is illustrated in (53). The final nasals in the oblique forms, which at one point were inflectional suffixes, were reanalysed as being part of the stem and consequently appeared in the nominative as well:[25]

(53)

Stage I:			Stage II:
Sg.	Nom.	grabe	Graben
	Gen.	graben	Grabens
	Dat.	graben	Graben
	Acc.	graben	Graben

The data in (51) are consistent with the claim that verbalization has only been possible before the historical suffix -n was reanalysed as part of the stem. In accordance with the generalization in (49) graben is a potential verb and hence may come into existence as long as there exists a word grabe functioning as the base. As soon as the noun grabe develops a final -n, graben ceases to be a potential verb, though it certainly may be and in fact is an actual verb.

The data in (52) clearly do not prove that verbalization was only possible prior to the reanalysis of the nasal suffix. More conclusive evidence in this matter would require access to the entire set of nominative forms as well as the entire set of verbs at every stage of the language. What can be shown is that

---

[25]The claim that the n appeared in the nominative forms because the oblique forms were reanalysed as simplexes presupposes that its morphological function failed to be recognized by learners. That failure would be expected if learners encountered the oblique forms far more frequently than the nominative forms. Such an asymmetry seems plausible in view of the fact that the leveling in (i) only affects inanimate nouns. According to Behagel (1928) inanimate nouns function generally as objects and consequently have oblique case marking whereas animate nouns are more likely to function as subjects and consequently appear in the nominative form. The direction of leveling is then due to frequency effects.

most verbs in (51) are already attested in MHG[26] and that for each verb which appears to be based on a homophonous noun synchronically, that noun goes back to a form with no stemfinal -n.[27] This observation is explained on the hypothesis that the generalization concerning potential verbs in (49) held true of earlier stages in the language as well. At the same time this hypothesis explains why the coinages in (45a) are not acceptable to native hearers of German.

In addition to the pairs in (51), Duden (1989) includes a total of sixteen pairs of etymologically related words, which show a phonological alternation similar to that between the starred nonce verbs in (45b) and the respective nouns to their left ("E": 'elevated', "A" 'archaic'):.

(54)	Atem 'breath'	atmen 'to breathe'
	Wappen 'coat of arms'	wappnen 'to prepare to face sth.
	Waffen 'weapons'	bewaffnen 'to arm'
	Orden 'decoration, medal'	ordnen 'to order, to arrange'
	Zeichen 'sign'	zeichnen 'to draw'
	Regen 'rain'	regnen 'to rain'
	Segen 'blessing; bliss'	segnen 'to make the sign of the cross
	eben 'level; flat'	ebnen 'to smooth; to level off'
	trocken 'dry'	trocknen 'to dry'
	eigen 'own'	eignen 'to be suited'
	offen 'open'	öffnen 'to open'
	gegen *prep.* 'against'	begegnen 'to encounter'
	vollkommen 'perfect'	vervollkommnen 'to make perfect'
	Willkommen 'welcome'	(E) bewillkommnen 'to welcome'
	Boden 'ground'	(A) verbodmen 'to pawn the cargo of a ship'
	Faden 'thread'	(A) auffädmen 'to string (beads)'

Except for vervollkommnen und bewillkommnen all verbs in (45) are already attested in MHG where they ended in two schwa syllables:

---

[26]The four paranthesized verbs in (52) are not listed in Lexers, but I assume that they were coined before leveling took place.

[27]The only counter-example to this generalization is the verb röntgen 'to X-ray' which however is based not on a common noun but on a name.

(55)	âtem 'breath'	âtemen 'breathe'
	wâpen/wâfen 'weapon'	wâpenen/wâfenen 'to arm'
	orden 'rule; order; law'	ordenen 'to order'
	zeichen 'sign, example'	zeichenen 'to put a sign on sth.'
	rëgen 'rain'	rëgenen 'to rain, to let rain'
	sëgen 'sign of the cross; blessing '	sëgenen 'to make the sign of the cross; to bless'
	ëben 'level, flat, equal'	ëbenen 'to level; to unite'
	trucken 'dry'	truckenen 'to dry'
	eigen 'own'	eigenen 'to acquire sth.'
	offen 'open, exlained'	offenen 'to open, to explain'
	gegen 'towards'	gegenen 'to come towards'
	bodem 'ground, floor'	bodemen[28] 'to make a wooden floor'
	vadem 'thread'	vedemen 'to thread'

The data in (55) are significant in that they show that the verbs in (45) came into existence before the reranking in (41) took place. That is, those verbs were coined before the prosodic restriction to a single final schwa syllable characteristic of NHG took effect. The verbs in (54) are therefore consistent with the generalization in (49), according to which a word can be verbalized only if suffixation of n or ən yields a form which satisfies the respective phonological wellformedness condition for verbs.

The existence of the noun-verb pairs in (51) and (54) consequently does not challenge the claim that words ending in a schwa syllable closed by a nonliquid are outside the domain of verbalization in NHG. Rather, the existence of those pairs challenges the view that the question of what the potential words of a language are can be decided on the basis of the synchronically existing 'alternations'. The assumption that a hearer will automatically pick up a rule for deriving new forms if there are enough recurrent pairs of words showing some phonological and semantic resemblance is inconsistent with the observation that the nonce forms in (45) are not acceptable. Rather, the acceptability of new verbs seems to be subject to prosodic wellformedness conditions for verbs interacting with phonological conditions on base relations.

4. Defining the set of potential words.

According to the generalization in (49) the potential formation of verbs depends on the wellformedness of the surface phonological structure of the derived verb. That generalization is therefore not consitent with the standard view on the morphology-phonology interface in Generative Grammar, according

---

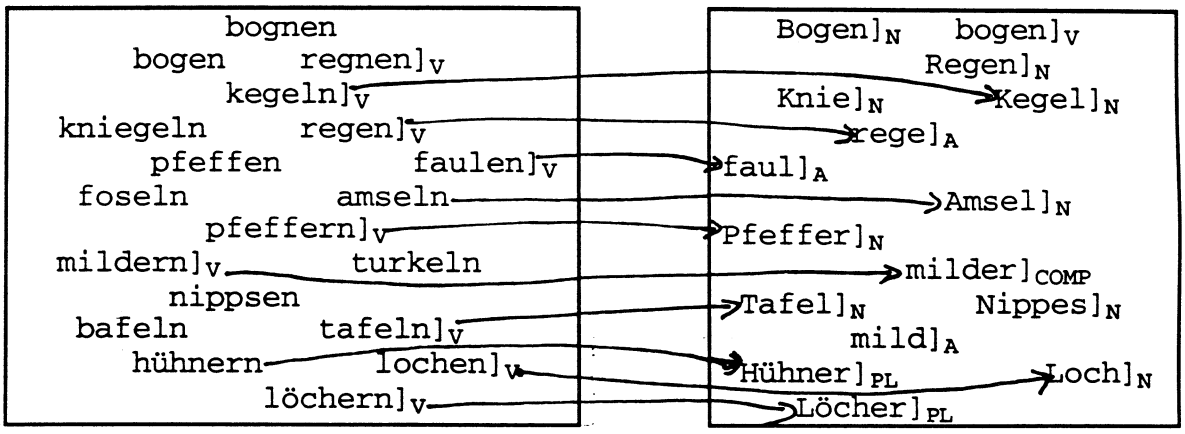
[28]This verb is not listed in Lexer's but in J. and W. Grimm's "Deutsches Wörterbuch" from 1854.

to which the phonology interprets morphological structure. Instead the generalization in (49) suggests that potential verbs are defined in terms of relations between two sets, the set "P_V" which consists of all nonce words satisfying the wellformedness conditions for verbs and the set of actual words. The set P_V includes all candidates evaluated as optimal on the basis of the constraint rankings given in (43), or (45) respectively. Potential verbs can then be defined as follows:

(56) The set of potential verbs consists of all strings [X(ə)n] included in P_V, for which there exists an actual word [X].

For speakers who find the verbalisations in (45a) marginally acceptable [X] can be substituted by [X((ə)n)] in (56). The unacceptability of the verbalisations in (45b) shows that the rule in (56) must refer to surface forms. The rule is illustrated with the figure in (57):

(57)



The leftmost set in (57) illustrates P_V in NHG. P_V thus includes almost all existing verbs[29] as well as an infinit number of nonce verbs. The rightmost set includes existing words of various categories (e.g. noun, plurals, adjectives, comparatives, etc.). The wellformedness conditions for verbs in conjunction with the rule in (56) rule out the possibility that items in the left set can be related to words ending in a schwa-syllable closed by nonliquids thus explaining the productivity gap.

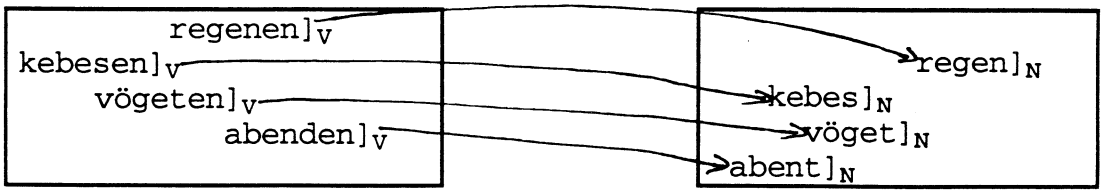
It is important to note that given the same input (e.g. any arbitrary string), the optimal candidates would differ in MHG and in NHG because of the difference in conastraint ranking. Under the assumption that the rule in (56) has not changed over time we would consequently expect that the productivity gap in question did not exist in MHG. The examples in (58) illustrate

---

[29]The existing verbs not included in P_V are the four verbs tun, sein, nähern, and wiehern mentioned earlier.

the claim that words, which fall outside the domain of verbalization in NHG, satisfy the conditions on potential verbs in MHG:

(58)



In fact, the data indicate that words ending in closed schwa syllables could be verbalized regardless of the sonority of the final consonants. The MHG verbs in (59) relate to bases ending in the sequence schwa plus nasal.

(59)

morgen]$_{Adv}$ 'tomorrow'	morgenen]$_V$ 'to procrastinate'
siben]$_{Num}$ 'seven'	sibenen]$_{Adv}$ 'to interrogate sb. in th presence of seven witnesses
zëhen]$_{Num}$ 'ten'	(ver)zëhenen]$_V$ 'to pay one tenth of a income'
brâdem]$_N$ 'steam'	brâdemen]$_V$ 'to steam'
kradem]$_N$ 'noise'	krademen]$_V$ 'to make a noise'
gadem]$_N$ 'one room house'	(be)gedemen]$_V$ 'to bring into a *gadem*'
krisem]$_N$ 'sacred unction'	krisemen]$_V$ 'to anoint with *krisem*'
ludem]$_N$ 'screaming'	ludemen]$_V$ 'to scream'
swadem]$_N$ 'steam'	swademen]$_V$ 'to steam'
mitten]$_{Adv}$ 'in the middle'	mittenen]$_V$ 'to sit down in the middle
kristen]$_A$ 'Christian'	kristenen]$_V$ 'to Christianize
û en]$_{Adv}$ 'out, outside'	û enen]$_V$ 'to divest oneself of sth.'
wolken]$_N$ 'cloud'	wolkenen]$_V$ 'to be full of clouds'
brëhen]$_N$ 'gleam, shine'	brëhenen]$_V$ to gleam, to shine'
keten]$_N$ 'chain'	ketenen]$_V$ 'to put in chains'
vesten]$_N$ 'fortress'	vestenen]$_V$ 'to build a fortress'
dëgen]$_N$ 'warrior, hero'	dëgenen]$_V$ 'to turn sb. into a hero'
tougen]$_N$ 'secret'	tougenen]$_V$ 'to keep secret'
trahen]$_N$ 'tear'	trahenen]$_V$ 'to cry'
lâchen]$_N$ 'medicine'	lâchenen]$_V$ 'to spread medicine on sb.
lêhen]$_N$ 'feoff'	lêhenen]$_V$ 'to enfeoff sb.'
bësem]$_N$ 'broom; rod'	bësemen]$_V$ 'to sweep; to whip sb. with a rod'
ougen]$_{N, Pl}$ 'eyes'	ougenen]$_V$ 'to show'
meiden]$_N$ 'stallion'	meidenen]$_V$ 'to castrate'
besamen]$_{Adv}$ 'together'	besamenen]$_V$ 'to gather warriors'
lougen]$_N$ 'denial'	lougenen]$_V$ 'to deny; to revoke'
widem]$_N$ 'the groom's dowry'	widemen]$_V$ 'to give a dowry'
solden]$_V$ 'to pay a soldier'	soldenen]$_V$ 'to pay a soldier'
biben]$_V$ 'to tremble'	bibenen]$_V$ 'to tremble'
bürden]$_V$ 'to give sb. a load to carry'	bürdenen]$_V$ 'to give sb. a load to car

	bederben]$_V$ 'to be useful'	bederbenen]$_V$ 'to be useful'
	rechen]$_V$ 'to count'	rechenen]$_V$ 'to count'

As a result of adjusting to the prosodic wellformedness conditions for verbs in NHG (i.e. morgenen > morgnen), the morphological status of the verbs in (59) changed. In particular, they lost their status of being potential verbs with the result that this group can only lose members but not gain new ones. In the examples listed in (59) either the base, the derived verb, or both have become obsolete in NHG. The only verbs of this type that are left in NHG are those listed in (54).

The examples in (60a,b) illustrate coinages based on words ending in the sequence schwa plus obstruent or schwa plus non-liquid cluster:

(60)a.	kebes]$_N$ 'concubine'	kebesen]$_V$ 'to commit adultery'
	houbet$_N$ 'head'	houbeten]$_V$ 'to decapitate'
	vöget]$_N$ ''	vögeten]$_V$ 'to protect'
	kachez]$_N$ 'roaring laughter'	kachezen]$_V$ 'to laugh loudly'
	market]$_N$ 'market'	marketen]$_V$ 'to trade'
b.	âbent]$_N$ 'evening'	âbenden]$_V$ 'evening comes'
	jugent]$_N$ 'youth'	jugenden]$_V$ 'to be youthful'
	tugent]$_N$ 'usefulness, virtue'	tugenden]$_V$ 'to show *tugent*, to lend sb.*tugent*'
	zehent]$_{ORD}$ 'tenth'	zehenden]$_V$ 'to pay a tenth'

The data in (60b) raise the question of why all verbs have disappeared in NHG. Note that the absence of dactylic verbs in NHG is not explained. The verb based on the noun Abend 'evening', for example, would clearly have two schwas according to the constraint ranking for NHG:

(61)			SON	*SCHWA	CONTACT	HEAD	CODA
a.		abndn	*!				
		abn.d[ə]n	*!				
		ab.nd[ə]n	*!				
		ab.n[ə].d[ə]n		**	*!		
		a.bn[ə].d[ə]n		**		*!	
	→	a.b[ə]n.d[ə]n		**			

Is the fact that abenden became obsolete in NHG accidental or was it "pushed out" by a so far unexpressed phonological requirement for maximally binary feet? Since only a handful verbs of that type existed in MHG (e.g. verbs with two

potential nonoverlapping sonority violations), the question will be left open.[30]

The data indicate then that the conditions for deriving new verbs in German have remained constant (cf.(56)). This assumption is supported by the fact that the rule shows much the same properties since MHG. Characteristic are not only the high degree of productivity but also the lack of sensitivity to the syntactic category and morphological complexity of the base.[31]

What has changed, are the phonological wellformedness condition for verbs. Given that the condition as stated in (56) is correct, words ending in a syllable closed by a nonliquid are expected to disappear from the domain of potentrial verb bases as soon as the reranking in (41) takes place. This dependence of potential verbalisations on the phonological surface form of the derived forms challenges the view of phonology as a rule system which merely interprets morphologically derived strings. Such a view could only be maintained if the productivity gap was encoded in terms of a prosodic subcategortization frame of the suffix -n. That analysis would fail, however, to explain why this prosodic requirement emerged at the same time when the phonological wellformedness conditions for verbs changed. Given the condition in (56) the productivity gap affecting verbalization in NHG is considered an epiphenomenon, rather than a property of the suffix n.

---

[30]One also needs to investigate the question of whether there are any independent clear examples of words becoming systematically obsolete because of phonological illformedness.

[31]In MHG. we even find verbs in the domain of verbalization (cf. the data in (cf. the last four examples in (59)). The fact that verbs cannot be verbalized in NHG is presumably due to their phonological form: verbs end in a sequence schwa plus nasal.

References:

Behaghel, O. (1928). Geschichte der deutschen Sprache. 5. Auflage. Berlin, Leipzig.

Duden (1989). Deutsches Universalwörterbuch Mannheim, Wien, Zürich: Dudenverlag.

Giegerich, H. (1987) Zur Schwa Epenthese im Standarddeutschen. In: Linguistische Berichte 112 (1987), 449-469.

Grimm, J, Grimm, W. (1854). Deutsches Wörterbuch. Leipzig.

Hall, T. A. (1992) Syllable structure and syllable-related processes in German. Tübingen: Niemeyer. (Linguistische Arbeiten 276).

Lexer, M. (1876) Mittelhochdeutsches Handwörterbuch. Verlag von S. Hirzel. Leipzig.

Mester, Arnim and Junko Itô (1994). Anaptyxis in Optimality Theory: the phonology and morphology of German schwa. Handout. Workshop in prosodic morphology, Utrecht.

Moser & Stopp (eds.) (1970). Grammatik des Frühneuhochdeutschen. Beiträge zur Laut- und Formenlehre. Heidelberg.

Neef, M. (1994). Wortdesign. Eine repräsentationelle Analyse morphologischer Regularitäten. Dissertation. Universität zu Köln

Paul, H., Wiehl P. and Grosse, S. (1989). Mittelhochdeutsche Grammatik. Max Niemeyer Verlag. Tübingen.

Phillips, B. (1984). Word frequency and the actuation of sound change. Language 60:320-342.

Prince, A & P. Smolensky (1993). Optimality Theory. Constraint Interaction in generative grammar. Ms. Rutgers University, New Brunswick, and University of Colorado, Boulder.

Raffelsiefen, R. (1994). Schwa patterns in German verbs. A Study in the phonology-morphology interface. Ms. Düsseldorf

- (1995). Conditions for stability: the case of schwa in German. Arbeiten des Sonderforschungsbereichs 282. Düsseldorf.

Ramers, K.H. (1988). Vokalquantität und -qualität im Deutschen. Tübingen: Niemeyer.

Sievers, E. (1901). Grundzüge der Phonetik. Leipzig: Breitkopf und Härtel.

Strauss, S (1982). A Lexicalist Phonology of English and German. Dordrecht: Foris.

Vennemann, T.(1982). Zur Silbenstruktur der deutschen Standardsprache. 261-305. In: Silben, Segmente, Akzente. Ed. by T. Vennemann. Tübingen: Niemeyer. (Linguistische Arbeiten 126).

- (1988). Preference Laws for Syllable Structure and the explanation of sound change. Berlin: Mouton de Gruyter.

Wiese, R. (1986) Schwa and the structure of words in German. Linguistics 24, 695-724.

- (1988) Silbische und lexikalische Phonologie. Max Niemeyer Verlag. Tübingen.

LEARNABILITY UNDER OPTIMALITY THEORY

Iggy Roca, University of Essex

0. Introduction

It is nowadays taken as commonplace that the success of a linguistic theory cannot be measured exclusively in terms of formal elegance or descriptive coverage, but must be subject to the litmus test of learnability. The reason for this is obvious: language is a cognitive object, and it is the (self-imposed, but eminently reasonable) ultimate task of the linguist to describe this object, rather than its textual manifestations. It follows from this that the architecture that a linguistic theory projects onto language must be learnable under the usual conditions of inevitability, spontaneity, speed, perfection, and irrelevance of environmental conditions other than availability of ordinary language input to the child. The question to be asked in this paper is therefore whether or not the conception of grammar underlying Optimality Theory meets such learnability criteria.

This paper summarises the issues as discussed in McCarthy & Prince (1993), Prince & Smolensky (1993), and Tesar & Smolensky (1993), and is organised as follows. In section 2 I introduce Optimality Theory ('OT') for the benefit of the uninitiated reader, availing myself of a fragment of the syllabification grammar of Imdlawn Tashlhiyt ('IT') Berber, originally analysed in Dell & Elmedlaoui (1985, 1988) and previously presented in section 1 here, and I compare the OT descriptive achievements with those of the standard theory. In section 3, I explore the mechanism for the establishment of underlying forms under OT. Finally, in section 4 I scrutinise the learnability of language-specific constraint ranking, the keystone of OT grammars.

1. IT Berber syllables

The Imdlawn Tashlhiyt dialect of Berber, described in Dell & Elmedlaoui (1985, 1988), presents some remarkable syllabification characteristics. Consider first the data in (1) (syllable division is indicated by a dot throughout):

(1) il.di 'he pulled'
ir.ba 'he carried on his back'
in.da 'he shook (milk)'
im.da 'he was born out'
iz.di 'he put together'
iẑ.la 'he got lost'
iɣ.za 'he digged'
is.ti 'he selected'
if.si 'he untied'
ix.si 'he went out (fire)'

Such syllables as in (1) are as ordinary as any syllables can be.

In particular, all the words in question exhibit a syllabic structure VC.CV. CV is of course the universal core syllable, from which VC is readily obtainable by onset deletion and coda addition, as discussed in Clements & Keyser 1983. Note that although such (word-initial, or, more accurately, phrase-initial) onsetless syllables are optionally provided with a glottal stop onset (cf. Dell & Elmedlaoui 1985: 127, fn. 20), this is a late phonetic process that plays no role in syllabification as such, and consequently it will be ignored here.

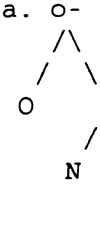
Consider now the forms in (2):

(2) tr.glt 'you(sg) locked'
ts.krt 'you(sg) did'
tx.znt 'you(sg) stored'
tz.dmt 'you(sg) gathered wood'
tl.bžt 'you(sg) stepped onto'
tr.kst 'you(sg) hid'
tn.šft 'you(sg) grazed (skin)'
tm.sxt 'you(sg) transformed'

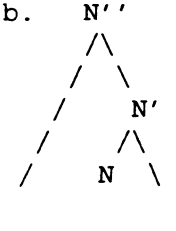
These words are remarkable in that they have no vowels (NB. some ultra-short transitional vowels can reportedly be heard, but these are totally predictable from the phonetic context and have no phonological significance; cf. Dell & Elmedlaoui 1985: 116 ff.). What is striking is that, not only are these words pronounceable, indeed totally ordinary in IT Berber, but they are syllabified into two syllables each. This is of course a truly bizarre situation from the perspective of the speakers of 'regular' languages like English or German, but it is strongly substantiated by evidence from speaker intuition, emphasis spread, gemination, intonation, versification practice, and prosodic morphology (cf. Dell & Elmedlaoui 1985, 1988, Elmedlaoui 1985).

As a background to our analysis of IT Berber syllabification, we now provide two alternative formal representations of the (universal) basic syllable structure. The first, in (3a), is traditional, and purely stipulative, while the second, in (3b), is grounded on X-bar theory, and was originally proposed in Levin (1985):

(3) a.



b.



On either representation, the core of each syllable is a vocalic nucleus, indeed the common situation universally. As is well-known, languages like English or German also allow sonorant consonants to constitute syllable nuclei under certain conditions

(cf. e.g. *litt*[l], *butt*[n], *Neb*[l], *werf*[n] etc.). A close examination of the data in (2) will reveal that some of the proposed syllables contain a sonorant consonant, which can therefore reasonably be construed as the nucleus of the respective syllable. In other cases, such a nucleus is, even less plausibly, apparently constituted by a fricative. Still, this situation is also not totally unfamiliar to us, since fricative nuclei are found in English informal, relatively fast speech (cf. e.g. [s]*port* 'support' vs. [s]*port* 'sport'). The designated nuclei are made explicit in (4) by means of capitalisation:

(4) tR.gLt 'you(sg) locked'
tZ.dMt 'you(sg) gathered wood'
tL.bẔt 'you(sg) stepped onto'
tS.kRt 'you(sg) did'
tX.zNt 'you(sg) stored'
tR.kXt 'you(sg) hid'
tN.s̈Ft 'you(sg) grazed (skin)'
tM.sXt 'you(sg) transformed'

From a traditional perspective, IT Berber could therefore be considered only to differ from such common-or-garden languages as English and French in its more liberal attitude to the nuclearisation of sonorant and fricative consonants. Consider, however, the forms in (5):

(5) mA.rA.tGt 'what will happen of you'
rA.tK.tI 'she will remember'
tF.tKt 'you suffered a sprain'

Here, alongside vocalic nuclei (mA, rA, tI), we find others consisting in an obstruent stop, whether voiced (tGt) or voiceless (tK, tKt). Clearly, we cannot go any further down on the scale of sonority, and consequently we must conclude that in IT Berber *all* segments qualify as syllable nuclei.

The apparently crazy situation just uncovered, where some (but obviously not all) of a word's consonants bear the syllable nucleus, may lead one to believe that this language simply has a few lexicalised consonantal nuclei, precisely as represented in (4) and (5) above. That this is not the case is forcefully brought out by the data in (6), which exhibit alternation with those in (1) above (non-nucleic first syllable consonant in (1) ~ nucleic first syllable consonant in (6)):

(6) tL.dI 'she pulled'
tR.bA 'she carried on her back'
tN.dA 'she shook (milk)'
tM.dA 'she was born out'
tZ.dI 'she put together'
tẔ.lA 'she got lost'
tY.zA 'she digged'
tS.tI 'she selected'
tF.sI 'she untied'
tX.sI 'she went out (fire)'

Indeed, were it to be just a matter of arbitrary lexicalisation, we would expect the distribution of such syllabic consonants to be random. However, as will now be shown, IT syllabification is constrained by the two principles in (7):

(7) IT syllabification principles:

i. no hiatus (i.e. *NN)

ii. maximisation of number of syllables

The principle of no-hiatus entails obligatoriness of onsets in all positions but phrase-initially. As a consequence, the parsings in (8a), but not those in (8b), will be legitimate (NB. /w/ = non-nucleic /u/; /y/ = non-nucleic /i/):

(8)	a. tI.wN.tAs	'you climbed on him'	b.	*tI.Un.tAs
	rA.yMm.$_{Y}$I	'he will grow'		*rA.Imm.$_{Y}$I

In turn, the principle of syllable maximisation implies both syllabification recursiveness and minimisation of coda construction (subject to some further constraints on the structure of the syllable). In a nutshell, the idealised effect of the IT Berber syllabification algorithm is as schematised in (9):

(9) IT core syllable: CV (subsequently incremented)

preferred syllabification: CV.CV.CV.CV.CV....
ON.ON.ON.ON.ON....

Underpinning the succession of core syllables CV.CV. ... is of course the sonority substance of segments and the universal principle of sonority dispersion (cf. Clements 1990), according to which sonority differences between onset and nucleus tend to be maximised, while sonority differences between nucleus and coda tend to be kept to a minimum (Ø is obviously the optimal such minimum). The role of sonority in IT Berber syllabification goes however beyond such universal effects, as illustrated in (10) (NB. onsetless stops are eventually desyllabified and incorporated into the adjacent syllables, as commented on in Dell & Elmedlaoui 1985: 127, fn. 16; we abstract away this eventuality for the sake of clarity and simplicity of exposition; note interestingly that the pre-desyllabification level is made use of in Berber poetry, as discussed in Dell & Elmedlaoui 1988):

(10)	a. T.zMt	'she is stifling'	b.	*tZ.mT
	rAt.lU.lT	'you will be born'		*rA.tL.wLt

Here, the parsings in a. and b. both comply with the basic syllabification requirements of the language (remember that onsets are not obligatory phrase-initially, since no hiatus results). This notwithstanding, only the parsings in a. are legitimate, for reasons relating to sonority, as will become

clear directly.

As a preliminary, I shall make explicit the (unremarkable) IT Berber sonority hierarchy (I ignore a few additional, exotic IT Berber consonants for the sake of graphic simplicity):

(11) IT Berber sonority ranking:

low vowel (a) >

high vowels (i, u) >

liquids (l, r) >

nasals (m, n) >

voiced fricatives (z, $\overset{Y}{z}$, $_{Y}$) >

voiceless fricatives (f, s, $\overset{Y}{s}$, x) >

voiced stops (b, d, g) >

voiceless stops (t, k, q)

Armed with such sonority-based grading of IT Berber segments, I will now provide the basic syllabification algorithm of the language, nucleus assignment being indeed algorithm-governed, and not lexical, as hinted at above:

(12) core syllabification algorithm (Dell & Elmedlaoui 1985, 111):

'associate a core syllable with any sequence (Y)Z, where Y can be any segment and Z is a segment of type T, where T is a variable to be replaced by a certain set of feature specifications [= descending sonority, IMR]'

The segment on the lefthand side, Y, can *only* be missing phrase-initially, as we already know, an important restriction which is however not made explicit by Dell & Elmedlaoui in the algorithm itself. The following derivations illustrate the workings of this algorithm (the forms in parentheses result from the operation of coda formation, not part of the core algorithm; nucleus parsing is symbolised by means of a vertical tree line):

(13)

trba	tzmt	ratlult	
/|	/|	/|	
trbA	tzMt	rAtlult	step 1
/|/|	|/|	/| /|	
tRbA	TzMt	rAtlUlt	step 2
		/| /|/|	
		rAtlUlT	step 3
	|/|\	/|\/|/|	
	(TzMt)	(rAtlUlT)	coda
'she carries on her back'	'you will be born'	'she is stifling'	

This algorithm is still ambiguous when applied to such underlying forms as /rksx/ 'I hid' and /bainn/ 'they (m.) appear', as illustrated in (14):

(14) a. R.kSx 'I hid' b. *Rk.sX
bA.yNn *'they (m.) appear'* **bAy.nN*

The difference between the two sets of outputs is a function of directionality (L-to-R in a. vs. R-to-L in b.), which must accordingly also be specified in the procedure (L-to-R). This enrichment is however still insufficient, and further conditions need to be imposed. For instance, we must rule out ambisyllabicity ($*(X(X)X)$), and prevent destruction of structure ((N)XX -/-> (ONC)), in line with the so-called Free Element Condition (cf. Prince 1985). Finally, while the high vocoids /i/, /u/ can be parsed in the onset when required by the algorithm, as we have seen, the low vocoid /a/ is only parsable in the nucleus, as illustrated by the following data from the similar Ait Seghrouchen dialect, analysed in Guerssel (1985):

(15) inna-ax 'he told us' -> in.na.**y**ax
a aryaz 'man (voc.)' -> a.**y**ar.yaz

As can be seen, hiatus is resolved by consonant epenthesis ([y], highlighted for convenience), rather than by the assignment of /a/ to the onset (cf. e.g. *In.nA.aX).

We shall now take brief stock and bring together the core syllabification machinery of IT Berber under the standard analysis of Dell & Elmedlaoui (1985):

(16) standard syllabification machinery of IT Berber:

i. algorithm (12)

ii. exhaustivity condition

iii. free element condition

iv. ambisyllabicity ban

v. /a/ -> Nuc (i.e. *.aN.)

The problem with this procedure concerns not so much its obvious complexity (only partially alleviated by the universality of some of its components), but specifically the apparent lack of connection between the various rules and conditions, the relation between which, if existent, is anything but obvious. As we will see in the next section, this difficulty is elegantly circumvented by OT.

2. Optimality Theory

The principles of OT are lucidly expounded in McCarthy & Prince (1993) and Prince & Smolensky (1993), and will now be summarised for the reader's convenience.

The basic mechanism of OT is extremely simple, and is made up of two components, viz. Gen and a set of constraints. I will comment on these in turn.

Gen (short for 'generator') is a device parsing each of a set of universal inputs into a (universally acceptable) set of outputs. Thus, for instance, given a string of segments, Gen will produce a sequence of universally (NB. not necessarily language-specifically) well-formed syllables (the question of which syllables are universally well-formed is itself of course still open, at least on the edges; the full answer to this question, whatever it may be, will thus simply be incorporated into the body of Gen, according to OT tenets). An important corollary of the restriction of Gen activity to parsing is that OT does not countenance physical deletion as such, and so any input will be contained in each of its outputs (the 'Principle of Containment').

If the grammar of all languages consisted exclusively of Gen, all languages would be identical. More precisely, there would literally be only one language, subject to random variation, given the relative unrestrictedness of Gen. Clearly, therefore, further principles are necessary to reflect both (relative) language-internal invariance and cross-linguistic variation. The key feature of the OT framework is that all such principles are couched in terms of (positive or negative) constraints (NB. not rules), which are moreover postulated to be universal, hence not learned (Gen is obviously also part of Universal Grammar). Again, given the universality of the constraints, the prediction is that all languages will be identical (equivalently, only one language will be in existence). This prediction is of course counterfactual, and OT consequently allows for the *ranking* of constraints according to language-specific stipulation, the universal, inviolable principle being that compliance with higher

ranked constraints takes precedence over compliance with their lower ranked counterparts. In this way, the preferred output will be the *optimal* output (i.e. the output that best meets the principle of higher ranking priority), rather than the *perfect* output (i.e. the output that violates no constraints), which is more often than not simply unobtainable (note also, and importantly, that, because all constraints apply to the output of Gen simultaneously, OT countenances no derivations).

I shall now illustrate the workings of this simple model by applying it to our familiar IT Berber data. Let us first formalise a couple of constraints playing a basic role in syllabification in general (for discussion of these constraints and other parts of the theory, and for a more extended OT analysis of IT Berber, see Prince & Smolensky 1993):

(17) two (output) constraints:

a. Onset = syllables must have onsets

b. Nuclear Harmony (NucHarm) =

if |x| > |y| then Nuc/x >- Nuc/y

The interpretation of these constraints is straightforward. In particular, Onset simply requires syllables to have onsets (remember that the most unmarked core syllable is CV, not V), in a manner equivalent to the Minimal Onset Satisfaction principle of Roca (1994) (in turn incorporating insights from Selkirk 1982, Steriade 1982, etc.). NucHarm dictates that, given two segments x and y, such that x is more sonorous than y ('|x| > |y|'), then x is a better (or 'more harmonic') nucleus than y ('Nuc/x >- Nuc/y'). Again, this is clearly a simple rephrasing of a universal principle of markedness.

Let us next look at the interaction between these two constraints in IT Berber. We shall postulate the ranking in (18):

(18) Onset >> NucHarm

In prose, satisfaction of Onset must take priority ('>>') over satisfaction of NucHarm.

In order to justify this ranking, we shall examine a tableau (= a table displaying a set of possible candidate parses output by Gen, and their respective fate under each constraint) for the underlying form /txznt/ 'you stored' (each * represents one constraint violation; an exclamation mark ! signals that the corresponding violation mark is fatal, i.e. that it effectively disposes of the candidate being evaluated; the optimal candidate is marked with an arrow head '>'):

(19)

candidates	*constraints*		*comments*
	Onset	NucHarm	
T.X.Z.N.T	*!***	nzxtt	NucHarm irrelevant
T.xZ.Nt	*!	n z t	NucHarm irrelevant
> tX.zNt		n x	optimal
Tx.zNt		n t!	|n| = |n|, |t| < |x|
tXz.nT		x! t	|x| < |n|, t irrelevant

Clearly, the two constraints in (18) will not be sufficient to account for all and only the existing types of syllables, both universally and in IT Berber. Accordingly, further syllabification-related constraints must be postulated, as follows:

(20) further syllabification constraints:

a. Parse =
segmental material must be incorporated into syllabic structure

b. Fill =
syllabic structure can only be built on segmental material

c. ~Coda =
there is no coda

d. ~M/a =
/a/ is not parsed in the syllable margin

Parse and Fill are the two 'faithfulness' constraints enforcing isomorphy between underlying and surface representations. ~Coda (or No-Coda) again aims at the universally unmarked syllable CV. Finally, ~M/a is an extreme instantiation of NucHarm, simply excluding the maximally sonorous segment /a/ from the syllable margins (thus forcing its parsing in the syllable nucleus).

I now illustrate the ranking of these constraints, and its consequences, in IT Berber. For convenience, I shall use the abstract underlying sequences /naa/, /nia/, /nai/, and /tk/ (dotted vertical lines between constraints in tableaux conventionally represent equality of ranking; a continuous line

indicates a left-to-right hierarchical relation):

(21)

	ONS	PARSE	$FILL^{nuc}$	~M/a	$FILL^{onset}$	NucHarm	~CODA
> nA.[]A					*	a a	
naA				*!		a	
nA<a>		*!				a	
nA.A	*!					a a	
> N.iA						a n	
nI.[]A					*!	a i	
nIa				*!		i	*
> nAi						a	*
nA.[]I					*!	a i	
naI				*!		i	
> tK						k	
t[]k			*!				*

As can be seen, all the facts of IT Berber syllabification are accounted for satisfactorily. The obvious advantage of this analysis over its counterpart in the standard theory lies in the simplicity and homogeneity of its machinery. In particular, the constraint inventory is universal, all the desired facts then simply falling out of a given language-specific ranking, as has been shown (note that the Onset >> NucHarm ranking needs reversing in phrase-initial position; this undesirable twist is however replicated in the unsightly standard condition '(Y) = ∅ only phrase-initially'). By contrast, the standard machinery displayed in (16) above is disturbingly diverse, as we pointed out at the time.

3. Underlying Representations in OT

As mentioned in the introduction, the acid test for theory evaluation is not so much descriptive success (which OT manifestly achieves for the data under scrutiny), but psychological plausibility from the perspective of ordinary language learning. This issue is directly addressed in Tesar & Smolensky (1993), who taxonomise knowledge of language under OT as in (22):

(22) knowledge of language under OT

i. Gen (= mapping of universal inputs onto universal outputs)
ii. constraints (on Gen outputs)

iii. underlying forms
iv. constraint ranking

On these, i. and ii. are part of Universal Grammar, and thus unlearned. The learning problem therefore only concerns iii. and iv., to which we now turn.

The issue of how underlying forms are set up by the learner, typically overlooked in the phonological literature, is specifically addressed in Prince and Smolensky (1993) under the label 'lexicon optimisation'. In particular, these authors propose to constrain underlying forms by means of the following principle:

(23) Lexicon Optimisation (Prince and Smolensky 1993:192):

> 'Suppose that several different inputs I_1, I_2, ..., I_n when parsed by a grammar G lead to corresponding outputs O_1, O_2, ..., O_n, all of which are realised as the same phonetic form ϕ - these inputs are all *phonetically equivalent* with respect to G. Now one of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labelled O_k. Then the learner should choose, as the underlying form for ϕ, the input I_k'.

This principle can effectively be interpreted as implementing respect for PARSE and FILL, the faithfulness constraints (so-called precisely because their role is indeed to ensure a faithful reproduction of the underlying representation in the surface form), by the learner setting up underlying representations, as we shall now see.

Take for example the sequence [CV], which constitutes the universally preferred syllable. Other things being equal, it of course makes sense to postulate /CV/ as the corresponding underlying representation, as any self-respecting phonologist undoubtedly knows. But how is this underlying form arrived at by the learner from the perspective of OT? Remember that the relevant constraints are PARSE and FILL, the faithfulness

constraints. I illustrate various logical possibilities in (24):

(24) possible sources of [CV]:

surface parse	hypothesised URs	constraint violations
CV	/CV/	
CV.<V>	/CVV/	*PARSE
<C>.CV.<V>	/CCVV/	**PARSE
<C><C>.CV.<V><V>	/CCCVVV/	****PARSE
C[]	/C/	*FILL
[]V	/V/	*FILL
[][]	/0/	**FILL
[]V<C>	/VC/	*FILL, *PARSE
<V>C[]	/VC/	*PARSE, *FILL
etc.		

As can be seen, all but the first of these UR candidates, /CV/, incur violations of the two given constraints. Consequently, by the Lexicon Optimisation Principle in (23), /CV/ will be selected as the underlying form of [CV], all according to common sense and phonologist's intuition, as pointed out.

The irrelevance of the remaining constraints to lexical optimisation, in the sense of (23) above, will now be demonstrated. Suppose that the input datum is [VC]:

(25) possible sources of [VC]:

surface parse	hypothesised URs	constraint violations
VC	/VC/	*ONSET, *~CODA
<C>V[]	/CV/	*ONSET, *~CODA, *PARSE, *FILL
<C>VC	/CVC/	*ONSET, *~CODA, *PARSE
[]C	/C/	*ONSET, *~CODA, *FILL
V[]	/V/	*ONSET, *~CODA, *FILL
<CC>V<V>C<C>	/CCVVCC/	*ONSET, *~CODA, ****PARSE
etc.		

As can be seen, any deviation of the UR from the surface form automatically results in the violation of the faithfulness constraints. Additional constraints are contravened by parsings yielding outputs at variance with universal unmarkedness, but this situation cannot be repaired by tinkering with the UR, since constraint violation is obviously computed on the surface form, and this is given. The upshot of the discussion is therefore that the Lexicon Optimisation Principle (23) will force the selection of URs identical to the surface realisation, a result no doubt highly encouraging for the hard-nosed empiricist phonologist all along suspicious of SPE-type abstractness.

The joy of such a phonologist will however be short-lived, since things become considerably more complex (and more lively) as soon as alternation (a fact of life for natural languages) is brought into the picture. In particular, Prince & Smolensky formulate a

Minimal Redundancy Principle disfavouring the presence of lexical material:

(26) Minimal Redundancy Principle (Prince & Smolensky 1993: 195):

> To the maximal extent possible, information should be excluded from the lexicon which is predictable from grammatical constraints

A more general (and more extreme), optimality-couched version of this principle is given in (27) in the form of a negative constraint:

(27) ~Spec (Prince & Smolensky 1993: 196):

> Underlying material must be absent

The relevance of these additional considerations will now be exemplified with the passive conjugation allomorphy of Maori, as described in the by now classic account of Hale (1973). The basic morphology is presented in (28):

(28)

UR		surface		
stem	affix	inflected	uninflected	
CVCV	+V	CV.CV.V	CV.CV	
kite	+a	ki.te.a	ki.te	'to embrace'
patu	+a	pa.tu.a	pa.tu	'to kill'

As can be seen, inflection consists of a suffix *-a*, added to the (CVCV) root. The URs postulated simply follow from the principle of Lexical Optimisation in (23), specifically from the action of the faithfulness constraints.

Consider now the forms in (29):

(29)

UR		surface		
stem	affix	inflected	uninflected	
CVCV	+CVV	CV.CV.CV.V	CV.CV	
wero	+hia	we.ro.hi.a	we.ro	'to stab'
tohu	+ŋia	to.hu.ŋi.a	to.hu	'to point out'
hopu	+kia	ho.pu.ki.a	ho.pu	'to catch'

The postulated URs are again faithful to the surface forms. The difference with (28) lies in the suffix, which is now CVV, with the additional complication that its initial consonant is seemingly unpredictable. Strict adherence to the faithfulness constraints will thus inevitably lead to the establishment of a sizeable number of conjugation classes, with the corresponding multiplication of underlying suffixes (i.e. one for each consonant), against the grain of the economy principles (26) and (27).

An alternative analysis circumventing both these difficulties is

however available:

(30)

UR		surface		
stem	affix	inflected	uninflected	
CVCVC	+VV	CV.CV.CV.V	CV.CV.<C>	
weroh	+ia	we.ro.hi.a	we.ro	'to stab'
tohuŋ	+ia	to.hu.ŋi.a	to.hu	'to point out'
hopuk	+ia	ho.pu.ki.a	ho.pu	'to catch'

What we are now doing is assigning the ostensibly suffix-initial consonant to the stem. The immediate consequence of this move is the reduction of the suffixal allomorphy to /a/ and /ia/. This remaining allomorphy is moreover reducible to rule (/a/ after a vowel, and /ia/ after a consonant), and consequently, we can do away with all conjugation classes.

We are still seemingly paying the price of a deletion rule disposing of the underlying stem-final consonant in uninflected forms. Prince & Smolensly, however, point out that such deletion will fall out of the (inviolable) syllabic template of Maori: (C)V. In particular, because codas are disallowed across the board in this language, the constraint ~CODA will be undominated, i.e. placed at the top in the ranking (this ranking is of course still unavailable at the time URs are being learnt: Prince & Smolensky are simply anticipating this result at this point; note however that the question still remains as to what makes the learner decide precisely for the desired UR in the absence of the relevant ranking information). Such a position in the ranking (crucially shared with FILL) will ensure deletion (more precisely, underparsing) of the underlying stem-final consonant of verbs word-finally, and thus no specific rule or equivalent will be necessary to achieve this result:

(31)

a.

	FILL : ~COD	PARSE	ONSET
> we.ro.hi.a	:		*
we.ro.h<i>a	:	*!	
we.ro.<h>i.a	:	*!	**
we.roh.i.a	: *!		**

b.	FILL	~COD	PARSE	ONSET
we.ro.h[]	*!			
> we.ro.<h>			*	
we.roh		*!		

Underlying forms such as /weroh/ are therefore optimal in the context of the observed alternation, even though they infringe the Lexicon Optimisation Principle (23). In particular, this solution is superior to the one that multiplies the UR of suffixes, examined above, as a consequence of the Minimal Redundany Principle (26) (or its bare bone constraint incarnation in (27). This means that, in the event of conflict between these two principles, the Minimal Redundany Principle (26) emerges victorious, since it is precisely this principle that licenses violation of the faithfulness conditions, and thus the existence of URs diverging from surface forms.

The implications of this scenario for learnability are obvious. We must assume that both the Minimal Redundancy Principle (26) and the Lexicon Optimisation Principle (23) are utilised by the learner as part of the general language learning algorithm, and that this algorithm awards greater weighting to the former constraint (itself crucially restrained by the caveat 'to the maximal extent possible', which obviously stands in the way of wild suppression of surface substance underlyingly). We have shown that, given this assumption, the desired results follow automatically from the set of available data. Note, however, that in the real world such data clearly do not become accessible instantaneously, and therefore the acquisition of URs will necessitate gradual exposure to a rich array of data over time.

4. Learnability of Constraint Ranking

We now turn our attention to the issue of learnability of constraint ranking, specifically addressed in Tesar & Smolensky (1993).

As will be recalled, Tesar & Smolensky (1993) assume that the basic material available to the learner are surface forms (given) and their corresponding URs (arrived at in the way described in the previous section). Thus, assuming the surface form [tola] for some hypothetical language L_1, the specific evidence directly available to the learner will be as in (32) ('[]' symbolises the abstract segment resulting from overparsing; we are obviously assuming that epenthetic consonants are realised as [t] in this language):

(32) positive evidence:

/VCVC/ ->	[]V.CV.<C>	L_1
olas	[]o.la.<s>	
	[to.la]	

The proposed underlying form /olas/ is of course not faithful to the surface form [tola], and therefore we must assume the existence of alternations in the general data body of L_1 motivating such a degree of abstractness, in line with our discussion in the previous section (remember, in particular, that infringements of Lexicon Optimisation must be offset by successes of Minimal Redundancy).

The parsing corresponding to L_1 in (32) is thus empirically legitimised. By contrast, all other parsings generated by Gen from the given UR will be in conflict with the facts. Such parsings (corresponding to languages L_2, L_3, L_4, etc., ad infinitum given FILL, all distinct from L_1), displayed in (33) below, constitute therefore negative evidence readily inferable by the learner (NB. vowel epenthesis is assumed to be implemented as [i] in these languages):

(33) negative (inferred) evidence:

V.CVC	L_2
o.las	
*[o.las]	
<V>.CV<C>	L_3
<o>.la <s>	
*[la]	
<V>.CV.C[]	L_4
<o>.la.s[]	
*[la.si]	
ad infinitum	⋮

As can be seen, the legitimate and illegitimate forms (all of them parse 'candidates') in (32) and (33), respectively, incur (or may incur) a number of constraint violations. The list of constraints being universal, and therefore unlearned, it will be possible for the learner readily to verify such violations, as we now represent in the table (NB. not tableau!) in (34):

(34) L_1 candidate evaluation:

	ONS	~COD	$FILL^{nuc}$	PARSE	$FILL^{ons}$
a. *V.CVC	*	*			
b. *<V>.CV.<C>				**	
c. *<V>.CV.C[]			*	*	
d. []V.CV.<C>				*	*

The table simply displays which of the various constraints are violated by each candidate. Such candidate evaluation is of course completely independent of constraint ranking (the constraints are obviously still unranked, the whole point of the exercise being precisely that of arriving at a ranking on the basis of the raw data).

The data in (32) and (33) above (all accessible to the learner, as we have seen) can readily be arranged as data pairs, as in (35):

(35) data pairs:

V.CVC -< []V.CV.<C>

<V>.CV.<C> -< []V.CV.<C>

<V>.CV.C[] -< []V.CV.<C>

In particular, each possible but empirically unsubstantiated parsing of the UR is stated as less harmonic ('-<') than the attested parsing (corresponding to L_1 in our example), as corresponds to the general scheme in (36) (in Tesar & Smolensky's terminology, 'suboptimal' refers to the parsings that yield illegitimate forms, and 'optimal' to the parsing corresponding to the attested form):

(36) $subopt_i$ -< opt

Specifically, each suboptimal parsing ('$subopt_i$') is less harmonic than its optimal ('opt') congener (indeed by definition!).

Now, such data pairs, automatically derived from the conjunction of the positive and negative evidence, as we have seen, contain the seed of constraint ranking. In particular, given the logic of

OT, where harmony is a function of constraint ranking, the suboptimal candidates can only be so if the constraints they violate (as manifested in the marks they incur) are ranked higher than the constraints violated by the optimal candidate:

(37) marks(subopt) >> marks(opt)

The next step in the procedure consequently involves the comparison of the constraint violations of each of the suboptimal candidates with those incurred by the optimal candidate, as illustrated in (38) (the labels a ... d refer to the lines in table (34) above):

(38) L_1 mark-data pairs:

	$subopt_i$ -< opt_i	marks(subopt)	marks(opt)
a -< d	V.CVC -< []V.CV.<C>	{*ONS, *~COD}	{*PARSE, $*FILL^{ons}$}
b -< d	<V>.CV.<C> -< []V.CV.<C>	{*PARSE, *PARSE}	{*PARSE, $*FILL^{ons}$}
c -< d	<V>.CV.C[] -< []V.CV.<C>	{*PARSE, $*FILL^{nuc}$}	{*PARSE, $*FILL^{ons}$}

As can be seen in (38), it is possible for the same constraint to be violated by both the optimal and the suboptimal candidate in the same line. Such a situation comes under the remit of the Cancellation/Domination Lemma of Prince & Smolensky (1993):

(39) Cancellation/Domination Lemma (Prince & Smolensky 1993):

> Suppose two parses B and C do not incur identical sets of marks. Then B >- C if and only if every mark incurred by B which is not cancelled by a mark of C is dominated by an uncancelled mark of C

In particular, harmony relations are, reasonably, only established on the basis of uncancelled marks. Equivalently, marks incurred on the same constraint by both candidates in the same line have no effect on their relative harmony. Accordingly, in the next step in the procedure, marks common to both candidates in each line are cancelled from the table of mark-data pairs (cancelled marks have been struck out, for greater visual clarity):

(40) common mark cancellation:

	$subopt_i$ -< opt_i	marks(subopt)	marks(opt)
a -< d	V.CVC -< []V.CV.<C>	{*ONS, *~COD}	{*PARSE, $*FILL^{ons}$}
b -< d	<V>.CV.<C> -< []V.CV.<C>	{*PARSE, *PARSE}	{*PARSE, $*FILL^{ons}$}
c -< d	<V>.CV.C[] -< []V.CV.<C>	{*PARSE, $*FILL^{nuc}$}	{*PARSE, $*FILL^{ons}$}

The table in (40), processed directly from the raw evidence, as we have seen, contains all the data on which the constraint ranking learning algorithm will operate.

In the initial state of this algorithm, all constraints are equally ranked (in fact, they are supposedly unranked), as corresponds to their neutral state in Universal Grammar (notice that this points to a Superset Principle, by which languages hypothesised at earlier learning stages are supersets of languages hypothesised later):

(41) constraint ranking learning algorithm ('H' = 'hierarchy'):

initialisation: $H = H_o$
{ONS, ~COD, PARSE, $FILL^{nuc}$, $FILL^{ons}$}

In the next step, the first line in table (40) (a -< d) is examined. $FILL^{ons}$ and PARSE are violated by the optimal candidate. In the logic of OT, this means that they must be out-ranked by the constraints violated by the suboptimal candidate, viz. ONSET and ~CODA. In the model of Tesar & Smolensky, this situation induces rearranging of the present constraint ranking. In particular, $FILL^{ons}$ and PARSE are demoted to the next lower rung of the hierarchy:

(42) constraint demotion (for a -< d):

a: {ONS,~COD, $FILL^{nuc}$} highest-ranked constraints

{$FILL^{ons}$, PARSE} not-yet-ranked constraints

Any suboptimal candidate incurring violation of one of the constraints ranked highest in (42) is automatically accounted for by the current ranking, for the simple reason that such a constraint is already ranked higher than $FILL^{ons}$ and PARSE, which have been demoted in (42). Consequently, any line containing a violation of any of the highest ranking constraints can be eliminated from the computation. This obviously disposes of line a. in table (40); also of line c., where the suboptimal candidate

violates highest-ranked $FILL^{nuc}$. The mark table is therefore reduced as in (43):

(43) reduced mark table:

	$subopt_i$ -< opt_i	$marks_{subopt}$	$marks_{opt}$
b -< d	<V>.CV.<C> -< []V.CV.<C>	{*PARSE, *PARSE}	{*PARSE, $*FILL^{ons}$}

Notice that PARSE, violated by the suboptimal candidate in the remaining line, is not included in the set of highest ranked constraints in (42).

As expected, the familiar demotion procedure is reapplied to the reduced mark table in (43):

(44) constraint demotion (for b -< d):

b:	{PARSE}	next-highest ranked constraints
	$\{FILL^{ons}\}$	not-yet-ranked constraints

The ranking of PARSE above $FILL^{ons}$ induces removal of line b in (43). After the removal of this line, the mark table becomes empty, in the obvious way:

(45) reduced mark table:

	$subopt_i$ -< opt_i	$marks_{subopt}$	$marks_{opt}$

Further application of constraint demotion to this table simply terminates the algorithm, since there are no constraints left to be demoted:

(46) constraint demotion:

$\{FILL^{ons}\}$	next-highest ranked constraints
{}	not-yet-ranked constraints

The resulting stratified hierarchy is thus as in (47):

(47) L_1 stratified hierarchy:

$$\{ONS, \sim COD, FILL^{nuc}\} >> \{PARSE\} >> \{FILL^{ons}\}$$

The crucial points are that the algorithm terminates, and that the outcome has been arrived at deterministically. The constraint ranking responsible for any particular form is therefore logically learnable.

A different consideration concerns learnability load. In particular a logically learnable ranking may not be learnable in real time, specifically in the real acquisition time pertinent to the acquisition of language.

Tesar & Smolensky (1993) contend however that this is not the case. The steps in their argument are as follows. First, each pass through the table of mark-data pairs must output at least one constraint. If so, the number of passes cannot be greater than the number of (universal) constraints ('N_{constr}'). Second, the number of steps in each pass cannot be greater than the number of uncancelled marks in the table, i.e. maximally N_{constr} x N_{pairs} ('N_{pairs}' = number of lines in the mark-data table). Consequently, the total number of steps involved in the implementation of the algorithm is as in (48):

(48) learnability load: $(N_{constr})^2 N_{pairs}$

The product of this equation is likely not to exceed a few thousand, a figure readily copable with by the neurological machinery of man. Consequently, not only is constraint ranking logically learnable, but it also appears to be learnable under the real-world conditions relevant to language. As things stand at the moment, however, this conclusion applies to lexical items piecemeal, and it remains to be demonstrated that the induction of the general constraint ranking relevant to any one grammar indeed is a feasible task. Crucially, learning of a language involves learning of its lexical items, in the sense that it cannot be said that the language is known unless the lexical items are known. Clearly, if the lexical items are known, the overall ranking, deterministically derived from such items, can be learnt.

Acknowledgements

Previous incarnations of this paper were presented at the University of Essex, the University of Stuttgart, and FAS, Berlin. I am grateful to the respective audiences for questions and comments, and to Greg Dogil and Richard Wiese for kindly inviting me to talk. I also thank Sylvia Löhken for giving me no choice about writing out the paper, and Martin Atkinson for comments on a prior version that considerably tempered my earlier sanguinity. All errors are mine.

References

Clements, G.N. (1990) 'The Role of the Sonority Cycle in Core Syllabification', in J. Kingston and M. Beckman, *Papers in Laboratory Phonology I*, Cambridge: Cambridge University Press, 283-333.

Clements, G.N. & S.J. Keyser (1983) *CV-Phonology*, Cambridge, MA: MIT Press

Dell, F. & M. Elmedlaoui (1985) 'Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber', *Journal of African Languages and Linguistics* 7: 105-130.

Dell, F. & M. Elmedlaoui (1988) 'Syllabic Consonants in Berber: Some New Evidence', *Journal of African Languages and Linguistics* 10: 1-17.

Elmedlaoui, M. (1985) Le parler berbère chleu d'Imdlawn: segments et syllabation, Doctorat de Troisième Cycle, Université de Paris VIII, Saint-Denis.

Guerssel, M. (1985) 'Glides in Berber and Syllabicity', *Linguistic Inquiry* 17: 1-12.

Levin, J. (1985) A Metrical Theory of Syllabicity, PhD dissertation, MIT.

McCarthy, J. & A. Prince (1993) Prosodic Morphology I, ms., University of Massachusetts at Amherst and Rutgers University.

Prince, A. (1985) 'Improving Tree Theory', in *Proceedings of the XIth Annual Meeting of the Berkeley Linguistics Society*, 471-90.

Prince, A. & P. Smolensky (1993) Optimality Theory, ms., Rutgers University and University of Colorado at Boulder.

Roca, I. (1994) *Generative Phonology*, London: Routledge.

Selkirk, E. (1982) 'The Syllable', in H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representations (Part II)*, Dordrecht: Foris, 337-83.

Steriade, D. (1982) Greek Prosodies and the Nature of Syllabification, PhD dissertation, MIT.

Tesar, B. & P. Smolensky (1993) 'The Learnability of Optimality Theory', ms., University of Colorado at Boulder.

Multiple constraint-rankings in Polish[1]

M.M. Verhijde - Research Institute for Language and Speech (OTS),
University of Utrecht, Trans 10, 3512 JK, Utrecht, The Netherlands
email: VERHYDE@RUULET.NL

1. Introduction

Optimality Theory proposes that constraints are universal, minimally violable and ranked in a language-specific way. Different grammars result from differences in constraint-ranking, a hypothesis that allows cross-linguistic typological issues to be stated in a straightforward way[2]. What happens when we encounter exceptions to the general phonological patterns? Are these exceptions to be expressed directly within the constraint-ranking, either by allowing morpheme-specific constraints to interact with more general ones within a single constraint-ranking, or by postulating multiple constraint-rankings? Or do we assert that constraints are inadequate of capturing exceptional phonological patterns, leaving little else to do but to assume such patterns in some shape to be part of the input (e.g. underlying representation, lexicon)? In other words, how do we account for morpheme-sensitive phonology within Optimality Theory? This is the main question of this paper.

Inkelas, Orgun, Zoll (1994) propose the following divisions in describing phonological patterns. Regular and subregular patterns are accounted for by distinct constraint-rankings (or cophonologies). The motivation for postulating a distinct constraint-ranking is productivity: cophonologies may be set up only if they are productive, e.g. morphologically[3]. In other words, constraint-rankings are postulated if two criteria are met: the regularity is supported by evidence from alternations and the class of morphemes belonging to the regularity is definable on independent grounds. Nonproductive phonological patterns may not be attributed to a separate constraint-ranking and are captured via prespecification of the phonological input.

The classification of regular, subregular and exceptional patterns in phonology seems to be less crisp than suggested in Inkelas, Orgun, Zoll (1994)[4]. In this paper I compare different strategies with respect to such fuzzy phonological patterns. In contrast with Inkelas, Orgun, Zoll's hypotheses I want to show that (i) due to the well-known tradeoff between phonological input and set of procedures or constraints, many positions may be taken, i.e. there appears to be no principled reason that decides which strategy is favourable above other ones[5];

[1] I want to thank Toni Borowsky, Jan Don, Chris Golston, Beth Hume, Sharon Inkelas, Uwe Junghanns, René Kager, Sylvia Löhken, Orhan Orgun, Wim Zonneveld and the audience of GGS 1995 at Jena for discussing various aspects that have found their way into this manuscript. "Multiple constraint-rankings in Polish" will appear as the third chapter of my forthcoming PhD-thesis, titled "Cycles, Relics and Scars". I will adress several objectives in this thesis. First, I aim to investigates the intimate relationship between the instrument of phonological cycle and any procedural model of phonology, which includes Cyclic Phonology, Lexical Phonology and theories of Prosodic Phonology. Second, I discuss more declarative perspectives on the phonological cycle, with special attention to Optimality Theory (OT). Third, within the framework of OT a number of declarative alternatives for core cyclic phenomena are provided, which includes analyses of Polish, Sanskrit and French. In chapter 1 the phonological cycle is introduced and motivated. OT is outlined in chapter 2; due to its declarative nature, cyclic phenomena are either irrelevant or extremely troublesome for OT. I propose a functional marriage between OT and Monotonic Cyclicity to overcome these attitudes with respect to cyclicity. Chapter 3 discusses the role of morpheme-sensitive phonology in OT; I argue against a derivational account of such phenomena, based on evidence from Polish vowel-zero alternations. In chapter 4 an OT approach on French phonology is given. Chapter 5 discusses nonderived environment effects in Sanskrit from an OT-perspective.

[2] Prince & Smolensky (1993:chapter 6) on typology in terms of different constraint-rankings.

[3] This is termed as the Alternation Criterion in Inkelas, Orgun, Zoll (1994), with a clear connotation to previous proposals formulated in Kiparsky (1973) among others.

[4] Inkelas, Orgun, Zoll (1994) proposal must be seen as a methodology to stop proliferation of cophonologies.

[5] Of course, we can rely on notions such as predictability, elegance of grammatical theory or statistical motivations. Notice, however, that these aspects are basically statements about the assumed phonological theories themselves.

(ii) vowel-zero alternations in Polish phonology as an example of a fuzzy phonological pattern is best analyzed in a model with multiple constraint-rankings that are motivated on the basis of distinct morpheme-sets.

The issues of prespecification and underspecification in phonological theory are important in a non-trivial way. Elsewhere I have argued that any procedural model of phonology assumes some sort of representation that forms the input to a set of procedures[6]. Prespecification and underspecification of phonological representations may be considered as distinct but related aspects of the dichotomy between phonological input and set of procedures or constraints. I have little to say about underspecification theories and Optimality Theory here; except that the most harmonic candidate selected by Eval should be fully specified for all phonologically relevant information, i.e. temporary underspecification must be resolved[7]. Thus it really makes no difference whether some aspect of phonology is absent from the underlying representation and provided by virtue of a Fill-violation, or appears to be present in the input but invokes underparsing. For discussion on underspecification theories the reader is refered to the insights of Mohanan (1991), Steriade (1994) and especially Inkelas (1994) and Smolensky (1994).

Prespecification, or lexical listing of phonological information as a theoretical concept resembles underspecification to a large degree, as the other side of the same coin. We hypothesize that an underspecified phonological input is guided by principles of predictability[8]. The fact that some element behaves in an unpredictable way forces by opposite reasoning some kind of prespecification of the phonological input. Many autors have considered the above line of argument as correct; for further discussion on the prespecification method see Kiparsky (1993), Inkelas & Cho (1993), Zoll (1993) and Inkelas, Orgun, Zoll (1994), among others. As stated above, Eval selects the most harmonic candidate which is phonologically fully specified and it is irrelevant whether or not this optimal output contains prespecified material as part of the underlying structure. The issue of lexical listing of phonological material is of some interest to the interplay between phonological input and set of procedures (or constraints), but crucially not in determining the most harmonic phonological output. Optimality Theory remains silent on the tradeoff between input versus set of procedures or constraints. We can choose to prespecify phonological information in the input and have Faithfulness constraints dealing with the listed information, ranked among other constraints, or we may account for the phonological behavior directly in the set of constraints themselves. In the latter case unpredictable phonological patterns are related to individual morphemes or to sets of morphemes and therefore constraints must be able to refer to them, for instance in a grammar that has multiple constraint-rankings or in a system that allows morpheme-sensitive constraints, conflicting with more general constraints.

While tentatively concluding that the prespecification method to exceptional phonological patterns is always available (and perhaps necessary), I describe phonological patterns in terms of constraints that are ranked differently, i.e. a grammar may exhibit multiple constraint-rankings, in contrast with the option to have morpheme-sensitive constraints under the hypothesis 'One grammar One ranking'[9]. There are a number of related issues involved, which will be discussed separately.

---

[6] See Verhijde (forthcoming), *Cycles, Relics and Scars*, PhD-thesis, chapter 2.

[7] Consider Smolensky (1994) on the impossibilities of computing the most harmonic condidate by Eval if underspecified information is allowed in the output.

[8] Steriade (1995) mentions the notion of Lexical Minimality: underlying representations must reduce to some minimum the phonological information used to distinguish lexical items. The notion originates in Halle (1959) and Chomsky & Halle (1968).

[9] Any nonlinguistic connotation is thought of as existing in the mind of the reader only.

First, I give some motivation with respect to multiple constraint-rankings over morpheme-sensitive constraints. As I demonstrate, the differences between these two approaches are less interesting than their similarities; both concepts use some notion of morpheme-set that allows for distinctions.

Second, if we deny the methodology outlined in Inkelas, Orgun, Zoll (1994) to postulate multiple cophonologies, then how do we justify a distinct constraint-ranking? In addition, we like to know exactly what it means to have more than one constraint-ranking. Constraints in Optimality Theory are highly conflicting in nature; therefore it must be the case that multiple constraint-rankings are always conflicting with each other. For example, are multiple constraint-rankings available that account for phonological patterns of derived words?

Finally, as I mentioned above, the familiar tradeoff between phonological input and set of procedures or constraints has been extremely influential in generative phonology, at least since Chomsky & Halle (1968). Optimality Theory as a more declarative approach to generative phonology may improve our understanding of the above interplay of input versus constraint-set, on the condition that it receives a non-derivational interpretation. To put it differently, any (pseudo-)derivational extension of the Optimality Theoretical framework will fail exactly in the tradeoff-theme. Examples of such extensions are, for instance, the introduction of cyclicity (Kenstowicz 1994) or the theory of correspondence relationships between input and output (McCarthy & Prince 1994, McCarthy 1995, Orgun 1994, 1995; Inkelas 1995).

The paper is organized as follows. An Optimality Theory overview of possible strategies with respect to morpheme-sensitive phonology is presented in section 2. I will focus on two issues here, namely phonology that is sensitive to individual morphemes or a particular set of morphemes, and multiple constrant-rankings within a single grammar. Vowel-zero alternations or yers in Polish phonology are the subject of section 3. As I will demonstrate, within Optimality Theory it appears to be irrelevant whether or not yers are represented as part of the phonological input; instead reference to a specific set of morphemes mst be made, which requires a distinct constraint-ranking. In section 4 the proposal is elaborated upon with references to other aspects of Polish phonology. Some remarks are discussed in section 5.

2. Morpheme-sensitive phonology

Leaving the prespecification method aside for the moment, the interplay of constraints and specific morphemes may be captured basically in two opposite ways. Either we allow constraints to become less universal by incorporating morpheme-specific information directly into them, or we maintain the hypothesis of universal constraints and allow for multiple constraint-rankings within a single grammar. In other words, we may choose to increase the complexity of constraints (section 2.1.) or to extend the number of constraint-rankings (section 2.2.). But first the notions 'morpheme' and 'set of morphemes' have to be clarified[10].

I consider a morpheme in the first place as a minimal meaningful element, in the structuralistic tradition that goes back to Bloomfield (1933). Of course, many different views may be positioned as to what the notion 'meaningful' means. In a given context a morpheme is considered to be a composite element that exhibits a number of distinct characteristics, such as semantic structure, grammatical function or phonological form. These

[10] My assumptions concerning morphemes and sets of morphemes are largely similar to the views discussed in Spencer (1991:4-8), although (i) I implicitely assume allomorphy to arise from constraint interaction and (ii) I do not discuss the topic of suppletion.

properties and their various combinations make up individual morphemes. While selecting a specific characteristic it is possible to group specific morphemes together as sets. In other words, a set of morphemes may be seen as a temporary and artifical generalization across individual morphemes that illuminate (at least) a single linguistic property.

The following examples illustrate the informal view given above. Prefixes are a set of morphemes that share the property that they are concatenated at the left edge of another set of morphemes (such as stems or roots). Nouns are a morpheme-set that may be characterized by the property of their grammatical function. As research from Siegel (1974) to Fabb (1988) indicate, suffixes in English may or may not be described as divided into two different sets of morphemes (or classes), largely depending on hypotheses concerning their distribution and phonological make-up. In principle any linguistic property can be isolated across individual morphemes to create a morpheme-set.

In addition, sets of morphemes may be thought of as containing more than a single linguistic property. It is here that the notion of a well-defined morpheme-set becomes less well-defined or fuzzy. An illuminating example is given in Itô & Mester (1994) in which Japanese morphemes, traditionally divided into four contrasting morpheme-sets (or strata), are analyzed as being constructed from a large number of interactive phonological constraints. Similar idiosyncratic morphological information like [+Latinate] in English (Chomsky & Halle 1968) or [Learned] in French (Walker 1975, Dell & Selkirk 1978) may be captured as single linguistic properties that divide morphemes into sets, but could perhaps be reduced to other characteristics, as it seems to be appropriate in the Japanese case.

2.1. Complexity of constraints

The introduction of Generalized Alignment (McCarthy & Prince 1993a) in Optimality Theory allows a direct reference to sets of morphemes and, according to some researchers, to individual morphemes. Generalized Alignment has been proposed as a way to capture effects of constituent-edges in phonological and morphological theory[11].

(1) **Generalized Alignment**

Align(Cat1, Edge1, Cat2, Edge2) $=_{def}$

$\forall$ Cat1 $\exists$ Cat2 such that Edge 1 of Cat1 and Edge 2 of Cat2 coincide

Where: Cat1, Cat2 $\in$ PCat $\cup$ Gcat; Edge1, Edge2 $\in$ {Right, Left}

As McCarthy & Prince suggest, PCat and GCat consist of sets of prosodic and grammatical (morphological or syntactic) categories provided by linguistic theory. The set of prosodic categories includes at least elements such as PrWd, F, σ, μ and segmental (featural) information such as Place features or Tone. With respect to grammatical categories a more restricted set of choices is proposed, namely MWd, Stem, Affix and Root.

[11] In essence the theory of Generalized Alignment is not limited to any specific phonological or morphological subtheory, as it is indicated in McCarthy & Prince (1993a:81).

Alignment interacts with other constraints in a language-specific constraint-ranking. A familiar example from Tagalog may be illustrative here. In Tagalog, the affix *-um-* wants to be concatenated as close as possible to the left stem-edge, provided that its final nasal consonant is not syllabified into coda position. Thus, whereas prefixation is preferable, infixation occurs under domination of the constraint NoCoda (McCarthy & Prince 1993a:79; dots represent syllable boundaries).

(2) *-um-* Infixation in Tagalog

u.ma.ral		'teach'
su.mu.lat	* um.su.lat	'write'
gru.mad.wet	* um.grad.wet	'graduate'

McCarthy & Prince account for this pattern via two interacting constraints. First, the wish for leftmost position is attributed to Align($[um]_{Af}$,L,Stem,L), satisfied in the output *umaral*, but minimally violated in *sumulat* and *grumadwet*. Due to the dominance of the prosodic constraint NoCoda possible alternatives such as * *umsulat* and * *umgradwet* are less harmonic[12].

The general schema of Alignment is provided by Universal Grammar. However, this need not be true for the possible arguments PCat and GCat. The theory of Generalized Alignment in itself does not provide a principled method to limit possible candidates, because this largely depends on available theories of prosodic and grammatical structures. Lack of restrictedness on the argument set allows for any kind of coincidence between prosodic and morphological information. Consider the following examples of alignment between prosodic and grammatical categories below.

(3) a. ALIGN(STEM, R, σ, R) (McCarthy & Prince 1993a)
in Axininca Campa, Lardil, Hebrew, Bedouin Arabic, Kamaiurá

b. ALIGN(STEM, L, PRWD, L) (McCarthy & Prince 1993a)
in Axininca Campa, Lardil, German, Polish, Malay-Indonesian, English

c. ALIGN(STEM, L, FT, L) (Kager 1994)
in Sibutu Sama

d. ALIGN(STEM, R, FT, R) (Kager 1995)
in Estonian

e. ALIGN(ROOT, L, PRWD, L) (Rowicka 1994)
in Polish

f. ALIGN(μ_H, R, MWD, R) (Zec 1995)
in Neo-Štokavian dialect of Serbo-Croatian

g. ALIGN($[\text{PRWD}]_{INF}$, R, [NUCLEUS], R) (Féry 1994)
in German

[12] Notice that NoCoda is violated once in the correct output candidate *sumulat* and twice in *grumadwet*. It appears to make a difference whether a constraint is addressed to in a nonderived or a derived form. McCarthy & Prince (1994) discuss this unexpected property that emerges from the very mechanisms of constraint interaction used in Optimality Theory.

All alignment constraints cited above are attested and proposed for various reasons. However, there are some remarks with respect to the interaction between phonology and morphology.

2.1.1. Morpheme-sensitive constraints

The use of the grammatical category $[um]_{aff}$ in the Tagalog Alignment constraint is confusing. McCarthy & Prince (1993a) undoubtely assume a distinction between GCat=Affix and a specific member of this category - *um*-. Due to the line of their argument it is necessary to show the behavior of an affix that ends in a consonant; only here the conflict with NoCoda becomes apparent. Consequently, the Alignment constraint is not morpheme-specific, because it does not refer to individual morphemes such as *-um-*.

Does Optimality Theory allow constraints to refer to individual morphemes or only to sets of morphemes? As pointed out in Russell (1995), it all depends on the definition of morphemes. It is perfectly plausible to view morphemes as *objects* of linguistic analysis, e.g. as representations that may be processed upon. Differences between morphemes can be contributed to differences in representations. This clearly emphasizes the prominence of phonological input.

The introduction of morpheme-specific constraints into the above representational approach suffers from two disadvantages. First, concerning the tradeoff between input and set of procedures (or constraints), it is redundant to add constraints that refer specifically to individual morphemes, hence burden both input and constraint ranking[13]. Second, it is not clear how to *rank* a constraint that refers to an individual morpheme. To illustrate this point consider an example from Inkelas (1994). Here several morpheme-specific Alignment constraints are used to capture the exceptionality of nonneutral morphemes in Turkish stress patterns[14]. Thus she proposes ALIGN(*mE*, L, σ', R), ALIGN(*Iyor*, L, σ'/Ft, L) and ALIGN(*penJere*, R, Ft, R) to account for the pre-stressing suffix */-Me/*, the initial-stressed suffix */-Iyor/* and the penult-stressed root */peJere/*, respectively. Notice that these Alignment constraints should outrank other constraints responsible for the (sub)regular stress assignment. However, the exceptional behavior of these morphemes is thus encoded twice, as individual arguments of the Alignment constraints and as a result of their ranking position.

As pointed out in Russell (1995), the opposite view is to treat morphemes as (clusters of) constraints, which specify what kind of properties the phonological representation must have. This approach is mainly developed in theories of Declarative Phonology, but has also been suggested within the Optimality Theory framework[15]. Russell illustrates this approach with an example of Nisgha coronal coalescence that may be accounted for in terms of a number of Alignment constraints that refer to sets of morphemes, such as the 3sg marker.

The position I am assuming here is that several individual morphemes may share any intelligible property which group them together as a set of morphemes. I do not see any motivation to limit possible criteria for the formation of morpheme-sets, which is in contrast with the position taken in Inkelas, Orgun, Zoll (1994). Below I discuss their objections against a similar unrestricted view to describe patterns in phonology.

---

[13] As pointed out to me by Jan Don, personal communication.

[14] "In the Alignment constraints we will invoke for exceptional stress, a morphological category - really, a specific morpheme - is aligned with a foot or a stressed syllable", i.e. ALIGN(morpheme$_i$ Edge$_i$ σ'/Foot, Edge$_j$) (Inkelas 1994:22).

[15] Theories of Declarative Phonology include Bird (1990), Scobbie (1991) and Russell (1993), examples of a declarative approach towards Optimality Theory are Russell (1995) and Hammond (1995).

I hypothesize that these sets of morphemes, instead of individual morphemes, function as categories of constraints, viz. Alignment constraints. Consequently, the possibility that constraints can refer to individual morphemes should be excluded.

2.1.2. Sets of morphemes

A possible set of morphemes can be defined with reference to Inkelas, Orgun, Zoll (1994), especially regarding their objections to allow separate cophonologies for nonproductive phonological patterns and exceptions[16]. Inkelas, Orgun, Zoll discuss data from Turkish phonology, which contains an example of a productive pattern in stress (Sezer stress), a nonproductive pattern in vowel harmony (Labial Attraction), and a case involving exceptionality (of a regular segmental rule of Coda Devoicing). Based on the Alternation Criterion as mentioned above, only Sezer stress is morphologically active, thus is captured within a distinct cophonology. Why are both Labial Attraction and Coda Devoicing exceptions denied a separate cophonology?

Labial Attraction is a root-structure constraint. If a vowel /a/ is followed by a labial consonant and a high back vowel, respectively, then that high vowel must be round, i.e. /u/. Inkelas, Orgun, Zoll refer to these patterns as /aBu/ sequences. The examples cited below are taken from their article.

(4) a. Some roots that obey Labial Attraction

karpuz	'watermelon'
sabun	'soap'
Habur	(place name)

b. Some roots that disobey Labial Attraction

kapI	'door'
KalamIS	(place name)
tavIr	'attitude'

c. Labial Attraction does not apply across morpheme boundaries

kitap	'book'	
kitab-I	'book-Accusative'	*kitab-U

Within Optimality Theory, the constraints responsible for Labial Attraction must have at least two properties. Labial Attraction is not active in derived environments. This suggests that we have here a case in which the active constraints need to be sensitive to a set of morphemes, namely roots. In addition, they must make reference to a separate group within these roots. In other words, it seems to be unpredictable whether or not a root shows Labial Attraction[17].

Coda Devoicing applies in a straightforward fashion to Turkish syllables. Inkelas, Orgun, Zoll (1994) show that plosives in coda position are devoiced. Voiceless obstruents do not alternate. Some examples are given in (5).

[16] A cophonology is defined in Inkelas, Orgun, Zoll (1994:5) as a ranked set of (universal) constraints. Here it is assumed that distinctions in phonological patterns may be attributed to distinct cophonologies, irrespective of how such cophonologies are structured.
[17] Orgun (1994) demonstrates that less than 75% of the roots that contain an /aBu/ environment actually show Labial Attraction.

(5) a. Coda Devoicing on voice obstruents

/kitab/	kitap	'book' (nominative)
	kitap-lar	'book-plural'
	kitap-tan	'book-Ablative'
	kitab-i	'book-Accusative'
	kitab-a	'book-Dative'

b. No Coda Devoicing on voiceless obstruents

/devlet/	devlet	'state' (nominative)
	devlet-er	'state-plural'
	devlet-i	'state-Accusative'

(6) Exceptional behavior of some roots to Coda Devoicing

/etüd/	etüd	'etude'
	etüd-ler	'etude-plural'
/katalog/	katalog	'catalog'
	katalog-dan	'catalog-Ablative'

Again it seems that the constraints which are responsible for the absence of Coda Devoicing in forms like *etüd* must refer to a separate set of morphemes.

Inkelas, Orgun, Zoll investigate if it is possible to capture the phenomena of nonproductive patterns (e.g. Labial Attraction and exceptions to Coda Devoicing) in a similar way as to productive patterns (e.g. regular stress and Sezer stress). They propose distinct phonological patterns to be a consequence of having multiple cophonologies; morphemes are subjected to these cophonologies[18].

(7) a. Productive Sezer stress

- Cophonology A: enforces Sezer stress - place names, derived, underived, borrowings
- Cophonology B: enforces word stress - all other words

b. Nonproductive Labial Attraction

- Cophonology C: enforces Labial Attraction - some roots
- Cophonology D: enforces no Labial Attraction - all other roots, plus derived forms

c. Coda Devoicing exceptions

- Cophonology E: enforces root-final coda devoicing - some roots
- Cophonology F: enforces no root-final coda devoicing - all other roots

The authors point out that defining phonological patterns by postulating different cophonologies for each individual pattern raises five objections, which I present as questions below.

[18] Notice that the approach outlined in Inkelas, Orgun, Zoll (1994) presupposes morphemes to be objects of phonology, i.e. representations that can be processed upon.

1. **Indeterminacy**: How do we classify morphemes that do not meet the structural description of the constraint that is reponsible for a separate cophonology?
2. **Uninteresting cophonologies**: How do we avoid establishing separate cophonologies to observed regularities that are applicable to all morphemes?
3. **Astronomical cophonology proliferation**: If one constraint may be suitable to set up a distinct cophonology, thus dividing the morphemes of the language, and if a morpheme may be subjected to several cophonologies, then how do we restrict the number of cophonologies of the language?
4. **Heterogeneous morphemes**: Are morphemes assigned exclusively to a single cophonology, i.e. how do we avoid violation of a distinctive constraint with respect to a morpheme?
5. **Heterogeneous words**: Are derived forms assigned exclusively to a single cophonology, i.e. how do we avoid violation of a decisive constraint with respect to a complex word?

On the basis of possible answers to these objections Inkelas, Orgun, Zoll (1994) propose to use (some form of) prespecification of Labial Attraction and Coda Devoicing exceptions. My aim is to show that this is not a necessary conclusion. The five objections are valid only within a particular view on Optimality Theory. By means of an alternative perspective, it will be shown that the objections are not as troublesome as they appear to be at first glance. There are two assumptions which may be interpreted in a different way, thereby providing an escape from prespecification.

First, I do not agree with the assumption expressed in Inkelas, Orgun, Zoll (1994) that the observed phonological patterns differ on single constraints only. Optimality Theory does not use inviolable constraints, but allows constraints to be dominated under certain conditions[19]. Instead of being the result of single constraint, it is assumed that any phonological pattern may emerge from a number of competing and conflicting constraints. For instance, Inkelas, Orgun, Zoll (1994) argue that on the basis of the derived form *tambura-m-dl* 'stringed instrument-1sg.poss-Past', we cannot decide whether or not it is subjected to Cophonology C, enforcing Labial Attraction, whereas nonderived *tambura* may belong to Cophonology C. However, exactly these phonological patterns are to be expected in an Optimality theoretic framework where the effects of constraints may be obscured[20]. Therefore, objections concerning Indeterminacy, Heterogeneous morphemes and similar words do not seem to be correct.

Second, it is true that within a framework that interpretes morphemes only as representations the problem of cophonology proliferation arises. A single morpheme may obtain membership of numerous cophonologies, each of them describing a single aspect of the phonological representation. Inkelas, Orgun, Zoll are right as they claim that this does not lead to extremely interesting phonological insights. However, the shift from a merely object-oriented view of morphemes to a more constraint-oriented one makes the whole question about proliferation redundant. In other words, I assume that constraints and their relative rankings are also

[19] Optimality Theory differs from other declarative phonology theories in the use of 'soft' constraints, i.e. constraints that need not be surface-true. In a slightly different way McCarthy & Prince (1994) discuss this theme, named as the Fallacy of Perfection (or FoP).
[20] The pattern of *tambura-m-dl* illustrates also another aspect of Optimality Theory, namely the occurrence of more unmarked structure in morphologically derived environments. Consider McCarthy & Prince (1994) on this phenomenon.

actively involved in describing the phonological representation of morphemes[21]. The objections regarding cophonologies are thus valid only from a particular point of view.

This eliminates the essence of the critique of Inkelas, Orgun, Zoll (1994) on establishing cophonologies for nonproductive and exceptional patterns. Notice that it does not remove the prespecification method from the grammar, as lexical listing of phonological information is always available as a last resort option. However, morpheme-sensitive phonology may be properly expressed by means of multiple cophonologies, exploiting constraints that are sensitive to any kind of morpheme-set.

2.1.3. Ranking and morpheme-sensitive constraints

Constraints in Optimality Theory are ranking in a language-specific order. How can we observe activities of morpheme-sensitive constraints in the ranking? I discuss two analyses that show Alignment constraints which have particular sets of morphemes as their arguments. In Kager (1994) Alignment constraints that are sensitive to word, stem and root morphemes govern the distribution of main stress and secondary stress. In Golston (1995) the phonological characteristics of roots and words are expressed in terms of Alignment. Notice that the constraints refer to specific morphemes such as stem and word which are part of a hierarchical structure. More concretely, due to the Alignment format we know that it is the phonology at the edges of these morphemes that is refered to. However, what happens if the phonological patterns at a similar edge are distinct? How does Optimality Theory account for these phenomena in derived forms?

Kager (1994) observes that stress in Sibutu Sama[22] is sensitive to morphological structure. Main stress is strict penultimate, whereas secondary stress in unprefixed words is initial.

(8) Stress in unprefixed words in Sibutu Sama

bissála	'talk'
bìssalá-han	'persuading'
bìssala-hán-na	'he is persuading'
bìssala-han-kámi	'we are persuading'

Kager proposes binary trochaic feet to account for the data. Main stress is distributed by ALIGN(PrWd, R, Ft, R) and dominates another Alignment constraint which is responsible for initial secondary stress. In prefixed words multiple secondary stresses are observed.

(9) Stress in prefixed words in Sibutu Sama

a.	màka-bissála	'able to talk'
	pìna-bìssalá-han	'to be persuaded'
	màka-pàgba-bissalá-han	'able to cause persuasion'

[21] Notice that I do not abandon the idea that morphemes may be seen as representations, merely that I do not see an absolute opposition between the two approaches.

[22] Sibutu Sama is an Austronesian language spoken in the Southern Philippines. Kager (1994) makes reference to the work of Allison (1979).

b. kà-pag-bissála 'able to talk to each other'

tà-pag-bìssalá-han 'the thing able to be spoken about'

ALIGN(Stem, L, Ft, L) explains the distribution of initial stress both on the stem and on prefixes (analysed as derived stems). In other words, whenever there is a stem, its left edge should coincide with the left edge of a stress foot. Notice the nice interaction between alignment of main stress (e.g. to the prosodic word edge) and alignment of secondary stress (e.g. to the stem edge). In (9b) there are three possibilities for application of ALIGN(Stem, L, Ft, L), twice at the prefix edges and once at the innermost stem edge. Due to FT-BIN, not every possible edge may coincide with a stressfoot. Kager assumes that the constraint ALIGN(Foot, L, PrWd, L) or ALL-FT-LEFT is responsible for the observed patterns.

However, Sibutu Sama also exhibits a fluctuation of secondary stress in prefixed words, which forces the introduction of another Alignment constraint: ALIGN(Root, L, Ft, L). The examples of the variation are given below.

(10) Variable secondary stress in Sibutu Sama

a. pà-missalá-han ≈ pa-mìssalá-han

'instrument for speaking'

pàg-bissalá-han ≈ pag-bìssalá-han

'the thing spoken about'

b. màka-pag-bìssalá-han ≈ màka-pà-bissalá-han

'able to persuade them'

tàpag-pa-bìssala-hán-bi ≈ tàpag-pà-bissala-hán-bi

'you (pl.) are able to make them persuade someone'

The root is interpreted as the innermost stem morpheme. For a better understanding of the issue involved, a more detailed account of the argument is necessary. First, note that ALIGN(Root, L, Ft, L) and ALIGN(Stem, L, Ft, L) will describe a similar pattern in cases of unprefixed words. The constraint ALL-FT-L is active in prefixed words, while it must dominate ALIGN(Root, L, Ft, L). To see why this is so, consider the tableaux below (see Kager 1994:7; morpheme boundaries indicated with square brackets, foot structure with round brackets, dots as syllable boundaries).

(11) Ranking argument between ALIGN(Stem, L, Ft, L) and ALL-FT-L

/ka-pag=bissala/	ALIGN(Stem, L, Ft, L)	ALL-FT-L	ALIGN(Root, L, Ft, L)
☞ [(kà-pag)=bis.(sá.la)]	**		*
[ka-(pàg=bis).(sá.la)]	**	KA !	*
[ka-pag=bis.(sá.la)]	***!		*

Notice that the two top candidates have a tie at ALIGN(Stem, L, Ft, L) and here ALL-FT-L is decisive. However, in the words that show fluctuation the constraints ALL-FT-L and ALIGN(Root, L, Ft, L) are violated in turns and

it is only in these words that the Alignment constraint that refers to the root morpheme is motivated. Kager therefore assumes that in Sibutu Sama two opposite rankings are active, (i) ALL-FT-L >> ALIGN(Root, L, Ft, L) and (ii) ALIGN(Root, L, Ft, L) >> ALL-FT-L.

(12) Ranking argument between ALIGN(Root, L, Ft, L) and ALL-FT-L

/pa=missalahan/	ALIGN(Stem, L, Ft, L)	ALL-FT-L	ALIGN(Root, L, Ft, L)
☞ [(pà=mis).sa.(lá.han)]	*		*
☞ [pa=(mìs.sa).(lá.han)]	*	PA !	
[pa=mis.sa.(lá.han)]	**!		*

The Alignment constraint that is sensitive to the root morpheme is dominated by a similar Alignment constraint on the stem morpheme. Its consequences will never be observed, except under special conditions (e.g. preceded by a monosyllabic prefix); then it may compete with other constraints.

Optimality Theory predicts that constraints, including morpheme-sensitive constraints, may be dominated. Whether we are able to observe the activities of a morpheme-sensitive constraint does not only depend on the kind of constraint, but crucially also on its position in the constraint-ranking.

Golston (1995) discusses phonological properties of roots and words in Sanskrit[23]. Verbal roots are monosyllabic and bimoraic. The bimoraic condition [μμ] may be captured as consisting of a single Foot. Then the constraints that reflect these properties are ALIGN(Root, σ) (or ROOT=σ) and ALIGN(Root, R, Ft, R). Some examples of possible roots are given below.

(13) Verbal roots in Sanskrit

aj	'drive'
gam	'go'
sta:	'stand'
band^{h}	'bind'
sa:d^{h}	'succeed'

Sanskrit phonology exhibits a number of neutralization phenomena word-finally. Consonant clusters are resolved, obstruents are devoiced, deaspirated and depalatalized.

(14) Neutralization word-finally in Sanskrit

a.	dant]$_{Rt}$	-	dan]$_{\omega}$	'tooth'
b.	jambh]$_{Rt}$-		jamp]$_{\omega}$	'chew up, cush'
c.	vac]$_{Rt}$	-	vak]$_{\omega}$	'voice'

[23] Golston (1995) crucially argues against multiple constraint-rankings in a language if constraints are allowed to make reference to (the edges of) specific morphemes.

Golston assumes that Alignment constraints can take *Feature specifications as their arguments[24]. The word-final neutralization facts follow from ALIGN(Wd, R, *CC, R), ALIGN(Wd, R, *LAR, R) and ALIGN(Wd, R, *HIGH, R) respectively.

What happens in derived forms that contain both a root and a word? Golston demonstrates that Sanskrit phonology may be properly described with the above mentioned constraints, that refer to different morphemes. The constraint-ranking is given below, the Alignment constraints that are sensitive to the word morpheme are abbreviated as in Golston (1995).

(15) Constraint-ranking of Sanskrit

ROOT=σ, ALIGN(Root, R, Ft, R), Parse >> *Hi]$_{\omega}$, *LAR]$_{\omega}$, *CC]$_{\omega}$

The role of the Parse constraint is crucial here: phonological material that is contained within the root morpheme is properly parsed, except in the situation that the right edges of root and word morphemes are the same. In that case the word-sensitive Alignment constraints become active and neutralization emerges.

Kager (1994) and Golston (1995) demonstrate analyses which refer to constraints that are sensitive to morpheme-sets. These constraints signal two aspects, namely (i) the existence of specific morphemes such as stem or root and (ii) the phonological pattern linked to such morphemes. Now what exactly does it mean to have morpheme-specific constraints ranked among other constraints? It seems that the position of these constraints with respect to other constraints presents an argument for conflicting constraint-rankings within a single phonology. In other words, a phonological system which has constraints that refer to sets of morphemes equals a model that makes the conflict in ranking position more explicit, by having multiple constraint-rankings.

Consider for instance constraint ALIGN(Root, L, Ft, L) (Kager 1994) which is dominated in virtually all contexts in Sibutu Sama, but whose activity can be observed under certain circumstances only. As pointed out in Kager, the root morpheme is actually nothing more than the innermost stem morpheme. Thus any violation of ALIGN(Stem, L, Ft, L) properly includes a single violation of ALIGN(Root, L, Ft, L). Now recall that it is the activity of another constraint, namely ALL-FT-L, that introduces the interesting phonological variation patterns. Under the influence of the root morpheme a bifurcation in the constraint-ranking occurs, i.e. ALL-FT-L is dominated and dominates ALIGN(Root, L, Ft, L). It all depends on the fact whether the root morpheme is available as a grammatical category for Alignment. To explain the fluctuation of secondary stress in Sibutu Sama we need two competing constraint-rankings.

Sanskrit phonology as outlined in Golston (1995) has to be accounted for in a similar way. We can observe that the featural distinctions at right edges of root and word differ dramatically, which is expressed in a set of Alignment constraints that refer to word-final position. Golston uses the Sanskrit case as an argument in favor of a single ranking, with morpheme-sensitive constraints. However, what is actually expressed is a phonological system that exhibits multiple rankings, motivated by the different phonological patterns at the edges of distinct morphemes.

[24] *Feature constraints are part of Don't Associate constraints, consider discussion of this particular group of constraints in Prince & Smolensky (1993). I think that the Alignment constraints that show the neutralization effects in Sanskrit are in fact better understood as NonAlignment constraints, i.e. Don't Align.

If it is true that the use of morpheme-sensitive constraints expresses the fact that there are instances of competing and conflicting constraint-rankings, then why should we not explicitely refer to multiple constraint-rankings? Such an approach has the advantage that it may decrease the complexity of constraints in Optimality; the consequences that follow from constraints that refer to sets of morphemes may be captured by a model that uses more general constraints active in several rankings. At the same time the disadvantage is that the decrease of complexity of constraints tends to run parallel to the increase of complexity in constraint-rankings.

2.2. Complexity of constraint-rankings

Constraint-rankings explain the phonological patterns of a language. If different patterns are correlated with different morpheme-sets, then we may want to express these distinctions more directly by postulating multiple constraint-rankings. In this section I demonstrate that (i) by their very nature, constraint-rankings must always conflict with each other, and (ii) consequently, we need to determine possible relationships between competing constraint-rankings. My aim is to offer a consistent view on the interaction that occurs in systems with multiple rankings.

2.2.1. Conflicting constraint-rankings

As has been observed by a number of researchers, it is usually the case that rankings within a single language tend to differ minimally. For instance, with respect to stress in Manam[25], Buckley (1994) motivates several constraint-rankings that are sensitive to morpheme-sets. Buckley assumes that the morphemes (or morphological levels) root, prefixed form, word and clitic form in Manam are interacting with ranked constraints. Each morpheme level is associated with a separate constraint-ranking. Below I give a rough and incomplete overview of the constraints involved[26].

(16) Constraints active in Manam (incomplete)

ALIGNHD	- ALIGN(PrWd, R, Head(PrWd), R)
FTBIN	- Feet are binary under moraic or syllabic analysis
*CLASH	- Clashing feet (stresses on adjacent syllables) are prohibited
FTONSET	- A foot must have either a phonological or a morphological onset
WDINTEG	- "Integrity of word constituency which is established at the previous level is respected"

In (17a-c) I focus on the differences in constraint-rankings as outlined in Buckley (1994:33-34)[27].

[25] Manam is an Austronesian language which is spoken in some parts of Papua New Guinea. Buckley (1994) refers to work of Lichtenberk (1983).

[26] I choose not to give an extensive overview of all constraints that are proposed in Buckley (1994), which would certainly provide a better understanding of Buckley's ingenious analysis, to which I refer for more details. For my argument, however, it is necessary only to show the ranking distinctions of relevant constraints.

[27] I do not take into account the indiosyncratic ranking for AP suffixes at the word level, nor the inherently stressed suffixes or roots.

(17) Multiple constraint-rankings in Manam

a.	Root level:	FTONSET, *CLASH	>>	ALIGNHD
	Word level:	ALIGNHD	>>	FTONSET, *CLASH
b.	Prefix level:	FTONSET	>>	ALIGNHD
	Word level:	ALIGNHD	>>	FTONSET
c.	Clitic level:	WDINTEG	>>	FTBIN
	Word level:	FTBIN	>>	WDINTEG

It must be clear that conflicting rankings are a necessary consequence of having multiple constraint-rankings within a single language. The abstract overview above illustrates the situation in Manam with respect to stress assignment, but any phonological system that exhibits more than one ranking invokes competing constraint positions. How do we relate multiple constraint-rankings to each other?

2.2.2. Serialism versus parallelism

There can be only two options for constraint-rankings to be related to each other. In a derivational approach we assume a relationship in terms of some linear ordering principle between distinct rankings. For instance, on the premise that there is a hierarchical structure of stems and words, we propose to order a stem constraint-ranking to precede a word constraint-ranking. Serialism has several important consequences, which I will discuss first. In a nonderivational approach constraint-rankings are not derivationally linked in any conceivable way. Due to the fact that constraint-rankings are always in conflict with each other, there is no automated outcome within a nonderivational view on the topic of priority between stem or word ranking. Parallelism is discussed in the remainder of this subsection. It is my aim here to argue against a serial and in favor of a parallel view on multiple constraint-rankings.

Buckley (1994) proposes that different levels are related in a derivational way[28]. Each morpheme level corresponds to a distinct constraint-ranking; its output is formed by a separate application of Gen. The output of one level is the input of another level, which resembles a kind of cross-level serialism. Although there are no derivational devices within Optimality Theory, the interaction between different constraint-rankings is assumed to be derivational.

In the tableau presented below the interplay between word level ranking and clitic level ranking is demonstrated[29]. Due to the dominant position of PARSEFT any stress foot that is part of the input of the clitic level must be in the output candidate. Only in the case of unfooted material a new foot is allowed[30] (while abstracting away from various important elements of Buckley's analysis, especially the influence of FTBIN and ALIGNHD, I indicate clitic boundaries as '=', square brackets indicate word level footing and round brackets signal clitic level foot structure).

[28] Other proposals are McCarthy & Prince (1993b), McCarthy (1994), Cohn & McCarthy (1994). Kenstowicz (1994) suggests that Gen may be applicable in a cyclic fashion, which would explain stress patterns in Indonesian, Polish and Shanghai Chinese.
[29] This tableau is adapted and adjusted from the one given in Buckley (1994:13).
[30] Buckley (1994) argues that the position of the constraint ParseFt describes the tendency for structure preservation in Manam and at the same time allows effects that formerly were attributed to the Free Element Condition (Prince 1985).

18) Clitic level ranking in Manam

input: wa[búbu], =a	PARSEFT	PARSESYLL
☞ wa[búbu]=a		**
(wà [bu) (bu]=a)	*!	

input: [bága]lo, =a	PARSEFT	PARSESYLL
☞ [bàga] (ló=a)		
[bága]lo=a		**!

All serial approaches towards multiple constraint-rankings known to me assume that phonological information established by Eval at a previous constraint-ranking influences the choice of input candidates for the next constraint-ranking[31]. There are striking similarities with the prespecification method mentioned in section 1 above. For instance, in an analysis of Turkish stress patterns Inkelas (1994) demonstrates that the output of the Sezer cophonology forms the input to the word cophonology. If a word contains a Sezer stem, the foot structure parsed at the stem constraint-ranking forms part of the input to the word constraint-ranking. Of course, other constraints at the word ranking will interact with the prespecified metrical structure.

As mentioned above, Optimality Theory and prespecification method do not have much in common. On the condition that Eval selects an optimal candidate that is phonologically fully specified, it does not make much difference whether or not the input candidates are prespecified for some kind of phonological information (or underspecified for that matter). Prespecification of intermediate phonological stages can easily be traced back to its origin, namely as an aspect of the familiar tradeoff between input and set of constraints (or procedures).

In fact, as the analysis of vowel-zero alternations in Polish will show, there appears to be a redundancy in a phonological system that postulates multiple constraint-rankings and states serialism to relate these rankings. In section 3 I present two competing analyses of Polish yers within Optimality Theory. Yers may be part of the underlying representation (e.g. as prespecified phonological information of some form) or inserted under conditions such as syllabic well-formedness (e.g. as epenthesis into final consonant clusters). I will demonstrate that the phonological patterns that follow from both assumptions are the same, which is an exciting result. However, the crucial observation is that yers or the absence of consonant clusters are part of the stem cophonology, whereas consonant clusters or the absence of yers are part of the word cophonology in Polish. Hence, in this case it is redundant to use a prespecification method, because it follows from the distinction between stem and word constraint-ranking. I like to extend this conclusion to the claim that any serial approach to multiple constraint-rankings pretends to do more than it actually does.

If serialism is incorrect, then how do we relate different rankings? I want to propose a nonderivational approach to this problem. To explain the mechanism involved I adopt the nonderivation model of Monotonic Cyclicity as developed in Orgun (1993, 1994)[32].

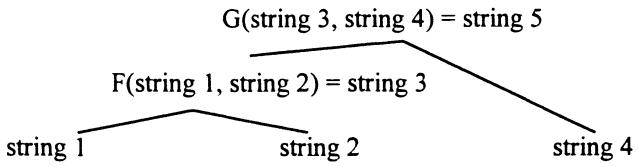
Monotonic Cyclicity as a model of phonology-morphology interaction is first proposed in Orgun (1993) to explain various cyclic effects in Turkish. Morphologically complex words are attributed with phonological

[31] Orgun (1995) proposes a novel extension of Correspondence Theory as described in McCarthy & Prince (1994). He states that Eval relates two strings, which are usually an input and an output. Faithfulness constraints FILL and PARSE are replaced in the new framework by constraints that specifically relate input and output, namely CORRESPOND and MATCH.

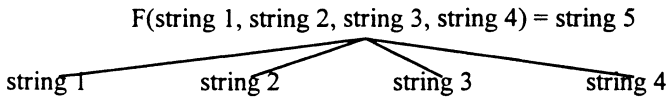
[32] There are great similarities with unification based theories (e.g. work of Gazdar et.al. 1985, Pollard & Sag 1992, Fillmore & Kay 1993) as well as with theories of Declarative Phonology (e.g. Bird 1990, among others).

constituent structure trees, such as given in (19) below. Each node of the structure represents a function of the nodes it immediately dominates. A node contains a complete phonological string, segmental and metrical structure included[33]. The following representations indicate that cyclic and noncyclic effects are available.

(19) a. Binary branching structure (cyclicity)

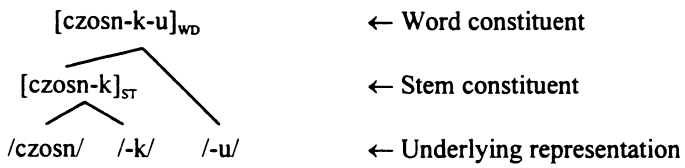


b. N-ary branching structure (noncyclicity)



An example from Polish may be illustrative. In (20) the phonological structure of the word *czosnku* 'garlic' (gen.sg) is given. I distinguish stem and word morphemes; the diminutive affix *-k* actively selects for a stem, while inflectional affix *-u* concatenates with a word.

(20) Morphological derived word from Polish



Note that affixes are represented as partial constituent structure trees and that affixation is interpreted as unification. In other words, affixation is an important means to activate a specific constraint-ranking[34].

I assume that Optimality Theory may provide different constraint-rankings that correlate with different constituents. In deviation from, among others, Inkelas (1994, 1995) I propose that Eval checks all constraint-rankings in parallel and simultaneously, selecting optimal candidates for each ranking that is activated. Due to the fact that constraint-rankings are in conflict with each other by default, it is predicted that what counts as an optimal candidate for one ranking surely does not need to be most harmonic for another, competing ranking. I hypothesize that beforehand there are no predictions available with respect to the exact output; this differs on a language-specific basis. However, the situation is not as unrestricted as it appears. Polish phonology offers a nice testing ground for a number of phonological patterns that are expected to occur in a parallel Optimality framework, while at the same time it proves that serialism of any kind is incapable to account for the facts.

[33] Terminal nodes of the tree are interpreted as the underlying strings supplied by morphemes, see Orgun (1993:9).
[34] Inkelas (1989) proposes to treat affixes as incomplete constituent structure, which makes it possible to view affixation as a process that actively creates new prosodic structure. A similar observation is made in Borowsky (1993).

Summary: I have shown that constraints in Optimality Theory may be sensitive to particular sets of morphemes, excluding the possibility to have constraints that refer to individual morphemes. Then it was argued against the hypothesis that nonproductive patterns and exceptions are denied separate cophonologies (e.g. constraint-rankings), leaving them as prespecified items. Instead, any kind of morpheme-set must be available as arguments for constraints. Two analyses that use morpheme-sensitive constraints gave evidence for the possibility to express this sensitivity to morphemes in a more direct way, namely by postulating multiple constraint-rankings within a single grammar. While discussing different constraint-rankings, I pointed out that due to their nature rankings must always conflict, which leaves open the question how to relate these competing constraint-rankings. I have argued against cross-level serialism within Optimality Theory and have offered a preliminary nonderivational approach to this problem.

3. Vowel-zero alternations in Polish phonology

The argument of this section concerns the behavior of the vowel-zero alternations or yers in Polish phonology. Previous attempts to account for this fuzzy phonological pattern can be grouped under two headings. Either we assume yers to be part in some shape of the underlying representation, or we assume that yers arise as a result of epenthesis, triggered by well-formedness conditions on mainly syllable structure. The first direction of research I will call the Underlying Representation approach to Yers, abbreviated as URYER. The second line is termed NOURYER, the abbreviation of the No Underlying Representation approach to Yers. I briefly show in section 3.1. some proposals with respect to both approaches. Many generative phonologists have been working in either uryer or nouryer and a large number of proposals exist that exclusively show URYER or NOURYER to be correct.

Optimality Theory shows that this issue is irrelevant, which seems to be a surprising consequence. In section 3.2. I present both analyses in terms of constraints and their ranking to prove this point. Instead, what is crucial for an adequate analysis of Polish yers is the recognition that there are two competing constraint-rankings in Polish phonology, which effects can be observed in different environments. Crucially, yers belong to the stem level constraint-ranking, whereas huge consonant clusters are part of the word level constraint-ranking. I use the model of Monotonic Cyclicity (Orgun 1993, 1994) as the framework in which the two conflicting constraint-rankings will operate. However, I demonstrate that only a nonderivational relationship between stem and word constraint ranking is capable of explaining the Polish facts.

Thus, I argue that the case of Polish yers as analyzed in this paper presents two important views on the theme of morpheme-sensitive phonology within Optimality Theory. First, fuzzy phonological patterns may be captured by means of distinct constraint-rankings. Second, multiple constraint-rankings may be related to each other in a nonderivational way. The following facts represent an overview of Polish vowel-zero alternations. Yers are denoted as capital E, as is common in Slavic literature.

(21) a. Noun inflection:

Neuter nouns	Nom.sg.	Gen.pl.
'apple'	jabLko	jabLEk
'box'	pudeLko	pudeLEk

'writing-table'	biurko	biurEk
'window'	okno	okiEn
Feminine nouns	Nom.sg.	Gen.pl.
'spring'	wiosna	wiosEn
'daughter'	córka	córEk
'mother'	matka	matEk
'aunt'	ciotka	ciotEk
Masculine nouns	Nom.sg.	Gen.sg
'boy'	chLopiEc	chLopca
'dog'	piEs	psa
'lion'	lEw	lwa

b. Derivational suffixes:

-k- (diminutive)	sarna 'roe-deer'	sarEnka
-n- (adjectival)	krfi 'blood' gen.sg.	krEvny 'relative'
-sk- (adjectival)	diabwa 'devil' gen.sg	diabElski
-stw- (nominalizing)	bLazna 'clown' gen.sg	bLazEnstwo

c. Prefixes (cf. Rowicka 1994)

ve+ss+aC 'suck in'	vEssem (1st sg. Pres)	vsYsaC (Sec.Imp.inf)

3.1. URYER versus NOURYER

The basic assumption are that yers are part of the phonological input (underlying representation, lexicon). All things being equal such an element will surface; but other conditions may prohibit the realization of yers. The following table lists some URYER proposals.

- Gussmann (1980) assumes a tense/lax distinction on [+high] vowels in underlying representation, plus a procedure Lower that either shifts high lax vowels to [e] in proper contexts, or deletes such vowels.
- Rubach (1984) proposes a similar distribution of [+high] vowels in UR, but divides the original procedure Lower into a procedure that alters such vowels to [e] in a cyclic manner and a procedure Yer Deletion that erases stray (high) vowels.
- Spencer (1986) argues for a distinction of specified and unspecified V-slots in UR. He sets up procedures that fill in segmental information in the course of the derivation and allows for special conditions like extrametricality on certain V-slots.
- Rubach & Booij (1990a,b) assume yers to be represented as floating feature matrices in underlying form. They have a procedure Yer Vocalization that links floating information to skeletal tier in certain contexts, together with a general procedure of deletion elsewhere.

- Szpyra (1989, 1992) argues in favor of a distinction of specified and unspecified root nodes, while she proposes procedures that provide segmental material in proper contexts and a procedure of deletion elsewhere.

The motivation for any URYER approach is that the occurrence of yers is not predictable, thus needs to be stipulated or prespecified in the input (cf. Szpyra 1992).

(23)	a.	kopEr	'dill' (nom.sg)	-	kopr-u (gen.sg)
	b.	kopr-a	'copra' (nom.sg)	-	kopr (gen.pl)

A number of researchers have seen that prosodic aspects may be relevant in the realization of yers, but other (non-phonological) factors govern these prosodic considerations. The URYER analysis abstracts away from such additional factors by allowing a difference in the phonological input.

The basic assumption of the NOURYER approach is that yers are not represented in underlying form, thus there is nothing that may surface, unless other conditions, especially syllable well-formedness, force the insertion of yers. The following researchers have proposed an epenthesis analysis.

- Both Gorecka (1988) and Czaykowska-Higgins (1988) assume that there are no yers in underlying form. Epenthesis is triggered due to two Coda constraints: on Sonority Sequencing Parameter and on coocurrence of two sonorant segments. The exceptions to Coda constraints are found in word-final position only.

The motivation for the NOURYER approach is that there are constraints on possible codas which are active in Polish phonology and they determine the position of yers. But unfortunately other conditions (e.g. phonological and/or non-phonological) interact with these constraints (cf. Szpyra 1992).

(24)	a.	walk-a	'battle'	-	walEcz-ny	'brave'
	b.	folwark	'farm' (noun)	-	folwarcz-ny	'farm' (adj.)

(25)	a.	wiosL-o	'oar' (nom.sg)	-	wiosEL (gen.pl)
	b.	pas-L	'he pastured'		* pas-EL

(26)	a.	miot-L-a	'broom' (nom.sg)	-	moit-EL (gen.pl)
	b.	miót-L	'he swept'		* miót-EL

In (25b, 26b) we can see the labiovelar glide as the preterite marker, in (26a) it functions as a nominalizing suffix. Notice that the contexts are similar in segmental and syllabic perspective.

Vowel-zero alternations as depicted above thus appear to be unpredictable on base forms, but quite regular in certain derived words. The strategy I suggest here is to locate the unpredictable property as part of the

morphological structure instead of viewing it as a phonological characteristic. More concrete, base forms that exhibit yers can be viewed as a distinct morpheme-set, apart from the morphemes that do not show this phonological pattern. In other words, the phonology of Polish must be able to refer to this separate morpheme-set. If we focus on the vowel-zero alternation the distinct morpheme-set appears to extend also to derivational structures, while the contrasting set of morphemes may be grouped with inflected forms.

3.2. Multiple constraint-rankings in Polish

Vowel-zero alternations in Polish reveal a complex interaction between phonology and morphology. With respect to the phonological aspects, something needs to be said about syllable structure and statements on well-formedness of codas. With respect to the morphological side, the distinction between stem constituents (e.g. derivation) and word constituents (e.g. inflection) appears to be relevant. My hypothesis regarding the interaction between phonology and morphology is that the stem level exhibits different phonological characteristics than the word level. More concrete, yers seem to be part of the stem level phonology, whereas consonant clusters at word-edges are part of word level phonology. The following morphological complex forms will be used in my argument.

(27) Morphological structures of [sarEn] 'roe-deer' (gen.pl), [sarna] (nom.sg)

(28) Morphological structures of [sarEnka] (nom.sg. dim) and [sarEnEk] (gen.pl. dim)

There is a crucial distinction between the statement that a morphological form is structured in a complex way and the fact that there may be constraint-rankings that are correlating with each node of such morphological complex structure. The model of Monotonic Cyclicity is capable of expressing the different phonological patterns (or cophonologies) that correlate with different nodes of the hierarchical structure. The special situation of Polish vowel-zero alternations illustrates the conflicting nature of constraint-rankings associated with stem and word level. I hope to show that any serial approach of relating such competing cophonologies must hopelessly fail; for this reason I assume a parallel application of Eval, checking all constraint-rankings at once.

First, I present the URYER analysis that assumes yers to be present (in some shape) in underlying representation. The analysis is largely based on Optimality Theory proposals of Zoll (1994, 1995) concerning latent segments and I refer to her work for more sophisticated details. After arguing for the precise formulation

of the constraint-rankings, I propose three hypotheses that account for the observed phonological patterns. Second, the NOURYER analysis is presented, assuming that yers are absent from the phonological input, but may be provided under certain conditions. I show the conflicting constraint-rankings and again point to three hypotheses that take care of the fuzzy data. A consequence of my presentation of the analyses URYER and NOURYER is that both approaches seem to be compatible and therefore equally adequate. I take this as an indication that from an OT perspective it does not matter whether or not yers are present in underlying representation.

3.2.1. URYER in Optimality Theory

Yers are represented as root nodes, indicated as "®". The Faithfulness constraint PARSE-® states that every yer-root must be assigned to a syllable. I assume that PARSE-® conflicts with another constraint that restricts the number of syllables in the output, *STRUC-σ, or "Have no syllables"[35].

(29)

input: sar®n (gen.pl)	PARSE-®	*STRUC-σ
sar<®>n	*!	*
☞ sarEn		**

In the diminutive form [sarEnEk] (gen.pl) both underlying yer-nodes are realized, against the penalizing constraint *STRUC-σ.

(30)

input: sar®n -®k (gen.pl)	PARSE-®	*STRUC-σ
sar<®>n<®>k	**!	*
sarEn<®>k	*!	**
sar<®>nEk	*!	**
☞ sarEnEk		***

A different optimal output form must result in the cases of [sarna] and [sarEnka]. In the form [sarna] (nom.sg) the yer-node is unparsed, that is, some constraint (for instance *STRUC-σ)[36] dominates the faithfulness PARSE-®.

(31)

input: sar®n -a (nom.sg)	*STRUC-σ	PARSE-®
☞ sar<®>na	**	*!
sarEna	***!	

In [sarEnka] the conflict between PARSE-® and *STRUC-σ is more complicated. If *STRUC-σ dominates PARSE-® then *no* yer-root will surface, which is clearly false. However, if PARSE-® dominates *STRUC-σ *all* yer-roots

[35]Consider Zoll (1994, 1995) for further motivation of this special version of Prince & Smolensky's general constraint *Struc.
[36]Notice that the actual choice of the constraint *Struc-σ is irrelevant here, as long as two aspects are expressed in the analysis: (i) Parse-® is dominated in one context, but in another context appears to undominated; (ii) these contexts are therefore correlating with conflicting constraint-interactions.

will surface, which is also the wrong output. In addition, PARSE-® may be applicable only on the outmost yer-root, which is not to be predicted on the basis of the current constraint-ranking.

(32)

input: sar®n -®k -a (nom.sg)	*STRUC-σ	PARSE-®
sar<®>n<®>ka	**	**
☞ ? sarEn<®>ka	***	*
sar<®>nEka	***	*
sarEnEka	****!	

I propose three hypotheses that will guide us to a more complete understanding of Polish yers.

- **Hypothesis I**: The solution to the paradox mentioned above is related to stem and word level phonology, that is, different stem and word ranking of the same constraints. PARSE-® is ranked above *STRUC-σ at the stem level, but vice versa at the word level.

(33) Multiple constraint-rankings in Polish

a. Stem ranking: PARSE-® >> *STRUC-σ

b. Word ranking: *STRUC-σ >> PARSE-®

- **Hypothesis II**: The following conditions regulate *satisfaction* of constraint-rankings.
  a. Eval checks a ranking of constraints iff there is a node in the morphological hierarchical structure. I assume such nodes to arise by means of derivational and inflectional morphology, corresponding with a stem cophonology and a word cophonology, respectively[37]. I also assume that an initial stem cophonology may be projected as the consequence of an idiosyncratic property of a morpheme-set[38].
  b. Eval checks all available constraint-rankings, i.e. maximizes cophonology satisfaction.
  c. In complex forms that contain both stem and word morphology, the constraint-rankings conflict. Here word cophonology as a consequence of overt inflection takes priority upon stem cophonology, which means that the word constraint-ranking must be satisfied, at the cost of the stem constraint-ranking. I hypothesize that this conflict is strictly local due to the fact that Eval maximizes satisfaction of constraint-rankings.
  d. Both constraint-rankings compete and produce separate optimal candidates if there is no overt inflection and no derivational evidence available. This predicts variation or fluctuation on nonderived forms without overt inflection.

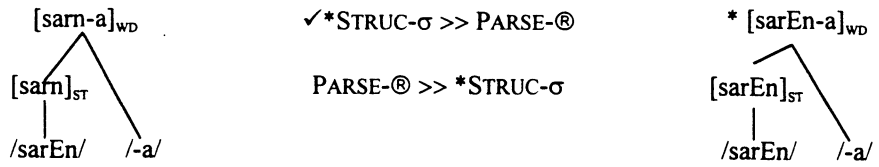
[37]In an informal manner, an active constraint-ranking should be understood as a node in the morphological structure that at least contains previous unprocessed material. Similar statements have been made in Inkelas (1989) with respect to affixation that actively constructs prosodic constituency structure, in Borowsky (1994) with respect to the postulation of prosodic and morphological structures in compounds, and in Inkelas & Orgun (1994) on Level Economy as a method to limit available cophonologies in Turkish.

[38]The influence of idiosyncratic morphemic information on the phonology is discussed in more detail in the next section.

- **Hypothesis III**: The paradox between stem and word cophonologies can be resolved only in nonderivational approach. Eval operates on stem and word constraint-rankings in parallel and simultaneously.

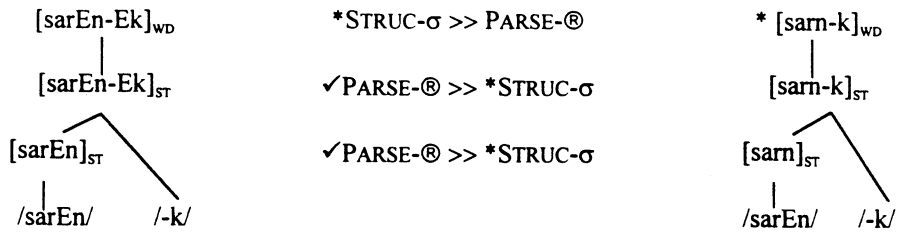
Below I present the relevant structure with conflicting constraint-rankings. The lefmost column contains the correct output, the middle column gives the conflicting constraint-rankings, where "✓" graphically signals satisfaction of ranking.

(34) Word ranking due to overt inflection: [sarna] 'roe-deer' (nom.sg)



The above structure shows that the stem constraint-ranking is not operative due to the fact that there is no derivational suffix available. The word cophonology dictates the absence of yers.

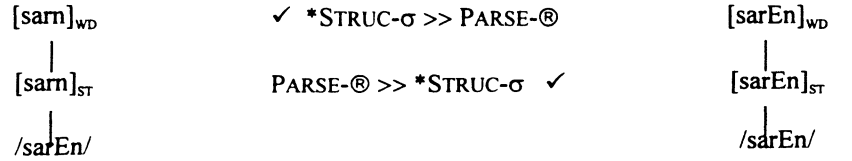
(35) Stem ranking due to overt derivational suffix: [sarEnEk] (gen.pl. dim)



In (35) above the diminutive suffix *-k* renders the stem constraint-ranking operative. The absence of any overt inflection results in the phonological pattern that contains yers. Notice that I assume the morpheme *sarn* to project an independent stem constituent.

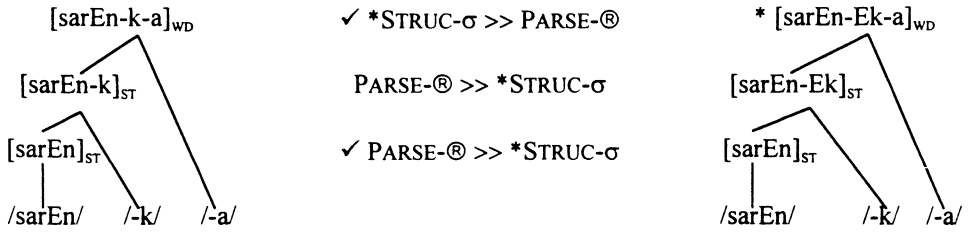
In the situation that no suffixation is available there appears to be a choice. Either the surface form exhibits the word level pattern (e.g. consonant clusters) or it may show the stem level phenomenon (e.g. yers). Consider the structures in (36) below, which are illustrative for the observed variation.

(36) Both word ranking and stem ranking are checked and provide optimal candidates, variation on nonderived stems without overt inflectional element: [sarEn] ≈ [sarn]



The fourth pattern is found in the cases in which both stem and word morphology is available. At first glance it seems as if the word constraint-ranking must obscure any effect of the embedded stem cophonology. But this is not the case, as can be seen in (37).

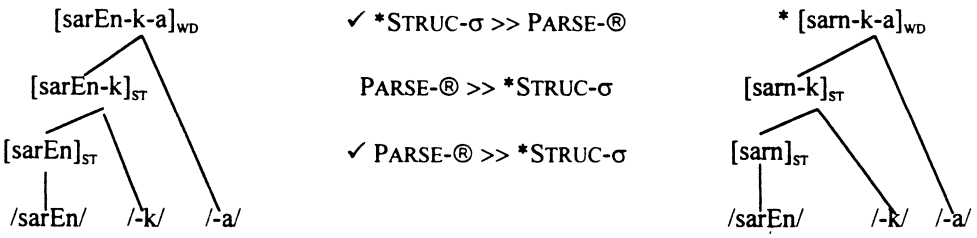
(37) Conflict between word and stem ranking, due to both word and stem elements.



While the word and stem cophonologies exclude each other due to their very nature (with the word cophonology as the winner), notice that the conflict appears to strictly local. In other words, the word constraint-ranking as activated by the inflectional suffix *-a* excludes the favourable phonological make-up selected on the basis of the stem ranking, but because there are multiple nodes that are correlated with stem constituents only the topmost constituent is overruled.

There are two other possibilities that relate these competing constraint-rankings. First, we could assume that due to the fact that there is a diminutive suffix the stem level cophonology must be applicable, that is, we expect some sort of preservation of previous information (or cycles). However, as can be seen from the righthand structure of (37), this cyclic approach makes a wrong prediction. Second, we could assume that due to the fact that constraint-rankings conflict, the ultimate surface structure should always pattern with the ranking of the topmost constitutent and we may formulate this as a kind of dominance relationship between ranking. Hence the word cophonology must always obscure any evidence from stem level constraint-rankings. To see why this approach is also wrong, consider the righthand structure below.

(37') Conflicting word and stem rankings



I think there is a straightforward answer as to why both the Preservation-option and the Dominance-option predict incorrect output candidates. These alternative options of relating constraint-rankings are not capable of expressing the correct phonological patterns, precisely because they assume a derivational relationship in the hierarchical structure, and thus between the constraint-rankings. Only a nonderivational approach is suitable to account for the Polish vowel-zero alternations. In fact, the above hypothesis that Eval maximizes the satisfaction of constraint-rankings already includes the nonderivational aspect of my proposal.

The above URYER analysis accounts for the observed phonological patterns in a straightforward way. Let us now consider the results of a NOURYER approach.

3.2.2. NOURYER in Optimality Theory

This analysis is based on the assumption that yers are interpreted as not available in the phonological input, and need to be provided by epenthesis, or in Optimality terms as a FILL-μ violation. The syllabic well-formedness constraint *COMPLEXCODA[39] dominates faithfulness FILL-μ and is thus responsible for the occurrences of yers.

(38)

input: sarn (gen.pl)	*COMPLEXCODA	FILL-μ
sarn	*!	
☞ sarEn		*

Since yers are not underlyingly present, the position of epenthesis has to be specified. I assume that the constraint ALIGN(Stem,R,σ,R) states the position of epenthesis. ALIGN-R is not ordered with respect to *COMPLEXCODA and FILL-μ. The stem-edge is graphically indicated as '|'

(39)

input: sarn (gen.pl)	ALIGN-R	*COMPLEXCODA	FILL-μ
sarn | E	*!		*
☞ sarEn |			*

Basically, these three constraints govern the surfacing of yers in Polish in a NOURYER proposal. In [sarEnka] the relevance of the syllable structure constraint ONSET is visible, dominating ALIGN-R.

(40)

input: sarn -k -a (dim, nom.sg)	ONSET	ALIGN-R	*COMPLEXCODA	FILL-μ
sarn | k | a		*	**!	
☞ sarEn | k | a		*	*	*
sarn | Ek | a		**!	*	*
sarEn | Ek | a		**!		**
sarn | Ek | a	*!		*	*

The form [sarEnEk] is problematic for the analysis presented so far because the ill-formed structure [sarnEk] is predicted to be more harmonic.

[39]Probably the constraint *COMPLEXCODA is itself constructed as a large cluster of independent constraints, which regulate the possible well-formedness of codas in Polish. In my argument I abstract away from these details.

(41)

input: sarn -k (dim, ge.pl)	ONSET	ALIGN-R	*COMPLEXCODA	FILL-μ
sarn | k |			**!	
sarEn | k |			*!	*
☞ ? sarn | Ek |		*	*	*
☞ sarEn | Ek |		*		**

Tableaux (40) and (41) point out an additional problem with respect to interpreting the most harmonic candidate. In both cases there are two distinct stem-edges which may be adressed to by ALIGN-R. Notice, however, that in evaluating *COMPLEXCODA the distinct stem constituents are also taken into account. In (41) the candidate [sarn | k |] is violating *COMPLEXCODA twice, once as [sarn] and once as [sarnk].

As in the URYER analysis described above I presume that three hypotheses bring a solution to the above problems. Actually, the versions of the hypotheses are quite similar, indicated here by means of a prime '.

- **Hypothesis I'**: The constraint *COMPLEXCODA is active on the stem level, but is dominated on the word level. It has been extensively documented that Polish allows huge consonant clusters at word-final positions (for example in Rubach & Booij 1990a,b). This observation can be captured while using multiple conflicting constraint-rankings at stem and word level.

(42) Multiple constraint-rankings in Polish

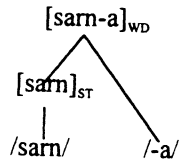
a.	Stem ranking:	*COMPLEXCODA	>>	FILL-μ
b.	Word ranking:	FILL-μ	>>	*COMPLEXCODA

- **Hypothesis II'**: The following conditions regulate *satisfaction* of constraint-rankings.
  a. Eval checks a ranking of constraints iff there is a node in the morphological hierarchical structure. I assume such nodes to arise by means of derivational and inflectional morphology, corresponding with a stem cophonology and a word cophonology, respectively. I also assume that an initial stem cophonology may be projected as the consequence of an idiosyncratic property of a morpheme-set.
  b. Eval checks all available constraint-rankings, i.e. maximizes cophonology satisfaction.
  c. In complex forms that contain both stem and word morphology, the constraint-rankings conflict. Here word cophonology as a consequence of overt inflection takes priority upon stem cophonology, which means that the word constraint-ranking must be satisfied, at the cost of the stem constraint-ranking. This conflict is strictly local due to the fact that Eval maximizes satisfaction of constraint-rankings.
  d. Both constraint-rankings compete and produce separate optimal condidates if there is no overt inflection and no derivational evidence available. This predicts variation or fluctuation on nonderived forms without overt inflection.

- **Hypothesis III'**: The paradox between stem and word cophonologies can be resolved only in nonderivational approach. Eval operates on stem and word constraint-rankings in parallel and simultaneously.

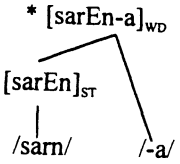
Below I give the relevant structures and indicate how the different constraint-rankings are interacting. It may become clear that, apart from the actual choice of constraints, the structures depicted in (43) to (46) bear a huge resemblance to the previous presented structures of the URYER analysis.

(43) Word ranking due to overt inflection: [sarna] 'roe-deer' (nom.sg)

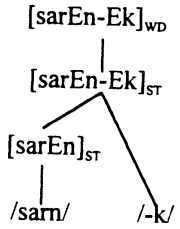


✓FILL-μ >> *COMPLEXCODA

*COMPLEXCODA >> FILL-μ



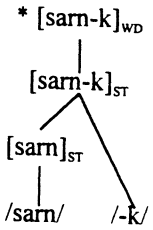
(44) Stem ranking due to overt derivational suffix: [sarEnEk] (gen.pl. dim)



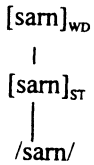
FILL-μ >> *COMPLEXCODA

✓ *COMPLEXCODA >> FILL-μ

✓ *COMPLEXCODA >> FILL-μ

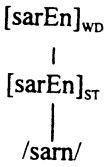


(45) Both word ranking and stem ranking are checked and provide optimal candidate, variation on nonderived stems without overt inflectional element: [sarEn] ≈ [sarn]

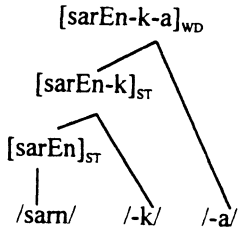


✓FILL-μ >> *COMPLEXCODA

✓ *COMPLEXCODA >> FILL-μ



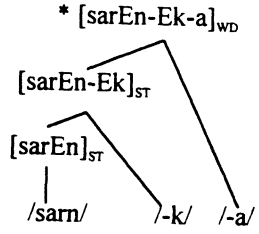
(46) Conflict between word and stem ranking, due to both word and stem "signaling elements".



✓FILL-μ >> *COMPLEXCODA

*COMPLEXCODA >> FILL-μ

✓ *COMPLEXCODA >> FILL-μ



All the arguments that I have given in the URYER analysis are applicable in the NONURYER account. For instance, notice the importance of a nonderivational approach to the phonological pattern in (46), which is stated in the hypothesis that Eval maximizes constraint-ranking satisfaction. It must be evident that the URYER and the NOURYER approaches lead to the same description of the vowel-zero alternations. Of course this result, surprising as it may be, must be traced back to the notion of what the approaches actually represent, namely

different choices with respect to the theme of the tradeoff between phonological input and set of procedures or constraints. I conclude therefore that the question addressed to in these approaches, do we assume yers to be phonologically underlying present or not, is irrelevant from the point of view of Optimality Theory.

Notice that at this point of my argument the URYER approach appears to suffer from an internal inconsistency: it seems as if the information about yers is encoded twice in the phonology of Polish. On the one hand I have argued for the fact that vowel-zero alternations are related to a stem level versus word level constraint-ranking, but on the other hand the URYER option chooses yers to be part of the underlying representation in a distinct group of morphemes. As further research may conclude, such redundancy should be excluded a priori. Although I assume in the remaining of this paper a NOURYER approach towards Polish yers, I do not motivate this choice, precisely because from an OT perspective that question seems to be of little relevance.

4. Some other phenomena in Polish phonology

Within the framework of Monotonic Cyclicity a distinction may be made between noncyclic and cyclic constituency. Noncyclic forms are represented as n-ary branching or flat structures, whereas cyclicity arises in cases of binary branching structures. In the first part of this section I discuss word level phonology in Polish. I will give attention to a typical word-level suffix, namely Comparative Degree Formation; the interaction of the constraints PARSE-segment and FILL-μ at the word level will thus be demonstrated. Then I show how cyclicity is expressed at the Polish stem level cophonology, while using binary branching structure. The interaction between noncyclic and cyclic structures as it happens to be in Polish phonology is a crucial aspect of my proposal. It can be concluded that the distinction between morphemes that exhibit yers and those that do not relates to an unpredictable morphological property, which results in an initial stem constituent.

4.1. Comparative Degree Formation as word level suffix

Comparative Degree Formation or CDF comes in two phonological shapes: as *-szy* and as *-Eszy*. As Szpyra (1992) and Bethin (1991) point out, the choice of the allomorph *-Eszy* arises in cases of an unsyllabified consonant[40].

(47)

gLup-i	'silly'	gLup-sz-y
tward-y	'hard'	tward-sz-y
manrd-y	'wise'	manrd-Esz-y
zimn-y	'cold'	zimni-Esz-y

The extended allomorph is also selected in cases of a morpheme that exhibits a yer, such as [pewiEn] 'certain' - [pewn-y].

[40]All data are from Szpyra (1992) and Bethin (1992).

(48)

pewn-y	'certain'	pewni-Esz-y	*pewiEn-sz-y
godn-y	'worthy'	godni-Esz-y	*godEn-sz-y

Also in cases of a derived adjective that in turn contains a yer-morpheme, such as [hanb-a] 'shame' - [haniEb-n-y] 'shameful', the extended allomorph is selected.

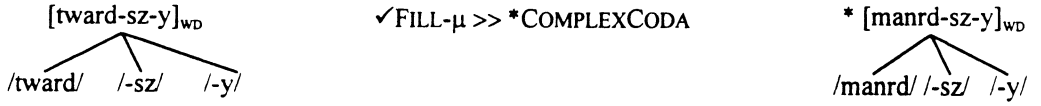
(49)

haniEb-n-y	'shameful'	haniEb-n-Esz-y	*haniEbEn-sz-y, *hanbEn-Esz-y

The analysis of the above data may be developed along the following lines. Let us assume that the enlarged version of the CDF-suffix results from epenthesis exactly in the cases in which morphemic material tends to remain unsyllabified. In OT-terms this can be expressed as the requirement that every segment must be faithfully parsed, forcing a violation of FILL-μ. What is important to note is that the constraint-ranking at the word level does not allow dominance of FILL-μ by the constraint *COMPLEXCODA, as I have explained in the previous section.

Optimality Theory assumes constraints to be violable under certain conditions. This is precisely what we encounter in the situation of Comparative Degree Formation. In other words, although FILL-μ is highly ranked in the word cophonology, it is itself dominated by the constraint PARSE-segment. I propose the following morphological constituents to account for the CDF-examples.

(50)



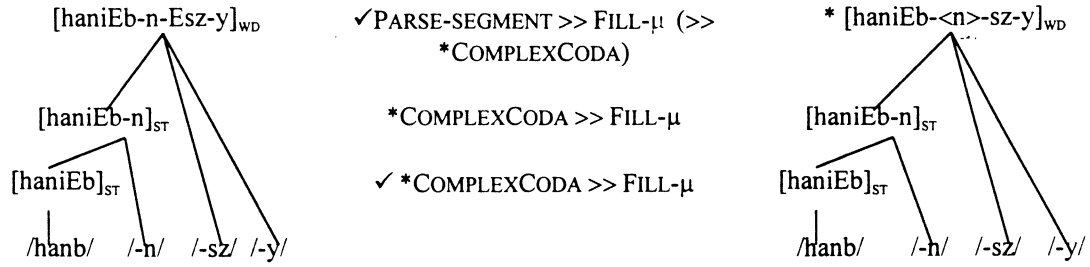
A closer look at the structure of $[manrd\text{-}sz\text{-}y]_{WD}$ shows that the final segment of the morpheme cannot be parsed into syllable position, therefore epenthesis occurs.

(51)

input: manrd -sz -y	PARSE-segment	FILL-μ
manr<d>szy	*!	
☞ manr.dE.szy		*

In the case of the derived base [hanb-n], an interesting conflict between stem and word constraint-rankings occurs. Notice that the constraint *COMPLEXCODA forces a FILL-μ violation on the stem [haniEb], but in the competing area of [haniEb-n-sz-y], the word level ranking does not allow epenthesis to occur twice, at the cost of leaving the nasal unparsed. The following morphological structure may express this observation.

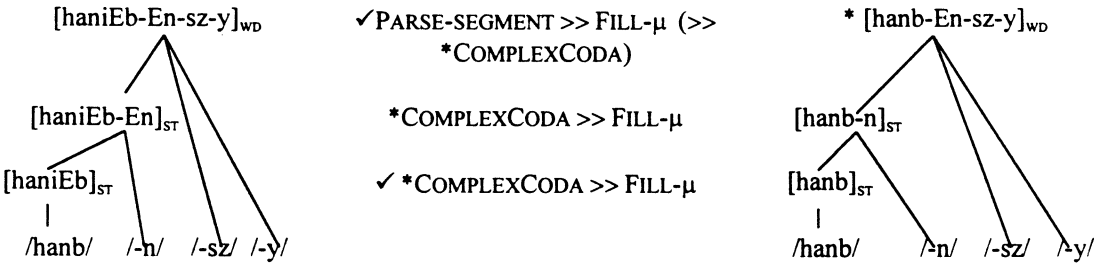
(52)



Why doesn't epenthesis occur in the position predicted by the stem level constraint-ranking? I propose that this possibility is excluded due to the fact that the stem cophonology of [hanb-n] is conflicting with the word level constraint-ranking activated by means of the inflectional marker *-y*, in the exact same way as constraint-rankings of stem and word level compete throughout the Polish phonology. In other words, the epenthetic site may not be positioned within the stem constituent, because the constraint *COMPLEXCODA is not outranking FILL-μ at the word level.

This interaction between constraints and their rankings can not be achieved in a derivational approach. For instance, as the lefthand example of (53) may illustrate, an account that allows information of embedded constituents to be preserved makes the wrong prediction that the allomorph *-szy* will be selected. In constrast, if the word cophonology is dominant on a global scale, then due to PARSE-segment an epenthetic segment will surface, but crucially in the wrong position. To see why this is the case, consider the rightmost structure of (53) below, motivated by the tableau in (54) (dots signal syllable boundaries).

(53)



(54)

input: hanb, -n, -sz, -y	PARSE-segment	FILL-μ	*COMPLEXCODA
hanb.<n>szy	*!		*
han.<bn>szy	**!		
ha.niEb.<n>szy	*!	*	
? ☞ han.bEn.szy		*	
☞ ha.niEb.nE.szy		**	
? ☞ ha.niE.bEn.szy		**	

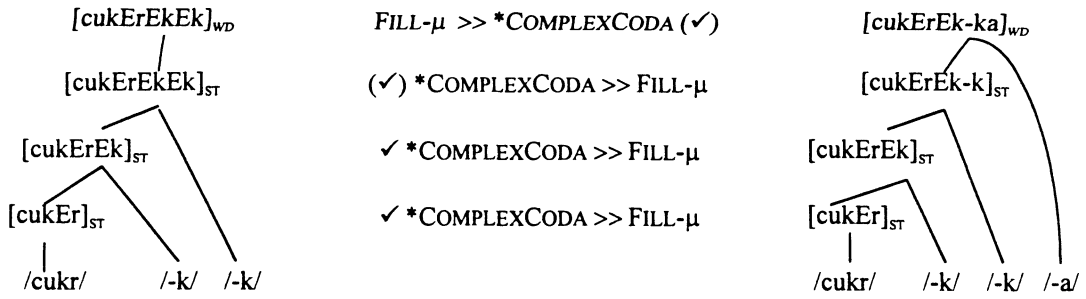
If the word constraint-ranking turns out to dominate all other cophonologies, the incorrect output candidate [hanbEnszy] will be selected as the most harmonic one, based on its number of FILL-violations. Notice that tableau (53) also allows an alternative candidate [haniEbEnszy] to compete with the actually optimal one. I think

that the emergence of such different possible output candidates are consequences of the statement that there is a single dominant constraint-ranking. Unfortunately, this assumption turns out to be wrong; it seems to me that any correct analysis can only come forth in an approach using multiple constraint-rankings in a nonderivational way.

4.2. Stem cyclicity

Below I show that the stem level must be binary branching, which expresses cyclicity. Consider the example [cukErECEk] 'candy' (dim. nom.sg) and [cukErECka] (gen.sg), as given at the lefthand and righthand side, respectively.[41]

(54)



A flat structure, representing noncyclicity, makes a complete false prediction here. Consider tableau (55) below that may illustrate this point. Tableau (55) results in the selection of incorrect output forms [cukErkEk] and [cukrEkEk], based on the number of Fill-violations. The actual output candidate can be achieved only on the condition that Eval checks the constraint-rankings on all depending stem constituents (dots represent σ-boundaries, " | " indicates morpheme boundary).

(55)

input: cukr -k (dim) -k (dim)	PARSE-segment	*COMPLEXCODA	FILL-μ
cuk<r> | <k> | <k>	***!		
cukr | <k> | <k>	**!	*	
cuk.r | Ek | <k>	*!		*
cukr | k | Ek |		*!	*
? ☞ cukEr | .k | Ek |			**
? ☞ cuk.r | E.k | Ek |			**
☞ cukE.r | E.k | Ek |			***

As can be seen in the topmost forms, any unparsed segment creates a final violation; observe that a violation of *COMPLEXCODA is also forbidden. Therefore, the number of FILL-μ violations must be decisive. A flat structure, processing the whole string in a single computation, cannot explain why epenthesis should occur three times, i.e. why does the correct output form contain three Fill-violations. Only the crucial assumption that the complex

[41] In the structures of (54) the palatalization of the velar is not taken into account.

word [cukErECEk] exhibits multiple stem constituents, each with their own stem level constraint-ranking and related by a nonderivational method, provides an adequate explanation for this unexpected redundant fact.

4.3. Stem constituents on morphemes

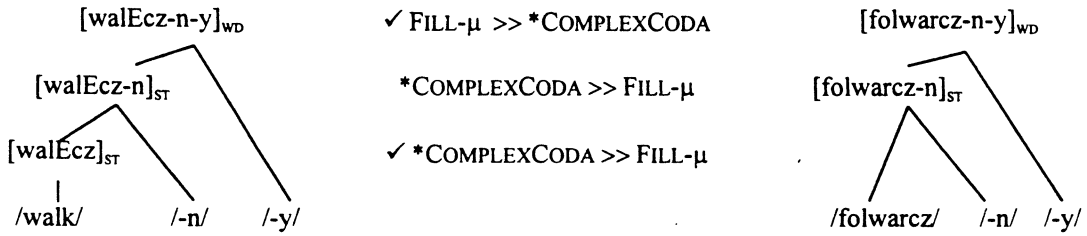
Polish phonology exhibits an unpredictable distinction between morphemes that show yer-phenomena and those that do not. In other words, there exists an separate morpheme-set which I will call yer-stems. A yer-stem constrasts with other morphemes because it projects an initial stem constituent that preceeds any morphology, i.e. it exhibits a distinct cophonology or constraint-ranking. This assumption is similar to the proposal of a stem 'cycle' as outlined in Inkelas (1989)[42]. This is expressed in the following hypothesis:

- **Hypothesis IV**: There is no distinction between an URYER and a NOURYER analysis. Yers are not specified phonologically, but follow from the constraint interaction on the stem level. Base morphemes that show yers are different, because they project an additional stem constituent, that correlates to an additional stem level constraint-ranking. Eval checks optimal candidates on all rankings, included this initial stem level.

I assume that the constraint-ranking of the yer-stem constituent is similar to the regular stem level ranking. This assumption allows for a consistent and elegant explanation of morpheme-sensitive phonology in Polish. Nonetheless, I do not exclude the possibility that further work on this language may reveal that such abstract similarities are not motivated; but these considerations are not explored in this paper[43].

The following structures illustrate the distinction in morpheme-set. On the lefthand side a yer-stem and its complex structure is given, whereas the righthand example contains a contrasting morpheme. It is important to see that the distinction is expressed in terms of absence versus presence of a stem constituent. Also notice that it is due to derivational morphology that a stem constituent arises. The tableau in (57) demonstrates that the initial stem level ranking precisely make up the difference in phonological patterning (only the stem constraint-rankings are given).

(56)



[42] I owe this suggestion to Sharon Inkelas and Orhan Orgun (personal communication);they refer to Perlmutter (1988) and to Henniss (1991).

[43] For instance, in a footnote above I have suggested that the constraint *COMPLEXCODA actually may be interpreted as a cluster of constraints. If this is the case, then we may expect this cluster to behave differently in different environments. The example cited in Prince & Smolensky (1993) of a constraint that actually consists of a large number of constraints is HNUC.

(57)

input: walcz, -n	*COMPLEXCODA	FILL-μ
walcz | n |	**!	
walEcz | n |	*!	*
☞ walEcz | En |		**

input: folwarcz, -n	*COMPLEXCODA	FILL-μ
folwarczn	*!	
☞ folwarczEn		*
folwarEczEn		**!

In this section three aspects of my proposal concerning vowel-zero alternations in Polish have been considered. I have argued for a distinction between noncyclic and cyclic structures that are both operative in the grammar of Polish. Noncyclic word cophonology and cyclic stem constraint-rankings conflict. Arguments that are crucially in favor of a nonderivational approach to the interaction between cophonologies have been brought forward in the cases of Comparative Degree Formation and cyclic stem constituents in multiple diminutive forms. Finally, I have proposed to account for the behavior of bare morphemes with respect to vowel-zero patterns in terms of an initial stem constituent.

5. Consequences and further remarks

Many prominent questions remain untouched in this paper. In this concluding section I merely indicate various aspects that need to be explored in future research. From a phonological point of view one could ask whether or not there exists additional phonological evidence for the proposed division between stem and word cophonologies. For example, Gussmann (1992) discusses several voicing phenomena in Polish, that relate to syllable structure and word constituency. Consonant clusters in word-final position show neutralization effects to the extend that any [+voice] specification on obstruents is lost. Thus the word constraint-ranking captures the fact that huge consonant clusters are allowed, but at the same time disallows contrasting featural voice specifications on these clusters. One may wonder whether the stem cophonologies differ in this respect.

Another interesting topic not covered in this paper concerns the phonological patterns in morphological complex structures that contain prefixes. As among others Gussmann (1980) shows, prefixed forms behave differently in a number of phonological respects such as vowel-zero alternations and palatalization effects.

From a typological perspective a large number of important issues can be stated which concern both the conceptual aspects of the Monotonic Cyclicity model applied to the Optimality Theoretic framework as well as the nonderivational nature as outlined in this paper. For instance, if morpheme-specific phonology in Polish can be described in the way I have proposed here, does this also imply that Inkelas, Orgun, Zoll (1994)'s division of the Turkish phonological patterns may be accounted for in the same way? It also raises questions about the common differentiation between phonological information that is supposed to be listed in the lexicon and similar information that is generated via application of a set of procedures or constraints.

References

- Bethin, Christina Y. (1992), *Polish Syllables. The Role of Prosody in Phonology and Morphology*, Slavica Publishers, Inc.
- Bird, Steven (1990), *Constraint-based phonology*, PhD-thesis, University of Edinburgh.
- Bloomfield, L. (1933), *Language*, Holt, New York.
- Borowsky, Toni (1994), Prosodic and morphological domains, ms. University of Sidney.
- Buckley, Eugene (1994), Alignment in Manam stress, ms. University of Pennsylvania.
- Chomsky, Noam & Morris Halle (1968), *The Sound Pattern of English*, Harper and Row, New York.
- Cohn, Abigail C. & John J. McCarthy (1994), Alignment and Parallelism in Indonesian Phonology, ms. Cornell University and University of Massachusetts, Amherst.
- Czaykowska-Higgins, Ewa (1988), *Investigations into Polish morphology and Phonology*, PhD-thesis, MIT.
- Dell, Francois C & Elizabeth O. Selkirk (1978), "On a Morphologically Governed Vowel Alternation in French", in: Keyser (ed.), *Recent Transformational Studies in European Languages*, MIT Press, Cambridge.
- Féry, Caroline (1994), The foot: Umlaut and Inflection in German, Handout Workshop in Prosodic Morphology, Utrecht.
- Golston, Chris (1995), Prosodic Constraints on Roots and Words, Handout CIP, FAS, Berlin.
- Gorecka, Alicja (1988), Epenthesis and the coda constraints in Polish, ms. MIT.
- Gussmann, Edmund (1980), *Studies in abstract phonology*, MIT Press.
- Gussmann, Edmund (1992), "Resyllabification and Delinking: the case of Polish Voicing", in: *Linguistic Inquiry*, volume 23-1:29-56.
- Hammond, Michael (1995), "There is no lexicon!", ms. University of Arizona.
- Henniss, Kathryn (1991), Stem allomorphy, level ordering and the representation of exceptions: Yiddish plurals and diminutives revisited, ms. Stanford University
- Inkelas, Sharon (1989), *Prosodic Lexical Phonology*, PhD-thesis, Stanford University.
- Inkelas, Sharon & Young-mee Cho (1993), "Inalterability as prespecification", in: *Language*, volume 69:529-574.
- Inkelas, Sharon (1994), The Consequences of Optimization for Underpecification, ms. UC. Berkeley/University of Utrecht.
- Inkelas, Sharon (1994), Exceptional stress-attracting suffixes in Turkish: representations vs. the grammar, ms. UC Berkeley.
- Inkelas, Sharon (1995), Dominant affixes and the theoretic status of morpheme-specific phonology, Handout CIP, Berlin.
- Inkelas, Sharon & Orhan Orgun (1994), "Level economy, derived environment effects and the treatment of exceptions", in: Richard Wiese (ed.), *Recent Developments in Lexical Phonology*, Theorie des Lexikons Nr. 56, Arbeiten des Sonderforschungsbereichs 282, Düsseldorf.
- Inkelas, Sharon, Orhan Orgun, Cheryl Zoll (1994), Subregularities as Cogrammars: the theoretical status of nonproductive patterns in grammar, ms. UC Berkeley.
- Itô, Junko & Armin Mester (1995), "Japanese phonology: constraint domains and structure preservation", in: John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, Basil Blackwell.
- Kager, René (1994), Generalized alignment and morphological parsing, ms. University of Utrecht.
- Kenstowicz, Michael (1994), "Cyclic vs. Noncyclic Constraint Evaluation", in: *Papers on Phonology and Morphology*, MIT Working Papers in Linguistics, volume 21.
- Kiparsky, Paul (1982), "Lexical Morphology and Phonology", in: I.-S. Yang (ed.); *Linguistics in the Morning Calm*, Seoul.
- Kiparsky, Paul (1985), "Some consequences of Lexical Phonology", in: *Phonology, 2-1*
- Kiparsky, Paul (1993), "Blocking in non-derived environments", in: S. Hargus & E. Kaisse (eds.), *Studies in Lexical Phonology*, Phonetics and Phonology, volume 4, Academic Press, San Diego.
- McCarthy, John (1993), The Parallel Advantage: Containment, Consistency, and Alignment, Handout Rutgers Optimality Workshop 1, Rutgers University.
- McCarthy, John (1995), Extensions of Faithfulness: Rotuman Revisited, ms. University of Massachusetts, Amherst.
- McCarthy, John & Alan Prince (1993a), "Generalized Alignment", in: Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1993*, Kluwer Academic Publishers, Dordrecht Boston.

McCarthy, John & Alan Prince (1993b), Prosodic Morphology I: Constraint Interaction and Satisfaction, ms. University of Massachussetts, Amherst and Rutgers University.
McCarthy, John & Alan Prince (1994), The Emergence of the Unmarked, ms. University of Massachusetts, Amherst and Rutgers University.
Mohanan, K.P. (1986), *The Theory of Lexical Phonology*, Reidel Dordrecht.
Mohanan, K.P. (1991), "On the bases of radical underspecification", in: *Natural Language & Linguistic Theory*, volume 9:285-325
Orgun, Orhan (1993), "Monotonic Cyclicity and Optimality Theory", in: *NELS, 24*
Orgun, Orhan (1994), Monotonic Cyclicity and Suspended Affixation, ms. UC. Berkeley.
Orgun, Orhan (1995), Correspondence and identity in two-level Optimality Theory, ms. Uc. Berkeley.
Perlmutter, David M. (1988), "The split morphology hypothesis: evidence from Yiddish", in: Hammond, Michaal and Michael Noonan (eds.), *Theoretical Morphology*, Academic Press, San Diego.
Piotrowski, Marek (1992), "Polish yers and extrasyllabicity: an autosegmental account", in: Fisiak, Jacek & Stanislaw Puppel (eds.), *Phonological Investigations*, John Benjamins Publishing Company, Amsterdam.
Prince, Alan (1985), "Improving tree theory", In: *Proceedings of Berkeley Linguistic Society*, volume 11.
Prince, Alan & Paul Smolensky (1993), Optimality Theory: constraint interaction in generative grammar, ms. Rutgers University and University of Boulder, Colorado.
Russell, Kevin (1993), *A Constraint-based approach to phonology and morphology*, PhD-thesis, Graduate School of the University of Southern California, Los Angeles.
Rowicka, Grazyna (1994), Prosodic Optimality and Prefixation in Polish, Handout Prosodic Morphology Workshop, University of Utrecht.
Rubach, Jerzy (1984), *Cyclic and Lexical Phonology: the structure of Polish*, Foris, Dordrecht
Rubach, Jerzy & Geert Booij (1990a), " Syllable structure assignment in Polish", in: *Phonology, volume 7-1*
Rubach, Jerzy & Geert Booij (1990b), "Edge of Constituent Effects in Polish" in: *NLLT, 8-3.*
Russell, Kevin (1995), Morphemes and candidates in Optimality Theory, ms. University of Manitoba.
Scobbie, James (1992), *Attribute Value Phonology*, PhD-thesis, University of Edinburgh.
Smolensky, Paul (1994), Harmony, markedness, and phonological activity, Handout Rutgers Optimality Workshop 1, Rutgers University.
Spencer, Andrew (1986), "A non-linear analysis of vowel-zero alternations in Polish", in: *Journal of Linguistics, 22.*
Spencer, Andrew (1991), *Morphological Theory. An Introduction to Word Structure in Generative Grammar*, Basil Blackwell.
Steriade, Donca (1995), "Underspecification and Markedness", in: John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, Basil Blackwell.
Szpyra, Jolanta (1992), "Ghost Segments in Nonlinear Phonology: Polish Yers", in: *Language, volume 68-2*
Walker,Douglas C. (1975), "Lexical Stratification in French Phonology", in: *Lingua 37.*
Zec, Draga (1995), Footed Tones and Tonal Feet: Rhythmic Constituency in a Pitch Accent Language, Handout HILP 2, Amsterdam.
Zoll, Cheryl (1993), "Ghost segments and Optimality", in: Spaelti et.al. (eds), *Proceedings of WCCFL*, volume 12.
Zoll, Cheryl (1993), Directionless Syllabification and ghosts in Yawelmani, ms. UC Berkeley.
Zoll, Cheryl (1994), An Unified treatment of Latent and Floating elements, ms. UC Berkeley.
Zoll, Cheryl (1995), Cheating Charon: Ghost that float, Handout HILP 2 conference.

" ...the *Chekists* said that everyone is a man of the White Guard, and if not, it must be proved in every separate case. Here the Soviet allegiance is a marked element... "

Jakobson (letter to Trubetzkoy (1930); in: Trubetzkoy (1975: 162f.))

Wolfgang Ullrich Wurzel

ON MARKEDNESS*

0. As is well known the preoccupation with the topic of markedness in linguistics is by no means new. But characteristically, this topic was not pursued in continuation since its first appearance; but disappeared again from the linguists' field of interest relatively soon. Since that time the topic has come up again at irregular intervals. In recent years the topic has been discussed vividly above all in connection with language change.

Of course, it is not possible to give a comprehensive history of the concept of markedness in linguistics here. However, some of its landmarks of it are to be mentioned :

- The first stage in the developement of the concept of markedness (in German: Markiertheit) was the concept of 'featuredness' (in German: Merkmalhaftigkeit), established by Jakobson and Trubetzkoy in the famous Prague Circle during the thirties. Its aim was to characterize the nonequivalence of the members in such oppositions, where one phoneme has a (positive) feature that the other phoneme does not have (cp. Trubetzkoy (1931)). Later the concept was transferred to the members of grammatical oppositions (morphological categories) by Jakobson (cp. Jakobson (1971a)).[1]

- The next step in the history of markedness is Jakobson's concept of phonological laws of founding (phonologische Fundierungsgesetze) presented in his book "Kindersprache, Aphasie und allgemeine Lautgesetze" (1941). Jakobson demonstrates the implicative structure of phonological systems (which represents a markedness structure) by analyzing facts from the linguistic areas of language acquisition, aphasia, and language specific phonological systems, interestingly enough without mentioning the term 'Merkmalhaftigkeit' (or 'Markiertheit').

- A very decisive contribution to the further development of the concept is represented by the famous 'Chapter Nine' in Chomsky/Halle's "Sound Pattern of English" (1968). The authors outline a formalized phonological markedness theory, consisting of a set of universal 'markedness conventions', and with it the first linguistic markedness theory at all. It is also noteworthy that Chomsky and Halle make a clear distinction between "the Praguian conception of markedness" - that is Merkmalhaftigkeit, and "our own (conception of markedness)" - that is Markiertheit (Chomsky/Halle (1968: 400, Footnote 4))[2]

- Then the works of natural phonologists of different schools like Stampe, Bailey and others in the first part of the seventies must be mentioned (cp. Stampe (1972) and Bailey (1973)). In these works the concept of markedness is extended to context sensitive markedness; the contradictions between context free and context sensitive markedness evaluations resulting from this are also discussed. A further topic is the 'place' of markedness theory in linguistic theory, i.e. the interrelations between markedness theory and the theory of grammar. It is also important to note that it was within natural phonology that the connections between markedness and language change were paid attention for the first time. Thus Bailey formulates the first version of a 'natural' change principle that says that non-socially conditioned change proceeds from more marked to less marked grammatical structures (Bailey (1973: 37)).

- In the late seventies and early eighties the school of natural morphology, represented by Dressler, Mayertahler, Wurzel and others, picked up the thread of natural phonology and transfered the concept

of markedness to morphology, thereby resuming many of Jakobson's fruitful ideas from the early thirties. A series of principles determining the markedness of morphological forms were postulated (cp. Dressler (1982), (1988); Mayerthaler (1981); Wurzel (1984); Dressler/Mayerthaler/ Pannagl/Wurzel (1987)).

- Finally Vennemann's concept of preference belongs to this tradition, presented in various publications from 1983 to 1990. Vennemann applies it to the complicated area of syllable structure. It is important that he formulates phonological markedness principles ('preference laws') which for the first time do not only distinguish between marked and unmarked, but assign gradual, relative markedness values to phonological entities (cp. Vennemann (1983), (1988), (1989)).

The concept of markedness presented here draws on within this linguistic tradition. It uses the findings of the different approaches to markedness without agreeing completely with any of them.

1. In present-day linguistic publications and discussions the terms markedness and marked occur relatively frequently. However, the term marked is often used simply to characterize linguistic entities, which are felt to deviate from the 'normal' in some sense. It is evident that this use of the term marked is pre-theoretical. The term as such doesn't explain anything. Before a linguist speaks about markedness, she or he has to say, in which sense she or he uses the term.

If one wants to clarify what markedness means, one has to make clear first of all that the phenomenon of markedness in grammar may and must be viewed on different levels of consideration, which must be distinguished carefully for methological reasons - one of the main points of this paper. Here the following three levels are relevant:

(i) the level of evidence: the facts;

(ii) the level of the actual theory: markedness theory;

(iii) the level of foundation: explanation of markedness theory by neighbouring disciplines of linguistics.

In a large part of the discussion on markedness, these levels are ignored and mixed up again and again, which leads to needless misunderstandings and confusion.

1.1. Let us start with the first level, the level of the facts underlying markedness theory. We will discuss a phonological and a morphological example, which we will follow up thoughout a series of linguistic areas of facts.We start with the phonological example. It concerns the relation between the front unrounded vowels /i/ and /e/ and the front rounded vowels /y/ and /ø/. Cp. the facts from the different areas:

(i) **The structure of language-specific systems**: Front unrounded vowels are found in all languages which differentiate front and back vowels (that means in almost all languages; the differenciation is for instance not found in Caucasian languages like Adyghe that only has /i̯/, /ə/ and /a/), front rounded vowels are found only in a little subclass of them[3] In other words: the occurrence of front rounded vowels is implied by the occurrence of front unrounded vowels, but not the other way round:

(1) $\begin{bmatrix} V \\ -\text{back} \\ +\text{round} \end{bmatrix} \supset \begin{bmatrix} V \\ -\text{back} \\ -\text{round} \end{bmatrix}$

Examples of languages without front rounded vowels are (as is known) English, Polish and Italian, languages with front rounded vowels German, French, and Hungarian. In languages with front rounded vowels their number is either equal to the number of front unrounded vowels (as in French with three each and Hungarian with two each) or

is smaller than the number of front unrounded vowels (as in Finnish with two front rounded and three front unrounded vowels and in Lezghian with one front rounded and two front unrounded vowels. Cp. the following vowel systems:[4]

(2) (a) Hungarian: (b) French: (c) Lezghian (d) Finnish:

i	y		u			i	y		u			i	y		u			i	y		u
e	ø		o			e	ø		o			e						e	ø		o
			ɒ			ɛ	œ		ɔ					a				æ		ɑ	
							a		ɑ												

(ii) **Language change:** There are many phonological changes in different languages where front rounded vowels become front unrounded vowels independently of the context, that is not conditioned by other segments.. Thus the Middle High German front rounded vowels have changed to front unrounded vowels both in most Upper and Middle German dialects as well as in Yiddish:

(3) (a) Middle High German *brücke* 'bridge' > Upper Saxon *bricke*, Yiddish *brik*
schoene 'nice' > *scheen*, *schejn*.

The same developement may be noticed in the history of English:

(3) (b) Old English *brydge* > New English *bridge*
soekan > *seek* .

In contrast, no context-free (nonassimilatory) transitions from front unrounded to front rounded vowels occur in language change. (Of course there exist assimilatory conditioned changes in this direction; we will come back to this below.) Thus we can record:

(4) $\begin{bmatrix} V \\ -\text{back} \\ +\text{round} \end{bmatrix} > \begin{bmatrix} V \\ -\text{back} \\ -\text{round} \end{bmatrix}$

(iii) **Language acquisition:** Front rounded vowels are mastered by the child after front unrounded vowels, as for instance, investigations of Dutch and French speaking children demonstrate (Jakobson (1971b: 365)). The same results hold for German-speaking children as well.

(iv) **Aphasia:** Speakers suffering from central speech disorders frequently loose front rounded vowels, whereas front unrounded vowels are retained as corresponding research has shown (Jakobson (1971b: 369)).

(v) **Slips of the tongue:** In languages with front rounded vowels, like German, slips of the tongue of the type [ti:r] instead of *Tür* and [le:zn] instead of *lösen* are found more common than slips of the type [ty:r] instead of *Tier* and [lø:zn] instead of *lesen*. (These generalizations hold independently of the dialectal origin of the speaker.)

The morphological example concerns the relation between the different types of category markers; we will restrict ourselves here to the main types, additive, modificatory and subtractive markers:

(i) **The structure of language-specific systems:** Inflectional forms with additive category markers (affixes and reduplication) appear in all languages that have an inflectional morphology, i.e. in all agglutinative and fusional languages. Inflectional forms with modificatory markers (vowel and consonant alternations, alternations of suprasegmental structures) are found only in a subclass of these languages, namely only in languages which are not strictly agglutinative like Turkish. Examples of languages with such modificatory markers include German, Latin and Arabic, but also in Finnish and Estonian. Remember that the appearence of forms with additive and modificatory markers is the basis for the classical typological classification of the languages. Inflectional forms with subtractive markers occur only in a small subclass of the languages with modificatory markers. However the occurrence of such subtractive markers is not systematic, but peripheral; cp. the often cited formation of Genitive Plural-forms in Russian (and other Slavonic languages) like *slovo* 'word' - Genitive Plural *slov* and *ryba* 'fish' - Genitive Plural *ryb* and the formation of Accusative Singular-forms in Old Norse (and New Icelandic) like *hundr* (New Icelandic *hundur*) 'dog' - Accusative Singular *hund*. As is known, there is no language type that is characterized by the occurrence of inflectional forms with subtractive category markers. That means the following implication between the occurrence of the three marker types:

(5) $\text{Marker}_{sub} \supset \text{Marker}_{mod} \supset \text{Marker}_{add}$

(ii) **Language change:** In the history of languages many morphological changes are known that proceed along this implicative chain from subtractive to additive symbolization of categories. Thus the subtractive G.Pl.-forms in Slavonian clearly show the tendency to be replaced by additive forms. In Sorabian this change has become regular, cp. Low Sorabian *słowo* - G.Pl. *słowow* and *ryba* - G.Pl. *rybow* (only of a small group of nouns zero forms are still possible in some contexts). In Russian a group of nouns make their G.Pl.-forms regular with an additive marker instead of the old subtractive one, cp. *oblako* 'apple' - G.Pl. *oblakov* and in colloquial Russian often all subtractive G.Pl.-forms are replaced by the corresponding *ov*-forms: *slovov*, *rybov*. In Continental Scandinavian (not in the more conservative Icelandic) the A.Sg.-forms with subtractive markers of the type *hund* versus N.Sg. *hundr* were removed as early as in tne Middle Ages in such a way that the accusative forms are transferred to the nominative, cp. later Old Swedish N.Sg. *hund* - A.Sg. *hund* (this levelling happens before all other nominative-accusative-levellings and before any other levellings between the case forms)[5]. A replacement of inflectional forms with modificatory markers with additive markers is the transition form strong to weak conjugation of German verbs that starts in Middle High German and is still continuing today; cp. older *bellen* 'bark' - preterite *(er) boll* > *(er) bellte* and *kreischen* 'screech' - preterite *(er) krisch* > *(er) kreischte*; more recent *melken* '(to) milk' - *preterite (er) molk* > *(er) melkte* and *gären* 'ferment' - preterite *(es) gor* > *(es) gärte*. Parallel changes are found in English and Continental Scandinavian. Thus the direction of the change in marker types is this:[6]

(6) $\text{Marker}_{sub} > \text{Marker}_{mod} > \text{Marker}_{add}$.

(iii) **Language acquisition:** Language acquisition also works along this chain of implication between the marker types. Thus Russian speaking children acquire not only the additive G.Pl.-forms like *stolov* from N.Pl. *stol* 'table' and *domov* from N.Sg. *dom* 'house' earlier than the subtractive ones like *slov* and *knig*, but also transfer such forms

to words of the type *slovo* and *kniga*. The same holds for the acquisition of the strong via transitory 'regularized' weak forms. So in German child language we observe verbal forms like *er gebte* and *er schwimmte* from *geben* 'give' and *schwimmen* 'swim', instead of the correct forms *er gab* and *er schwamm* in German child language.

(iv) **Aphasia:** In aphasic speech inflectional forms with subtractive and modificatory markers are disturbed more often than forms with additive markers. Here also the mastery of the strong and weak verbs in German (as in other Germanic languages) is a good example. With aphasics the formation of strong verb forms is often disturbed, whereas the formation of weak verb forms is totally intact.[7]

(v) **Slips of the tongue:** In producing slips of the tongue significantly more inflectional forms with subtractive and modificatory markers are replaced by forms with additive markers than the other way around. Thus, in Russian current slips are *slovov* instead of *slov*, but not **stol* instead of *stolov*, and in German *ratete* from *raten* 'advise', instead of *riet,* as well as *greifte* from *greifen* 'grasp' instead of *griff*, but not **wiet* of *waten* 'wade' instead of *watete* as well as **riff* from *reifen* instead of *reifte*.

These facts from five different linguistic areas show that grammatical entities of the same class, in this case of the phonological class of front vowels and of the morphological class of category markers respectively are not simply equivalent, 'equally good' for the speaker. Certain grammatical entities are obviously dealt with more easily by speakers than other grammatical entities, which is illustrated by the three psychogrammatical areas language acquisition, aphasia and slips of the tongue, and they are obviously prefered to other grammatical entities by the speakers, which is illustrated by the structure of language specific systems and language change - two sides of the same coin. All languages are full of such relations in all parts of their systems. This may be interpreted to the effect that there exist markedness relations between the corresponding grammatical entities. In the examples discussed front rounded vowels are marked and front unrounded vowels are unmarked; subtractive category markers are more marked than modificatory ones, and modificatory markers are more marked than additive ones. From the facts presented above we can conclude that

- the existence of (more) marked grammatical entities in a language system implies the existence of their less marked/unmarked counterparts[8],
- (more) marked grammatical entities are replaced by their less marked/unmarked counterparts in language change[9],
- (more) marked grammatical entities are acquired before their less marked/unmarked counterparts in first language acquisition,
- (more) marked grammatical entities get lost before their less marked/unmarked counterparts in aphasia,
- (more) marked grammatical entities are more likely to undergo slips of the tongue than their less marked/unmarked counterparts.

What can we tell about markedness at this level now which I called above the level of facts or of evidence? We may now answer the question which effects markedness has, but we have not yet answered the question what markedness is. Since - to come back to the first example - the front rounded vowels are not marked, because they behave as stated, but they behave as stated, because they are marked. Their specific behaviour within the scope of facts from language systems, language change, language acquisition, aphasia, and slips of the tongue is an epiphenomenon of markedness. What markedness is, does not result directly and theory-independently from the facts, but just from a theory on the facts. Of course this should be trivial, but the use of the notion of markedness in present-day linguistics is pretheoretical in this sense for a large part.

It is important to note that already the facts at this level of consideration show that markedness (Markiertheit) cannot be identified with 'featuredness' (Merkmalhaftigkeit), which is done especially in publications written in English frequently, but not only there. 'Featuredness' of a grammatical entity means that this entity has a feature (in the broadest sense), which is absent from another entity of the same class. In our example of rounding of front vowels markedness and 'featuredness' agree: The front rounded vowels have one more positive feature than their unrounded counterparts, namely the rounding of the lips. The relations between the nonlow back vowels are quite different, however. Here, of course, the round vowels /u/ and /o/ are 'featured' (merkmalhaft) compared with their unrounded pendants /ɯ/ and /ɤ/ . But in the class of nonlow back vowels the rounded ones are unmarked and the unrounded ones are marked, which is suggested by phonetic facts discussed below. The distribution of markedness and 'featuredness' for the vowels mentioned above is as follows:

(7)

[- round] 'unfeatured'		[+ round] 'featured'		[- round] 'unfeatured'
i	y	u		ɯ
e	ø	o		ɤ
unmarked		marked		unmarked
marked				

The same can be easily demonstrated for morphology by the plural formation of the English nouns. As is well known the normal plural formation is carried out by the additive category marker *-s,* cp. *dog - dogs* and *cat - cats*. But there exists a small group of animal names with zero plurals like *sheep - sheep* and *fish - fish*. It is evident that the English plural forms with the marker *-s* are 'featured' and the forms without the marker are 'unfeatured'; the relevant feature is just the *-s* .But in English plural formation with a category marker is unmarked, and plural formation without a marker is marked as we will see later on in detail.[10] Again, the values of markedness and 'featuredness' do not coincide.

1.2. This brings us to the second level of consideration, the level of markedness theory. Markedness theory lays down the markedness relations between grammatical entities of the same class by assigning markedness values to them. (As is known the phonological markedness theory of Chomsky/Halle (1968: 401ff.) operates in a different manner on that score.) A markedness theory consists of universal principles or laws, which (at least in some cases) may be ordered hierarchically. These principles are to be called markedness principles here (they are called 'marking conventions' by Chomsky/Halle (1968) and 'preference laws' by Vennemann (1983, 1988, 1989)). These markedness principles evaluate grammatical entities. They do so not generally but always regarding certain parameters. Markedness theory - in this point misunderstandings also appear frequently - describes the markedness relations in the language system, it does not explain them (cp. Vennemann (1983: 13)).

Markedness is fixed theory-internally. That does not mean, however, that this fixation may occur arbitrarily. It is hedged in a twofold manner:
- firstly, markedness theory has to explain the independently given facts;

- secondly, markedness theory itself has to be explainable independently (at least in principle).
A markedness theory, also one that meets these conditions, may be outlined in different ways. We will soon come back to the question of what such a theory could look like.

1.3. The next level of consideration, is that of founding markedness theory in terms of support from adjacent sciences. Not only the examples stated above, but masses of grammatical facts from every known language system suggest that markedness reflects grammatical complexity that strains the human language capacity. In other words (more) marked grammatical entities strain the language capacity stronger than their less marked/unmarked counterparts. It is in this sense that grammatical entities are not equivalent or 'equally good' for the speaker. Less marked/unmarked grammatical entities may be acquired and dealt with more easily by the speakers and are hence unconsciously preferred by the speakers. Markedness, or more precisely: the degree of markedness of a grammatical entity, is thus the relative measure for the straining of the language capacity regarding a certain parameter.

In the case of our phonological example this is understandable without difficulties: Front vowels are articulated with a tongue position in which unrounded (spread) lip apperture may be executed easier than round lip apperture; rounding of the lips requires an additional articulatory effort. The vowels /y/ and /ø/ are articulatorily more complex than the vowels /i/ and /e/ and strain language capacity stronger. In contrast, the nonlow back vowels are articulated with a tongue position in which a rounded lip apperture may be executed easier than an unrounded one and the avoidance of lip rounding requires an additional articulatory effort. Thus, the unrounded vowels /ɯ/ and /ʏ/ are articulatorily more complex than their rounded pendants /u/ and /o/ and strain the language capacity stronger. In sum, phonological markedness is founded on phonetics, that means the articulatory and/or auditive complexity of the phonological entities. Thus phonetics can give justified statements, if /i/, /e/ or /y/, /ø/ and if /u/, /o/ or /ɯ/, /ʏ/ is articulated easier, independently of any markedness theory.[11]

Also, the three mentioned types of category markers in our morphological example strain the human language capacity to different degrees. They are handled differently easy by the speakers and especially by the hearers because of their particular sign shapes. In the case of additive markers the category is symbolized directly by a special formal entity, a morpheme of its own, within the word. In comparison, modificatory markers symbolize their categories indirectly by a formal change of the base morpheme; this morpheme then symbolizes its lexical meaning and the category together. Finally, subtractive markers symbolize the category indirectly by the absense of a morpheme present in the base form; in contrast to the other marker types there is no formal part of the word at all to which category semantics is limked. The sign relations in words with the three marker types differ in their complexity. Morphological markedness is based on semiotic complexity, more precisely: on the complexity of the mutual assignment of semantic and formal elements within the scope of the largest morphological sign, the word. Also, semiotics can give justified statements, for instance if a morphological sign with additive category symbolization like the Russian genitive plural form *stolov* from *stol* 'table' or a morphological sign with subtractive category symbolization like the genitive plural form *knig* from *kniga* 'book' is more complex regarding its sign relations. Again this is valid totally independent of the existence of a markedness theory. It shall be added that syntactic markedness is also based on semiotic complexity, i.e. the complexity of the mutual assignment of semantic and formal elements in the scope of the syntactic signs, the syntactic constituents.

Finally the semantic markedness of grammatical categories (number, case, tense, mood and so on) is founded on the cognitive complexity of the underlying concepts. Also, corresponding independent statements are possible here, such as, if the concept 'more than one', underlying the category of plural, or the concept 'more than one and exactly two', underlying the category of dual, is more complex. Thus the category of dual is semantically more marked than the category of plural (cp. the implicative relation between the occurrence of these categories in natural languages).

It must be conceded that the situation is not so clear in all cases, especially if it concerns semiotic and cognitive complexity, but it is decisive that grammatical markedness relations may be reduced in principle to independently given facts and explanations of adjacent disciplines of linguistics and can be explained itself in terms of their theories.

Now we can answer the question what markedness really is: markedness is nothing but straining of the human language capacity, conditioned by the articulatory-auditive, semiotic and cognitive complexity of the respective grammatical entities.

2. As promised we'll now return to the question, what a markedness theory could look like. A markedness theory is a grammatical evaluation theory consisting of universal markedness principles which assign markedness values to grammatical entities. I will assume that the general form of the markedness principles is as following:

(8) **General form of markedness principles**
A grammatical entity G_j is the less marked regarding a markedness parameter M_i the stronger the degree of its property P_k is.

Grammatical entities in the sense relevant here are:
- in phonology: segments, segment clusters, syllables, phonological words, suprasegmental structures;
- in morphology: morphological markers, morphemes, inflectional and derivational forms;
- in syntax: syntactic phrases and sentences;
- in semantics: inflectional and derivational categories.

As already stated it is decisive that markedness principles do not evaluate the grammatical entities in general, but regarding certain markedness parameters that make up essential aspects of their structure. There is no 'markedness as such', and the statement 'G_j is marked' is either an abbreviation or simply meaningless. Markedness parameters refer to certain properties of grammatical entities. Markedness then results form the realization of these properties. (The relation between a markedness parameter and the corresponding property is a practically and theoretically interesting point which unfortunately can't dealt with here for different reasons.) The properties relevant for markedness evaluation are frequently not found in binary realization (G_j has the property P_k or not), but exist in gradual realization; cp. the example of different marker types. Therefore markedness itself also must be understood to be basically gradual, the common distinction 'marked versus unmarked' is not sufficient. Thus, relative markedness evaluations of the type 'G_1 is more/less marked than G_2 regarding a markedness parameter M_i' arise. As as a rather simple example of a markedness principle, the phonological principle relating to the rounding of front vowels independent of context may be quoted:

(9) **Rounding of front vowels**
A front vowel is unmarked regarding rounding, if it is unrounded, and marked if it is rounded.

Here the property relevant for markedness only occurs in binary distinction; we only have rounded and unrounded vowels.Therefore the distinction 'marked versus unmarked' is sufficient in this special case. But there are many cases also in phonology which unambiguously indicate that this binary distinction is not enough. One of them concerns the phonological substance of unstressed syllables in accent-counting languages (like English and German among others); cp. the following markedness principle:

(10) **Phonological substance of unstressed syllables in accent-counting languages**
In accent-counting languages a phonological word is the less marked regarding the phonological substance of its unstressed syllables, the less phonological substance these syllables have.

The following facts demonstrate the markedness gradation with an example of phonological realizations of the word *geben* 'give' in different stages of German language history and in Modern German varieties From left to right there is a continuous reduction of markedness in the different forms:

(11)

Germanic	Old High German	Middle/New High German	Colloquial New High German
*'*gebanan*	> *'geban*	> *'geben* ['gebən] ['ge:bən]	> ['ge:bm] > ['ge:m]

In reconstructed Germanic words of this type still have two unaccented syllables and thus are relatively strongly marked regarding principle (10). In Old High German there is only one unaccented syllable with a full vowel left. In Middle High German and in Standard New High German these words show even only one unaccented syllable with the reduced vowel [ə], and in colloquial varieties of New High German the unaccented syllable consists only of a nasal consonant or has disappeared altogether. The resulting monosyllabic form ['ge:m] is completely unmarked regarding the principle.

Let us return to our morphological example now, the markedness relations between the three discussed types of category markers. The relevant morphological markedness principle can be formulated in the following way:

(12) **Category marker types**
A morphological category marker is unmarked regarding its marker type, if it is additive; it is the more marked, the stronger it deviates from the additive type.

Following the arguments for the different complexity of the three marker types, it is plausible to assume that a subtractive marker deviates stronger from the unmarked additive type than a modificatory one and therefore is more marked than a modificatory marker.

It is substantial for the understanding of markedness (and also for its consequences for the theory of language change) that different markedness principles may assign varying, contradictory markedness values to the same grammatical entity. Thus, to continue our example of front vowels, besides the principle (9), there is another markedness principle concerning the rounding of front vowels in a certain context. It can be formulated in the following manner:

(13) **Rounding of front vowels preceeding rounded consonants**
A front vowel preceeding a rounded consonant is unmarked regarding rounding if it is rounded and marked, if it is unrounded.

With that we get the constellation that, preceeding a rounded consonant, a front unrounded vowel is unmarked according to the context-free principle (9) and marked according to the context-sensitive principle (13), whereas a front rounded vowel is marked according to the context-free principle (9) and unmarked according to the context-sensitive principle (13). It is easy to see that unmarkedness according to both principles at the same time is not possible. The example shows that a phonology that is optimal in every respect can't be achieved; the same is valid for the whole language system. Here a markedness conflict results. Markedness conflicts of this type (there also are other types) in phonology are solved in general by following the maxim that a markedness principle applying to a larger grammatical entity overrides a principle applying to a smaller entity. That means that there is a hierarchical order between the corresponding markedness principles. In our case this means that the context-sensitive principle (13) is stronger than the context-free principle (9). That this assumption is correct may be seen for instance in the Berlin dialect that (unlike Standard German and most other languages and dialects) has a rounded palato-alveolar fricative [ʃʷ] and consequently shows [y] and [ø] instead of Standard German [i] and [e] preceeding this round consonant); cp. (the examples meaning 'fish', 'mix'; 'ash tree' and 'laundry'):

(14) Berlin dialect: [fyʃʷ], [myʃʷn] - Standard German: *Fisch, mischen* ([i])
[øʃʷə], [vøʃʷə] *Esche, Wäsche* ([e])

There is strong evidence that such hierarchical relations also exist between morphological markedness principles. This is expressed by the following two principles:[12]

(15) **Constructional Iconicity**
A semantically more complex, derived morphological form is unmarked regarding constructional iconicity, if it is symbolized formally more costly than its semantically less complex base form; it is the more marked, the stronger its symbolization deviates from this.

(16) **System Adequacy**
A morphological form is less marked regarding system adequacy the more it corresponds to the system-defining structural properties of the respective morphological system.

Principle (15) states (among other things) that in noun inflection the forms of the derived (nonnominative) cases should be symbolized formally more costly than the nominative form. This should be valid also for the dative and accusative forms in German.However, this is correct only partially, cp. the type *(der) Bär* 'bear' - *(dem/den) Bären vs.* the type *(der) Hund* 'dog'- *(dem/den) Hund, (die) Kuh* 'cow' - *(der/die) Kuh, (das) Pony* 'pony' - *(dem/das) Pony* and so on. Thus, dative-accusative forms like *(dem/den) Bären* with case suffix are unmarked regarding constructional iconicity, dative-accusative forms like *(dem/den) Hund* without case suffix are marked (whereby the degree of markedness is irrelevant here).

In German noun inflection only case symbolization by the inflected article is system adequate, and symbolization by suffixes on the noun is not system adequate

(concerning the subparameter of marker type).Thus, principle (16) implies that dative-accusative forms like *(dem/den) Hund* are unmarked, whereas forms like *(dem/den) Bären* are marked regarding system adaquacy.[13]

From this it follows that the German inflectional forms cited above have contradictory markedness values regarding the two parameters: Forms like *(dem/den) Bären* are unmarked regarding constructional iconicity and marked concerning system adequacy, and forms like *(dem/den) Hund* are unmarked regarding system adaquacy and marked regarding constructural iconicity. Both types of forms are marked, but in a different manner. If we want to know which type of markedness 'weighs heavier', i.e. which markedness principle is stronger, we may look for historical changes in this area. And indeed, such changes do occur: Presently, the case forms of the type *(dem/den) Bären* are replaced by forms like *(dem/den) Bär* without case suffixes, but there are no changes from *(dem/den) Hund* to **(dem/den) Hunden* or another corresponding form. This confirms our claim that the principle of system adequacy overrides the principle of constructional iconicity.

Also, in morphology there obviously exist hierachical relations between markedness principles reflecting the quantity of the domain relevant for markedness evaluation. Thus for markedness evaluation regarding system adequacy the whole morphological system of noun, verb etc. is relevant, for markedness evaluation regarding constructional iconicity only two morphological forms are relevant, the respective base form and the derived form. There are good reasons to assume that due to its maximal domain the principle of system adequacy is the strongest morphological principle whatsoever (Wurzel (1984: 186f.)). As for the hierachical relations between the other morphological principles there exist some more far-reaching hypotheses that cannot be discussed here (cp. Wurzel (ibid.) and Dressler (1985), (1988)). In this area, much work has yet to be done.

Let us conclude with the assessment that markedness principles form a specific class of language universals which may be characterized as evaluation universals. Whereas the theory of universal grammar (UG) explicates what kind of grammatical entities must occur, may occur or cannot occur in natural languages, markedness theory (MT) explicates what kind of grammatical entities from the universal 'offer' are prefered or more or less disprefered by the speakers and thus by the natural languages.This way MT explains the systematic asymmetry in the use of universally available grammatical means, observable in the structure of all languages and in language development. Markedness is not only a descriptive device in linguistics that a linguist may use or ignore ad libitum; it is one of the most important properties of natural language structure and belongs - to take over a formulation of Roman Jakobson, one of the 'pilgrim fathers' of the markedness concept - to the "essence of language".

Footnotes

* This paper is dedicated to Manfred Bierwisch - who led me on the path of markedness 30 years ago (time flies!) - on the occation of his 65th birthday.

1 An interesting account of the role of markedness (understood in Jakobson's and Trubetzkoy's sense) in natural language and beyond it in all semiotic systems is given in Waugh (1982).

2 A comprehensive theory of markedness based on Chomsky/Halle's concept of phonological markedness is outlined in Kean 1981).

3 For the statistical distribution of front unrounded and front rounded vowels in the languages of the world cp. Chrystal (1993: 167).

4 Vowel systems are based on Trubetzkoy (1939: 86ff.) and Lass (1984: 134ff.).

5 In Swedish nouns all morphological distinctions between nominative and accusative (and dative as well) were levelled. But whereas the levelling between nominative and accusative in paradigms of the *hundr -hund* type already took place in Old Swedish, the nominative-accusative distinction in paradigms without subtractive accusative symbolization (for example in the type *bonde* 'farmer' - accusative *bonda*) was still intact in the 16th century and was not levelled until the emergence of New Swedish (Wessén (1969: 137ff. and 185ff.)).

6 This implication gives only the direction of the changes in marker type, not nessessarily their stages. It does not claim that a subtractive marker cannot be replaced directly by an additive one. - Of course there are also transitions from additive to modificatory and to subtractive markers in language history, but they are always conditioned by phonological changes. Cp. the development of German nouns like *Apfel* 'apple': Old High German *apful* - N.Pl. **apfuli* > *aphul - ephili* (in normal Old High German orthography) > Middle High German *apfel - epfele* > New High German *Apfel - Äpfel* and the development of nouns like *Ring* ' ring' in South East Thuringian dialects: Middle High German *rinc* [riŋk] - N.Pl. *ringe* [riŋgə] > [rink] - [riŋə] > [riŋk] - [riŋ]. All the relevant transitions are the result of phonological reductions. For a parallel case in Upper Hessian cp. Schirmunski (1962: 417).

7 These observation are based on research in the former working group of aphasiology (headed by Egon Weigl) at the Berlin Academy of Sciences.

8 Of course this is valid only for 'natural grammatical change', i.e. change that is not initiated socially. The role of markedness in language change is treated in detail in Wurzel (1994).

9 Vennemann (1988: 3f.) points out that there may be certain exceptions to this under special, restricted conditions. "Nethertheless the normal situation occurs with sufficient frequency..."

10 The English plural forms without markers are marked both regarding constructional iconicity and system adequacy; cp. markedness principles (15) and (16) below.

11 As is known, in terms of phonetics front unrounded vowels belong to the class of primary cardinal vowels, the class of front rounded vowels to the class of secondary cardinal vowels.

12 Cp. this markedness principle with principle (12) concerning marker type: Whereas principle (15) refers to the whole morphological construction, principle (12) only refers to the category marker within this construction. For the principle of constructional iconicity and its consequences cp. Mayerthaler (1981: 23ff.).

13 For the notion of system adequacy and the markedness relations regarding system adequacy in German noun inflection cp. Wurzel (1984: 81ff.)).

Literature

Bailey, C.-J.N. (1973), *Variation and Linguistic Theory*, Arlington: Center for Applied Linguistics.

Chomsky, N./M. Halle (1968), *The Sound Pattern of English*, New York/Evanston/London: Harper and Row.

Chrystal, D. (1993), *Die Cambridge Enzyklopädie der Sprache*, Frankfurt/New York: Campus.

Dressler, W.U. (1982), On word formation in Natural Morphology; *Wiener Linguistische Gazette* 26, pp. 3-14.

------ (1985), Typological Aspects of Natural Morphology; *Acta Linguistica Academiae Scientiarum Hungaricae* 35 (1-2), pp. 51-70.

Dressler, W.U./W. Mayerthaler/O. Panagl/W.U. Wurzel (1987), *Leitmotifs in Natural Morphology,* Amsterdam/Philadelphia: Benjamins.

Jakobson, R. (1971a), Zur Struktur des russischen Verbums; R. Jakobson, *Selected Writings II, Word and Language*, The Hague/Paris: Mouton, pp. 3-15.

------ (1971b), Kindersprache, Aphasie und allgemeine Lautgesetze; R. Jakobson, *Selected Writings I, Phonological Studies*, The Hague/Paris: Mouton, pp. 328-401.

Kean, M.L. (1980), On a Theory of Markedness: Some General Considerations and a Case in Point; A. Beletti, L. Brandi, L. Rizzi (eds.), *Theory of Markedness in Generative Grammar*, Pisa: Scuola Normale Superiore.

Lass, R. (1984), *Phonology*, Cambridge: Cambridge University Press.

Mayerthaler, W. (1981), *Natürliche Morphologie*, Wiesbaden: Athenaion.

Stampe, D. (1972), *How I Spent my Summer Vacation*; Doctoral Thesis, Ohio State University, Columbus.

Schirmunski, V.M. (1962), *Deutsche Mundartkunde*, Berlin: Akademie-Verlag.

Trubetzkoy, N.S. (1931), Die phonologischen Systeme; *Travaux du Cercle Linguistique de Prague* 4, pp. 96-116.

------ (1967), *Prinzipien der Phonologie*, 4.Aufl., Göttingen: Vandenhoeck und Ruprecht.

------ (1975), *Letters and Notes*, ed. by R. Jakobson, The Hague: Mouton.

Vennemann, Th. (1983), Causality in language change: Theories of linguistic preference as a basis for linguistic explanations; *Folia Linguistica Historica* 6, pp. 5-26.

------ (1988), *Preference Laws for Syllable Structure and the Explanation of Sound Change*, Berlin/New York/Amsterdam: Mouton/de Gruyter.

------ (1989), Language change as language improvement; V. Orioles (ed.), *Modelli esplicativi della diacronia linguistica,* Pisa: Gardini, pp. 11-35.

Waugh, L.R. (1982), Marked and unmarked: A choice between unequals in semiotic structure; *Semiotica* 38-1/2, pp. 299-318.

Wessén, E. (1969), *Svensk språkhistoria; Ljudlära och ordböjningslära*, 18. uppl., 2. tryckn., Stockholm:

Wurzel, W.U. (1984), *Flexionsmorphologie und Natürlichkeit*; Studia grammatica XXI, Berlin: Akademie-Verlag.

------ (1994), *Grammatisch initierter Wandel*. Unter Mitarbeit von A. und D. Bittner, Projekt 'Prinzipien des Sprachwandels', Bochum: Brockmeyer.