Introduction to Optimality Theory in Phonology and Syntax

Gisbert Fanselow and Caroline Féry, University of Potsdam

June 2000, Course material for ESSLLI
First week:

1st session  
Precursors and background: transformationalism and derivations

2nd session  
Precursors and background: the answers to derivationalism

3rd session  
Outline of OT

4th session  
Examples of markedness hierarchies: conflicts between markedness and faithfulness

5th session  
The third kind of constraints: Alignment

Second week:

6th session  
Output-output correspondence

7th session  
Input and ineffability

8th session  
Opacity (Counterfeeding, Counterbleeding), Sympathy, Constraint conjunction

9th session  
Serial and cyclic OT

10th session  
Applications: optimal parsing, learnability

Suggested Readings
1. Precursors and Background

A Conflicts in Phonology
A.1.1 Overview of ordered rules
A1.2 Which conflicts arise in which part of the phonology.
A.1.3 Conflict resolution in OT

B Syntax
B1 Precursors: Representations
B2 Precursors: Conflict Resolution

A Conflicts in phonology

Resolution of conflicts are now considered by many linguists as lying at the heart of linguistic generalizations as well as typological variations, though generalizations and variations were not always analyzed in these terms. It will be shown in this first section how the precursor theories of OT addressed the conflicts raised by phonological facts. The first section focusses on derivationalism and the second on the answers to derivationalism. It will be shown which kinds of conflicts arise and how derivational and non-linear theories, the immediate precursors of OT, solve them.

As in syntax, phonology has been envisaged in two different kinds of approaches: derivational and representational. The notion of conflicts, central to OT, has not really been identified as such in older theories, though it is evident that conflicts and their resolution played at all times an important role. In a nutshell, derivational theories used the device of rule ordering to solve potential conflicts, and representational theories formulated constraints on the kind of phonological structures licensed by Universal Grammar.

Historically, however, the two approaches are not contemporary. Derivationalism came first, itself as an answer to structuralism, and representationalism was an answer to a crucial problem of derivationalism, namely the duplication problem, to be addressed below.

Before we do so, let us go back in time and take a brief look at the way structuralism whose main representants are Trubetzkoy and Jakobson envisage phonology. For them phonology was not a discipline whose main task is to resolve conflicts, but rather they thought that phonology should aim at developing a system of phonological oppositions both at the level of the phoneme (Trubetzkoy) and at the level of features (Jakobson). Emphasis was put on the notion of contrast rather than conflict. In structuralism, phonology consisted of three levels: a phonemic level, which is the level specifying which surface sounds are parts of the phonemic system of the language, thus the phonological level, a morphonological level including all idiosyncratic phonological information and a phonetic level which is relevant for the acoustic and articulatory stream of sounds, thus the result of phonology without being itself phonological.

The transformationalists, like Chomsky & Halle (Sound Pattern of English, SPE, 1968), proposed to eliminate the phonemic level of structuralism, since, according to them, a grammar consisting basically of a phonemic and a phonetic level could not account for the difference between free variation and contrast (but see Postal 1968). More abstractness was required. The result was a considerable extension of the power of phonology in allowing the successive transformation of underlying structures (the mentalist view of phonology). This is achieved by means of ordered rules deriving surface structure. The primary concern of transformationalists was not to resolve conflicts, but rather to account for observed alternation
in the phonological structure of roots, like English electric[k] ~ electricity[s], fable ~ fabulous (but obiːse ~ obiːsity). Such alternations arise not as consequence of conflicts but as the natural consequence of ordered generalizations and principles which, when applied in the correct order, ineluctably lead to these alternations. Of course, the ordering of rules is itself a resolution of conflicts, as elaborated below, but it was not conceived in these terms by transformationalists. The format of the transformationalist approach forced phonologists to make choice as to the underlying form of words. Fable and electric are underlying, or at least closer to the underlying forms, and fabulous and electricity are derived by Trisyllabic Laxing and Palatalization respectively. The choice of the underlying form is a logical consequence of the fact that the longer forms are morphologically derived from the shorter ones. But also in cases the derivations do not deliver a morphologically complex form it has to be decided which form is underlying, like in deriving the segmental structure or the metrical pattern or in noun to base derivations (tórmentyN ~ torméntV) and so on.

In Lexical Phonology, Kiparsky and Mohanan went a step further in grouping sets of rules in ordered levels, inside of which rules were also ordered. The levels are determined by the morphology-phonology interface and follow the linear organization of morphemes. Level I is the level of irregular morphological formation and of heavily phonologically biased derivation (instrument-al). Level II is the level of regular and not so phonologically influenced derivation (instrument-less) and composition (instrument store), and level III is the level of inflection (instruments, instrumentals). In Lexical Phonology, conflicts are thus solved in two different ways. First, rules are strictly ordered inside of each level and second levels are also ordered.1

1 Overview of ordered Rules

Cycles, feeding, counterfeeding, bleeding and counterbleeding were devices aimed at the description of phonological alternations from the point of view of ordered rules. The basic insight was that if rules had been ordered in a different way than they in fact were, things could have been otherwise. Important for this view is that rules CAN in principle be ordered in a different way. It looks as an accident that they come in the order they do. Locally, dialects can show reversed applications of two or three rules, like Final Devoicing and g-deletion in German (see below). Two different languages can also differ on how they order rules, but this was never a primary concern of derivational approaches (but see Kisseberth and Kenstowicz & Kisseberth 1977 for some remarks in this direction).

Rules do not seek to attain universality or even to describe unmarkedness. They are just a descriptive apparatus which can have a very limited range of application or a very large one. In other words, they either account for alternations of sounds which can be attested in just one environment in one language, or which are universally valid. The derivational approach does not make any difference between these two kinds of rules.

To sum up the review of conflict resolution in derivational phonology: conflicts arise only locally and are resolved by means of rule ordering. A different order leads to a different derivation. Typological variations is accounted for by different, possibly disjuncted sets of

1 OT concentrates on the conflicts appearing inside of each level, and until now has not really offered an alternative to the ordering of levels (though alignment facts as well as the different approaches to opacity, like output-output correspondence or sympathy, can account at least in part for some of the insights of Lexical Phonology). In several papers, Kiparsky defends a view of OT in which Lexical Phonology is an integral part. Ordered rules inside of levels are replaced by constraints hierarchies, but levels are maintained. Prince & Smolensky (1993) and McCarthy & Prince (1993) are agnostic as to the question whether levels exist. According to them, this question is orthogonal to the main concern of OT.
rules. The conflict between faithfulness and markedness of OT is not addressed. Rules can deliver indifferently more marked structures or less marked ones.

Problems related to the generative view of phonology are the following:
- Unconstrainedness: Rules are too powerful and too unconstrained. They allow too many derivations, also impossible ones because unnatural or illegal.
- Duplication problem. Assuming a level at which phonotactic generalizations are stated, rules derive only structures compatible with these phonotactic generalizations, and no others. Stated otherwise, rules are redundant. This can be considered as the opposite of conflict. Rules and phonotactics work hand in hand.
- Conspiracies of the rules. Rules aim at something. Different rules can have the same aim. In a rewriting rule A → B/ C _ D, the aim is CBD. We will see the example of hiatus in some detail, but the point can be made with many different phonological facts, like the unmarked syllable structure, achievement of unmarked segments, unmarked metrical structure, the best sonority structure etc. However, nothing in the format of the rules specifies what the rules try to achieve. Thus, one can formulate a rule B → A/ C _ D, whose aim is CAD, the reverse of the rule above. If one looks at derivations in the SPE style it is sometimes difficult to figure out what the aim of some rules is.

Of these three problems, only the last one is directly related to the view of grammar as a conflict resolving device. Different languages achieve the same goal by different means: hiatus can be resolved by epenthesis, deletion or gliding. The conflict in such a case can be stated as follows: all languages have a constraint against hiatus (two adjacent vocalic nuclei are forbidden) because hiatus is a universally dispreferred structure. But all languages have faithfulness constraints which require identity of the segmental organization between input and output. In particular, some faithfulness constraints can aim at preservation of an input hiatus. These faithfulness constraints take on different forms: some militate against epenthesis, some against deletion, other ones against gliding of vowel, etc. Because of the different weighting of constraints, languages differ as to whether and how they resolve hiatus.

**Representationalism: Non-linear phonologies**

The non-linear phonologies like autosegmental and metrical phonologies, resolved some of the problems raised by generative phonology, most of all the first one mentioned above, the problem of the unconstrainedness of the rules. If features are organized in a feature tree, some assimilations can not happen because they would deliver an ill-formed tree. A laryngeal feature like [voiced] cannot assimilate to a place feature, like [labial] or [dorsal] for instance, since these features appear at two different places in the feature tree. However, this theory just restricts the possibilities of rule applications, but it has nothing to say about the basics of conflicts, namely whether an assimilation happens or not, and whether a certain assimilation, deletion of a feature or of a class node will reduce the overall markedness of the structure.

We will see below that metrical theory, the other nonlinear branch of phonology, was confronted with a lot of conflicts and how these conflicts were resolved.

**Constraint-based phonologies**

The duplication problem is addressed by the so-called monosstral theories, like declarative phonology (Bird, Scobbie a.o.), harmony phonology (Goldsmith) and constraint and repair phonology (Paradis). All these models assume that there is only one level of phonology, the surface one, and that all alternations can be describe at these levels. However, no model can
escape opacity problems. Repair strategies are derivation-like. The monostratal approaches need different kinds of representations which interact by means of rules.

2. Which conflicts arise in which part of the phonology: the answers to derivationalism

Once you start looking for conflicts, you see them all over the place. Up to now, OT has been addressing just a small part of the facts that can be described in terms of conflicting tendencies. This section lists some of the places where conflicts in phonology arise and shows how older theories accounted for them. The solutions offered by OT are summarized briefly for each case and discussed from a broader perspective in part 3. This section is organized in the way OT resolves the conflicts mentioned. Basically, OT uses three kinds of constraints which can conflict with each other: markedness, faithfulness and alignment constraints, both across the families and inside of the families.

2.1 Conflicts between markedness and faithfulness.  

Phonological outputs are torn between the desire to be unmarked, and so fulfill the basic phonology of the language, and the desire to be faithful to their inputs, which can be badly marked. These basically conflicting tendencies is found in all domains of phonology. Beside the hiatus example, some additional examples are the following:

- Syllable structure: the most unmarked syllable structure is an open syllable with an onset (CV), but many languages are confronted with inputs which cannot be syllabified straightforwardly in a series of CV syllables. Some languages are faithful to there inputs (an extreme case is German with Strumpfs ‘sock, gen.’), some others chose the way of unmarkedness (the word Arbeit ‘work’ in Japanese is aribaatu, and the expression happy Christmas in Hawaiian is meli kirimaseki). For German-like tolerant languages, older theories account for the syllable structure in the phonotactic component or in prosodic phonology: the syllables template is so designed that it can include all edge consonants and no rule for deletion of segments is necessary. The restrictive Hawaiian option, on the other hand, is accounted for with ordered rules of deletion and epenthesis, beside the ones accounting for segmental transformations. Of course, some intermediate solutions between German and Hawaiian are available, like the prohibition against complex onsets and codas or the prohibition against just certain classes of segments in the syllable edges, and, important for the phonological theory, the fact that some sonority hierarchy violations are possible in some languages but not in others. All these facts have been accounted for in the derivational approach with rules, and in later, non-linear approaches, with templates of the syllable structure.

In OT such typological differences are accounted for by just one set of constraints. All languages have markedness constraints requiring an onset, prohibiting codas and complex margins. All languages have also different kinds of faithfulness constraints, like MAX and DEP prohibiting deletion and insertion of segments respectively. The typological difference is made by the difference in the ranking of these segments.

- Prosodic Morphology: the best Prosodic Word (PW) of a language is often a bisyllabic word, usually a trochee, like in German. However, in the largest part of morphology the unmarked PW is not visible, since faithfulness to inputs is more important than prosodic unmarkedness. In some cases, though, which have been called Emergence of the Unmarked by McCarthy & Prince (1994), the unmarked PW is required. In German this happens in some derivations like sonnig, lustig…, in inflection (segeln, wandern …), in frozen expressions (fix und fertig, *fertig und fix ‘exhausted’, Müller 1998) and most clearly in clippings and i-
formations, like Fabi and Kathi. Thus we find a conflict between the tendency to achieve the best PW and the one to be faithfull to an input in all domains of morphology/phonology interface but the conflict can be resolved in different ways in different parts of the morphology. Beside this intralinguistic variation, we also observe that different languages make different choices. Some languages do not even allow words lighter than a bimoraic syllable or foot (Minimal Word), some others allow all kinds of words, even monomoraic ones. To return to the i-formations of German, they are accounted for in the transformational phonology in terms of deletions of segmental material up to a certain point, often described as the "maximal syllable" (Kenstowicz 1994). The derivational suffix -i is added to the maximal syllable. Other solutions, like the Prosodic Circumscription (McCarthy & Prince 1986), have been proposed for similar prosodically-based morphological processes. In this model, a prosodic constituent, usually a foot or a syllable, is first isolated inside of a word. In a second step, a phonological operation is executed on one part and the last operation consists in putting the pieces together again, or not, in the case of the i-formation.

In OT, faithfulness of the inputs is ranked higher than the constraints responsible for prosodic unmarkedness, at least as far as ordinary outputs are concerned. However, faithfulness of some morphologically derived or inflected forms to the input or to the underived output is ranked below prosodic unmarkedness. This pattern (Faith-IO >> Prosodic Unmarkedness >> Faith-OO) is an instance of Emergence of the Unmarked.

- Some segments have redundant or unmarked values for some features. For instance, obstruents prefer to be voiceless and sonorants voiced. In German, voiced obstruents are allowed, though only in perceptually prominent positions (syllable onset), but voiceless sonorants are not: the Welsh word Lloyd is pronounced with a voiced [l], though it is voiceless in the original language. In this case, too, ordered rules were heavily used in the past, though the rules themselves could not show why it is better for an obstruent to be voiceless and for a sonorant to be voiced. To answer this problem, underspecification was introduced into the phonological theory. Some segments are unspecified for some features, rendering them redundant. Such predictable features are introduced late in the derivation or even on the phonetic level. But the arbitrariness problem mentioned above remains: pure representational models can not express why segments are unspecified for some features. In OT, there is a universal constraint against voiced obstruents and another one against voiceless sonorants. If an input contain segments with the wrong value for voice, two options are available: keep the value (faithfulness >> markedness) or change the value (markedness >> faithfulness).

Examples bearing on the conflict between markedness and faithfulness can be found on other parts of the phonology, as well. The solution is essential always the same. Markedness compete with faithfulness.

2.2 Conflicts between markedness and markedness: Different modules of phonology make different predictions

One aspect of phonology can require some kind of unmarked structure and another aspect some other unmarked structure for the same segmental or syllabic material. This point can be illustrated with metrical phonology. A canonical trochee (σσ) consists of two equal syllables, in which both σs have the same weight, mono- or bimoraic for instance. In contrast, a canonical iamb (σ σ_i) consists of unequal syllables: σ_i is lighter than σ. (Hayes 1993 a.o.). This requirement is conflicting with the requirement that all syllables have the structure CV. (It is also often conflicting with the faithfulness of the segmental material, but this is put aside here). The way languages resolve this kind of conflicts is again not uniform. Among iambic
languages, French lengthens the last syllable but Beduin Arabic reduces the first syllable. In the trochaic languages, languages can choose to ignore the difference in weight in syllables and build trochees on all kinds of syllables. This is what happens in the so-called quantity-insensitive languages. But in quantity-sensitive languages, ill-balanced trochees are avoided and different kinds of "repairs" arise. The heavy syllable is made lighter (by vowel reduction for instance) or no trochee is erected on certain sequences of syllables, and so on (see Hayes 1995 for examples). Additional conflicts arise from the fact that one cannot expect that all words of a trochaic language will consist of an even number of perfectly balanced syllables. First, words can be three or five syllables long. Second, syllables do not always meet the expectations of the perfect trochee. Again, these kinds of conflict were not perceived as such in older theories. Metrical theory was loaded with repair devices aimed at improving the syllable structure of ill-balanced feet. Some of the devices and parameters introduced by metrical phonology are discussed in section 2.3. However, it must be noticed that Metrical Phonology made huge progress when non-linear phonology was introduced which laid the focus of the theory on representations and no longer on rules. In this respect, non linear phonologies paved (part of) the way to constraint-based approaches in phonology.

Also very important for the new development in phonology was the representational model of syllables by means of CV-positions, X-positions, or moras, standing for their weight properties. In this kind of approach, templates play an important role. The best syllable or foot is the one which best matches the canonical template. Conflicts between input material and syllable templates were resolved by construction rules. Syllables or feet are constructed on segmental material step by step: first, nuclei are projecting a syllable node, then onsets, then codas. Consonants which cannot be adjoined to the syllable immediately because they find no docking place in the template, are either eliminated or included to the prosodic structure in a non-standard way, as an appendix to the syllable, as adjoined to a higher prosodic constituent, as a semi-syllable, etc. Stray consonants are ultimately deleted. The problem of such an approach is that syllable-building rules must be formulated on a language-specific basis because languages make different choices for the way they treat consonants which are not readily syllabifiable.

Conflicts involve alignment constraints: Too many or too few elements for the structure
- Phonology can require that two different phonological (or grammatical) constituents align at one or both edges. It can also require that two phonological entities, like features, tones or stress are associated with just one bearing unit. Consider the mapping between lexical tones and syllables in tone languages. The ideal autosegmental representation is the one in which there is exactly one tone per syllable and reversely, when melodies are described in terms of sequences of tones, there should be ideally one syllable per tone. However, this state of affair is not always fulfilled. Melodies, say HLH (high tone - low tone - high tone) in Igbo or Mende, can be associated with one-, two-, three-syllables or longer words. The ideal tone-to-syllable association will be present only in case the word is three syllables long, otherwise something will happen as a consequence of the conflict between number of tones and number of syllables. Again, different solutions are envisageable: in cases in which there are too few
syllables for too many tones, languages can choose to crowd tones at the end of the word, thus on the word final syllable (which is also the initial one in monosyllabic words). They can also choose to let the remaining tone(s) unrealized. In the opposite case, thus when there are too many syllables for the tones, the solutions are different. Often, tone association takes place left-to-right on a one-to-one basis until no tone is left. Then the last tone is spread across the remaining syllables. Here, too, other solutions exist. It can be the case that the last tone must be associated with the last syllable of its associating domain. In this case, another tone spreads. Or the medial syllables are left unspecified for tones. The last solution, again chosen by many languages is that the syllables for which there is no tone left take a default tone, mostly the low tone.

Alignment constraints, enriched with no crossing prohibition constraints or a constraint against multiply associated tones or syllables and so on, predict the way tones associate with syllables and specify whether tones can crowd on syllables, of whether it is better to leave some tones or syllables unspecified. It must be clear at this point that the older theories had to formulate different rules for all different cases, and that that OT can account for typological variations with one set of constraints ranked differently according to the option chosen by the language.

- An interesting source of typological variation that OT resolves to a large extent with alignment constraints is metrical structure. Consider as an example Maranungku, Weri: Warao and Choctaw which are organized in syllabic trochees, but with slight differences in each cases. ‘V stands for a stressed vowel.

(1) Type 1 (Maranungku): 'V  V  'V  V  'V  V  'V  V  'V
Type 2 (Weri):
Type 3 (Warao):
Type 4 (Choctaw):

The metrical phonology introduced different parameters to account for the different criteria entering the best metrical structure. These parameters are for example: boundedness (feet are binary or not), head (feet are left-headed or right-headed), directionality of feet assignment (feet are assigned to syllables from left-to-right or from right-to-left), obligatory branchingness (the stressed syllable of a foot must be heavy), extrametricality (the last syllable can be invisible for the metrical structure) and so on. The result was a theory which can deliver many different patterns, according to the settings of all these parameters. However, it was clear that only a few of the possible structures are actually present in the languages of the world.

OT replaces parameters with constraints and the assignment of feet with alignment constraints.

3. Conflict resolution in OT

We saw in this short overview that OT in phonology has three kinds of constraints: markedness, faithfulness and alignment. Examples were given for each kind of constraints, and it was shown how older theories explain the conflicts accounted by the three kinds of constraints. OT resolves all kinds of conflicts by ranking constraints.

It remains to be seen whether deeper generalizations can be made which explain why some parts of the phonology prefer one kind of constraints over the other.
B: Precursors of OT in the Syntax

B1. Representations

1. Identification of the place of OT among syntactic theories
2. Brief summary of two major arguments against explicit rules in syntax
3. Discussion of the pros and cons for representation-based explanations in syntax

1.0 Our discussion of OT syntax begins with a brief reflection on its predecessors. OT is a representational model of syntax, and variation between languages is described in terms of differences in the resolution of conflicts between very general principles of Universal Grammar. These two properties of OT are independent of each other - therefore, OT models have been proposed which are not representational in nature (e.g. the proposals of Gereon Müller), and the Government & Binding - approach to syntax (Chomsky 1981) is a representational theory but denies the existence of conflicts among the principles.

1.1. A grammar has to specify the set of all and only the grammatical sentences of a language. A generative grammar carries this specification out by defining an algorithm for generating this set of grammatical sentences. In a derivational theory, the explanatory force of the grammar lies in its rules to produce sentences and the way they are applied. The early generative models like the so-called standard theory (Chomsky 1965) were purely derivational models.

In a representational model, less or no emphasis at all is given to rule application. Rather, the rule component may overgenerate, because there is a set of principles by which the products of the generative component of grammar are evaluated - only those that fulfill the principles of representation are considered grammatical.

The difference is easy to exemplify. That every clause has a subject is guaranteed in a derivational model by assuming a rule like

\[ S \rightarrow \text{NP VP} \]

and by not assuming any rule that deletes the subject. In a representational model, rules such as

\[ X \rightarrow Y Z \]

may generate syntactic trees freely, and movement and deletion rules operate in a a fairly unrestricted way, but all representations that are generated must fulfill the Extended Projection Principle EPP:

Every clause has a subject!

in order to be grammatical.

1.2. As we have just mentioned, modern syntax started out as a derivational theory, with explicit construction-specific rules. There were explicit phrase structure rules for VPs, explicit passive transformations, etc. In that respect, phonology and syntax had a very similar beginning. The first major step towards representationalism was made when theories became modular/ constraint-based. Beginning with Chomsky (1964), and, in particular with Ross (1967), it became obvious that there are interesting
generalizations on the form of rules that go beyond the individual constructions in a certain language, and that may make up the universal core of syntax. Thus, Ross discovered that no transformation can extract anything out of a subject clause in English, and later research revealed that many languages share this property.

- who do you think that she loves?
- *who does that she loves annoy you
- *a girl who that you love annoys Bill
- he is taller than you think that she is
- *he is taller than you think that Bill is annoy us

Similarly, construction specific rules do not explain why categories are mostly endocentric, a theory working with construction specific rules seems too unconstrained, and it needs to duplicate information.

1.3. The constraints on syntax/ the principles of UG may figure in various ways in grammatical theory. They may function as constraints on the formulation of individual syntactic rules - a view of grammar which still subscribes to the idea that a grammar consists of a set of construction-specific rules - but what form these rules may take is severely restricted. Thus, we might say that any explicit phrase structure rule that we find in natural language must be a special case of

\[ XP \rightarrow \ldots X \ldots \]

or we might say that

\[ X' \rightarrow X_W Y \ldots Z \]

can be a PS-rule of a language only if the subcategorization frame W of X = Y...Z.

In the Chomskyan tradition, a different approach has been pursued: the constraints were always considered active in the derivation: the X-bar statements just introduced were taken to be the only mechanism by which structures are generated, or constraints were viewed as being applied in the course of derivations working with very general rules: the rule relates two positions, but whether two items may be affected by a certain process (like, say, a movement transformation, or Case assignment) or not is expressed in terms of construction-unspecific constraints on rule application.

1.4. The discovery of such formal constraints on syntactic rule application led to a considerable renaissance of syntax research in the seventies and the eighties. Constraint-based syntax started out as a derivational model, but more and more, representational aspects were integrated into the theory. Chomsky (1981), the Government and Binding Theory, was already mostly representational, and Chomsky (1986), the Barriers model, is one of the first generative syntax theories that claims to be only representational.

Arguments for representationalism often presuppose some background theory of what can be done by a derivational rule in syntax. If certain kinds of grammatical information need to be compared, and if this information is located in positions that cannot be linked by derivational rules of grammar, this may be seen as evidence for a representational account. Non-coreference facts concerning pronouns seem to be a case
in point: a rule assigning a disjoint-reference index would have to violate e.g. the coordinate structure constraint.

If the phrases in italics are intended to be coreferential, we obtain the indicated pattern of grammaticality

(2)  
*John likes him
   *John expects him to win
   John expects Mary to invite him
   John hopes that he will win the race
   *John saw pictures of him
   John saw Mary's pictures of him

We could try to express this by a rule that assigns a disjoint reference index, in addition to the referential index. If a phrase carries i in its disjoint reference index, this means that it must not corefer/be bound with/by any phrase carrying the referential index i. Such a rule of assigning a disjoint reference index would have to violate the Coordinate Structure Constraint, which seems to be respected by ALL syntactic processes: no rule can affect an element that is part of a coordinated structure.

(3)  
You saw Mary and someone else
   *who did you see [Mary and ___]
   *John saw him and Mary
   *Hans sah ihn und Maria

Therefore, it is better to express disjoint reference as a constraint on representations.

Likewise, it is hard to see how statements like, say, the Case Filter (each NP must be Case-marked) or the ECP (each trace must be properly governed) could be expressed derivationally.

On the other hand, the conspiracy problem familiar from phonology (several rules "conspire" so that all representations have a certain property P, but this property does not figure at all in the rules, so that P appear to hold by "chance") played little role in the decision for representational models, although it exists in syntax as well. Chomsky & Lasnik (1977) introduced the idea of syntactic filters like, say, the that-trace filter (an overt complementizer must not be followed by a subject gap,
   *who do you think that loves her?)

and their arguments were conspiracy based. In many other domains, obvious reference to surface constraints was avoided. Thus, the fairly widespread verb-second phenomenon is a clear example of a number of rules (verb movement, focus/topic/wh-movement, expletive insertion, stylistic movement) conspiring to produce a constant result, but attempts to account for verb-second phenomena tried to avoid reference to a surface constraint.

1.5. The rise of the Government- and Binding Model (Chomsky 1981) proved that it is possible to describe large areas of syntax in a purely representational way. There are at least two provisos that must be added, though. First, it seems that the subjacency phenomenon (no movement rule may cross more than n "bounding nodes", 1 ≤ n) shows
that movement dependencies between two positions that are far away from each other as in (4) have to be established in terms of a succession of shorter links. Likewise, the analysis of the proper interpretation of the reflexive pronoun himself in (5) seems to presuppose that which picture of himself has a connection to the object position of show. The two problems can be addressed easily in a derivational model (in which, say, the binding of himself is checked before it is moved into the matrix clause), but a purely representational description is possible as well if one assumes that the derivational history is itself coded in the representation: by traces in the original sense, or by multiple copies of the moved phrase (of which only one is pronounced), as in more recent approaches.

(4) who do you think _ that she hopes _ that you kiss _
(5) which picture of himself do you think she hopes that Bill will show _ to Jane

B2. Conflict Resolution

1.6 Discuss conflict resolution as rule ordering
1.7 Discuss the disguised conflict resolution component of the GB-theory
1.8 Discuss the elsewhere and minimality approach of conflict resolution in the GB-model
1.10 Discuss the conflict resolution in the MP

1.6. In a derivational account of language, conflicts arise whenever two rules that imply different outcomes are in principle applicable at a certain derivational stage. Such conflicts can be solved in terms of rule ordering. The interaction of Case marking and movement is an example: NPs get their Case in certain configurations (say, when they are governed by a verb), but long distance wh-movement may place the noun phrase out of the context in which it may receive Case. Thus, in German, Case must be assigned to a noun phrase BEFORE it is moved to sentence initial position in a question ...

(1) Wen denkst du dass sie liebt
   who-acc think you that she loves
because the noun phrase would be too far away from the verb after it has been preposed. This looks like intrinsic rule order in phonology: the NP would get no Case at all if rules were applied in reverse order, and the Case Filter would be violated. Note however that noun phrases sometimes (have to) pick up Case after wh-movement in French ...

(2) *je crois Jean être intelligent
   qui crois-tu être intelligent
or in Hungarian, in Quechua ...

In German, accusative Case assignment must come later than passivization, because otherwise, the wrong Case pattern would be generated

(3) dass der Mann angerufen wird
   that the-nom man called up is
   *dass den Mann angerufen wird

In Ukrainian, Hebrew, North Russian and other languages, accusative Case may, however, show up in passives. We thus need extrinsic rule ordering, too

German: passive > accusative
Grammatical differences are, thus, the result of different rule orderings.

In a representational model, there can of course be no rule ordering problem. Case marking can be carried out after wh-movement within trace theory, because one can assume that Case is assigned to the trace in (4), which transmits it to the preposed wh-phrase. The descriptive categories just mentioned then have to be described in different terms: if a phrase forms a chain with its traces, which position of the chain is relevant for, say, Case marking? There have been few attempts only to systematically account for the variation that we find in the interaction of Case and other grammatical features with movement, and that certainly goes beyond the "official" position that Case is always assigned to the root of a wh-chain.

(4) who did you see t

1.7. The Government and Binding framework had no conflict resolution component, because it assumed that a sentence could be grammatical only if it respected all principles of Universal Grammar. A sentence that could not fulfill a certain principle because it tried to respect all other constraints was thus simply ungrammatical.

GB had, however, a disguised conflict resolution component, viz. its architectural makeup. Consider e.g. the requirement that the arguments of a verb must be part of its maximal projection, that seems to stand in obvious conflict with the requirement (EPP) that the subject position Spec,IP (outside VPI) be filled overtly (or by a trace). The conflict is avoided by assuming that the grammatical description of a sentence must be distributed among various "levels of representation" (as in (6)), and that the EPP does hold not at all levels, in particular not at D-structure, which can therefore satisfy the argument requirement in a conflict-free fashion (5a,c). At S-structure, the EPP and the argument requirement must be respected, and they can be so because of traces: t "receives" takes over the argument role of the verb (the "thematic role"), and transmits it again to the moved phrase.

(5) a. e was invited he
b. he [was invited t]
c. e seems e to appear e to be invited Bill
d. Bill seems t to appear to be invited t

(6) D-Structure \(\rightarrow\) S-Structure \(\rightarrow\) LF
    \(\downarrow\) PF

GB-theory thus could avoid clashes between constraints because (a) constraints did not hold at all levels and because (b) it felt free to insert empty elements (traces, pro, PRO) whenever a constraint would not have been respected at the level otherwise.

1.8. In a very restricted domain, viz. the interpretation of pronouns, conflicts between means of expression and a conflict resolution strategy was admitted into GB. Consider (6).
(6) John prefers PRO going to the movies
*John prefers his going to the movies

John and his cannot be coreferent in the second example - unlike what we have in: John prefers his cat. Given the corefence option in the simpler case, Chomsky could not make the binding theory proper responsible for (6). Rather, he assumed the AVOID PRONOUN PRINCIPLE APP, which states that one should not use an overt pronoun when the same meaning could be expressed with a covert one. A similar effect can be found in (7).

(7) il veut PRO venir
   il veut qu’il vienne

The two il cannot be coreferential in the second structure because of APP relative to the blocking function of the infinitive.

A further instance of the APP can be seen in the well-known paradigm (8).

(8) John saw himself/*him in the mirror
    John laughs at himself/*him
    John expects himself/*him to win
    John expects Mary to love him/*herself
    John expects that he/*himself will win

The use a pronoun is blocked whenever the same meaning could be expressed by a reflexive. Chomsky’s own solution avoids reference to conflict resolution between two kinds of expressions of coreference/blocking: there are two independent principles:

Principle A:
A anaphor must be bound in the local domain S

Principle B:
A pronoun must NOT be bound in the local domain S

Denis Bouchard’s solution is a conflict based one: He just assumes PRINCIPLE A plus the AVOID PRONOUN PRINCIPLE extended to anaphors. This answers:

a) why the local domains for bound anaphors and disjoint pronouns are typically the same

b) why we get overlaps in the distribution, as in they saw their/each other’s wives.

1.9. In later versions of GB-theory, a minimality approach to government was formulated which may also be interpreted as a special way of conflict resolution in the syntax. Suppose Case is checked or assigned to a noun phrase under "government", a relation between a certain syntactic head and a noun phrase. Government was locally restricted by assuming that X could govern its m-command domain only (i.e., a head X can govern everything that is present in its maximal projection XP), and by assuming that certain syntactic configurations are "barriers" for external government (say, X cannot govern into a subject YP even if YP is part of XP). In addition, a minimality condition was imposed on government: X cannot govern α if there is a "closer" governor Y for α. In other words, if X and Y could in principle determine, say, the Case of α, only the closer element is eventually able to do so. The discussion of the "relativization" of minimality effects (Chomsky 1986, Rizzi 1990, Fanselow 1991) might thus be reconstructed as a
discussion of different strategies for conflict resolution.

1.10. In the late eighties, syntacticians became more and more aware of the need to give an account for the question of WHY rules apply. Given certain representational constraints, one has to wonder how and when they should be met in a wellformed structure. The first reflections in this domain, preparing the Minimalist program, were again concerned with order or rule application:

  According to Pesetsky 1990, rules repairing a constraint violation should be applied as early as possible to meet requirements
  According to Chomsky 1991, repair driven movement applies preferentially after S-structure. Language particular processes (such as do-insertion) are last resort process, that apply only if nothing else helps ...

The Minimalist Program (Chomsky 1995) assumes that phrases should not be moved (ECON). This constraint can be overriden by the need to check features (LEG). Movement is visible, however, only if the pertinent features are strong. Strength determines if LEG or ECON wins. Thus, the movement-triggering feature of TENSE is strong in English (the subject moves there visibly), while the movement-triggering feature of v (AGR-O) is weak - the object moves in the covert component of grammar only.

  [he [TENSE [v [likes Mary]]]]

The MP thus is a conflict resolution model, too. It differs from OT in (a) assuming an unmarked resolution of the major conflict, the one between inertia and the need to check certain features, and it (b) expresses language variation (the ranking of ECON and LEG) in terms of strength.

2. Outline of OT

A2.1 The components of OT
A2.2 Comparison between rule ordering and OT

B2.1 Introduction and discussion of key concepts of an OT-syntax
B2.2 The GEN component
B2.3 Inputs
B2.4 Constraints
B2.5 Faithfulness & Markedness

OT concentrates on surface forms. For this reason it has the following properties:
- No derivations
- No rules
- No ordering of the rules (obviously)
- Generation of surface forms (candidates, outputs)
- Evaluation of different candidates to deliver the optimal output

1. The components of OT

OT consists of two steps: Gen and Eval. First, Gen is a function generating candidates for every input. It works alike in all languages.
(1) Gen(\text{in}) = \{ \text{cand}_1, \text{cand}_2, \ldots \}

There are some limits to the generative power of Gen, in particular it cannot generate structures which are impossible in all languages, like the following:
- Prosodic structures which do not respect the prosodic hierarchy.
- Impossible feature configurations, like a laryngeal feature (voiced, aspirated…) associated with the Place of articulation node.

Eval is the function that evaluates the candidates. It differs from language to language, since it is dependent on the constraint hierarchy.

(2) Eval( \{ \text{cand}_1, \text{cand}_2, \ldots \} ) = \text{out}_{\text{real}}

Two different views of OT have been proposed.

Containment Theory
According to Prince & Smolensky (1993) and McCarthy & Prince (1993), every output contains its input. In (3) it is shown how every segment in the input is contained in the candidates. The constraint NOHIATUS prohibits hiatus, Fill (which was replaced by Dep in the correspondence theory) says that each segment of the output must correspond to a segment in the input. In other words, it militates against epenthesis. PARSE, which was replaced by PARSE, says that all segments of the input correspond to a segment of the output.

\begin{align*}
\text{(3) Containment} \\
\begin{array}{c|c|c|c}
\text{/le/+/ami/} & \text{NOHIATUS} & \text{DEP (FILL)} & \text{MAX (PARSE)} \\
\hline
\text{\textasciitilde a. l<e>ami} & & * & \\
\text{\textbf{b. l e ami}} & *! & & \\
\text{\textbf{c. l e ami}} & & *! & \\
\end{array}
\end{align*}

Correspondence Theory
According to McCarthy & Prince (1995), Gen is completely free. This is called Freedom of analysis (or Richness of the Base) and it posits that you may add as much structure to inputs as you want. The relationship between input and candidates is expressed with the help of subscripts added to the individual segments, as illustrated in (4). Constraints like DEP or MAX check the correspondence between the individual segments, and militate against epenthesis and deletion. IDENT(F) is responsible for the featural structure of the segments. Features cannot be changed.

\begin{align*}
\text{(4) Correspondence} \\
\begin{array}{c|c|c|c}
\text{/l\text{e}_2/+/a:m:i_is/} & \text{NOHIATUS} & \text{DEP} & \text{MAX} \\
\hline
\end{array}
\end{align*}
Eval evaluates all candidates against a unique input. We will see some extensions of Eval like the following later on:
- Output-Output Correspondence: candidates are evaluated against the output of another evaluation.
- Sympathy Theory: candidates can also be evaluated against other candidates.

**Inputs**

The standard view about inputs is that they are equivalent to the underlying forms of the Generative Phonology. We will follow this position for the moment…and adopt a more elaborate view of inputs later on.

**Candidates**

Generation of candidates is either constrained or unconstrained. If it is constrained, some hard constraints are needed, which posit which kind of structure may be generated and which may not. Hard constraints have an effect at the level of Gen, soft constraints at the level of Eval. Assuming the correspondence view, all kinds of candidates can be generated, also the absurd ones, like those with lots of epenthetic segments, for instance. They are eliminated by high-ranking constraints. The number of hard constraints is reduced, and some soft constraints are assumed to be unviolable in all languages.

**Constraints**

- constraints are ranked and violable
- violation is minimal
- well-formedness is comparative
- all rankings of rankable constraints are assumed to stand for possible grammars. This delivers the so-called factorial typology (see below).

**Universal grammar**

Universal grammar in OT consists of the set of all constraints. However, the definition of this set is an extremely difficult task. In phonology we have some clues of how to do that: constraints should be phonetically grounded, and in agreement with the results of typological studies.

**Grammar of the single languages**

Grammars of individual languages consist of orderings of the constraints. The grammar of each language (dialects, sociolects, idiolects) is defined by just one non-permutable hierarchy of the constraints.

**Factorial typology**

Permutations of the constraints result ideally in possible (but not necessarily existing) grammars. Not all constraints are permutable:

Some constraints have a universal, fixed ordering (markedness hierarchies)
Faithfulness and markedness constraints
Faithfulness constraints require identity between input and output.
In Correspondence Theory (from now on we will use only this version of OT), following families of constraints are active:

(5) Examples of Faithfulness constraints
MAX: No deletion of segments or features.
DEP: No epenthesis of segments or features.
IDENT(F): No change in the featural make-up
CONTIGUITY: Contiguous segments in the input are contiguous in the output.
HEAD-MATCH: A head in the input is a head in the output.

Markedness constraints require unmarkedness of the output.
Markedness constraints depend heavily on results of the markedness theory and typology.
Even though it is not always clear how to decide what is unmarked. There are unclear and even contradictory cases (more on markedness later).

(6) Examples of markedness constraints
ONSET: Syllables have onsets.
NOCODA: Syllables don’t have codas
NOCOMPLEXONSET: Onsets are not complex.
NOCOMPLEXCODA: Codas are not complex.
NOVOICEDOBSTRUENT: Obstruents are voiceless.

In OT the universal tendencies are accounted for by the interaction of universal violable constraints. The constraints define the grammar.

2. Comparison between rule ordering and OT

In order to illustrate the difference between the traditional generative theory (transformationalism) and OT, it is useful to concentrate on a concrete example, like hiatus avoidance, the avoidance of two syllabic peaks in a row. All languages avoid hiatus, but they differ in at least two dimensions in the way they do that. First, in the quantity of hiatus avoidance. Some languages do not tolerate hiatus at all, or just very few (German for instance), some others tolerate a great number of them (Hawaiian, French). Second, languages also differ in the way they ‘repair’ their unallowed hiatuses. An output structure which avoids hiatus can be obtained in different ways. Frequent strategies are vowel deletion, consonant epenthesis, and glide insertion.

(7) Hiatus avoidance by vowel deletion in French
   *le éléphant _ l’éléphant ‘the elephant
   *ce est quelque un _ c’est quelqu’un ‘it is somebody’

(8) Hiatus avoidance by vowel deletion in Yoruba
   ra ɔgèdè  →  rɔgèdè ‘to buy bananas’
   lo asa  →  ʃaʃa  ‘to use a towel’

(9) Hiatus avoidance by consonant epenthesis in German:
   Chaot → Chaʔot ‘chaotic person’, Ruin → Ruʔin ‘ruine’, Beamte → Beʔamte ‘civil servant’

(10) Hiatus avoidance by consonant epenthesis in Axininca Campa
In a traditional derivational framework, rules of the following kind are needed. These rules present all problems that were identified before, like arbitrariness, no visible aim, duplication, conspiracy and the like. In case these rules interact with further processes, like compensatory lengthening in Luganda, for instance, the ordering is ‘intrinsically’ imposed.

(11) Rules of hiatus avoidance
a. Vowel deletion: \( V \rightarrow \emptyset / \_ V \)
b. Consonant epenthesis: \( \emptyset \rightarrow C / V \_ V \)

In OT, hiatus avoidance is expressed completely differently. Only one markedness constraint is needed (ONSET or NOHIATUS). How languages achieve resolution of hiatus depends on the ordering of several faithfulness constraints and just one (ideally) markedness constraint. The following tableaux, a factorial typology, give just an overview of the technique.

(12) Hiatus resolution by vowel deletion (French, Yoruba)

\[
\begin{array}{|c|c|c|}
\hline
/le+/éléphant/ & NOHIATUS & DEP \\
\hline
\text{a. l’éléphant} & \star & \\
\text{b. le éléphant} & \star! & \\
\text{c. le téléphant} & \star! & \\
\hline
\end{array}
\]

(13) Hiatus resolution by consonant epenthesis (German, Axininca Campa)

\[
\begin{array}{|c|c|c|c|}
\hline
/Beamte/ ‘civil servant’ & NOHIATUS & MAX & DEP \\
\hline
\text{a. Be?amte} & \star & \star & \\
\text{b. Beamte} & \star! & \star! & \\
\text{c. Bamte} & \star! & \star! & \\
\hline
\end{array}
\]

(14) No resolution (Maori)

\[
\begin{array}{|c|c|c|}
\hline
/puea/ ‘be avenged’ & DEP & MAX & NOHIATUS \\
\hline
\text{a. puea} & \star & \star & \star! \\
\text{b. puteta} & \star! & \star & \star! \\
\text{c. pea} & \star & \star! & \star! \\
\hline
\end{array}
\]

**OT in Syntax**

2.1. OT Syntax is, in a sense, the most conservative successor of the Government & Binding Model, because it concentrates on constraints on surface structures, and specific derivational rules play little (if any) role at all. Unlike the MP, OTS does not deny the existence of an autonomous syntactic component.

As is phonology, surface forms generated by a GEN component are subjected to an EVAL component. The optimal candidate is the only grammatical structure.

2.2. As in phonology, there is a GEN (generator) component that generates the set of all possible candidate relative to a given input. Ever since Chomsky 1972/Chomsky 1981, generative syntacticians have believed the generative component of grammar to be trivial: there is a simple structure building component, and a simple movement transformation. Up to now, there is no generally accepted view of what GEN should do in OTS, but all options share the general generative background just mentioned.

If we think of a simplified version of what the Minimalist program suggests, then GEN
might consist of two processes:

- **MERGE** items from a given set of lexical items
- **MOVE** items from one slot to another

Grimshaw 1997 proposes that we distinguish the operations

- **MERGE** items from a lexical array & **MOVE**, from
  the addition of functional material/ functional heads not present in the input

In such a model, the presence/absence of functional projections is directly subject to the evaluation procedure. What else could be missing? Perhaps the operation "Delete!":

(1b-d) may have been derived from (1a) by deletion.

(1)   a. a man who        that we like
       b. a man who        we like
       c. a man          that we like
       d. a man          we like

If we work with deletion would be then imply the postulation of a **CORRESPONDENCE APPROACH** to syntax. What would be a **CONTAINMENT** approach? One that expects "phonology" to do the job of preventing certain items from being pronounced.

An even less trivial version on GEN might assume that **MERGE** respects Predicate-Argument Structure: A can merge with B only if A "checks a feature" of B. This rules out

(2)   *He is sleeping a cat

but maybe the sentence DOES make sense, and should therefore be generatable! An even less trivial version could say that **MERGE** respects X-bar-Theory ... but we might wish to allow for headless phrases! Woolford goes much further: in her treatment of Case, she assumes that GEN is more or less a GB style syntax. Pesetsky claims that **MOVE** respects constraints on movement. This would imply why (3) is not grammatical.

(3)   *who did you weep because she left

   *which city did you find Bill's letter from in the waste paper basket?

2.3. The disagreement on GEN is already a bit frustrating ... but the nature of inputs may be a much less understood issue in OTS. A first option would be

- Input = a sets of lexical entries

But this solution runs the risk of

(4)   who did Mary kiss

being blocked by

(5)   who kisses Mary

because the movement of who is shorter in (4) than in (5). Optimization is not blind with respect to meaning!

The second option is (suggested by Grimshaw)

- Input = Lexical items + the correlation of predicates with arguments
  
  (**PREDICATE ARGUMENT STRUCTURES**)

This solves the problem with (4) and (5). But perhaps, more complex theories are
necessary -- we might need an indication of the intended scope of quantifiers (e.g. suggested by Legendre & Smolensky), so that we might, at the end, come to the conclusion that the input consists of kernel sentences in good old Harris sense.

2.4. Assume we have made some decision concerning GEN and inputs. Then we need to make up our minds what kind of entities the candidates emanating from GEN should be. Of course, the most straightforward and most traditional version of OT syntax will have surface structure representations as candidates. And these will more or less identical to the S-structures of the GB-model. Given that monostratal versions of GB have been proposed (Koster 1988), we can see again why OTS is an elaboration of the GB model.

But again, more complex options for candidates are conceivable:
- D-S, S-S, LF triples
- Derivations ...

2.5. We have now generated our candidates, and they must be evaluated on the basis of constraints. In OT, constraints are universal restrictions on the well-formedness of structures. Again, OTS is like GB (and quite unlike the MP). Constraints affect surface structures only, at least in conservative versions of OT. This is different from canonical GB, but it resembles monostratal GB-approaches. Constraints are violable, if that is in the interest of other principles. This has never been dreamt of in GB-theory (but it assumed principles were parametrized, and worked with various tricks like empty pleonastics ...)

In very early OTS papers, the principles looked much like what they would have looked liked in GB. Nowadays, this seems no longer accepted (and Grimshaw explicitly argues against it) - constraints should be simple, the complexity of natural language lies in the interaction of the constraints, not in the constraints itself. Note this is the idea that GB started out with (although it ended at quite a different point).

The rest is as in phonology. Constraint violation is minimal in the sense that a constraint may be violated by a grammatical structure only if that is the only means of avoiding the violation of principle with a higher rank. Universal Grammar is the set of all constraints. There are no parameters left! Two more differences to GB are notable. First, in OTS, the grammar determines the nature and set of possible lexical entries, not vice versa. X is a possible lexical entry in a language if the features of X can surface, given the ranking of constraints. Language variation is NOT the result of language L having a certain kind of lexical entry, and language L* lacking it.

In phonology, the groundedness of constraints constitutes an interesting perspective on language. In syntax, GB denied that its principles might be grounded in other domains, and given its dominance in the field in the eighties and early nineties, interesting proposals concerning the grounding of syntactic principles such as the one proposed by Erteshik-Shir were not pursued. GB being abandoned, syntacticians seem to have become more open-minded in this respect, and Legendre seems to particularly argue for
a high degree of interaction between syntax and its neighbors. It is conceivable that some of the principles may be semantically grounded (as proposed by Szabolcsi), some may be grounded in pragmatics (Erteshik-Shir), others still may be grounded in perception (Hawkins).

2.6. Grammars of individual languages consist of hierarchical orderings of the constraints. Recall there is no parametrization, and lexical differences are a consequence of grammars.

Given a hierarchy $H$ of constraints, and a set $C$ of candidates generated from the input $I$, $c$ in $C$ is grammatical iff there is no $d$ in $C$ such that $d$ violates the highest constraint on which $c$ and $d$ differ less often than $c$. Conflict resolution is thus lexicographic.

2.7. As in Phonology, we may distinguish Faithfulness Constraints (Inputs and outputs should be identical) from Markedness Constraints.

These are potential Faithfulness Constraints:
- **Full Interpretation (FI)**
  - Don't insert expletives
- **Parse Scope!**
  - The syntactic position of a quantifier must correspond to its scope!

These are potential Markedness Constraints
- **EPP**: All clauses have a subject
- **HdLeft**: Heads are rightmost in functional projections

Whether a given constraint is an instance of a faithfulness requirement is of course a function of what counts as an input ... so our answers cannot be definitive!

3. Examples of markedness hierarchies: conflicts between markedness and faithfulness

**A3.1** Markedness in Phonology

**A3.2** Markedness and faithfulness in OT phonology

**B3.1** Grimshaw’s account of do-support in English: the interaction of faithfulness and markedness

**B3.2** Pesetsky’s deletion theory

**B.3.3** A reflection on markedness hierarchies in the syntax

1. Markedness in phonology

Markedness is a rather vague term which has received several different definitions like the following:
A marked segment is less natural, more complex, less common, not basic, not stable, it appears later in language acquisition, in fewer languages, in less positions, it is subject to neutralization, harder to articulate, perceptually more salient, etc. (see also Rice 2000)

- Nasal vowels are more marked than oral ones (all languages that have nasal vowels also have oral ones)
- Voiced obstruents are more marked than voiceless ones (obstruents have a tendency to be voiceless).
- Voiced sonorants are more marked than voiceless ones (sonorants have a tendency to be voiced).
- Open syllables are less marked than closed ones.
- Sonority hierarchy (maybe this is something else than markedness)

Some diagnostics for markedness:
- Implication: the marked feature or segment implies the unmarked one.
- Frequency: unmarked features are more frequent than marked ones.
- Processes: the unmarked features emerge only under special circumstances (otherwise emergence of the unmarked), the unmarked feature neutralizes more easily.

2. Markedness and faithfulness in OT
Since markedness requires simple structures and faithfulness requires that inputs always be realized the way they come, regardless of their complexity, the two kinds of constraints often impose conflicting demands on outputs.
Depending on which win, structures are faithful or unmarked. The constraints requiring more unmarked structures and more faithful ones are conflicting. They can be differently ranked depending on the languages, leading to factorial typologies

2.1 Simple factorial typology 1
Syllables without onsets and with codas are tolerated in many languages, but only as the marked kind of syllables. French, German and English tolerate syllables without onsets and with codas, but only if there is no way out, that is if no other syllabification is better.

- In German, a placeless consonant (glottal stop) is inserted before a stressed vowel if there is no input consonant:

(1) Onsets in German
   a. Cha²dot, Ru²tin etc. vs. Muséum, Koréa.
   b. Monomorphemic VCV are always syllabified as V.CV (Ute, Era, Asa, …) or VCamhysylV (Ebbe, Asse …), but never as VC.V

Other languages do not have syllables without onsets, or with codas, and use strategies like epenthesis to enhance ‘bad’ syllables, for instance in loanwords, like in Japanese: kurimasu for Christmas and arubatu for Arbeit, a German word for ‘work’.
How does OT account for these observations?

Prince & Smolensky (1993) on syllabification:
(2) Markedness Constraints: ONSET, NOCODA, NOCOMPLEX
(3) Faithfulness Constraints: MAX, DEP

Through permutation of these constraints, different kinds of syllables are obtained.
(4) Prince & Smolensky’s solution

<table>
<thead>
<tr>
<th>/V/</th>
<th>DEP (FILL)</th>
<th>MAX (PARSE)</th>
<th>ONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>ø =&gt; V</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ø</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>CV</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/V/</th>
<th>DEP (FILL)</th>
<th>ONSET</th>
<th>MAX (PARSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>ø =&gt; ø</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>CV</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/V/</th>
<th>ONSET</th>
<th>MAX (PARSE)</th>
<th>DEP (FILL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ø</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>ø =&gt; CV</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

With 3 constraints, 6 permutations are possible, but only 3 deliver different results. Given 4 constraints: ONSET, NOCODA, MAX, and DEP, what are the possible grammars? 24 (4!) possible permutations of them. But not all 24 are interesting, because some constraints never conflict with each other, like ONSET and NOCODA, for instance (they both impose V.CV.)

Another case is the following: if both MAX and DEP dominate ONSET and NOCODA, the ordering of the faithfulness constraints is indifferent: MAX >> DEP >> {ONSET, NOCODA} and DEP >> MAX >> {ONSET, NOCODA} are equivalent. MAX conflicts with DEP only if ONSET or NOCODA dominates one of them.

Notice that, given the set of constraints, no language can exist that has only syllables with codas but without onsets. This has nothing to do with the orderings, but with the kinds of constraints used.

Given the fact that all languages have Cs and Vs, there is also no language without CV
syllables. And VCV will always be syllabified as V.CV, as desired.

2.1 Simple factorial typology 2: Syllable peaks
The second factorial typology considers syllable peaks. French has only vowels as syllable peaks. In German and English vowels and sonorants can also be syllable peaks.

(5) Syllable peaks in French: only vowels
   ocre ‘ochre’, siècle ‘century’

(6) Syllable peaks in German and English: vowels and sonorants
   G: Himmel ‘sky’, eben ‘even’, nieder ‘low’
   E: eagle, meter, button

The most extreme case is exemplified by Imdlawn Tashlhiyt Berber which tolerates all kinds of segments as syllable peaks, also obstruents.

(7) Imdlawn Tashlhiyt Berber (Dell & Elmedlaoui 1985, P&S)
   Syllable structure       Morphological structure
   Voiceless stop .rA.tK.tL.  ra-t-kti ‘she will remember’
   Voiced stop .bD.dL.       bddl ‘exchange’
   Voiceless fricative .tF.tKt. t-ftk-t ‘you suffered a strain’
   .tX.zNt.                  t-xzn-t ‘you stored’
   Voiced fricative .txZ.nAkk.". t-xzn#nakk"‘she even stockpiled’
   Nasal .tzMt.              t-zmt ‘it (f.) is stifling’
   Liquid .tR.gLt.           t-rgl-t ‘you locked’
   High vowel .tL.dL.        i-ldi ‘he pulled’
   .rAt.lUlt.               ra-t-lul-t ‘you will be born’
   Low vowel .tR.bA.         t-rba ‘she carried on her back’

The domain for syllabification in Berber is the Phonological Phrase. All syllables have an onset, except for the phrase initial ones, which can be onsetless. This is accounted for by an alignment constraint which is not important here.

Prince & Smolensky propose the following margin and peak hierarchies, which are universal rankings on the sonority of the segments: [a] is universally the best peak, and voiceless stops are universally the best margins.

(8) Universal margin and peak hierarchies (Prince & Smolensky 1993):
   a. *M/a >> *M/i >> *M/l >>… *M/t  ([a] is not allowed to be a margin, …)
   b. *P/t >> … >> *P/l >> *P/i >> *P/a  ([p] is not allowed to be a syllable peak, …)

Languages will choose a place in these hierarchies, from which on no segment can be a peak or a margin. An example is given in (9), which is the place chosen by Berber: Every segment can be a margin, except for [a]. In other words, the maximum sonority of possible onsets cannot be higher than the sonority level |i|. The ranking in (10) expresses the fact that the segments which are more sonorous than a nasal cannot be margins. French, German and English chose different places in the hierarchies for MAX, but the markedness hierarchy does not change.

(9) Untenable margins:
   *M/a >> MAX >> *M/i >> *M/l >>… *M/t
(10) Untenable peaks:

\[*P/t >> *P/n >> \text{MAX} >> *P/l >> *P/i >> *P/a\]

Other active constraints require a correct syllable structure, like ONSET, NOCODA, NOCOMPLEX, MAX, DEP.

An input /ta/ is always syllabified as [.ta.] (in all languages, especially in Berber).
An input /iat/ is syllabified as [.iAt] in Berber, because [i] can be a margin (this candidate is better than an onsetless syllable, or an epenthetic onset).

### B3. Faithfulness and Markedness in Syntax

3.1 As a first example, we will briefly present Grimshaw's account of English basic clause structure. Our goal is the explanation of the basic patterns of English sentence structure, with a particular emphasis on the issue of the application of do-support and the addition of functional projections, two violations of faithfulness.

Consider first simple assertions. From an input such as 

\{John, left\}

GEN will generate the following candidates (and others)

- \[VP \text{ John left}\]
- \[IP \text{ Infl } [VP \text{ John left}]\]
- \[IP \text{ John Infl } [VP \text{ t left}]\]
- \[IP \text{ John left } [VP \text{ t t}]\]
- \[IP \text{ John did } [VP \text{ t leave}]\]
- \[IP \text{ Did } [VP \text{ John leave}]\]
- \[IP \text{ Left } [John \text{ t}]\]

These candidates should be evaluated by the following constraints:

**Obligatory Heads (OblHd)**

A projection must have a visible head.

Recall that the input consists of PAS only. The addition of functional layers above VP is also a faithfulness violation.

**Full Interpretation**: Do not insert meaningless elements/PAS must be respected.

FI is also a Faithfulness principle

**Econ/Stay!**: Trace is ungrammatical/Do not move!

Econ/Stay may be considered a faithfulness constraint, too, because traces are not in the input. On the other hand, it may also be profitable to view it as a markedness constraint (for derivations).
As the tableau indicates, a sentence cannot (and does not) project beyond VP when no markedness constraints have a chance of being applicable.

The situation is different in a simple negative clause, i.e., when the input is something like \(\text{John, not, left}\). Some of the candidates we generate are:

\[
\begin{array}{l}
\text{not [John left]} \\
\text{[Neg] John [not [t left]]} \\
\text{[IP John left [not [t t]]]} \\
\text{[IP John Infl [not [t left]]]} \\
\text{[IP John did [not [t leave]]]} \\
\text{[IP Did [not [John leave]]]}
\end{array}
\]

The faithfulness principles considered so far alone would not yield the correct result. We need to take into account a few markedness principles that govern English syntax:

**Subject (vulgo: EPP)**
The highest A-specifier of a sentence must be filled!

We do not make reference to Spec,IP, because the subject may in fact end up in a lower (or higher) position. SUBJ implies that the subject is the highest element in the clause --but in the end not every structural position will do. Grimshaw therefore proposes CASE, a second markedness principle, which restricts the structural options for subjects to Spec,VP and Spec,IP (there is some hacking in this, isn't there?)

**Case (vulgo: Case Filter)**
The highest position of a chain must have Case!

Do-insertion is then forced by the assumption (with a high rank assumed for English only) that lexical movement is a particularly bad form of movement.

**LexEcon:*Lexical Trace!**

With the ranking indicated in the following tableau, we arrive at the correct result:

<table>
<thead>
<tr>
<th></th>
<th>LE</th>
<th>Case</th>
<th>OblHd</th>
<th>subj</th>
<th>FI</th>
<th>Econ</th>
</tr>
</thead>
<tbody>
<tr>
<td>[not [John left]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Neg] John [not [t left]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[IP John left [not [t t]]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[IP John Infl [not [t left]]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[IP John did [not [t leave]]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[IP Did [not [John leave]]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.2. We can now attempt a factorial typology:

If LE > FI: No movement of the verb, expletive insertion
  John did not kiss Mary

If FI > LE No expletives but verb movement
  Jean n’embrasse pas Marie

If Econ > OblHd Neither expletives nor movement
This looks like Danish (embedded) clauses
  ---- COMP Jan inte har drukket kaffee
  lit: John not has drunk coffee

More systematically, we observe:

<table>
<thead>
<tr>
<th></th>
<th>LEFT &gt; OblHd &gt; FI</th>
<th>LE &gt; FI &gt; OblHd</th>
<th>FI &gt; LE &gt; OblHd</th>
<th>FI &gt; OblHd &gt; LE</th>
<th>OblHd &gt; LE &gt; FI</th>
<th>OblHd &gt; FI &gt; LE</th>
</tr>
</thead>
<tbody>
<tr>
<td>English type system</td>
<td>English type system</td>
<td>Danish type system</td>
<td>Danish type system</td>
<td>French type system</td>
<td>English type system</td>
<td>French type system</td>
</tr>
</tbody>
</table>

3.3. FI and Econ disfavor the insertion of elements which are not in the input. There should also be principles that require that input material be realized. These are principles of the Parse-family. They have e.g. been studied by Pesetsky,

Before we can consider deletion, we should identify a further markedness principle:

**LECP** Complement clauses are introduced by a complementizer

*Je crois que Pierre a faim*

*Ich denke dass Peter Hunger hat*

I think that Peter is hungry

This is a special case of

Left Edge of Functional (Sub-)Trees (LEFT)

Projections of functional categories are introduced by their heads

If Left > FI, we get the right results for complement clauses and simple relative clauses:

<table>
<thead>
<tr>
<th></th>
<th>LEFT</th>
<th>FI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Je crois que Pierre a faim</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>*Je crois Pierre a faim</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LEFT</th>
<th>FI</th>
</tr>
</thead>
<tbody>
<tr>
<td>L’homme que je connais</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>*L’homme qui je connais</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>*L’homme qui que je connais</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
BUT: We must not delete the relative pronoun when it is part of a PP - then, the complementizer has to disappear:

l'homme avec qui j'ai dansé

*l'homme que j'ai dansé

The principle REC (faithfulness) disfavors deletion.

REC (Pesetsky):

A unit with semantic content must be pronounced unless it has a sufficiently local antecedent

If Rec > Left > Tel, and if the preposition is not recoverable, we again get the correct result:

<table>
<thead>
<tr>
<th></th>
<th>REC</th>
<th>LEFT</th>
<th>FI</th>
</tr>
</thead>
<tbody>
<tr>
<td>L'homme avec qui j'ai dansé</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*L'homme avec qui que j'ai dansé</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>*L'homme que j'ai dansé</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*L'homme j'ai dansé</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What is recoverable and what not seems to be dealt with differently in different languages. Further faithfulness principles would be "Parse (wh)" (wh-features of the input must be realized in the output), ParseScope, Parse (Case) (if an exceptional Case is mentioned in the lexical entry). Further markedness principles are:

Scope: If X c-commands Y, then X takes scope over Y
Gov: Traces must be governed
Agree! The verb agrees with the nominative DP

3.4. Markedness Hierarchies have figured prominently in OT Phonology since its inception. In syntax, markedness is a concept employed in the typological tradition, while GB and other OT precursors tried to avoid reference to such a concept. Yet, a number of observation must be accounted for, and they seem to involve the same dimensional values over and over again.

The hierarchy

nominative < accusative < dative

first exerts its influence in the constitution of Case patterns in German and other languages: the regular Case frames always work with the least marked Case combination (nom, nom acc, nom acc dat). Across languages, we can observe that if a verb can agree with an NP with case c, it also agrees with the less marked cases, but not necessarily vice versa (German: nom agreement, Hungarian: nom and acc agreement, Basque, nom, acc and dat agreement). In some languages, only nominatives can be antecedents of anaphors (French?), in others nominative and accusative NPs are licensed antecedents (German), in yet other languages, NPs with all cases can bind a reflexive. The Keenan-Comrie Hierarchy implies that mobility (say, for question or
relative clause formation) also follow the markedness scale: in some languages, only nominative NPs (subjects) may be questioned (say, Malay languages), other restrict movement to subjects and direct objects (certain Hebrew relative clause like structures), etc.

4. The third kind of constraints: Alignment

A Phonology
A4.1 Separation of domains as an effect of alignment: Syllabification in German
A4.2 Linear organization of constituents as an effect of alignment: Place of affixes
A4.3 Falling together of constituents as an effect of alignment: position of stress

B Syntax
B4.1 A simple illustration of alignment effects on word order
B4.2 A simple illustration of other structural effects of alignment
B4.3 A discussion of fin-second effects in the spirit of Geraldine Legendre

Alignment constraints have been first proposed by McCarthy & Prince (1993b). They differ both from markedness and from faithfulness, and at the same time have similarities with both. They require that edges of morphological, syntactic or phonological (prosodic) constituents fall together (which makes them like faithfulness constraints) or that constituents appear in certain specified positions (which makes them like markedness constraints). Three kinds of application can be distinguished:

1. Separation of domains: for example „crisp“ syllabification
2. Concatenation of constituents: prefixation, suffixation and order of words
3. Falling together of different kinds of entities. Position of heads in phonology and in syntax.

Before these three kinds of application can be considered in some details, alignment must be defined.

(1) Generalized Alignment (adapted from McCarthy & Prince 1993b)
Align (Cat₁, Cat₂, (Edge))
Cat₁ and Cat₂ are prosodic and grammatical categories.
Edge is Left or Right
In order for (1) to be true, Cat₁ must fall together with Cat₂ or one edge of Cat₁ must fall together with one edge of Cat₂.

The two categories mentioned in the definition are variable and taken from a universal inventory of prosodic, morphological or syntactic constituents (it could be the case that future research shows the need to enlarge the class of categories predicted to take part in alignment relations). The order of the arguments in an alignment constraint is not indifferent. There is a universal quantification on the first argument and an existential quantification on the second:

For all Cat₁ (∀Cat₁) there is a Cat₂ (∃Cat₂) such that,…

We will see examples of the three effects of alignment in turn. First, consider a case in which separation of domains is expressed by means of Align constraints.

1. Separation of domains as an effect of alignment: Syllabification in German
SUFFIXATION IN GERMAN IMPLIES SYLLABLE BOUNDARIES BETWEEN STEM AND SUFFIX, EXCEPT IN CASE THE STEM ENDS IN A CONSONANT AND THE SUFFIX BEGINS WITH A VOWEL, IN WHICH CASE THESE TWO SEGMENTS ARE SYLLABIFIED TOGETHER.
(2) Suffixation in German
C+C: faul/Faul-heit fau.l.hajt ‘lazy-laziness’
V+V: Ruhe/ruh-ig ru:.iç ‘quietness-quiet’
V+C: froh/fröh-lich frœ.liç ‘joyful-joyful’
BUT:
C+V: Kind/kind-isch kœ.n.d∫ ‘child-childish’

To account for this in OT, two constraints are needed, formulated in (3).

(3) a. ONSET: Syllables have onsets
b. ALIGN(stem, syllable, R): the right edge of a stem falls together with the right edge of a syllable (for all right edges of stems there is a right edge of syllable, so that both edges fall together).

In suffixation, the need to satisfy the unmarked syllable structure is higher than the need to separate morphemes. In the case of Faulheit, ruh-ig and fröhlich, ONSET and ALIGN(stem, syllable, R) make the same prediction. In ruhig, the second syllable has no onset, but a syllable boundary separates the two morphemes. In Faulheit and fröhlich, morpheme structure and syllabification fall together. In kindisch (and also sonnig ‘sunny’, Ladung ‘cargo’, lachen ‘to laugh’ and so on) there is a conflict between ONSET and ALIGN(stem, syllable, R). ONSET is ranked higher than ALIGN.

A consequence of t ALIGN(stem, syllable, R) is that two adjacent vowels that could fuse together in a diphthong refrain from doing so because of morpheme edges.

(4) No coalescence across morpheme boundaries
ruh-ig [u.i], prosa-isch [a.i] ‘prosaic’ (vs. Pfui [u_i] ‘yuck’ and Fleisch [qi] ‘meat’)

Turning now to prefixation and compounding, the need for ‘crisp’ syllabification (term from Ito & Mester 1994) is even higher. Even when the stem ends with a consonant and the suffix begins with a vowel there is no resyllabification across the morpheme boundary.

(5) Prefixation in German
verärgern [vg.g.gn] * [vø.rg.gn] ‘to annoy’
unartig [on.agtiç] * [u.nagtïç] ‘naughty’

(6) Compounding in German
Stockente [tœ.kœntœ] * [tœ.kœntœ] ‘mallard’
Seeadler [ze.aedle] ‘sea eagle’

A second Alignment constraint is needed which aligns the stem with a syllable but from the left side.

This new constraint ALIGN(stem, syllable, L) is higher ranking than ALIGN-R.
ALIGN-L >> ONSET >> ALIGN-R. Both prefixes and compounds have a clear syllable boundary (though some lexicalized elements seem to be able to trigger resyllabification in fast speech, as in a word like unerhört ‘unheard of’ which can be pronounced as u.ner.hört).

At the beginning of a foot, an onset is obligatory in German. If the input does not provide any, a glottal stop is inserted that functions as the onset of the syllable falling together with the left edge of the foot.

In a word like Krähe ‘crow’ no glottal stop is inserted before an onsetless schwa syllable. This syllable is in the foot weak position ([h] is not pronounced before a schwa).

(8) ALIGN(C,Foot, L) >> DEP

<table>
<thead>
<tr>
<th>Align (C, Ft, L)</th>
<th>DEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rightarrow (\tilde{E}u\ell)e_{F} )</td>
<td>*!</td>
</tr>
<tr>
<td>((E\ell)e_{F})</td>
<td>*</td>
</tr>
<tr>
<td>( \rightarrow (K\ddot{r}\ddot{a}h)e_{F} )</td>
<td></td>
</tr>
<tr>
<td>((K\ddot{r}\ddot{a}?e)_{F})</td>
<td>*!</td>
</tr>
</tbody>
</table>

In German, the domain of syllabification is the Prosodic Word. Each stem is a Prosodic Word (we can replace stem with PW above).

In French the Phonological Phrase is the domain of syllabification:

*Les oiseaux se sont en*volés cet après-midi.* ‘The birds flew away this afternoon.’*

Alignment effects are usually not felt below the PhPh in French, but recall the anti-coalescence effect (*biannuel, antialcoolique,…, transatlantique, anti-alcoolique ([i.a] vs. piano [ja]). Different processes involving resyllabification can have different domains of applications.

2. Linear organization of constituents as an effect of alignment: Place of affixes

In German, English, as well as in all languages without infixation, the alignment constraint responsible for the position of the prefixes and the suffixes are always higher-ranking than any prosodic constraint that would cause a shift of an affix.

*Instrument-al, instrument-less, im-polite, un-fair*

*Ge-länd-e, kind-isch, un-art-ig*

In Tagalog (McCarthy & Prince 1993a) the infix *-um* is located after the onset of the first syllable, if there is one.

<table>
<thead>
<tr>
<th>Root</th>
<th>um + Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>aral</td>
<td>um-aral</td>
</tr>
<tr>
<td>sulat</td>
<td>s-um-ulat</td>
</tr>
<tr>
<td>gradwet</td>
<td>gr-um-adwet 'to teach'</td>
</tr>
<tr>
<td>*um-sulat</td>
<td></td>
</tr>
<tr>
<td>*um-gradwet 'to graduate'</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>sulat + um</th>
<th>NOCODA</th>
<th>ALIGN(Aff, PW, L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>um.su.lat</td>
<td>**!</td>
<td></td>
</tr>
<tr>
<td>su.mu.lat</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>su.um.lat</td>
<td>**!</td>
<td>**</td>
</tr>
<tr>
<td>su.lu.mat</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>aral + um</th>
<th>NOCODA</th>
<th>ALIGN(Aff, PW, L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>u.ma.ral</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>a.um.ral</td>
<td>**!</td>
<td></td>
</tr>
<tr>
<td>a.ru.mal</td>
<td>*</td>
<td>**!</td>
</tr>
</tbody>
</table>
3. Falling together of constituents as an effect of alignment: position of stress

Stress is often peripheral: final, penultimate or initial, which speaks for an analysis in terms of alignment.

Stress is grouping of constituents, at the lower level, syllables are grouped into feet. Feet are trochaic (left-headed) or iambic (right-headed). This alignment is trivial, since feet consist of two syllables.

More interesting is what happens at higher levels, in a prosodic word or a phrase. There, too, stress can be interpreted as standing for groupings of constituents, and there, too, it is peripheral. But, since feet are the relevant constituents, when feet are trochaic and stress final, stress is penultimate.

Two examples:

**English stress** (McCarthy & Prince 1993b)

ALIGN(PrWd, Ft, L): all Prosodic Words start with a left-aligned foot

ALIGN(Ft, PrWd, R): all feet are right-aligned with the right edge of the word

(Tâta)ma(góuchee) *Ta(tâma)(góuchee)

Main stress is determined independently.

**German stress**

ALIGN(PrWd, Ft, R): all Prosodic Words end with a right-aligned foot.

ALIGN(PrWd, Ft, L): all Prosodic Words start with a left-aligned foot.

These two constraints give a stress pattern with exactly two stresses, one initial and one final. The final one is the main stress.

<table>
<thead>
<tr>
<th>/kveþbis/</th>
<th>FOOT-FORM(TRO)</th>
<th>FT-BIN</th>
<th>ALIGN-FOOT - RIGHT</th>
<th>ALIGN-FOOT-LEFT</th>
<th>PARSE-SYLLABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( x . )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-&gt; Kürbis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ( x )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kür bis</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>c. ( x )</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Kür bis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ( . x )</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Kürbis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### B4. Alignment in Syntax

4.1. The primary function of alignment constraint in syntax seems to be the establishment of linear order facts. If inputs are nothing but PAS, then no serialization is specified there, and all linearization must be primarily effected by the evaluation procedure.

A simple example comes from the serialization of heads in phrases. Here, we may assume the following principles:

- **Head Left!** (a generalization of LEFT): The head X is leftmost in XP
- **Head Right!** : The head X is rightmost in XP
- **SpecFirst!** : The specifier of XP is the leftmost element in XP.

With these three principles, we can derive the basic word order typology of the majority of the world’s languages:

<table>
<thead>
<tr>
<th>/apoteoza/</th>
<th>NON HEAD (a)</th>
<th>NO CLASH</th>
<th>FOOT-FORM (TRO)</th>
<th>FT-BIN</th>
<th>ALIGN-FOOT - RIGHT</th>
<th>ALIGN-FOOT - LEFT</th>
<th>PARSE-SYLLABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>(x .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-&gt; Apo the ose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>x</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x .) (x)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apo the ose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td></td>
<td>(x .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>(x .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A pothe ose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x . .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x .)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apothe ose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SpecFirst > Head Left**

- **SF**
  - SVO
    - \*!
  - VSO
    - \*!
  - SOV
    - \*\*!

With this ranking, we therefore capture SVO languages like English. A reranking of the two constraints yields VSO languages such as Irish.
Head Left > Spec First

\[
\begin{array}{ccc}
\text{HdLft} & \text{SF} \\
\text{SVO} & *! & !
\end{array}
\]

If Head Right has a higher rank than HeadLeft, its order relative to SpecFirst appears to be irrelevant: we always get SOV order, because it violates none of the two principles.

\[
\begin{array}{ccc}
\text{SF} & \text{HdRt} \\
\text{SVO} & * & *
\end{array}
\]

Two violations in a cell arise when there is a distance between the item in question and the position targeted in the constraints of two elements. If there is binary branching, both X and Y \([x \text{ subject} [y \text{ object Verb}]\) violate HeadLeft, so that counting the number of categories in which HdLeft or HdRight is violated is an easier accurate means of determining how "severe" the violation of the ordering principle is.

The principles Head Left and Head Right can of course also be formulated in a very direct way in terms of alignment

\[
\begin{align*}
\text{Align (Phrase, Head, Left)} &= \text{HdLeft} \\
\text{Align (Phrase, Head, Right)} &= \text{HdRight}
\end{align*}
\]

If either of the two principles is undominated by other alignment constraints we get languages with „Cross-Categorial Harmony“, in which heads have the same position in all categories. Only fifty percent of the languages are that symmetric. Deviations from the harmonic ideal can be captured when we have alignment principles which (a) have more specific values in the first two parameters, and when these dominate the two general alignment constraints. Thus, if Align (Verbphrase, Verb, Left) dominates Align (Phrase, Head, Right), we arrive at a language in which which verb phrases are headinitial, while other categories are headfinal

It is also conceivable that the order of the head may be different in different projections of the same head. As the following tableaus shows, we can thereby encode languages in which the objects appear on a different side of the verb than adjuncts and specifiers. Such a language seems to be Chinese.
It is reasonable to suspect that the more important specific alignment constraints get, the more marked the language is. One may wonder whether the specific alignment principles are part of the grammars of all languages (with a very low rank normally), or whether we may assume that progressive specification of general alignment principles may be a second source of language variation in OTS.

4.2. SpecFirst from above can easily be reinterpreted as an alignment principle as well.

\[ \text{SpecFirst} = \text{Align (Specifier, XP, left)} \]

Whether this is the optimal move or not remains unclear. Unlike what holds in head serialization, there seems to be no principle-counterprinciple pair, or it is at least doubtful that there is one: There appears to be no SpecLast effect e.g. for wh-phrases: wh-movement always goes to the left, the two exceptions to this hopefully only being apparent. If we deny the existence of Align (Specifier, XP, right), we must have something different to say about rare VOS or OVS languages, and such models not assuming phrase-final specifiers even in these languages have been developed. Likewise, LEFT does not really seem to have a mirror-constraint.

4.3. Here are two more examples for alignment. V-to-I movement, as we have it (arguably) in Jean embrasse souvent Marie can be forced by Align (Infl, Verb, left), but only in languages in which OblHd dominates Stay/Economy.

If (1) > (2), we predict that Focus goes to the slot immediately preceding the verb in German

(1) Align ([+v]-P, Head, right)
(2) Align (CP, Focus, right)

4.4. Consider now the statement Align (IP, specifier, left). This alignment principle guarantees that each IP begins with a specifier. This looks like the EPP, or Grimshaw’s principle Subject. The EPP may thus simply be a consequence of a particular alignment. Whether that is the optimal solution is not clear, however, One may wonder why there seems to be no mirror-constraint: Align (IP, specifier, right). Furthermore, OV languages do not seem to respect the EPP at all. And one may finally wonder why there are no similar alignment constraint for other categories: there are no forced expletives in Spec,DP or Spec,AGR-O-P. Headedness, on the other hand, may be understood as a consequence of the alignment principles. This could then take over the function of OblHd - but only if OTS is applied cyclically (see below), because unalignable traces fulfill OblHd, too.
4.5. This session concludes with a detailed discussion of linearization - it illustrates a certain widespread phenomenon that is fairly enigmatic: noninitiality effects. Recall that the head of a clause comes first in a language like Irish:

D’eirigh Ciaran
rose-3sg Ciaran
duirt Seán go-bhfuil Cathal ag rince
said John that-is Charles -ing dance

Geraldine Legendre describes this as Edgemost (Fin, LEFT), which is nothing but Align (clause, head, left).

Breton does not differ from Irish in embedded clauses:

Kredin ran en deus aret Yann e bark
believe do-I 3sg has ploughed Y. his field

But in the matrix clause, the finite element must NOT be initial.

Lennet en deus Yann al levr
read 3-have John the book

*En deus lennet Yann al levr

The additional constraint Legendre proposes is NonInit(Fin, root): the finite element must not be aligned with the left edge of the root.

With NonInit(Fin, root) > Align (clause, head, left), the second position of the finite auxiliary in Breton is derived. The Non-Initiality Constraint can be fulfilled by placing a focus constituent into first position, as in

Yann en deus aret e bark
JOHN has ploughed the field

e bark en deus aret Yann

But when there is no special focus in the matrix clause, the Non-Initiality Constraint must nevertheless be respected. Consequently, the non-finite verb is placed in first position - a construction much reminiscent of so-called stylistic fronting in Icelandic, a Germanic finite-second position language:

Fram hefur komidh adh fiskadh hefur veridh í leyfisleysi
forth has come that fished has been illegally

German differs from Breton in few respects only. First, the finite verb is aligned at the right edge in simple embedded clauses.

Dass er das Buch sieht
that he the book sees
Align (clause, head, right)

The main clause observes a finite-second constraint, i.e. it is characterized by

Align (root, fin, left)
NonInit(Fin, root)

German also differs from Breton in that the specifier of the root clause need not be focal. It can e.g. be a subject pronoun or an expletive.
es kommt der Pfarrer
there comes the priest

Basque differs from German in being verb final in the root clause as well, as
Jonek hori daki
John that knows
illustrates. But certain root clause types, viz. questions and focus clauses, again obey a
finite second constraint:
nork daki hori
who knows that
JONEK daki hori

In the latter respect, Basque is like English, where verb-second effects are restricted
to questions …
Who did you see?

4.6 One interesting aspect of second position phenomena is that they are not restricted
to finite verbs. For example, all types of clitics must strictly appear in second position
on Croatian.
Ivan kaze da mu je Marija jučer dala auto
Ivan said that him-is Maria yesterday given car
*Ivan kaze da Marija mu je jučer dala auto

Clitic Placement in Polish shows that the Non-Initiality Effect is an independent
phenomenon: clitics may go to any position as long as they do not end up in a clause-
initial slot:
My znowu wczoraj poszli smy do parku
We again yesterday went are to park
My znowu wczoraj smy poszli do parku
My znowu smy wczoraj poszli do parku
My _smy znowu wczoraj poszli do parku
*Smy My znowu wczoraj poszli do parku

5. Output-Output Constraints

A5. Allophonic variation: Output-output correspondence
A5.1 Output-output correspondence
A.5.2 A First excursion: Emergence of the Unmarked

B5. Syntax, Morphology and Coordination: OO-correspondence in the syntax

B5.1 Output-output correspondence
B5.2 A First excursion: Emergence of the Unmarked

One of the main tasks of phonology is to account for allophonic variation. Different
morphological instances of a morpheme can have different phonological forms. Some well-
known examples for English and Germanic are listed in (1) and (2).

(1) Allophonic variations in English
a. Palatalization of velars after high vowels: electric-electricity
b. Trisyllabic laxing: fable-fabulous
c. Flapping plus schwa/full vowel: atom-atomic
d. Allophony of the plural morpheme: cats, dogs, judges

(2) Allophonic variations in German
a. Final Devoicing: Hund- Hunde
b. Velar nasal: Diphthong-diphthongieren

In the transformationalist approach, the existence of allophony is expected since the surface forms (the outputs in OT) are derived by means of ordered rules and rules apply at different stages of the derivation in order to deliver different forms: first, different from the input, and second different from related forms. In the classical SPE each variant of a morpheme evolves independently, and the derivation can result in very different surface forms, according to the rules which have applied. In Lexical Phonology, since derivations are ordered in time, differences between the surface realizations of a morpheme are predicted to be somewhat reduced. But in principle, nothing blocks the creation of two morphologically related forms which have very different phonological outputs.

However, in OT, there is just one grammar, and ideally, we should not find allophonic variation at all, or only very few. This is because all forms underly the same grammar, and, crucially, all forms are subject to the same faithfulness constraints. As a consequence, it is now allomorphy which is the exceptional case. The fact that OT does in principle not allow as much variations as earlier grammars is basically a desirable result. However, since allophony is an integral part of all phonological systems, the grammar must be flexible enough to account for it. We will sum up here some of the methods which have been proposed in the OT literature to account for allophony, as well as a new one.

1. Output-output correspondence

Output-output correspondence was introduced by McCarthy & Prince (1995) to account for some morphologically-based phonological effects. Instead of taking an input as a reference, a morphological operation applies to a ready output, a form which has already been through phonology.

Standard OT (the containment OT of Prince & Smolensky 1993 and McCarthy & Prince 1993) only allows input-output relationships. This view was strongly influenced by the derivational phonology. However, because it is output-oriented, the admission of related forms as references is a natural expansion of OT. As a consequence, a phonological form can have two ‘bases’, its input and another, related, output form. In a first step, the input-output relation was called faithfulness and the relation existing between two related morphological forms was called correspondence. Faithfulness to input is ranked differently than correspondence between two outputs. However, in a second step, the two kinds of relations are fused together. Both use the same kind of methods to compare related forms, namely correspondence, formulated as in (3).

(3) Definition of Correspondence (McCarthy & Prince 1995:262)
Given two strings $S_1$ and $S_2$, correspondence is a relation $\mathcal{R}$ between the elements of $S_1$ and those of $S_2$. When $\alpha \mathcal{R} \beta$, the elements $\alpha$ of $S_1$ and $\beta$ of $S_2$ are called correspondents of each other.

The notion of correspondence is utterly vague. The correspondence relation takes its substance from a series of constraints implementing the kind of relation needed in each case. Some of them are well-known, and some are new.
To express the fact that two forms are in a correspondence relation, their segments are subscripted, and the constraints in (4) check the correspondence between each segment wrt different properties, like their featural make-up, their linearity etc. Since the grammar is now allowed to relate not only inputs to outputs but also outputs to each others, it is much more elaborated.

(4) Correspondence constraints: MAX (no deletion), DEP (no epenthesis), IDENT(F), LINEARITY, CONTIGUITY, ANCHOR-Edge, HEAD-MATCH …

(6) Input-Output faithfulness and Output-Output correspondence

\[
\text{Input} \quad \text{?} \quad \text{Output1} \quad \text{Output2}
\]

Corresp.

Three standard kinds of phenomena can be used to illustrate the need of output-output correspondence: reduplication, affixation and hypochoristics

1.1. Reduplication
Reduplication is a morphological operation (often expressing plural, iteration, habituation, intensification …) consisting in copying (reduplicating) part or whole of a stem. According to McCarthy & Prince, only ‘authentic’ prosodic constituents can be reduplicants.

(7) Ilokano Reduplicant Template: Heavy syllable (McCarthy & Prince 1995)

\[
\sigma \\
/ \backslash
\mu \mu
\]

(8) Examples of Ilokano reduplication
   a. Reduplicant consists of a closed syllable
      \[
      \text{tra.ba.ho} \quad \text{trab} - \text{tra.ba.ho} \quad \text{‘work’}
      \]
   b. Reduplicant consists of a syllable with a long vowel
      \[
      \text{ró\^ot} \quad \text{ro:} - \text{ró\^ot} \quad \text{‘litter’}
      \]

A second example comes from Lardil.

(9) Lardil Reduplicant Templates (McCarthy & Prince 1995): Binary feet

\[
\text{F} \quad \text{F} \\
/ \backslash \\
\sigma \sigma \sigma
\mu \mu
\]

(10) Examples of Lardil reduplication
   a. Reduplicant consists of two syllables
      \[
      \text{[kele-th]} \quad \text{kele} \quad \text{kele-kele} \quad \text{‘to cut’}
      \]
      \[
      \text{[pareli-th]} \quad \text{pareli} \quad \text{parel-pareli} \quad \text{‘to gather’}
      \]
   b. Reduplicant consists of a heavy syllable
Until now, a simple relation between input and output is sufficient to account for the reduplications. Why, then, did McCarthy and Prince introduce the powerful method of output-output correspondence into reduplication? The answer is that in some languages, the segmental make-up (so-called melody - a misnomer) of the reduplicant copies the segmental make-up of the full form, rather than taking its raw material from the input. Two cases show that.

- overapplication: a phonological process has seemed to apply, though its context of application is not visible at the surface (counterbleeding, non-surface apparent)
- underapplication: a phonological process does not apply, though its context of application is present at the surface (counterfeeding, non-surface true).

Let us first take a look at overapplication with one of McCarthy & Prince’s original examples, Javanese h-deletion. This language deletes [h] intervocalically. In the reduplicant, [h] is also deleted, even if the context of deletion is not present, since [h] is not intervocalic.

(11) Javanese h-deletion
Root Root+my Root+Dem.
aneh aneh-ku ane-e ‘strange’
arah arah-ku ara-e ‘direction’

(12) Reduplication in Javanese
bedah bedah-bedah beda-beda-e ‘broken’
dajøh dajøh-dajøh dajø-dajø-e ‘guest’

Turning now to underapplication, we can also take an example from McCarthy & Prince: palatalization in Akan. In this language, some sequences of consonant + vowel are not allowed, like those shown in (13). Rather consonants palatalize before a front vowel. However, unallowed sequences appear just the same when they are the result of reduplication. The reduplicant has a prespecified vowel [i] and copies the consonant of its base. It happens to be [k], the sequence [ki] is maintained, although it is changed into [t∫i] in all other contexts.

(13) Palatalization in Akan
t∫e *ke ‘divide’
dΩ *de ‘receive’
çi *hi ‘border’

(14) Reduplication in Akan
ki-ka? *t[i]-ka? ‘bite’
hi-haw? *çi-haw? ‘trouble’

In sum, over- und underapplication show that a reduplicant sometimes copies the material of its related base, an output, rather than taking the input of the morpheme. Output-output correspondence allows straightforwardly for such cases.

1.2 Different levels of affixation
It has been observed that affixes appear in a certain order, and that they behave as classes w.r.t. this property.
In English, besides other morphological operations like compounding and inflection, two levels of derivational affixation have been identified.

- Level I affixes which influence the phonology of the stem: -ic, -ation, -al
- Level II affixes which do not: -less, -ness, -y, -ing

Level II affixes are peripheral to Level I affixes (but see Fabb 1988 who showed that more restrictions are at play than just ordering). To account for this, Kiparsky, Mohanan and others developed a model of Lexical Phonology, in which morphology and phonology are interleaving. After each morphological operation of type I, the whole phonology, consisting of a set of ordered rules, applies. After completion of phonology of level I, morphology II (level II affixation) applies, then the whole phonology applies again. Level II phonology has no access to morphological information provided at earlier levels (and vice-versa): we thus have a cyclic model of the morphology-phonology interactions (but see Mohanan who allows loops in Malayalam).

When all levels have been completed (there may be more than two), the so-called post-lexical phonology applies, which is the sentence-level phonology. This phonology is automatic, applies in all contexts, and doesn’t care about levels. Final Devoicing in German is an example of this type.

Because of this extreme ordering of morphological and phonological events, OT has problems to imitate the results of Lexical Phonology. OT can account for the set of ordered rules inside of each level, at least to a certain extent, but the levels themselves are more difficult to reproduce.

In many instances of affix ordering, it seems to be the case that one level is base-like in the sense that one instance of the allomorphic variation is similar to the underived form. This is illustrated by the following words, from Benua(1995).

(15) Level ordering of affixes (Benua 1995): New York-Philadelphia dialects (æ-tensing: E is tense)

<table>
<thead>
<tr>
<th>Unaffixed</th>
<th>Class 1 Affix</th>
<th>Class 2 Affix</th>
</tr>
</thead>
<tbody>
<tr>
<td>class [klEš]</td>
<td>classic [klɐ:sik]</td>
<td>classy [klE.ʃi]</td>
</tr>
<tr>
<td>mass [mEš]</td>
<td>massive [mæ.siv]</td>
<td>massable [mE.s...]</td>
</tr>
</tbody>
</table>

The standard kind of OT cannot account for the different vowel in the stem of these words, due to the different kind of affixation. The problem is that the tensing rule expressed in the form of a constraint in (16) is syllable-based. However the alternating vowel is always in the same kind of syllable, as can be seen from the transcriptions.

(16) æ-tensing

\[ *\text{æC}_a \]

This constraint cannot be ranked as to deliver all forms properly.

The solution that Benua (1995) proposes for this puzzle is to account for level II affixes with correspondence to the underived outputs, in the examples above class, pass, and so on. Level I affixes take the input as input, and level II affixes takes the ready outputs of their bases.

The faithfulness to the output, when relevant, is assumed to be greater than the faithfulness to the input. This explains why level II affixes do not trigger much phonological changes in the stem.
1.3 Hypochoristics
A third kind of morphological process for which OO-correspondence makes good predictions is hypochoristic formation. Here, too, the crucial cases are those for which input-output faithfulness delivers another form as output-output correspondence. Two different cases will be considered, one in which the prosodic form is involved and another one concerning the segmental material of the short form.
In many languages, the prosodic form of hypochoristics is the unmarked foot of the language, generally a syllabic trochee or – much more rarely - an iambic foot. It must be emphasized that the unmarked foot of a language is different from the minimal word, though these two notions are rarely kept apart in the literature. The minimal word is the smallest kind of word, in general a foot, that a language allows for its content words. French, for instance, has no word minimality (content words’ rimes can consist of just a single short vowel, as in *eau ‘water’, *bu ‘drunk’, *là ‘there’, *riz ‘rice’) but it has a disyllabic iamb as its unmarked prosodic word (17), which is also the unmarked hypochoristic form, besides other possibilities. Some examples of hypochoristics are listed in (18).

(17) Prosodic Constraint in French
Hypocoristics = F = [σ] or [σσ]

(18) Hypochoristics:
Véronique  Véro
Dominique  Domí, Dom, Do
Bénédicte  Béné
Elisabeth  Zabeth, Babé, Babette, Beth

French also has reduplications, like those shown in (19), which are highly limited in use. Reduplications like those in (19) are either lexicalized or only used when talking to babies or dogs, thus a kind of motherese.

(19) a. Reduplications (Echo-words) = σ σ] built from σ]
    \(\land\)
    (C)V
    b. père -> pépère, ours -> nounours, main -> main-main, os -> nonos

The input to these reduplications is a monosyllabic word and the result is an iamb. Since an onset is obligatory, when no consonant is present that could be used as an onset, an
epethetical [n], probably in analogy with liaison [n] of un ‘a’, is inserted. However, this segmental particularity won’t bother us here and we concentrate on the prosodic structure. To account for the pattern in (19a), output-output correspondence is needed, since the syllabification is not part of the input: it is an added structure appearing on the output. As a consequence, the reduplicated form needs to refer to the output syllabification of the full form before it is possible to decide whether it is a possible base for reduplication or not.

(20) IO-Faithfulness >> Prosodic Constraints >> BT-Faithfulness (Base-Truncation)

First excursion: Emergence of the Unmarked
The prosodic pattern of hypochoristics, not only in French but in many other languages as well, is characterized by the fact that the least marked properties of prosodic constituency become apparent. This has been called The Emergence of The Unmarked (TETU). The prosodic constraints in the middle of (20) are responsible for the unmarked pattern of the language: bisyllabic foot, trochaic pattern, open syllables… Full forms have only few possibilities to display really unmarked properties, because they have to obey IO-faithfulness and are thus subject to all kinds of idiosyncrasies (as far as they are allowed in the language at all) to which they have to be faithful. But the truncated, or reduplicated forms in French are faithful to their bases only to a certain point. More important than faithfulness is the obedience to prosodic unmarkedness.

The so-called i-formations in German could also be used to illustrate the same point. The unmarked prosodic word in German is a syllabic trochee, which emerges in the i-formations (Hansi, Kathi, Studi, Andi, Ulli etc), as well as in a series of other morphological operations (as in infinitives and some suffixation with –ig).

Second excursion: Ineffability
A second remark which can be made at this place is that if the full form has more than one syllable, no reduplicated correspondent is possible. This is a case of ineffability (see below).

The second case of hypochoristics in which it appears that output-output correspondence is needed is truncation in English (Benua 1995). A first name like Larry can be shortened to Lar. However in such syllables, closed by [r], the quality of the vowel [æ] is usually replaced by [α], as one can see from words like car and lark. The fact that [æ] is possible in Lar is a case of underapplication of a phonological process since the operation / æ/ _ [α] does not apply. The reason is that Lar wants to be as similar to Larry as possible. Some other cases are listed in (21).

(21) Truncations in English

<table>
<thead>
<tr>
<th></th>
<th>Larry [læ.ri]</th>
<th>Harry [hæ.ri]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lar</td>
<td>[lær]</td>
<td>[hær]</td>
</tr>
<tr>
<td>car</td>
<td>[kær]</td>
<td>[lærk]</td>
</tr>
<tr>
<td>lark</td>
<td>(*[kær], *[lærk])</td>
<td></td>
</tr>
</tbody>
</table>

In this case, the opposite effect from the one observed for French is observed. Output-output correspondence is higher-ranking than IO-Faithfulness.

(22)

<table>
<thead>
<tr>
<th>Input: /lar/</th>
<th>OO-Faith</th>
<th><em>ær</em></th>
<th>*α</th>
<th>IO-Faith</th>
</tr>
</thead>
<tbody>
<tr>
<td>a [lar]</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b [lær]</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion and open problems
1. Since correspondence is a vague notion, all kinds of forms should be able to enter into a correspondence relation. How can we delimit the desirable correspondence relations from the undesirable ones?

2. OO-constraints lead to an explosion of the constraints.

3. OO-correspondence needs an existing output in order to be workable. In some cases, surface forms seem to be faithful to a form which is never realized as an output. In those cases, we have opacity. Lexical Phonology, as well as all models using ordered rules have no problems with opacity. The existence of intermediate forms, neither inputs nor outputs, is a natural consequence of rule ordering. OT has big problems with those. Since no derivation enters phonology, no intermediate step should ever be needed.

Before turning to opacity and the solutions which have been proposed to account for it in OT, let us briefly mention Burzio’s gradient attraction approach, which is an extension of the basic insights of output-output correspondence, and a rejection of this approach at the same time. Burzio applies Flemming’s (1995) notion of perceptual contrast to allomorphy. He claims that similar representations attract each other and that they do so gradiently. The more similar they are, the greater the attraction. Allomorphs are similar because they consist to a large extent of the same segmental material and because they have (partly) the same semantic representation. But they also contrast with each other in order to be distinct. Burzio rejects output-output correspondence on several grounds, the most relevant one being the fact that allophonic variations of complex words can be triggered not only by the stem but also by the affix(es). The word *titánic*, for instance is stressed on the penultimate syllable, because of the propensity of the suffix -ic to position stress on the syllable preceding it. In Burzio’s terms, *titanic* is thus not only ‘attracted’ by *títan*, but also by words like *barbáric* and *dynámic*. Likewise the stress on *módernist* is influenced by *módern* and not by the one of *modérnity*, because -ist adjoins only to adjectival bases.

**B5. OO-Correspondence in the Syntax**

5.0. Possible evidence for Output-Output-Correspondence in the syntax comes from at least two domains: the syntax-morphology interaction, and coordinate structures.

5.1. In the passive, syntactic differences to the active counterpart tend to be minimized. In a representational model, this suggests an influence of OO-correspondence. Case marking is a good example for this. The following statements are plausible Case rules for the active construction:

- **Nom:** NPs bear nominative case
- **Acc:** NPs that are not the highest argument bear accusative case
- **Dat:** NPs that are neither the highest nor the lowest argument bear dative case

One would expect these rules to be applicable to the passive as well. Note that the active subject is suppressed in the passive construction ... What we find as a passive case grid is:

```
Es wird ihr gegeben
it-nom is her-dat given
```
By the Case rules for the active construction, we should rather get (e.g., the dative has no chance of being assigned when there are two arguments only ...)

*sie wird es gegeben
she-nom is it given

It would certainly not be a good move to postulate additional Case rules for the passive. A possible account might, however, invoke OO-correspondence between the active and the passive, maximizing faithfulness between the active and the passive. Obviously, a ban against accusative case showing up in the passive has to outrank OO-correspondence, however.

Instead of invoking OO-correspondence, the passive facts can be accounted for in terms of rule ordering:

1. The Case potential is determined on the basis of an active thematic structure (in the lexicon)
2. „Absorption“ of accusative Case and of the thematic role of the subject.

5.2. ECM constructions and their behavior with respect to Case might be a second domain for OO-correspondence. This can be argued for in the domain of Case agreement. In German, we find Case agreement with some predicate nominals, and with certain adverbial modifiers.

Ich bin ein Esel
I-nom am a-nom donkey

er grüsst die Männer einen nach dem anderen
he greets the-acc men one-acc after the other
die Männer grüssen ihn einer nach dem anderen
the men greet him one-nom after the other

Consider now ECM constructions. Pure Case agreement would lead one to suspect that both the subject and the predicate nominal/the adverb appear with accusative case in an ECM construction. However, there seems to be a dialectal split in German in this domain:

1. Ich lasse ihn einen Helden sein
I let him-acc an-acc hero be
2. Ich lasse ihn ein Held sein
I let him-acc an-nom hero-nom be

1. can be characterized as a maintenance of Case agreement, 2. suggests that the nominative case can be maintained, too. Similar facts can be observed for the adverbials:

Ich lasse die Männer einen nach dem anderen ankommen
I have the men one-acc after the other arrive
Ich lasse die Männer einer nach dem anderen ankommen
I have the men one-nom after the other arrive

The behavior of case agreement constructions in ECM-contexts thus allows two solutions:
(a) OO-Correspondence between the finite clause and the infinitive implies the maintainance of the second nominative
(b) Case determination for the predicate nominal before nom > acc change in the subject position of the infinitive guarantees the maintainance of the second nominative

5.3. Certain parallelism effects in coordinate structures might also be captured by postulating OO-faithfulness constraints. Scope is a case in point. Normally, the presence of two quantifiers in a clause introduces a scope ambiguity, as in
I introduced one of the boys to every teacher,

which can be interpreted as either involving the scope relation ONE > EVERY or EVERY > ONE. When two clauses are coordinated, and when there is some phonological reduction involved, the two conjuncts must fulfill a certain parallelism requirement, however. The sentence
I introduced one of the boys to every teacher, and Bill did, too

is just two-ways ambiguous, not four ways, as one might expect (because each of the (unreduced) conjuncts allows two readings). More examples for parallelism are:
an American runner seems to have won a gold medal, and a Russian athlete does, too

The relation of the indefinite relative to seem must be the same in the two clauses of
one guard was seen in front of every building, and a policeman was, too.

same scope relation necessary, again

In other words, in an ellipsis/coordination reduction construction, the scopal relations among the elements in clause A must be identical. In the so-called Y-model of grammar,

D-Structure \rightarrow S-Structure \rightarrow LF \rightarrow PF

in which phonology and semantics do not communicate with each other, this is difficult to account for. OO-faithfulness between the two conjuncts might be helpful. The alternative would be to develop a (complicated?) concept of "across-the-board rule application".

5.4. Many syntactic approaches assume more than one level of representation ... surface structure, Logical Form, Argument Structure, etc. A special form of OO-correspondence might be seen in a set of constraints that try to minimize differences between these levels. One would assume that several levels are generated independent of each others from a single input - and "economy of derivation" may then reduce to OO-faithfulness (a suggestion attributable to Edwin Williams). Minimal Link (=superiority) effects may be the best examples reflecting the attempt to minimize structural differences between levels of representation (Müller, Williams).
5.5. It has been argued that the sluicing construction

\[ \text{Someone had met someone in this park, I do not recall who whom} \]
\[ \text{Jemand hat hier jemanden getroffen, aber ich habe vergessen wer wen} \]

involves the fronting of all wh-phrases in the second clause, followed by the deletion of the remainder of the CP. If correct, this suggests that all wh-phrases of a multiple question may be moved in German and English under certain circumstances (just as they do quite generally in Bulgarian). By being fronted, the wh-phrases go to their scope position. In this sense, sluicing illustrates the emergence of the unmarked in syntax. Sluicing is also a potential example for OO-correspondence: the wh-phrases should correspond to indefinites in similar structural positions.

\[ \text{*Jemand hat jemanden geküsst, aber ich weiss nicht wer wann} \]
\[ \text{someone has someone kissed, but I know not who when} \]

6. INPUTS & INEFFABILITY

SYNTAX

A6.1. The structure of an input
A6.2. The ineffability problem and bidirectional optimization
A6.3. Other ways of solving ineffability

PHONOLOGY

B6.1. Ineffability
B6.2. Ties

6.0 In this section, three ways of understanding inputs in OT syntax will be discussed. Special emphasis will then be given to the issue of ineffability. Various alternative accounts of how to handle "absolute" ungrammaticality in OT will be discussed.

6.1 As we have already seen, the simplest assumption concerning inputs seems to be that inputs are just sets of words. This view is untenable in the light of certain markedness principles, like the MLC/the principle implying superiority. When there are two candidates for fronting in a multiple question, only the higher one usually moves:

\[ \text{Who saw what} \]
\[ \text{*what did who see} \]

This constraint may be implemented in various ways - but if the input is just a set of words, some of these formulations of the MLC will imply that

\[ \text{who kissed Jane?} \]
\[ \text{which involves a very short movement of who will incorrectly block} \]
\[ \text{*who did Jane kiss involving longer movement and an FI violation. This unwelcome consequence can be avoided if one assumes assume that the input is a predicate-argument structure like e.g.} \]

\[ \text{Kiss agent = Jane} \]
\[ \text{patient = who} \]
If faithfulness to input-PAS overrides markedness (in English, but not in KwaKwala or Malagassy ...), then the MLC violation in who did Jane kiss will be compensated for by the need to respect PAS.

6.2. A comparable markedness consideration seems to show that PAS are not sufficient either.
what did she tell me that he bought
she told me what he bought

These two sentences involve the same PAS, to which both are faithful. The key difference lies in the length of the movement path, so that the lower sentence should block the upper one. This is an incorrect consequence, though. In this particular case, closer inspection may reveal that the two sentence differ in terms expressible by PAS: tell selects a wh-clause in the second example, but not in the first one!! Thus, the object argument differs categorially -- the PAS are not identical, and the two sentences are not in the same competition.

Although the examples just discussed can be explained away, more complex structures reveal that PAS are insufficient: There are no MLC-effects in multiple questions whenever we get a semantic difference! Consider what happens in the embedded question in the following examples:
Who wonders who bought what
Here, the MLC is respected ... but interestingly,
Who wonders what who bought
is grammatical, too. We can have an MLC violation if who takes wide scope! In this respect, the embedded clause differs sharply from
*what did who buy

The point seems to be that we get an MLC effect whenever the two clauses would be semantically indistinguishable, but whenever it makes a difference (as it does in embedded questions whenever the wh-phrase in situ may take wide scope) which phrase moves, the MLC can apparently be violated.

This suggests an even more complex view of inputs: Input = PAS + indication of final scope. Our example
who wonders what who bought
is then a perfect sentence if ParseScope >> MLC.

6.3. If one allows scope indicators to enter input specifications, a serious problem appears to arise. There are inputs such as
PAS
wonder: you, wh-clause
fix: MANNER: how
OBJECT: what
scope(how) = low clause
scope (what) = high clause
that can easily be expressed as

\[ \text{what do you wonder how to fix} \]

but the slightly different input

\[ \begin{align*}
\text{wonder:} & \quad \text{you, wh-clause} \\
\text{fix:} & \quad \text{MANNER: how} \\
\text{OBJECT:} & \quad \text{what}
\end{align*} \]

\[ \text{scope(how) = high clause} \]

\[ \text{scope (what) = low clause} \]

cannot be expressed as

\[ *\text{how do you wonder what to fix} \]

nor by any other grammatical sentence of English. But in the logical structure of OT, this is a situation that should not arise: for every input, there must be an optimal way of expressing it—contrary to what holds in English. This is the ineffability problem.

But we have overlooked a possibility—what if the ban on long adjunct movement (the displacement of how) has a higher rank than Parse(Scope)? Then

\[ \text{you wonder how what to fix} \]

has a better profile than

\[ \text{how do you wonder what to fix} \]

even when the input assigns wide scope to how! And if the ban against multiple wh phrases/the need to fill the matrix wh-position also has a higher rank than Parse(Scope), then

\[ \text{what do you wonder how to fix} \]

has the best profile for our input. Thus, ineffability appears to have a simple solution ... There are many aspects of the input to which the output cannot and need not be faithful ... and if scope can be among these aspects, a certain meaning simply cannot be expressed.

The picture is not complete, however!

\[ \text{what do you wonder how to fix} \]

may be the optimal way of dealing with an input in which how has wide scope—but the sentence does not have that interpretation!

6.4 This problem can be addressed in a theory of bidirectional optimization! The grammatical evaluation is conceived of as going from meaning to form and from form to meaning at the same time.

\[ \text{STEP 1: Given an input, what is the optimal way of expressing it?} \]

\[ \text{STEP 2: Given an output, what is its optimal interpretation?} \]

An interpretation is then ineffable if its optimal expression has a different optimal interpretation! Bi-directional optimization may have a parallel in phonology ... certain possible UR-s cannot enter the lexicon ...

6.5. Other ways of dealing with ineffabilty have been proposed in the literature, too. Unfaithfulness to scope is just one example of what aspects of the input one can be unfaithful to, a further example would be unfaithfulness to featural specifications. Thus, Irish forbids multiple questions -- the optimal candidate may then be one in which
a wh-feature is not parsed, and the input is realized as a simple indefinite. One can also
deal with absolute ungrammaticality by having the "null-parse" (a candidate in which no
input material is parsed) win.

Ineffability may also be dealt with by enlarging the expressive power of GEN. Thus,
whenever case incompatibility blocks the use of a free relative clause in German, as in
the second example of

Wer ihn kennt, liebt ihn
who-nom him knows loves him
*wenn er kennt liebt ihn
who-acc he knows loves him
we may have a full relative clause
jeder, der er kennt, liebt ihn
everyone who he knows loves him

figure as the optimal realization of the input. Absolute ungrammaticalness may also be
the result of heavy restrictions on what GEN can do ... if, say, the island conditions are
part of GEN, EVAL need not account for them, and

*who do you weep because left you
would be ruled out as absolutely ungrammatical ...

But this line of reasoning faces a couple of problems

a) are there really constraints that are not violated in any language?
b) constraint violation need not uniform across categories

Pesetsky proposes that we confine OT to matters of pronunciation economy in syntax,
because of ineffability ... but this idea faces similar problems! We need to invoke
parameters etc., which we wanted to avoid ...

B. Ineffability, ties in phonology

1 Ineffability

Ineffability is the property of some potential phonological forms to be unpronounceable:
different parts of the grammar make irreconcilable requirements.
In this text, we will consider first some interface cases between phonology and morphology,
and afterwards, a pure phonological case will be addressed.

1. An English case: suffixation with deadjectival –ize
A first example, from Raffelsiefen (1996), cited in Kager (1999), is suffixation of -ize to
adjectival stems. The roots must have the right stress pattern in order to be good candidates
for suffixation. In particular they must not be finally stressed. Thus randomize and foreignize
are possible –ize formations, but *corruptize is not because corrupt is finally stressed.

2. A German case: suffixation with the diminutive suffix –chen
A similar case comes from productive umlaut in German, which is triggered by suffixation of
the diminutive suffixes –chen and –lein to a finally stressed stem (to the exception of a
metrically invisible schwa syllable).

(1)  a. Jahr → Jährchen, Woche → Wöchlein,
b. Bruder → Brüderchen, Mauer → Mäuerchen

The diminutive suffix has a fronting effect on the vowel immediately preceding it, but only in case this vowel has primary stress. In words like Europa and Monat we are confronted with an unsolvable conflict. The final vowel of the stem is not mainly stressed and the mainly stressed vowel is not adjacent to the suffix. In such a case, most German speakers prefer not to diminutivize the stem. An input consisting of Monat + chen is just not realized.

One solution, applicable to both cases just discussed, is to implement a violable constraint M-PARSE, which, appropriately ranked, will block the ineffable cases (McCarthy & Prince 1993). In the case of German, M-PARSE is ranked below the constraints requiring umlaut to occur in the right prosodic conditions. If the constraints on prosodic structure cannot be fulfilled, the morphemes are not realized and M-PARSE is violated. Since the other candidates are eliminated by higher-ranking constraints, violation of this constraint no realization is the best option.

(2) M-PARSE: Realize the morphemes’

(3) Jährchen

<table>
<thead>
<tr>
<th>/Jahr+chen/</th>
<th>No unstressed ü/ä/ö</th>
<th>UMLAUT</th>
<th>NOCROSSING</th>
<th>M-PARSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jähr-chen</td>
<td>!</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jahr-chen</td>
<td>!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ø</td>
<td>!</td>
<td>!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(4) *Monätcchen

<table>
<thead>
<tr>
<th>/Mónat+chen/</th>
<th>No unstressed ü/ä/ö</th>
<th>UMLAUT</th>
<th>NOCROSSING</th>
<th>M-PARSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monat-chen</td>
<td>!</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mönat-chen</td>
<td>!</td>
<td></td>
<td></td>
<td>!</td>
</tr>
<tr>
<td>ø</td>
<td>!</td>
<td></td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

Kager (1999) lists the following problems which show up with this approach:

1. Many suffixes influence the stress pattern of the stem they adjoin to. Why do –ize or –chen not do that?

2. Stress clash is tolerated in English in some cases. Why not here? The same tolerance is found in German for unstressed umlauted vowels, as well as for stems suffixed with –chen and without umlaut: Frauchen, Blondchen, …. Why not Wodkächen or Wodkachen?

Conclusion: OT alone cannot account for the different behavior of –ize and a stress-shifting suffix like, -ic (remember titánic).

Another solution is the one advocated by Orgun & Sprouse (1997) who propose a filter, called Control, located between Gen and Eval, and consisting of unviolable constraints. This filter blocks the formation of the words not listed in the lexicon.

Here, too, problems immediately arise, namely with the kinds of constraints which are supposed to be unviolable: *STRESSCLASH is not unviolable in English (see gymnast or Chinése), neither are NOUNSTRESSED ü/ö/ä (compare möblieren) and UMLAUT in German (Frauchen).

3. Prefixation of ge- to a participle

Prefixation of German participle ge- illustrates a different kind of ineffability. This prefix is
only present when the verbal root has initial stress. But the existence of a participle is independent of the presence of ge-. This can be analyzed as a case of partial ineffability. Only the prefix is ineffable.

(5) ge- in German
   a. geärbeitet, gegészten, getrómmelt, [ge.[(lacht)]_{PW}]_{PW}
   b. gewallfahret, gefürbüstuckt, géöhrfeigt, gekémonzeichnet
   c. spaziert, trompétet, verpásst, prophezéit

Compare also:

(6) More participles
   a. untergetaucht, angekommen (separable particle)
   b. überholt (*übergeholt), entfallen (*entgefallen) (inseparable particle)
   c. missverstanden (*gemissverstanden, *missgverstanden, *missvergestanden), berücksichtigt (*begerücksichtigt, …) (inseparable particles)

The relevant constraint is MAX(Aff). It is just the morpheme which is not parsed, since the participle exists. This constraint is violated in the best candidate when the prosodic constraints on the base are not fulfilled.

4. Ineffability in segmental phonology

In every language, there is a ranking of the constraints on admissability of segments. In OT, it is not necessary to list an explicit inventory of segments. The inventory of allowed segments is a consequence of the grammar. In other words, whether a segment emerges in a certain language or not, is not an independent part of the grammar but is determined by the ranking of the constraints on admissability of segments, like the one in (7) (tentative and partial for English).

(7) *Ω >> *ɛ >> *fi >> ∫ >> γ >> Ω >> ∫ >> etc.

The voiced palatal fricative [Ω] is not a canonical segment of English, but is nevertheless realized in some environments. It is a kind of link between unallowed and allowed segments. The segments to the left of the voiced palatal fricative in (8) are not legal segments of English and are eliminated by the markedness hierarchy. The segments to the right of [Ω] are licit segments of English and are dominated by Faithfulness constraints. Thus, *∫ is dominated by MAX[∫] for instance.

What happens to an input like /ɛ fi ∫/ which is a possible input because of Richness of the Base which says that inputs are unconstrained? The answer is probably: nothing good can emerge from such an input. It is unpronounceable and is consequently ineffable.

II. Ties

Some forms are in free variation. A phenomenon like free variation implies that the grammar must be flexible enough for allowing competing expressions to emerge. Consider first a few examples.

(8) Free variation in the syntax of French
   Tu as vu qui? ‘Who did you see?’
   Qui as-tu vu? ‘Who did you see?’
   Qui est-ce que tu as vu? ‘Who did you see?’
(9) Free variation in the phonology of German
[taç] and [ta:k] for Tag ‘day’
[gero:n] and [ge:reo:n] for Gereon ‘a name’
Téléfon vs. Telefón

(10) Free variation in French
a. ancien ‘old’: [ãʃjɛn] and [ãʃjɛn] in the masculine liaison case
b. ouest [west] ~ [uest] ‘west’, nuage [nu.aO] ~ [nuaO] ‘cloud’, piano [pjano] ~ [piano]

Let’s concentrate on the last cases, listed in (10b) where the high vowel can be realized facultatively as a glide. There is a free alternation between the two realizations [nu.aO] and [nuaO] for instance. An OT analysis of these cases concentrates on the conflict between reduction of hiatus avoidance and avoidance of complex syllable margins. In a word like nuage, which alternates, the pronunciation with a glide results in a strange complex onset [n]. In contrast, preservation of the full vowel results in a hiatus, a state of affair which is universally avoided. A simplified OT tableau is given in (11).

(11) Free variation in nuage

<table>
<thead>
<tr>
<th>/nuaO/</th>
<th>NOHIATUS</th>
<th>NOCOMPLEXONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>nuaO</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>nuaO</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

The debate of how to account for free variation in phonology centers around two possibilities: co-phonologies, often dismissed because too permissive, and ties, which are unranked constraints. At just one place in a constraint hierarchy, constraints can be ranked in one way or another. The result is two different optimal outputs. The rest of the hierarchy is identical.

### 7. Opacity

**Constraint conjunction, sympathy, turbidity**

A7.1  Opacity
A7.2  Opacity in Standard OT
A7.3  Constraint Conjunction
A7.4  Sympathy
A7.5  Turpidity

B7.1  Constraint Conjunction in Syntax
B7.2  Ties in Syntax

### 1. Opacity

Output-output correspondence is subject to some limitation, since it is restricted to cases in which one of the allomorphs is identical to an existing grammatical form, thus to a realized output. Beside these cases, derivational phonologists have described an important number of phonological processes which require reference of output structures to non-existing forms,
thus to intermediate steps between the inputs (the underlying representations of the derivational phonology) and the outputs (the surface structures): this has been called phonological opacity.

(1) Opacity (Kiparsky 1973)
A phonological rule $\mathcal{R}$ of the form $A \rightarrow B / C \_ D$ is opaque, if there are structures with any of the following characteristics:
   a. instances of $A$ in the environment $C \_ D$
   b. instances of $B$ derived by $\mathcal{R}$ that occur in environments other than $C \_ D$.

In general, two phonological processes interact in an opaque way when the result of one operation conceals the result of another operation. The two situations mentioned in (1a) and (1b) can be described derivationally as follows.

- the environment of a process is given, but the process does not take place. This situation arises when the the environment of the process arises as a consequence of a later process. This is a case of non surface-true counterfeeding operation (underapplication). In (1a), a structure $A$ occurs on the surface, although there is a rule $\mathcal{R}$ transforming instances of $A$ into $B$.

- A phonological process applies in a certain context, but this context is destroyed by the application of another, later process. This is a case of non surface-apparent counterbleeding operation (also called overapplication). In (1b), a structure $B$ occurs on the surface, which has been derived by rule $\mathcal{R}$ although the context for application of the rule is not apparent.

For illustration, let’s take a look at some concrete cases. In order to make the terms counterbleeding und counterfeeding more accessible, it is useful to look first at the simpler cases, feeding and bleeding.

First case: feeding
In the case of feeding, a rule ‘feeds’ a second rule in creating its environment (so called structural description). The derivation (2) shows an abstract case in which $R_1$ feeds $R_2$ and (3) a case of feeding in Lomongo, a Bantu language.

(2) Underlying /CAD/ $\mathcal{R}_1$: $A \rightarrow B / \_ D$ CBD $\mathcal{R}_2$: $C \rightarrow E / \_ B$ EBD

(3) Feeding in Lomongo (Hulstaert 1957, Kenstowicz & Kisseberth 1977:156)
/lo-en-a/  ‘you see’
lw-en-a $\mathcal{R}_1$: [o] $\rightarrow$ [w]/ $\_ V$
d’w-en-a $\mathcal{R}_2$: [l] becomes [d’] before a glide
[d’wena] Surface

Second case: counterfeeding
In counterfeeding the rules apply in the opposite order: there is thus no feeding relation. In the example (5) from Arabic, the first rule ([a] becomes [i] in non-final open syllables) applies before the second one, glide vocalization. Rule 2 would feed Rule 1, would they appear in the reverse order. Since they apply in the order they do, however, Rule 2 counterfeeds Rule 1. In more general terms, the application of a second rule creates the structural description of a first rule, but this first rule cannot apply any longer. In (5), the word /badw/ is pronounced [badu], although the context for $a$-raising is present on the surface.
(4) Underlying /CAD/
\[R_1: C \rightarrow E / _B \]
\[R_2: A \rightarrow B / _D \]
CBD (There is a sequence CB in spite of Rule 1)

(5) Arabic (McCarthy 1998)
‘Beduin’ ‘he wrote’
Underlying /badw/ /katab/
\[R_1: a\text{-}raising \_ /kitab/\]
\[R_2: Glide vocalization /badu/ \_\]
Surface [badu] (*[bidu]) [katab]

Third case: bleeding
In bleeding, a rule destroys the environment of another rule. The second rule could only apply if it were ordered before the first one, though there is no feeding relation here. Either R1 applies, or R2.

(6) Underlying /CAD/
\[R_1: A \rightarrow B / _D \]
\[R_2: D \rightarrow \_ / A _ \]
CBD (*CA)

(7) Northern German realization of the sequence [\text{\textligth}_+g]
/lapg/
\[R_1: \text{Final Devoicing} \_ /la\text{j}k\]
\[R_2: \text{g-deletion} \_\]
Surface [la\text{j}k]

Fourth case: counterbleeding
Now to the final case, counterbleeding. The rules apply in the opposite order relatively to bleeding. In the example (8), the second rule would block the first one if they would apply in the other order. The crucial fact for opacity is that it is no longer visible why epenthesis has applied in the surface form [de\text{\textligth}e].

(8) Underlying /CAD/
\[R_1: \_ \rightarrow E / A– D \]
\[R_2: D \rightarrow \_ / \_\]
CAED

(9) Counterbleeding in Tiberian Hebrew [de\text{\textligth}e] ‘tender grass’ ((McCarthy to appear))
Underlying form /de\text{\textligth}?/
Vowel epenthesis between two consonants de\text{\textligth}?
Glottal stop deletion de\text{\textligth}e
Surface [de\text{\textligth}e]

As illustrated in (2) to (9), these cases present no problem in a derivational approach. Rule ordering makes them easy to cope with.

2. Opacity in standard OT

Feeding and bleeding present no problem for OT, as shown in (10) and (11). The markedness constraints M-Constraints, *[l] +Glide and *HIATUS, rank higher than the IDENT-Constraints.
(10) Feeding in Lomongo

<table>
<thead>
<tr>
<th>/lo--n-a/</th>
<th><em>[l]</em> +Gl</th>
<th>*HIATUS</th>
<th>IDENT (o)</th>
<th>IDENT (l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. d’w-en-a</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. lw-en-a</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. lo-en-a</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(11) Bleeding in German

<table>
<thead>
<tr>
<th>/la-g/</th>
<th>*[-voiced Obstr]_e]</th>
<th>*tautosyllabic [-ng]</th>
<th>MAX-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lan-k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. lan</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>c. lang</td>
<td>*!</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

Counterbleeding und counterfeeding are not as straightforward. As a matter of fact, every ranking of the relevant constraints (12) delivers contradictory results in standard OT. To see this, we can use the examples (5) and (9) above. First the example for counterbleeding [de\-e]. The optimal form is candidate a. in (13). There is no possibility in standard OT to make this candidate better than b. The reason for this is that candidate b violates a subset of constraints of those violated by candidate a. Candidate b. is the so-called transparent candidate. Under every thinkable ranking of the necessary constraints candidate b will always wins the competition. The necessary constraints as well as their hierarchization are in agreement with the remaining phonology of the language.

(12) OT-Constraints

a. CODA-COND: [?] is not a possible coda.

b. MAX-C: No consonant deletion.

c. DEP-V: No vowel epenthesis.

(13) Counterbleeding in Tiberian Hebrew

<table>
<thead>
<tr>
<th>/de-e]/</th>
<th>CODA-COND</th>
<th>MAX-C</th>
<th>DEP-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. de-e</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Transparent candidate b. de-l</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. de-e?]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. de-?]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the case of counterfeeding problems are similar. The ranking of *\[a\]_o and IDENT(high) shown in (13) is motivated by the general phonology of Arabic. Under this ranking, badu can never win over bidu.

(14) More OT-constraints:

a. *COMPLEX: Codas are not complex.

b. *\[a\]_o: No \[a\] in a non-final open syllable.

b. IDENT(high): [high] in the input corresponds to [high] in the output.

(15) Counterfeeding in Beduin Arabic

<table>
<thead>
<tr>
<th>/bad-w/</th>
<th>*COMPLEX</th>
<th>*[a]_o</th>
<th>IDENT(high)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. badw</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Transparent candidate b. bidu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. badw</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The insight that opacity is not solvable in a classical kind of OT has been immediately noticed by several phonologists. Different solutions have been proposed to account for the opacity problem, to which we now turn.

3. Constraint Conjunction
Constraint conjunction was first proposed by Smolensky (1995). In his proposal, two or more constraints are conjoined or and the result is a new constraint. A particular case of constraint conjunction is self-conjunction. In order to make the effect of a conjoined constraint visible, it must be ranked higher than the single constituting components (C₁ & C₂ >> C₁, C₂).

Ito and Mester (2000) have proposed to apply constraint conjunction to the opacity cases, in particular to German g-spirantization. According to them, this is an opacity problem, since it is not visible on the surface why [ç] occurs from an input consisting of [g]. In standard OT an output [k] is always better than [ç], since it satisfies Final Devoicing without violating an additional faithfulness constraint. Ito & Mester’s approach makes use of the following conjoined constraints.

(16) a. *VC: [*Voiced Obstruent & *Coda] (Final Devoicing)
   b. *CD: *Dorsal & *DorsalPlosive (No [g], no [k])
   c. [*CD & Ident(voice)]: *Dorsal & *Dorsal Plosive & Ident(voice)

(17) Spirantization in German (Ito & Mester 1998)

<table>
<thead>
<tr>
<th>Input</th>
<th>*VC</th>
<th>*CD &amp; Id(voi)</th>
<th>Max</th>
<th>Id(cont)</th>
<th>*CD</th>
<th>Id(voi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ho:nig]</td>
<td><img src="image" alt="" /></td>
<td>*!</td>
<td>*</td>
<td>---------------</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>[ho:niv]</td>
<td><img src="image" alt="" /></td>
<td>*!</td>
<td>*</td>
<td>---------------</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>[ho:n]</td>
<td><img src="image" alt="" /></td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[ho:nk]</td>
<td><img src="image" alt="" /></td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[ho:nç]</td>
<td><img src="image" alt="" /></td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

An important aspect of the constraint conjunction is the fact that their domain of activity must be restricted. This can be illustrated with Final Devoicing in (16a). This constraint prohibits voiced obstruents and codas. In order to eliminate all voiced obstruents from the coda, it must be very high ranking, since Final Devoicing is obligatory in German. However, voiced obstruents and codas are generally allowed to co-occur in domains larger than the coda, in a syllable (Bein ‘Leg’ is an example) or in words. The domain of activity of (16a) must thus be restricted to the segment. A segment is not allowed to be at the same time a coda and a voiced obstruent. But now, notice that if we have conjoined constraints which have a specified domain of activity, it should be possible to find all kinds of conjoined constraints in all kinds of domains, and this leads to a further explosion of the number of constraints.

3. Sympathy

A further approach is the Sympathy Theory of McCarthy (1998) and McCarthy (2000). In Sympathy, one has to identify a sympathy candidate to which the optimal candidate tries to resemble. The sympathy candidate is chosen on the base of its faithfulness to input (w.r.t. a certain constraint, called ‘selector’). Another constraint, the sympathy constraint, chooses the optimal candidate.

In the following tableau IDENT-ADJ(Dur) is the selector, and SYM(PATHY) the sympathy constraint (●). Because (kát).bat is not so similar to (ki.tí).bat as (k.tí).bat, (k.tí).bat is the optimal candidate. Dur stands for duration.

(18) Sympathy in Tiberian Hebrew

<table>
<thead>
<tr>
<th>Input</th>
<th>CODACOND</th>
<th>NoComplex</th>
<th>●Max-V</th>
<th>MAX-C</th>
<th>DEP-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opaque a. deʃe</td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Transparent b. deʃ</td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

59
Sympathetic

*! | √ | *

Faithful

*! | *! | √ | *

(18) Sympathy in Arabic (McCarthy 2000)

<table>
<thead>
<tr>
<th>/katab-at/</th>
<th>GRPHARM</th>
<th>SYM</th>
<th>*</th>
<th>IDENT-ADJ(Dur)</th>
<th>REDUCE</th>
<th>IDENT(Dr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opaque</td>
<td>a. (k.ti).bat</td>
<td>*</td>
<td>*</td>
<td>*!</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Sympathy</td>
<td>b. (ki.ti).bat</td>
<td>*!</td>
<td>√</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Transparent</td>
<td>c. (kát).bat</td>
<td>**!</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

(20) Sympathy in Arabic

<table>
<thead>
<tr>
<th>/badw/</th>
<th>*COMPLEX</th>
<th>IDENT(high)</th>
<th>IDENT(high)_IO</th>
<th>DEP-µIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. badu</td>
<td>*</td>
<td>!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. bidu</td>
<td>!</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>c. badw</td>
<td>*!</td>
<td>!</td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

Sympathy-with-Cumulativity (McCarthy 2000) restricts the possible sympathy constraints and at the same time the possible candidates. The candidates are compared w.r.t the unfaithful Input-Output-Mappings. When a candidate K has a superset of the violations that the sympathy candidate has, then these two candidates are in a relation of cumulativity.

4. Turbidity

Turbidity is a new proposal by Goldrick and Smolensky (Goldrick 2000). Instead of modifying CON, the set of constraints in order to account for opacity, Turbidity provides additional to the outputs. One example from Goldrick (2000) is compensatory lengthening in Luganda, illustrated in (21) and (22) in a derivational model. In an underlying string VVC each vowel projects first its own mora. The first vowel is deleted and this deletion causes a delinking of the association line between vowel and mora. In a next step, the resulting floating mora is associated with the remaining vowel with the effect that the vowel is lengthened.

(21) compensatory lengthening in Luganda

/k/a + tiko/ → katiko ‘mushroom’
/k/a + oto/ → ko:to ‘fire place, dim.’
/k/a + ezi/ → ke:zi ‘moon, dim.’

(22) Derivation

/V_1 V_2 C/ → [V_2: C]

Underlying

Project Hiatus- Reassociation Surface
-µ resolution

According to Goldrick this derivation causes opacity, because Project-µ is counterbled by Hiatus resolution. In other words, if the hiatus would be first be resolved, Project-µ would not apply. Vowel length is a result of Project-µ and thereason for the additional mora is not visible on the surface.
Now to Turbidity: The two autosegmental relations of $\mu_1$ are reinterpreted in the following way:
The line barred by the equality symbol shows projection (\(\uparrow\)), an abstract relation between mora and vowel. The (\(\downarrow\)) characterizes pronunciation.
Present moras must be pronounced which means that they must be linked to the root node by means of an association line.

(23) 
$$\mu_1 \mu_2 \\
\uparrow \downarrow \\
V_1 V_2$$

(24) Turbidity in Luganda

<table>
<thead>
<tr>
<th>/ $V_1 V_2$/</th>
<th>*VV$\uparrow$</th>
<th>MAX</th>
<th>V-WT</th>
<th>PRONOUNCE-$\mu$</th>
<th>$\mu$+RT</th>
<th>PRONOUNCE- RT $\downarrow$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [V$_2$]</td>
<td>$\mu_2$</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\vert$ V$_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [V$_2$]</td>
<td>$\mu_2$</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\vert$ V$_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. [V$_2$]</td>
<td>$\mu_2$ $\mu_2$</td>
<td>$\uparrow$</td>
<td></td>
<td>$\vert$ V$_2$</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. [V$_1$, V$_2$]</td>
<td>$\mu_2$ $\mu_2$</td>
<td>$\uparrow$</td>
<td></td>
<td>$\vert$ V$_2$</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>e. [V$_2$]</td>
<td>$\mu_2$ $\mu_2$</td>
<td>$\uparrow$</td>
<td></td>
<td>$\vert$ V$_2$</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Candidates a. b. c. and e. realize only $V_2$. Candidate d. in contrast realize both vowels and is eliminated by the first constraint which prohibits hiatus. Candidate a. is eliminated by MAX. The second candidate is eliminated by Vowelweight, the constraint requiring that vowels project moras, thus that vowels possess some weight. Candidate c. is also readily eliminated, because the mora projected from the first vowel is not pronounced. The last candidate is the only one which survives, in which both vowels project a mora and in which both moras are pronounced. This candidate violates two constraints. The first, Reciprocity, says that when a vowel projects a mora, then this mora must be realized by the same vowel and the last constraint posits that a present root must be pronounced. However, these two constraint are low ranking and their violation is not fatal.

5. Summary so far

Let us take a critical look at the different approaches to opacity just discussed.

5.1 Output-Output-Faithfulness
Output-Output-Faithfulness is in our view a useful extension of Standard OT, which can be applied in various domains of the grammar. However, it can only be used when comparison is made with existing output forms, since it evaluates candidates on the base of a grammatical output. Burzio’s gradient attraction approach is subject to the same limitation. Only
grammatical outputs can function as attractors. But we saw in the counterfeeding and counterbleeding cases that an intermediate step of derivation is sometimes needed, which does not surface. In principle, it could be possible to generalize the outputs serving as s to cases of non-existing intermediate steps. But then, the problem immediately arises as to which of the possible forms are the relevant ones. Another remark is that this solution is equivalent to the sympathy theory.

5.2 Constraint conjunction

In the Constraint Conjunction approach, two problems can be identified. First, constraint conjunction leads to an exponential explosion of the power of the theory though some constraints are just never needed and may exist only as potentialities. It is not clear how to limit this power. The counterfeeding and counterbleeding cases that we saw above and that were shown to induce real opacity cannot be easily accounted for with constraint conjunction. In the case of [de\text{\text influenced}e], for instance, it does not seem possible to account for the epenthesis and the deletion of glottal stop at the same time with the help of conjoined constraints. CODA-COND and DEP or MAX will just not do. The fact is that Constraint Conjunction has been mostly successful for cases like chain shifts (a \rightarrow e, e \rightarrow i, Kirchner 1997), which have nothing to do with opacity. It is thus not clear how Ito & Mester’s analysis of German spirantization can be applied to all cases of opacity.

3. Sympathy

Sympathy can account for the real opacity cases. The first version of sympathy can even account for too many cases, also the so-called. Duke-of-York plus Feeding, which McCarthy show not to exist.

(21) Quasi-Yokuts (not plausible)

<table>
<thead>
<tr>
<th>Underlying</th>
<th>/maat/</th>
</tr>
</thead>
<tbody>
<tr>
<td>R₁: Epenthesis</td>
<td>maati (the trimoraic syllable is resolved)</td>
</tr>
<tr>
<td>R₂: Palatalisation</td>
<td>maat\text{i} ([t] \rightarrow [t\text{\text influenced}] before [i]: R₁ feeds R₂)</td>
</tr>
<tr>
<td>R₃: Apocope</td>
<td>maat\text{i} ([i] is deleted word finally)</td>
</tr>
<tr>
<td>R₄: Shortening</td>
<td>mat\text{\text influenced} (the trimoraic syllable is resolved)</td>
</tr>
<tr>
<td>Surface</td>
<td>[maat\text{i}]</td>
</tr>
</tbody>
</table>

Relatively to maat, mat\text{\text influenced} has changed in two respects: first, the vowel has been shortened and second, there is an unmotivated palatal fricative on the surface. This derivation contains two different processes to the same effect, namely reduction of a trimoraic syllable. Such a derivation should be made impossible both in OT as in every other theory. But Sympathy can account for something like that without ado. McCarthy (2000) restricts the power of Sympathy to cases in which the output has a superset of the violations of the sympathy candidate (Cumulativity). Only faithfulness constraints are possible restrictors and sympathy constraints. Sympathy is successful because it imitates derivations. However, sympathy suffers the same drawback that was mentioned w.r.t. derivations, namely its arbitrariness. In principle all kinds of faithfulness constraints can be violated and all kinds of candidates can serve as sympathy candidates. Moreover, as Ito & Mester observes, sympathy is in contradiction with the spirit of Richness of the Base since it requires faithfulness to a fully specified input (Ito & Mester’s ConstraInt Conjunction approach has exactly the same drawback, by the way).

4. Turbidity
Turbidity is a return to the containment theory in which all information of the input is kept in the candidates. The candidates which have cancelled crucial information of the input are eliminated. The distinction between markedness and faithfulness is blurred. It is not clear how *defe* and *badu* can be accounted in turbidity, though we are confident that there are several ways to do that.

Until now we saw several solutions to the opacity problem in OT, and the observer may get the impression that different kinds of problems call for their own, different, solutions. However, there is a deep pattern to be found in the diversity of the solutions proposed. All solutions use correspondence and comparisons between members of an input/output paradigm. The most important difference between this approach and the older generative view has to do with the fact that OT allows for surface forms and intermediate forms to be visible, whereas derivational theories consider just one form in turn: When a new step arises in the derivations, all older steps are erased or made invisible. Cyclicity is based on exactly this kind of organization. In OT, all forms, inputs, outputs, related forms and sympathy candidates are available for comparison, and do not disappear.

The different OT approaches to opacity look at related forms in different ways. First, output-output faithfulness uses comparisons at the level of outputs, whereas constraint conjunction and turbidity use only inputs and outputs. Finally, sympathy needs an intermediate level of correspondence, similar to the derivational model.

In all extensions of OT that we have considered so far, some additional machinery has been needed as compared to standard OT. Output-output faithfulness, as well as Sympathy, need to expand what counts as the reference correspondents. Not only inputs can be examined for evaluation, but also other output forms and/or other candidates. The difference between output-output faithfulness and Sympathy is that output-output faithfulness (and Gradient Attraction) takes existing forms whereas Sympathy takes intermediate forms equivalent to the derivational steps of the older generative phonology. Constraint Conjunction increases the kinds of constraints needed by the theory, and finally, Turbidity adds structure found on the output candidates.

A further thinkable solution has to our knowledge never been considered so far. This solution consists in taking Prince and Smolensky’s notion of Richness of the Base seriously and manipulate the input appropriately. Until now nearly nothing has been said in this script about the input. It has been tacitly assumed to be equivalent to the underlying form of the generative phonology, thus just one form, rather abstract, possibly underspecified and ready to experience all kinds of phonological alternations. This is the view usually encountered in the literature. The solution that we envisage here is the following: Instead of allowing for just one input for each optimal output – something reminiscent of transformationalism – the alternative is to include all kinds of allophonies into the input and let the constraint hierarchy choose between the options. Richness of the Base says that inputs are free.

However, even if inputs are free, outputs are not. In order to be an acceptable item in a certain language, an output has to correspond to an element in the lexicon of this language. An input like /plik/, submitted to the English constraint hierarchy, may well have an optimal output, the segmental and prosodic composition of which are in agreement with the English phonology, but since it has no correspondent in the lexicon (no category, no semantics, etc…) it fails as an English word. Thus, the importance of the lexicon must not be denied, even if it is not clear how to implement it in OT.

Let’s turn to concrete examples that illustrate how Richness of the Base in the inputs can account for opacity. First, consider a simple allophony without opacity, as in the alternating French nasals. In an adjective like *bon* ‘good’, the masculine form ends in a nasal vowel and
the feminine has an oral vowel and a nasal consonant. Two different allophonies are present, one involving the vowel, which can be oral [ɔ] or nasal [n] and one on the consonant, which alternates between [n] and zero. All possibilities are listed in the input, and the constraint hierarchy takes the decision as to which output pair is optimal, as usual. It must be observed that it can be the case that allophones arise, which are never present on the surface, in the present case, /bɔn/ with both a nasal vowel and a nasal consonant.

(1) Allophony in French nasals

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Fem(C)} & \text{*MAX(Nasal)} & \text{*VN}_\sigma & \text{*V} & \text{NOCODA} \\
\hline
\text{bɔn} & \text{☆} & \text{☆} & \text{☆} & \text{☆} \\
\text{bɔ/ bɔ} & \text{☆} & \text{☆} & \text{☆} & \text{☆} \\
\text{bɔn/ bɔn} & \text{☆} & \text{☆} & \text{☆} & \text{☆} \\
\text{bɔn / bɔn} & \text{☆} & \text{☆} & \text{☆} & \text{☆} \\
\text{bɔ/ bɔn} & \text{☆} & \text{☆} & \text{☆} & \text{☆} \\
\hline
\end{array}
\]

Consider now the cases of opacity that we saw before, and first the counterfeeding case [badu]. The only relevant allophones contain only [a], since the vowel height does not alternate in this morpheme. Thus even if the constraint hierarchy contains a relatively high-ranking constraint *a]_α, this constraint has no effect in this particular case because it is ranked below the constraint requiring identity of vowel height between input and output. In an alternating word, one which has allophones with a high vowel as well as some with a low vowel, IDENT(high) applies vacuously, and it is the constraint *a]_α which takes the decision.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{/badw/}, \text{/badu/} & \text{*COMPLEX} & \text{IDENT(high)} & \text{*a]_α} & \text{DEP-μ} \\
\hline
\rightarrow a. \text{badu} & & & \text{☆} & \text{☆} \\
\hline
b. \text{bidu} & & & \text{☆} & \text{☆} \\
\hline
c. \text{badw} & & & \text{☆} & \text{☆} \\
\hline
d. \text{bidw} & & & \text{☆} & \text{☆} \\
\hline
\end{array}
\]

The next case to be considered here is the difference between Honig and Musik, accounted by Ito & Mester with Constraint Conjunction. The crucial difference between the two words is the fact that the final segment in Honig alternates in its continuancy as well as in its voiceness, whereas the last segment of Musik never alternates. IDENT(cont) can have no effect in Honig since this word has allomorphs in which the dorsal is a fricative and others in which the dorsal is a stop. The decision of which of the inputs of Honig emerges as the best candidate is taken by the ranking between *k and *ç. The final segment in Musik has only one allophone, a stop and a candidate like [muziç], which is possible from the phonotactics fails because the dorsal violates Ident(cont).

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{/honig/}, \text{/honiç/}, \text{/honik/} & \text{ALV} & \text{MAX-Seg} & \text{*γ} & \text{IDENT(cont)} & \text{*k} & \text{*ç} \\
\hline
\end{array}
\]

64
Also the compensatory lengthening in Luganda, that Goldrick explains with turbidity, is readily accounted for with allophony. The allophony consists in the vowel quality which varies between [eo] and [e:]. There is thus no input with a short vowel, to which a candidate could be faithful. Thus, a candidate with a short vowel violates the constraint MAX-µ which is high-ranking. The constraint MAX-Seg is violated in the optimal output, but since this constraint is lower ranking than NOHIATUS and MAX-µ, the candidate with a long vowel and no hiatus is better than the candidate with two vowels, which violates NOHIATUS and also than the candidate with a short vowel.

Now to counterbleeding in Tiberian Hebrew. The proposal consists in listing all allophonies in this case, too, which consist in absence or presence of the final vowel and absence or presence of the final glottal stop. Here, however, we have a problem, since [de∫] will be better than [de∫e] in all rankings, as shown in the next tableau.
In this case, one candidate can be singled out and provided with a subscript. The candidate which is rendered special in this way is the one with the maximal number of segments, before deletion and after epenthesis, so to say. One faithfulness constraint is referring to this candidate, which guarantees that it will be better than the transparent candidate.

<table>
<thead>
<tr>
<th>/de?/, /de|, /de(e)/, /de?(e)/</th>
<th>CODACOND</th>
<th>MAX-Seg₁</th>
<th>DEP-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>-&gt; a. de(e)</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. de(e)?</td>
<td>!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. de(e)?</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. de(e)</td>
<td></td>
<td></td>
<td>**!</td>
</tr>
</tbody>
</table>

The solution which consists in enlarging the set of inputs to a grammatical output allows a simple and homogeneous analysis of all cases of opacity in phonology that have been discussed above. Moreover it is in agreement with the basic principle of Richness of the Base. In a nutshell, an opaque form is faithful to allophones present in the set of inputs, but different allophones can be distributed in different part-inputs. If there is no allophony (as in the case of [k] in Musik), faithfulness prevails. In the case of allophony, markedness constraints decide of which allophones is the best one.

It might be asked at this point whether this step does not represent a dramatic impoverishment of the grammar, since allophony is not decided by the grammar alone anymore, but by both the input and the grammar. Thus the role of the grammar is more restricted in this view. The grammar consists in choosing the right one among a certain number of inputs. Allophony is not the result of the grammar but the result of an underspecification of the input, as well as the influence of the context. The answer to this question is twofold.

First, the conception of the input as a completely determined form, subject to underly changes under the influence of rules or constraints is a relict of a purely derivational approach, which views the input as an abstract object, the exact format of which is chosen to make it agree with the rules or constraints of the grammar. The framework of OT allows us to get rid of this conception and to regard the input as a more plastic object, thus coming in different shapes, which is in narrow relationship with the output. Smolensky’s work on the double relationship between input and output, as well as between output and input should be taken seriously. The different inputs give rise to different outputs, but each input is related to all outputs. And vice-versa each output is in relation with all inputs. However, in a particular context, only one output at a time can be realized. this is the result of the constraint hierarchy, thus of the grammar proper.

Second, trivialization of the grammar is not really a concern. Ideally, grammar should be intuitive and plausible. And the idea that one abstract form is stored for each morphologically related n-tuple, which must be processed each time we want to utter an instance of this n-tuple, is not very appealing as a good cognitive process. Much more plausible is the interrelation of several objects differing from each other in term of allophony, which is something that the grammar has to account for anyway, and the retrieving of the correct form according to the grammatical environment.

An ot constraint hierarchy is in this conception the device allowing us to make the right choice in each case.
In OT, the only way for two forms to be very different from each other is to have very different inputs. If their inputs are identical, as predicted by early transformationalism, or just similar, as predicted by OT for related forms, their outputs are bound to be identical or similar, too, since the same phonological material is submitted to the same grammar. In other words, the architecture of the OT grammar makes some predictions about the way in which different but related outputs may differ from each other.

B7 Constraint Conjunction & Ties in Syntax

7.0 In this session, a couple of proposals will be discussed that modify a central claim of OT - that conflict resolution is strictly lexicographic. The first topic is constraint conjunction in the strict sense, the second part is concerned with different ways of dealing with optionality.

7.1 Recall that conflict resolution in OT is lexicographic: there is a hierarchy $H$ of constraints, and $C$ is better than $D$ relative to $H$ iff $D$ violates the highest constraint on which $C$ and $D$ differ more often than $C$. The need for adding a quantitative dimension may e.g. be exemplified with Chomsky's „Barriers“. In this 1986 book, Chomsky proposed that we measure the distance between a phrase and its trace in terms of the number of barriers that have been, crossed.

$\begin{array}{ll}
0 \text{ barriers} & \text{what do you fix} \\
1 \text{ barrier} & \text{what do you wonder how to fix} \\
2 \text{ barriers} & ??\text{what do you wonder how one should fix}
\end{array}$

If it is true that languages/ construction types may differ as to how many barriers may be crossed, then we need to be able to rank e.g. the PARSE constraint between

$k \text{ barriers crossed} > \text{PARSE} > k-1 \text{ barriers crossed}$

CERTAINLY: "$k \text{ barriers crossed}" should not be an atomic constraint. There is an easy way of constructing such constraints from simple ones: Self conjunction of constraints. We say that $\text{CON}^k$ is violated if $\text{CON}$ is violated at least $k$ times. It seems necessary to assume that $\text{CON}^k \gg \text{PRIN} \gg \text{CON}^{k-1}$. By this move, we distance ourselves from lexicographic conflict resolution, however.

It seems more adequate to say that $\text{CON}^k$ is violated if $\text{CON}$ is violated at least $k$ times by the same element/in the same domain! Obviously, two violations by SAME element, as in

*?What do you wonder who bought

may lead to ungrammaticality, while two violations by DIFFERENT elements may be tolerable, as in

what he wonders how to fix has an influence on what I wonder when to fix
7.2. It is a sensible step, then, to assume that different constraints may be conjoined as well, yielding a new and higher ranked constraint. Compare once more

What do you wonder how to fix t
*how do you wonder what to fix t

Adjuncts have to fulfill stricter locality requirements than arguments. This has been expressed by assuming a principle like REF that forbids adjunct chains

REF: a chain is not headed by an adjunct

and by conjoining this with a barriers constraint. Thus, BAR _ & REF means: an element must not be an adjunct and cross two barriers. By assuming

BAR1 & REF > ParseScope > BAR1

we can derive the contrast introduced above between adjunct and argument movement.

7.3. By allowing constraint conjunction, the weighting (compensatory) type of conflict resolution can be represented in OT. OT gets less restrictive, which is certainly not a welcome consequence. The alternative is obvious: Counting the number of barriers crossed by movement need not be the optimal way of capturing islands for movement. It is more likely that we are confronted with a markedness hierarchy!

7.4. But perhaps, there are compensatory effects in syntax. Q-Scope in German may be an example. It has been proposed that it is governed by the following principles

PREF: a quantifier in the prefield takes wide scope
NOM: A nominative quantifier takes wide scope
DIST: Inherently distributive quantifiers take wide scope

There is only one interpretation ∀∃ for

\textbf{Jeder Pianist hat eine Fuge gespielt} every pianist has a fugue played

because the three principles have the same consequences. The following two examples are also not ambiguous, a fact suggesting that neither NOM nor DIST alone can win over PREF.

\textbf{Jede Fuge hat ein Pianist gespielt} ∀∃
\textbf{Ein Pianist hat jede Fuge gespielt} ∃∀

But when both NOM and DIST counteract PREF, ambiguity seems to arise -- there seems to be a cumulative effect of NOM and DIST violations, in contrast to what standard OT implies

\textbf{Eine Fuge hat jeder Pianist gespielt} ∀∃ and ∃∀

There may be a factual problem with this argument, because I find examples such as

\textbf{jeden Studenten hatte ein Pianist aus Polen empfangen} each-acc student had a pianist from Polen received

ambiguous. In general, it seems that cumulativity determines how easy an ambiguity can be detected, perhaps it is not related to the grammaticality of an interpretation at all.
7.5. Sometimes, constraint violations seem to have equal importance. The EPP/SUBJECT can be satisfied in English by movement

\[
a \text{moose has been shot}
\]

or by expletive insertion

\[
\text{there has a moose been shot}
\]

It does not seem to matter whether Econ/Stay or FI is violated in order to respect the EPP. The two principles seem tied. The same seems to hold for Spec,CP in German

\[
der \text{Pfarrer ist gekommen}
\]

\[
\text{the priest has come}
\]

\[
es \text{ist der Pfarrer gekommen}
\]

\[
\text{there has the priest come}
\]

7.6. These two cases illustrate the general observation that OT seems to predict that there should be exactly one optimal candidate, but in reality, we often observe that more than one grammatical output seems to correspond to one input. How can we deal with such optionality?

In the first place, optionality may be only apparent. There IS a meaning difference between movement and expletive structures, and if this difference figures in the inputs, the two structures

\[
der \text{Pfarrer kommt}
\]

\[
es \text{kommt der Pfarrer}
\]

simply do not compete with each other.

Second, EVAL will not always be able to differentiate between the options. Suppose that complementizers do not induce an FI violation, then it will not matter for whether the complementizer is projected or not. Likewise, in a purely representational version of OT, the two structures

\[
\text{That there was a moose shot suggests that a problem will arise}
\]

\[
\text{that a moose was shot suggests that there will arise a problem}
\]

have identical global profiles: both involve one relevant \textsc{econ} and one relevant \textsc{stay} violation each.

Finally, ties between constraints may be introduced into the theory. Ties can be interpreted in various ways. Consider first a Pesetsky-style treatment of complementizers, working with the constraints LE(CP) and TEL

\[
\text{LE(CP): A CP must begin with a complementizer}
\]

\[
\text{(Align (CP, COMP, left))}
\]

\[
\text{TEL: Do no pronounce function words}
\]

A tie between the two constraints will again yield that the complementizer is optional, as the tableau shows.

\[
(1) \quad \text{I think that he is a fool}
\]

\[
(2) \quad \text{I think he is a fool}
\]
In this concept of a tie, the number of violations of both constraints taken together is relevant for evaluation. This gives the correct result for relative clauses:

(3) a man who that I like  
(4) a man who I like  
(5) a man that I like  
(6) a man I like

Wh-expletive insertion in German seems to be another case in point: any combination of Stay and FI violations yields a grammatical result:

Wen denkst du t dass sie meint t dass Fritz liebt  
who think you that she believes that Fritz loves  
was denkst du wen sie meint t dass Fritz liebt  
was denkst du was sie meint wen Fritz liebt  
was denkst du t dass sie meint wen Fritz liebt

In German, complement clauses may either be introduced by a complementizer or involve verb-second order:

Ich denke / I think  
1. dass der Fritz nicht geschlafen hat  
   that the Fritz not slept has  
2. der Fritz hat nicht geschlafen

1 violates FI  
2 violates Stay (Comp) and Stay (prefield)

A Pesetsky-style tie would incorrectly favor 1. The data may thus illustrate a second kind of tie, in which a tie implies a bifurcation into two hierarchies - S is grammatical if optimal relative to one of the hierarchies.

Scrambling may be seen as an instance of such multiple hierarchies (Uszkoreit). We seem to have the constraints:

nom > acc/dat  
animate > inanimate  
definite > indefinite
for word order in German. A sentence is grammatical if it satisfies at least one constraint:

- Dass der Mann ein Buch liest
  - that the man a book reads
  - NOM, DEF, ANIM
- *dass ein Buch der Mann liest
  - *NOM, *DEF, *ANIM

the only ungrammatical option, because it violates all serialization principles!

- dass ein Mann das Buch liest
  - that a man the book reads
  - NOM, ANIM,*DEF
- dass das Buch ein Mann liest
  - *NOM, *ANIM, DEF

A Pesetsky style concept of a tie would incorrectly predict that structures with n violations are blocked by structures with n-k violations.

8. Serial and cyclic OT

8.0 One of the central differences between OT and the Minimalist program lies in the fact that the former is representational, and the latter derivational. In this session, we will discuss whether optimization can or has to be applied to derivations as well.

8.1 The following contrasts suggests that the idea that FI and Stay are „tied“ is too simplistic:
- there seems t to be a unicorn in my garden
- *there seems a unicorn to be t in my garden
- a unicorn seems to be in my garden

The "associate" a unicorn of there cannot move. How can this be explained? Suppose that there insertion is „cheaper“ than movement (if semantically licensed), i.e. STAY > FI. Globally, both candidates have one FI and one Stay violation. But if we build up the tree from bottom to top and compare
- there to be a unicorn in the garden
- a unicorn to be t in the garden

then the former has a better constraint violation profile. Local optimization seems to predict the immobility of a there-associate.

One problem with that argument is that other types of expletives behave differently, recall, e.g.,

- was denkst du wem sie meint t dass er t liebt

Other accounts for there are conceivable ... and we do not really have evidence against the idea that overt XPs simply are not licensed at the surface in the relevant slot!

8.2. The it-expletive may also be the basis of a further argument for local optimization. The ungrammaticality of

- *Bill seems that it was told t that it is raining
may be understood as a Minimal Link Effect (Bill crosses it). The optimal candidate
it seems that Bill was told that it is raining
does not differ globally from the ungrammatical one, except in terms of the MLC. An
obvious question arises: Why can the MLC not be overridden? Suppose, again, that
optimization applies locally when the tree is built up
it was told Bill that ...
Bill was told t that

Apparently, the Case Filter violation forces Bill to move locally. After that, it may be
inaccessible to further extraction. Local optimization seems to presuppose that we also
have some idea about the accessibility of non-peripheral material in a local domain.
Finite clauses (properly understood) are indeed islands for Case and related processes!

8.3. Shape Conservation effects/MLC may be a further argument in favor of local
optimization. Shape conservation/MLC implies that only the upper of the two objects
may become the subject in a passive (at least in British English), because only then,
their hierarchical relations remain unchanged: the overall shape of the clause is
preserved.

He was given t a book
*a book was given him t

But subsequent wh-movement may then prepose the lower object, as in
what was he given?
These are the hands we're given
In the final representation, c-command relations are thus not preserved. If we try to
preserve them in the final representation, ungrammaticality arises.

*what was given him

Stepwise local optimization can solve the problem. As long as it is possible (the passive
step) to respect shape conservation, one preserves it. Otherwise, we may need traces
to account for effects of intermediate stages of a derivation, and make reference to
them when we formulate the MLC.

8.4. Recall our parallelism constraint on coordinate structure interpretation:

-a boy admires every teacher
ambiguous with respect to scope
-a boy admires every teacher, and Mary does, too
unambiguous!

The unavailability of a bound reading for her in the first example suggests that
quantifiers may be moved to higher scope positions only if this has a semantic effect -
but there cannot be scope relative to Bill, i.e., a proper name cannot be crossed.

#Bill wants to date every girl in his class, and has already asked her out
A student wants to date every girl in his class, and has already asked her out
every man [some woman loves t]
*every man [Bill loves t]
Apparently, whether this constraint banning quantifier movement without a semantic effect is fulfilled or not is determined individually in each conjunct, and NOT globally!! (There would have been a semantic effect in the coordinate examples!).

8.5 Thus, taking various observations together, several constructions seem to find an easy treatment if the grammatical account may make reference to the derivation of the structure. This would be accounted for if optimization is local & cyclic.

Cyclic Optimization proceeds as follows

Take a numeration for a simple PAS.
Have GEN generate the candidates.
Have EVAL select the best one.
Take the result and feed it into GEN once more ....

Restrictions on what can be done to already evaluated substructures are obviously necessary, because otherwise the final representation would also have the final word --- Candidates for the cycles to which optimization applies are:

- every maximal projection
- every „phase“ (vP and CP)
- every bounding node (DP, IP/CP)
- every clause ...

8.6. There are other domains in which global optimization seems to be called for. Recall we get a local MLC violation in the complement clause as in

*I wonder what who bought
whenever this is justified by the final interpretation of the matrix clause. Perhaps, local optimization can be maintained, if we relativize the MLC to XPs with the same scope features.

Standard cyclic movement as in

who do you think t that he likes
may also be a problem (there is no need for movement in the lower cycle!) unless we find a way of forcing cyclic movement locally. Several proposals in this direction have been made.

A case picked up when a phrase moves higher (Quechua, Hungarian)

Mariya Xwancha-q-ta-n muna-n platanu ranti-mu-na-n-ta
Mary John-gen-acc-af wants-3sg banana exchange
is also not easy to account for (what licenses the genitive in the lowest cycle?). Indeed, we find many different types of examples in which the case of a noun phrase originating in a low cycle is determined in a higher cycle only! Similarly, the applicability of certain operations may depend on where the clause is finally inserted into (there is no topicalization in (derived) subject clauses) - which is easy to handle in a global evaluation procedure, but not in a local one.
9. Take the Best (from left to right)

9.0 Is OT a psychologically realistic model? The standard answer is: No! Why? Because the architecture of OT seems to force us to compare millions of candidates if it is applied in online-processing - and it is not likely that we do so. We think the standard answer is incorrect. Parsing with an OT syntax allows to predict a fair amount of results w.r.t. parsing preferences.

9.1. The conclusion can be arrived at within principle-based parsing: The abstract principles of syntax are applied on-line when we parse a sentence. Grammar plays a role in parsing. There is no pre-compilation. Standard examples involve the Theta-Criterion:

θ-Criterion (non-standard formulation)
Each argument expression (e.g. each noun phrase) must be linked to an argument place of a verb, and each argument place of a verb must be linked to an argument expression.

The claim is that "the θ-criterion attempts to be satisfied at every point during processing [...]" (Pritchett 1992:12). We thus predict the preference for the complement clause interpretation of the clause introduced by that in

He told the girl that ...
the father had kissed the child (preferred)
the father had kissed the story (dispreferred)
because in the upper example, the theta-criterion is satisfied earlier (by that) than in the second one (by the story)

One of the standard problems of principle based parsing is that principles that are surface true do not have enough predictive power for making really interesting predictions on parsing.

9.2 In OT, we can claim that the grammar and the parser have close to identical structures

Grammar: The top-down perspective
Parser: The left-to-right perspective

Recall OT assumes very general (but violable) principles with high predictive power - not just for grammaticality, but also for parsing, when we apply them from left to right.

Here is an (oversimplified) example. In German, the segment

welche Frau ....
Which woman

is locally ambiguous between a nominative and an accusative reading - but the sentence is always parsed with a subject preference. We can derive that preference from an immediate application of the EPP. Note a GB approach could not do the same - because it could not assume that the EPP is satisfied in German, in the light of

mir wird geholfen
9.3. The key issue in OT parsing is, of course: How many representations can we evaluate in parallel? We need to identify an algorithm for constructing the best candidate (Tesar 1995).

There are three operations for an input (segment of a clause)
- Parsing an input (the words are placed into syntactic structure)
- Underparsing an input (Never chosen (incrementality))
- Overparsing an input
  Postulate nodes for items that have not yet been part of the input string

**Overparsing** builds up structure in the absence of words filling it - this additional structure encodes grammatical decisions concerning the status of the words -- i.e., the overparsed structure may represent parsing preferences! Overparsing is applied repeatedly as long as the constraint satisfaction profile improves thereby. One of the enemies of overparsing: Fill (no nodes not dominating words) = OblHd.

9.4. An example for overparsing: Suppose we postulate

\[
[\text{DP welche Frau}]
\]

as the initial structure for welche Frau. This structure violates:

- Wh-In-Spec: A wh-phrase must be in the specifier position of a CP

If we want to avoid a wh-in-Spec violation, we need to overparse:

\[
[\text{CP [DP [\text{CP [DP welches} welche [NP Frau]]] [Comp [Inf]]] [Comp [Infl]]] [Infl [VP]]]}
\]

Is this step of overparsing justified? OblHd counteracts overparsing, but it seems OblHd is dominated by Wh-in-Spec. Note that

- CaseFilter: The Case of a Noun Phrase must be checked!
- must also be respected. For this, we have to postulate

\[
[\text{CP [DP [\text{CP [DP welches} welche [NP Frau]]] [Comp [Inf]] [Comp [Infl]]]} [Comp [Infl]] [Infl [VP]]]}
\]

which violates OblHd twice. Lacking evidence to the contrary, we postulate Case Filter >> OldHd.

The Bijection-Principle

There is a one-to-one correspondence between operators (e.g. wh-phrases) and variables (e.g. traces) forces us to assume the following representation

\[
[\text{CP [DP [welches} welche [NP Frau]]] [Comp [\text{CP [Infl [VP]]}]]]
\]

a representation of the subject preference.

9.5. What we expect to happen in OTP is the following: a low constraint might have a considerable effect on left to right parsing because the syntactic information that forces it to be overridden comes very late. Such an effect CANNOT arise with precompiled grammars. For Lyn Frazier, such configurations are the testing ground for the competing models.
Here is the abstract structure of the example we need:

Two principles P and Q, with P > Q
a kind of construction with a left-to right structure, such that at the point X, Q is only relevant
.... X ....
Q favors S, P irrelevant
and a later point, when Y forces the considerations of high-ranked P-
.... X ... Y....
P favors S’ >> reanalysis effects

This translates into the following concrete structure of the example we tested:

..DET N Relative pronoun .... Verb
Case Agreement Govt.

The case of the relative pronoun is subject to two influences
GovCase: NP realizes the case it is governed for by V, P, etc.
    which CAN be checked only at the end in verbfinal relative clauses!!
AgrCase: If NP1 and NP2 are (semantically) coindexed, they agree in Case
    which can be checked immediately after the head noun has been parsed!

When the case of the head noun is determined, a case ambiguous relative pronoun can be shown to temporarily take this case over (as predicted by AgrCase) - so reanalysis effects arise whenever the late Case governor turns out to require something else! If our interpretation of the experimental results is okay, the results support OTP in a strong way.

With few exceptions, parsing preferences can be derived by simply working with the principles of an OT grammar in a left to right fashion. An exception is
a gift to a boy in a box
What we need also need for a complete picture of processing with OT:

a theory of perceiving ungrammaticality (conflict of markedness constraints with constraints that require the analysis to be faithful to the input)
a theory of parsing difficulty
a theory of reanalysis

Suggested Readings
Barbosa, Pilar, Danny Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (eds.) (1998) Is the Best Good Enough?
   Cambridge: MIT Press.
   University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

References


