Poznań Studies in Contemporary Linguistics 44(3), 2008, pp. 401–421

doi:10.2478/v10010-008-0020-0

LINEARIZATION IN BARE PROSODIC STRUCTURE

MARTIN HAIDEN
Université François Rabelais de Tours
haiden@univ-tours.fr

ABSTRACT

This paper defends the claim that the linear order of morphemes in words is determined by templates. Templates are defined as parametrized parsing rules over sets of CV syllables. The article addresses three potential challenges for this view: (i) templatic order may violate a certain interpretation of the Mirror Principle, (ii) templates (viewed as parametrized rules) are not immediately compatible with the view that parameters are set in the lexicon, and (iii) templates (viewed as phonological structure constraining the distribution of syntactic terminals) constitute a case of systematic look-ahead from syntax into phonology. With respect to (i) the paper argues that the Mirror Principle is a generalization over hierarchical, rather than linear order; concerning (ii), it presents inter-species comparative evidence suggesting that templatic ordering rules are learnable. (iii) is identified as a general property of morphological structure that needs to be addressed by any theory of the phonology-syntax interface. The paper presents a theory of parallel derivation that appears to be particularly well suited to this end.

KEYWORDS: PF interface; template; bare prosodic structure.

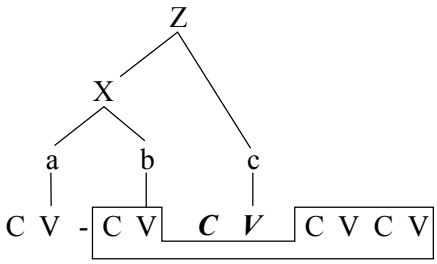
1. Problems with templates

Linear order is not always in strict correlation with hierarchical structure. In syntax, the most evident case at hand is phrasal movement: a fronted wh-constituent, for example, remains linked to the position of its first merger (and it may be reconstructed into this position), even if it is pronounced in another one. In word-structure, head movement may change the linear order of morphemes, but it is usually assumed that this rule is much more constrained than phrasal movement. In particular, Baker's (1985) Mirror Principle is used to guarantee that the linear order of morphemes in words is the exact mirror image of syntactic heads in phrases. Numerous problems have been identified for this strict identification of morpheme-order and syntactic hierarchical structure (Anderson 1992). They notably include non-concatenative morphological marking. The present paper does not share Anderson's skepticism with respect to the segmentability

of markers in non-concatentive systems. Following a sizable literature on the phonology of mainly Afroasiatic languages (Bendjaballah 1999; Bendjaballah and Haiden 2008; Guerssel and Lowenstamm 1990, 1996; Lowenstamm 1996; Lowenstamm and Kaye 1986; Ségéral 1995), it assumes that so-called "non-concatenative" marking of morphological features is contingent on the presence of designated positions in a template. In other words, I assume that there is no truly non-concatenative morphological marking. However, a concatenative analysis of so-called non-concatenative phenomena does not answer to the challenges these phenomena pose for the Mirror Principle.

Consider as an example the configuration in (1), from Guerssel and Lowenstamm's (1990) seminal discussion of Classical Arabic. In this configuration, an affixal marker *a* is concatenated with a root *b* to form constituent X. X is then further concatenated with an infixal marker *c*, to form category Z. Linearly, *c* is closer to the root than *b*: *c* included in the string *b*, *a* is left-adjacent to it. Hierarchically, the inverse is true: *c* c-commands *a*.

(1) Problematic templatic representation

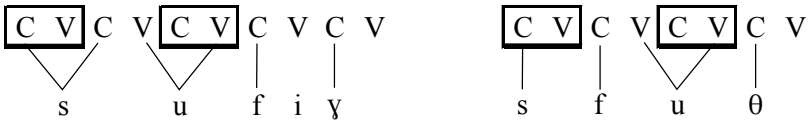


The Berber causatives in (2) exemplify the configuration in (1). Bendjaballah (2006) identifies three morphemes in these forms, a root, the causative prefix *s-/ss-* and the vocalic alternation that marks aspect. She argues that all of these morphemes can be associated with segmentable phonological structure: the first CV unit encodes the causative, the third one encodes aspect, and the remaining positions host root segments.

(2) Berber causatives (Bendjaballah 2006)

Aorist	Perfective	Gloss
ss-afəɣ	ss-ufəɣ	'make fly'
s-faθ	s-fuθ	'make miss'
s-xiḍ	s-xaḍ	'make sew'
s-ʁiwəl	s-ʁawəl	'make hurry'
ss-ali	ss-uli	'make go/climb up'
ss-ajəs	ss-ujəs	'make hurry'

(3) Linearization of markers



If Bendjaballah's description is correct, then Berber causatives instantiate the problematic representation (1): the higher aspectual marker is surrounded by root positions, while the lower causative one is left-adjacent to the root. In this sense, the aspectual position is "closer" to the root than the causative one. If we assume that the linear order of morphemes reflects their hierarchical arrangement in syntax, then we have a problem.

(4) The template problem for linearization
In templatic forms, syntactic hierarchy is not uniformly mapped on linear order.

The problem with templates is serious, because templates are not exceptional. At least in some languages, as in our Berber example, they make up the core classes of productive morphology. So templates must be allowed by Universal Grammar. But how should they be formalized?

The formalization of the notion of a template must take into account the fact that the positions in templatic forms do not necessarily encode the same feature in all contexts. In many cases, they are just designated sites that make it possible to isolate a marker.[1] Take the boxed positions in (3). Bendjaballah (2006) argues that the internal position can encode the aorist, the perfective, and the intensive. In the initial position, we find the causative, the intensive, the reciprocal and the passive. Similar observations can be made about the Classical Arabic templates that will be discussed in Section 3. There, the second position may host either causative or reciprocal markers. In the German examples in Section 5, the prepenultimate position encodes either tense, or causative.

The generalization across all of these cases is that the positions in templates are not markers by themselves; rather, they make it possible for the parser to find the markers in a non-concatenative form. Consequently, templates must be abstract morphophonological structure. Unlike syntactic structure, though, templates are obviously learned as part of the grammar of a particular language. This raises our next problem. According to the dominant view about parameter setting in recent generative grammar,

[1] As pointed out by A. Holmberg (p.c.), this observation can be compared with the reasoning that led to the assumption of the Extended Projection Principle (Chomsky 1981): clauses contain a designated position that hosts subjects. The subject position may be filled by various syntactically and semantically different objects like external arguments, raised DPs, clauses, or expletives, but it is always present as such.

parameters are set in the lexicon, as features of lexical items. If templates are structure affected by linguistic variation, then they apparently challenge the dominant view of parameter setting.

(5) The template problem for parameter setting: learned rules
Templates are learned structure.
Parameters are set in the lexicon (Borer 1984; Chomsky 1995; Manzini and Wexler 1987).

There is a final problem with templates. We would like to maintain that morpho-syntactic features are essentially manipulated in syntax. This position, expressed in Baker's Mirror Principle, is shared by other frameworks like Distributed Morphology. Now, templates are *phonological* structure that governs the distribution of morpho-syntactic features. This should not be possible, if the phonological features of lexical items are accessed after the syntactic derivation has been completed. Templates thus constitute a case of systematic look-ahead from syntax into PF.

(6) The template problem for the interface: look-ahead
Templates are *phonological* structure that constrains the distribution of morpho-syntactic features.
Syntactic Hierarchical Structure All the Way Down (Halle and Marantz 1994).
Late Insertion (Halle and Marantz 1994).

The paper addresses these three problems in turn. Section 2 argues that look-ahead is a general property of morpho-phonology. Therefore, it cannot be used as an argument against the claim that templates are structures. Section 3 turns to learnability and argues that the type of structure instantiated by templates is easily learnable. If templates (viewed as learned structure) challenge the dominant view of parameter setting at all, then the latter, not the former is in trouble. Section 4 presents a positive theory of the phonology-syntax interface that predicts the existence of constraints over the number and type of affixes. Section 5 applies the theory to German *umlaut* causatives, and it elaborates on the relation that holds between templatic patterns and hierarchical syntactic structure. Section 6 extends the scope of templatic parsing rules to capture syntactic parameters, and it places them in the larger perspective of the acquisition of parameter values. Section 7 is the conclusion.

2. Look-ahead: Borer's (2003) indexed phonological representations

If syntax does morphology and lexical insertion is late, as it is assumed in Distributed Morphology, then morpho-syntactic objects must be generated without reference to their phonological properties. Hagit Borer argues that this is impossible. The syntactic

properties of derived nominals in English and Hebrew can only be predicted, if the overtness of the derivation is taken into account. The syntactic properties of derived nouns cannot be predicted from lexical-semantic features alone. Consider the contrast between a complex event nominal as in (7a), and a result nominal as in (7b):

(7a) The instructor's (intentional) examination of the student.
(7b) The instructor's examination/exam.

Only complex event nominals project argument structure. Result nominals don't. So this is clearly a syntactic property. In addition, only those nouns that are overtly derived from a verbal or adjectival source can be complex event nominals. The problem is that, by hypothesis, the syntax cannot know if a given derivation is overt or not. The argument is as follows.

Take a lexical item with the semantic properties of [*transform*]. On standard assumptions, this lexical item can be inserted into a verbal syntactic frame, to yield the verb /transform/, as in (8).

(8) $|\textit{transform}| + \text{V} \rightarrow [_{\text{V}}\ \textit{transform}]$

Then, still without accessing any phonological information, the verb [transform] can be inserted into a nominal frame, to yield a derived noun as in (9).

(9) $[_{\text{V}}\ \textit{transform}] + \text{N} \rightarrow [_{\text{N}}\ [_{\text{V}}\ \textit{transform}]]$

Finally, the derived noun $[_{\text{N}}\ [_{\text{V}}\ \textit{transform}]]$ is sent off into PF. The lexicon is scanned for an element with the semantic properties of [*transform*] and the syntactic category N, and it returns a list like the one in (10). This list will include *transformation*, but also nouns like *metamorphosis*, *shift*, *turn*, and so forth.

(10a) $[_{\text{N}}\ [_{\text{V}}$ transform]] $\rightarrow$ *transformation*
(10b) $[_{\text{N}}\ [_{\text{V}}$ transform]] $\rightarrow$ *metamorphosis*
(10c) $[_{\text{N}}\ [_{\text{V}}$ transform]] $\rightarrow$ *shift*
(10d) $[_{\text{N}}\ [_{\text{V}}$ transform]] $\rightarrow$ *turn*

By hypothesis, the forms in (10) all have the same derivational history and the same core meaning. The only difference between them is that the output form in (10a) contains several overt morphemes, one of which is a verb; the other output forms are opaque. The syntax must have access to this information (only *transformation* projects argument structure), but how can this happen? If lexical insertion is late, then the syntax must look ahead to PF.

It might be objected at this point that there is a slight semantic difference between *transformation* and *metamorphosis*, *shift*, *turn*, and that it is this semantic difference

that determines the difference in argument structure. This is a bad argument because, by Late Insertion, any event-concept should be permitted in a V-to-N derivation, and thus take arguments. However, even if the argument is admitted as valid, it faces immediate counter-examples. Borer gives the following one. Hebrew has two nouns that express the concept *transformation*. One of them is a native root, the other one is an English loan. There is no semantic difference between these two nouns, but there is a morpho-phonological one. The native Hebrew noun has a verbal counterpart; it is overtly marked as a nominalization. The loan is opaque. As expected, only the overtly nominalized Hebrew noun can take arguments (11a). The opaque borrowed noun can only be used as a referential noun, as in (12b).

(11a) ha-šinui šel merkaz ha-'ir 'al yedey ha-'iriya
the-transformation/change of center the-city by city hall

(11b) ha-šinuy haya madhim
the-change/transformation was amazing

(12a) *ha-transformacia šel merkaz ha-'ir 'al yedey ha-'iriya
the-transformation of center the-city by city hall

(12b) ha-transformacia hayta madhima
the-transformation was amazing

Once more, it seems necessary that the syntax can distinguish these lexical items. Since the only difference between them resides in overt morphology, Borer concludes that the syntax must have access to overt morphology by means of *indexed phonological representations*. Take π to be a reference to an indexed phonological representation that is associated with every lexical item by lexical stipulation. Then, derivations like (10b–d) can be ruled out on the basis that π in (13) refers to the phonological shape *transformation*, not to *metamorphosis*, *shift*, or *turn*.

(13) ([$_N$ *transformation*], π) → **metamorphosis*/*shift*/*turn*

Once the syntax can make reference to indexed phonological representations, look-ahead is general. For each lexical item, there is an indexed phonological representation that the syntax can access. Notice that the argument is not contingent on Borer's example. It can be reproduced with different data, discussed in different frameworks. For example, the phonological rules of stratum 1 in Lexical Phonology feed the morphological rules of stratum 2 (Kiparsky 1982). If morphological structure is built in syntax, then stratum 2 syntax must look ahead to stratum 1 phonology.

In sum, look-ahead is a very general property of morphology, if morphological structure is generated in syntax. Eliminating templates (viewed as parametrized mor-

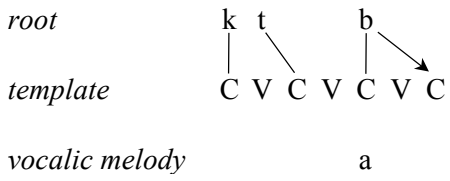
pho-*phonological* structure) from our theoretical vocabulary would not help us solving this problem. I therefore maintain the notion of templates as parametrized morpho-phonological structure and turn to the question whether they can be learned.

3. Templates and rule learning

Let us take a short historical detour and look at the way templates were thought of at a time when phonology was assumed to be a system of grammatical representation, rather than a shallow sound dispenser and filter. In the late 1970s, John McCarty (1979) proposed a theory that allowed the spell-out of morpho-syntactic categories by means of phonological configurations. In his theory, the rules of autosegmental phonology are sufficient to derive the correct output forms, once a root has been concatenated with a CV-template. The theory worked amazingly well in most cases, but it had one major problem that was discussed in Lowenstamm and Kaye (1986) and much subsequent work by these authors: the problem is medial gemination.

Consider as an example the formation of the causative/intensive (form II) and of the reciprocal (form III) in Classical Arabic, exemplified with the root *ktb*: In form II, the root appears as *kattab* 'to make write'; in form III, it appears as *kātab* 'to correspond'.[2] The principles of autosegmental phonology stipulate that melodic elements are associated with skeletal positions one by one, from left to right. Therefore, there should only be final gemination, as in the ungrammatical form (14). The fact that the form II requires the gemination of the medial root consonant must therefore follow from non-phonological conditions.

(14) *katbab (Lowenstamm 2003a (22e))

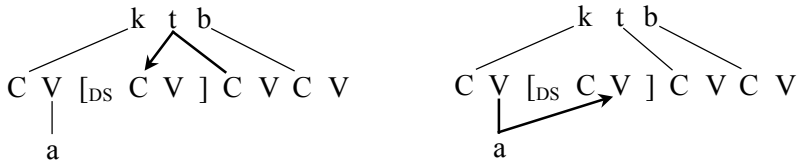


Based on a comparable argument for Tiberian Hebrew, Lowenstamm and Kaye (1986) propose that templates may contain empty positions that become activated for segmental identification after root segments have been linked to root positions. Guerssel and Lowenstamm (1990) apply this reasoning to Classical Arabic. In (15), the CV-position marked as *DS* (*derivational syllable*) is invisible for the association of root consonants with the skeleton. Once the root consonants are linked to the skeleton, a derivational

[2] The discussion will be based on perfective active forms, abstracting away from agreement affixes.

head merges with V in morpho-syntax. At this point, the derivational CV position becomes visible for segmental association, and it is immediately identified. In form II (15a), identification is from right to left, and this yields a geminated medial consonant. In form III (15b), identification is from the left to right, and the result is vowel lengthening.

(15a) form II: *kattab* 'make write' form III: *kātab* 'correspond'



(Guerssel and Lowenstamm 1990 (5a,b); Lowenstamm 2003a (31a,b).)

So much is well known in the literature on Semitic. It is less known that the labeled CV positions in templates like (15) and (1) are amazingly similar to the patterns of CV syllables used by Fitch and Hauser (2004) to test the rule learning capacities of humans and non-human primates. In their experiment, Fitch and Hauser used two distinct sets of CV syllables: one set is spoken by a male voice, the other set by a female voice. The distinction of different classes of syllables allowed them to place tokens of the two classes into specific patterns that correspond to a rule. For example, if one syllable of class A is always immediately followed by one syllable of class B, the pattern exemplifies a Finite State Grammar AB^n. By contrast, if n tokens of class A are followed by the same number of tokens of class B, the pattern exemplifies the recursive rule of central embedding A^nB^n.

Fitch and Hauser familiarized populations of cotton top tamarins and human adults with the two types of rule, and then tested whether their subjects recognized violations of the respective rule. They found that humans learn both the recursive pattern A^nB^n and the Finite State pattern AB^n. The tamarins learned only the Finite State pattern.

Turning back to our Classical Arabic examples, the labeled CV position in (16) can easily be identified by a Finite State rule like (18).

(16) Marker positions in Classical Arabic forms II, III
ABAA
where A is a CV position with root value, B is a derivational infix

This rule differs slightly from the ABAB pattern tested by Fitch and Hauser, but it is still a finite state rule. Viewed from this perspective, it is rather surprising, why abstract rules should pose a learnability problem. Quite the opposite, we would expect to find

templatic rules much more often, in particular in the concatenative type. As a matter of fact we do. Conditions that constrain the number and size of affixes abound in natural language. They are not limited to the non-concatenative type.

4. Formalizing templates

There is of course one major difference between the stimuli used by Fitch and Hauser in their experiment, and the templates of Classical Arabic. In the experiment, classes of syllables were distinguished by different speakers. In Classical Arabic, and in natural language in general, they are distinguished by morpho-syntactic features. We need a theory of the phonology-syntax interface that makes this possible. The ingredients of this theory are a *bare prosodic structure* that is derived like *bare phrase structure*, and a single derivation that spans over the two bare structures.

4.1. Bare prosodic structure

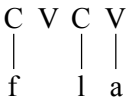
If we take the existence of templates seriously, then the first thing we are forced to accept is a level of grammatically significant phonological representations. This is the level at which templatic rules are defined. A second thing to observe is that templatic ordering rules cannot simply be reduced to syntactic hierarchical structure. They do not obey the Mirror Principle (i.e., observation (4) above). A third thing to notice is about prosodic structure. Prosodic structure can be quite complex, and it is clearly hierarchically organized. In contrast to syntax, though, the widely used prosodic hierarchy (Nespor and Vogel 1986) is not a bare structure: its constituents are structural primitives, introduced by derivation. This is in fact one of the main reasons why PF does not fit into Chomsky's *bare phrase structure* model. The prosodic hierarchy violates the inclusiveness condition.[3]

Of course, defining the constituents of the prosodic hierarchy as primitives is a theoretical choice, not an empirical necessity. Indeed, much work in the framework of Government Phonology that was conducted during the 1980s and 90s was dedicated at deriving particular prosodic construction types from universal principles of syllable-structure. Pushing this reasoning further, Lowenstamm (1996: 419) proposes that "for all languages closed syllables, geminate consonants, and long vowels must be reanalyzed in terms of sequences of light open syllables". On this assumption, a branching onset is represented as a sequence of two light open syllables, where the first V-position

[3] The inclusiveness condition is introduced in Chomsky (1995) as a desirable for phrase structure theories: if possible, phrase structures should be generated as configurations of lexical items, without derivationally created objects like bar levels or indices. A phrase structure that satisfies the inclusiveness condition is called a *bare structure*. Section 4.1 of the text argues that prosody can be analysed as a bare structure as well.

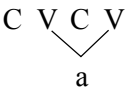
is phonetically unrealized (17a).[4] A closed syllable is represented as a sequence of two CV positions with an empty final V position (17b). A branching nucleus (17c) is represented as a sequence of two CV units with an empty C position between the two vocalic positions. A geminate consonant is linked to the C positions of two subsequent CV units (17d).

(17a) branching onset

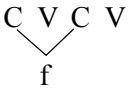


(17b) closed syllable

(17c) branching nucleus

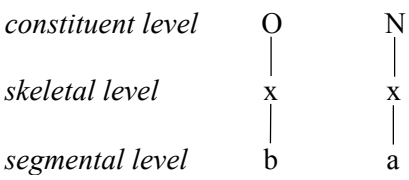


(17d) geminate

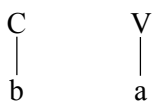


Since neither onsets nor nuclei branch, there is a one-to-one correspondence between syllabic constituents and skeletal positions. This allows us to reduce the skeletal level to the syllabic constituents C and V in strict alternation. The traditional representation in (18a) is replaced by (18b).

(18a)



(18b)



(18b) does not contain derivationally created objects. This means that the inclusiveness condition is satisfied at the level of the skeleton.

[4] Notice that the correctness of the structure in (17a) depends on the status of a given branching onset with respect to the calculation of phonological length. If it introduces a mora of its own, as it does in some languages, then (17a) is the correct analysis. If it does not, then the branching onset must be analyzed as a contour segment, cf. Lowenstamm (2003b) and references cited there.

For higher prosodic structure, the same reasoning can be applied. For example, Vergnaud (2003) derives prosodic constituents as *occurrences* of basic timing units at different levels.[5] In this model, one would have to think of, say, a foot as an *occurrence* of a syllable that embeds a lower *occurrence* of itself, plus a weak syllable. So PF is a recursive, bare structure. There is nothing in the way of integrating it into the general minimalist theory of bare structure.

4.2. Spell-out in a parallel model

Armed with *bare prosodic structure*, we can now address the look-ahead problem. In the standard model, lexical insertion applies at spell-out, long after *merge*. Under this assumption it is necessary that the syntax makes reference to indexed phonological representations – to look ahead.

One way of solving the "look-ahead problem" is to generalize it. We might want to reject the assumption that spell-out applies after syntax. This may sound outrageous, but is it? In fact, at one point Chomsky (2004) proposes something quite similar:

> Assume that all three components are cyclic […]. In the worst case, the three cycles are independent […]; **the best case is** that there is a single cycle only [...] and the three components of the derivation of <PHON, SEM> proceed cyclically in parallel.
>
> (Chomsky 2004:107)[6]

If we take the "best case" seriously, then we will have to say that the phonological and the syntactic-semantic features of lexical items merge – in their respective component – at the same stage of the derivation. How can this be made to work?

A single cycle spanning over multiple components can only work if the derivations in the different components converge. Let us say that derivations in PF and syntax converge, if, and only if, the concatenation of lexical items is transparent across the two domains. In other words, concatenation must be defined in exactly the same way, and it

---

[5] The assumption is not at all a novelty. Any theory that uses metrical grids (cf. Halle and Vergnaud 1987) designates strong beats by copying them into the next higher level. The stars in a given column correspond to the *occurrences* of the respective syllable in Vergnaud's (2003) model. Every *occurrence* is the head of a constituent that includes itself and the weak beats of its domain.

[6] A *PSiCL* reviewer remarks that parallel derivation is defended in Jackendoff's (2002) and work cited therein. Notice, though, that Jackendoff's model constitutes what Chomsky refers to as the "worst case": independent cycles operating in different components or modules. In Chomsky's work, the best case is approached by the (successive) application of spell-out at the phase level, rather than once and for all at the root. A similar direction is pursued in Grohmann's work (cf. Grohmann 2007 for details and references to related claims). The present paper proposes to break down spell-out domains even further into individual lexical items. In this sense, the present claims pursue minimalist concerns to a radical end, rather than being "incompatible with Chomsky's view", as the reviewer writes.

must have the same properties, in PF and in syntax. In still other words, the mapping between syntax and PF must be structure-preserving for concatenation.

Now, it is well known that hierarchical structure in syntax and prosody may differ quite radically. Take for example a VP: *eat an apple*, *have a drink*. In syntax, the verb c-commands its object: the object is structurally lower than the verb. In prosody, the noun is more prominent than the verb.[7] This means that the PF-syntax map cannot be structure preserving for hierarchical structure; and this makes it necessary to separate hierarchical structure from concatenation.

(19) eat an apple, have a drink, ...

Obviously this is quite different from how Chomsky defines *merge*. For Chomsky, the merger of two elements α and β automatically assigns a label to the new object, either α or β. The label is the head of the new object. So *merge* automatically creates hierarchical structure. I doubt the desirability of this formalization.[8] First, because Chomsky's *merge* must be defined as a conjunction of two conceptually and formally quite different operations in any event: pair/set formation and assignment of a label. And second, because of the empirical reason just mentioned: flat merge allows us to formalize structure preservation at the PF interface. Chomsky's merger does not.

Let us say that *flat merge* (20a) selectively introduces phonological and syntactic/semantic features into the respective component. At each time *flat merge* applies, the newly formed object is *parsed* (20b): its PF-side is mapped into syntax in order to check, if the derivation in the two components is still based on the same lexical items – if it converges. Hierarchical structure is introduced by a *labeling* operation (20d) that replaces a non-projecting constituent by a constant. *Labeling* is not structure-preserving under *parse*: hierarchical structure may be different in syntax and PF.

(21) illustrates the general architecture of the model. In this model, look-ahead from syntax into PF does not arise as a problem, because the syntactic and phonological features of lexical items are accessed and procesed in their respective component at the same stage of the derivation.

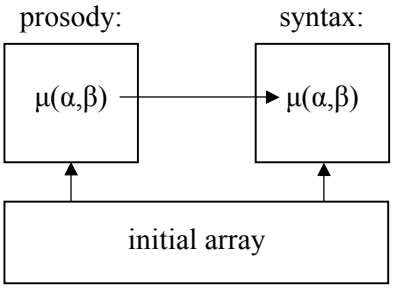
(20) Toolkit (Bendjaballah and Haiden 2003; Haiden 2005, to appear)

flat merge (selective feature insertion):	$\mu(\alpha,\beta) = \alpha\beta$
parse/vocabulary access:	f: $PHON \rightarrow SEM$
structure preservation:	$\mu(f(\alpha),f(\beta)) = f(\mu(\alpha,\beta))$
labeling:	$\Sigma_{\beta \in B}\ \mu(\alpha,\beta) = \alpha K$

[7] For a formalization of prominence on prosodic structure cf. Halle and Vergnaud (1987).

[8] Cf. Section 6 for a formalization of the head parameter that approaches Chomsky's *merge*, restricted to overt categories.

(21) Parallel merge and structure preservation



4.3. The place of templates in the grammar

In the theory (20, 21), templates have a very natural place. They are parsing rules that allow the identification of lexical items in a phonological string.

(22) Template
A template is a parsing rule.

Two questions remain to be addressed. First, how are language-specific templates compatible with the hypothesis that parameters are stored in the lexicon, and second, why don't we find templates in syntax? I postpone the first question to Section 6, and start with the second one. As the averted reader might have expected, there are templates in syntax; but in order to recognize them, we first need to know how they can access hierarchical structure.

5. Marking by occurrence: German causatives

This is where German umlaut-causatives come into the picture. To cut the argument short: German strong verbs have an infixal marker position like Arabic verbs, but this infixal position encodes both root and affixal values. At the level of prosodic terminal nodes, all positions are part of the root. No position is left out, as it was in Lowenstamm's Arabic examples. At a higher level of prosodic structure, an occurrence of the pre-penultimate CV syllable is identified as an infix.

Let us look at the facts. German has a set of verbs that are called *strong*. These verbs do not mark the past tense by an affix, but by an apophonic stem alternation that is exemplified in (23).

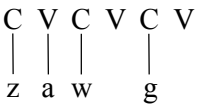
(23) German verbs (Grimm 1819; Halle 1953; Ségéral and Scheer 1998)

	Inf.	Past 3sg.	Past participle	Gloss
strong:	dr**i**ng-en	dr**a**ng	ge-dr**u**ng-en	'penetrate'
	f**a**hr-en	f**u**hr	ge-f**a**hr-en	'go/drive'
	fl**ie**ßen	fl**o**ss	ge-fl**o**ss-en	'flow'
	s**au**g-en	s**o**g	ge-s**o**g-en	'suck'
weak:	schauspieler-n	schauspieler-**te**	ge-schauspieler-**t**	'play-act'

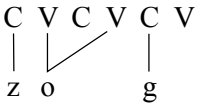
The vowel that is affected by the apophonic alternation is always linked to the prepenultimate CV position of the stem. This means that strong verbs have a common metrical structure. Weak verbs are not constrained in this way.

(24a) Strong – stable template (Bendjaballah and Haiden 2002)

saug-(en) 'suck.inf.'



sog 'suck.past'



(24b) Weak – no constraint

schauspieler-(n) 'play-act'

Apart from tense, the prepenultimate CV position of strong verbs may host markers for agreement, the conditional, the past participle, and the causative. (25) gives examples of deverbal *umlaut* causatives.[9]

[9] The base of *umlaut* causatives can be adjectival, nominal, or verbal (25), and the interpretation of the derived form is always causative, sometimes it has an additional intensive flavor.

(25) *Umlaut* causatives (Bendjaballah and Haiden 2002; Haiden 2005, to appear)

Base	*Umlaut*-causative	Gloss
bitten	beten	'ask' – 'pray'
dringen	drängen	'penetrate' – 'push'
fahren	führen	'drive/go' – 'lead'
fallen	fallen	'fall' – 'fell'
fließen	(ein)flößen	'flow' – 'pour/force into'
liegen	legen	'lie' – 'put down'
saufen	(er)säufen	'drink' – 'drown (caus.)'
saugen	säugen	'suck' – 'breast-feed'
schinden	schänden	'maltreat' – 'violate'
schwimmen	schwemmen	'swim' – 'rinse'
schwinden	(ver)schwenden	'fade' – 'waste'
schwingen	schwenken	'swing' – 'wave/toss'
sinken	senken	'fall' – 'lower'
sitzen	setzen	'sit' – 'put'
springen	sprengen	'jump' – 'detonate'
trinken	tränken	'drink' – 'make drink'
zwingen	zwängen	'force' – 'squeeze'

If you compare the *umlaut* causatives in the second column with their base, a generalization emerges (27). The base is always strong, the derived causative is always weak. (26a) is a typical strong base, (26b) the derived causative. In (26a), the past is marked by a change of the vowel, in (26b) it is marked by a suffix. This generalization follows, if strong verbs and *umlaut* causatives are templatic – if they contain an infixal position that may host different markers, but always just one at a time.[10]

[10] There is one apparent counterexample to this generalization: *wiegen* 'weigh' (as in *weigh 5 kilogramms*) vs. *wägen* 'weigh' (as in *weigh 5 apples*). The causative member of this pair remains strong. However, this pair is exceptional in more than one respect. First, the causative reading of *wägen* can already be observed for the base *wiegen*, as in (i).

(i) Er wiegt die Äpfel.
he weighs the apples

(ii) Die Äpfel wiegen 1 Kilo.
the apples weigh 1 kg

Second, the I/E alternation appears to depend on number agreement. A Google search suggests that causative *wiegen* is common if its subject is plural, but rare if the subject is singular. The stem *wägen* is rare alto-

(26)

	Inf.	Pres. 2sg.	Past 2sg	Gloss
(26a)	sink-en	sink-st	sank-st	‘fall’
(26b)	senk-en	senk-st	senk-te-st	‘lower’

(27) *Umlaut* causatives

(27a) A verbal base can be *umlaut*-causativized, if it contains an available infixal position (i.e., if it is *strong*).

(27b) Causativization by *umlaut* renders the infixal position inaccessible; *umlaut* causatives are weak.

The generalization conveniently extends to pairs like (28, 29), where umlaut is not visible in the infinitive because of the quality of the vowel /e/.

(28)

Base	Causative	Gloss
erschrecken	erschrecken	‘get a fright – frighten’
(er/ver)löschen	löschen	‘go/die out – put out’

(29a)

Inf.	Past	Gloss
erschrecken	erschrak	‘get a fright’
(er/ver)löschen	(er/ver)losch	‘go/die out’

(29b)

Inf.	Past	Gloss
erschrecken	erschreck-te	‘frighten’
löschen	lösch-te	‘put out’

Up to now, (27) looks like a standard templatic effect, just like the marking patterns we know from Classical Arabic: the presence of a marker in a position excludes other

gether, unless it appears with the particle *ab* in *ab-wägen* ‘to weigh’. I conclude that the pair *wiegen–wägen* does not encode a regular causative alternation.

markers from that same position. The German facts are more complicated, because the marker position does in fact have more than one value. There is causative or tense, but then there is always a root value, too. We can find lexical oppositions like (30) for well-behaved strong verbs.

(30) *fallen* 'to fall' – *voll* 'full' – *Fell* 'fur'
liegen 'to be lying down' – *lügen* 'to tell a lie'

This pattern can be derived, if the template of German strong verbs can recognize hierarchical prosodic structure. At the level of prosodic terminal elements – the bottom line of CV positions in (31, 32) – the pre-penultimate is part of the root. But at a higher level (the top line of CV positions), an occurrence of the prepenultimate encodes an affixal value: causative or tense.[11]

(31) *sank* 'fall/sink.past' (Bendjaballah and Haiden 2003)

```
C V              ~>      T
C V C V
C V C V C V      ~>      V
| | |   |
z a n   k
```

(32) *senk-te* 'lowered'

```
C V              ~>  Caus
C V C V
C V C V C V ~>   V        C V ~> T

z ɛ n   k                t ə
```

The templatic parsing rule in (33) recognizes hierarchical prosodic structure.

(33) German strong verbs
$[_A$ (CV*) $[_B$ CV] CVCV]
where A includes root segments, B hosts an affix

[11] For expository reasons, syntactic merger has been ignored in (35,36).

6. Parametric syntax and the lexicon

Once we have allowed templatic parsing rules to recognize hierarchical structure, it seems that there is nothing in the way anymore against an application in syntax – templates that map a phonological string on a syntactic hierarchical structure.

The head-parameter is a good example of such a mapping. The head parameter states that a head of category X takes complements to its right or to its left. Assume it is set to head-final. As a templatic parsing rule, it would have to look like (34). This template states that a prosodic string is to be parsed as a constituent of type X, with a final head X and a left-hand side complement Y. This template is in fact a restricted instance of Chomsky's *merger*: X merges with Y to form X. In addition to Chomsky's generalized operation, (34) is parametrized for linear order, it can be parametrized for both category (the value of X) and subcategorization (the value of Y), and it applies to overt material only - silent categories are not linearized.

(34) head-final
$[_X ([_Y CV^*]) [_X CV^n]]$

Let us return to acquisition. If templates are so pervasive, where are they stored? The model presented in this paper places a lot of explanatory burden on the acquisition of rules, as opposed to properties of individual lexical items. But is this warranted in the literature on acquisition? It turns out that it is.

Consider for example a study conducted by Pléh et al. (2003) on the acquisition of Hungarian postpositions by people with Williams syndrome. In Williams syndrome, spatial orientation is severely impaired, and so is the spatial lexicon. Hungarian postpositions are interesting, because they encode spatial relations, but at the same time they have specific grammatical functions in case marking. So if the lexicon of spatial relations is impaired, then the associated grammatical properties should be absent. However, Pleh and colleagues observe that the grammar of postpositions is fully intact in their population. They conclude that the rule system must be dissociated from the mental lexicon – or, we might want to qualify, that the conceptual component must be dissociated from a grammatical lexicon that admits rules as entries.

This latter approach seems to be adopted at least implicitly even in Wexler's recent work, when he argues that parameters are correctly set at the earliest observable stages. Wexler's basic parameters like OV/VO and V2 encode generalizations over lexical entries, rather than specific features of token lexical items. New lexical items (that are acquired in large numbers after the parameter has been set) follow the pattern rather than altering it. So if Wexler's basic parameters are set "in the lexicon", then the lexicon must be a place where abstract patterns and structural generalizations can be stored; and if this place exists, then of course it should admit templatic parsing rules too.

(35) Very Early Parameter-Setting (VEPS) […]
Basic parameters are set correctly at the earliest observable stages, that is, at least from the time that the child enters the two-word stage around 18 months of age (Wexler 1998).

7. Conclusion

Let us recapitulate. This contribution addressed three problems that templates pose for the theory of grammar: they linearize markers in apparent disrespect of the Mirror Principle, they are learned rules, and they constitute a case of systematic look-ahead from syntax into PF.

It was argued that look-ahead should not be considered a problem, because it is a general property of morphological derivations. In the proposed parallel architecture, the effects of "look-ahead" are accounted for by parallel merger in a single cycle.

Concerning disrespect of the Mirror Principle, it must be noted that the model in (20, 21) actually instantiates a very strict correspondence between morpho-phonological and syntactic derivations. The only part of the Mirror Principle that is lost is its interpretation as a generalization over linear orders. This part has been replaced by templatic parsing rules, and such rules can be used to encode parameters both in morpho-phonology and in syntax.

For the learning problem, it was argued that it is a non-issue. Rules can be learned. They are stored in the same place as Wexler's "basic parameters". In fact, it would be tempting to state Wexler's basic parameters as templates.

Let us conclude with a final word about syntax. The theory outlined above does not order PF after syntax. This entails that in real time, syntactic structure cannot be generated bottom-up, as it is normally assumed. The strictly parallel architecture commits us to the incrementality hypothesis of Colin Phillips, restated as a corollary.

(36) Incrementality Corollary
"Sentence structures are built incrementally from left-to-right" (Phillips 2003).

REFERENCES

Anderson, S. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Baker, M. 1985. "The Mirror Principle and morphosyntactic explanation". *Linguistic Inquiry* 16. 373–415.

Bendjaballah, S. 1999. Trois figures de la structure interne des gabarits. [Unpublished PhD dissertation, Université Paris 7.]

Bendjaballah, S. 2006. "Positions and markers in templates". Paper presented at *XXXII Incontro di Grammatica Generativa*, Florence.

Bendjaballah, S. and M. Haiden. 2002. “Templatic architecture and German strong verbs”. Paper presented at *4èmes journées internationales du GDR CNRS 1954 « phonologies »*, Grenoble.
Bendjaballah, S. and M. Haiden. 2003. “Templatic architecture”. *Recherches Linguistiques de Vincennes* 32. 157–168.
Bendjaballah, S., and M. Haiden. 2008. “A typology of emptiness in templates”. In: Hartmann, J., V. Hegedus and H. van Riemsdijk (eds.), *Sounds of silence. Empty elements in syntax and phonology*. Amsterdam: Elsevier.
Borer, H. 1984. *Parametric syntax*. Dordrecht: Foris.
Borer, H. 2003. “Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon”. In: Polinsky, M. and J. Moore (eds.), *Explanation in linguistic theory*. Stanford: CSLI. 31–67.
Bródy, M. 2000. “Mirror theory: Syntactic representation in perfect syntax”. *Linguistic Inquiry* 31. 29–56.
Chomsky, N. 1981. *Lectures on government and binding*. Berlin: de Gruyter.
Chomsky, N. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
Chomsky, N. 2004. “Beyond explanatory adequacy”. In: Belletti, A. (ed.), *Structures and beyond. The cartography of syntactic structures*. Oxford: Oxford University Press. 104–131.
Fitch, W.T. and M. Hauser. 2004. “Computational constraints on syntactic processing in a nonhuman primate”. *Science* 303. 377–380.
Grimm, J. 1819. *Deutsche Grammatik.* (4 Theile.) Göttingen. [Reprinted 1837].
Grohmann, K. 2007. “The road to PF”. In: Agathopoulou, E., M. Dimitrikapoulkou and D. Papadopoulou (eds.), *Proceedings of the 17th International Symposium on Theoretical and Applied Linguistics*. Thessaloniki: Monochromia. 94–104.
Guerssel, M. and J. Lowenstamm. 1990. The derivational morphology of the Classical Arabic verbal system. [Ms., Université Paris 7.]
Guerssel, M. and J. Lowenstamm. 1996. “Ablaut in Classical Arabic measure I active verbal forms”. In: Lecarme, J., J. Lowenstamm and U. Shlonsky (eds.), *Studies in Afroasiatic grammar*. The Hague: HAG. 123–134.
Haiden, M. 2005. *Theta theory*. Berlin: Mouton de Gruyter.
Haiden, M. to appear. “On bare prosodic structure and the spell-out of features”. In: Grohmann, K. (ed.), *Phase theory: Features, interpretation at the interfaces*. Berlin: Mouton de Gruyter.
Halle, M. 1953. “The German conjugation”. *Word* 9. 45–53.
Halle, M. and J.R. Vergnaud. 1987. *An essay on stress*. Cambridge, MA: MIT Press.
Halle, M. and A. Marantz. 1994. “Some key features of distributed morphology”. *MIT Working Papers in Linguistics* 21. 275–288.
Jackendoff, R. 2002. *Foundations of language*. Oxford: Oxford University Press.
Kiparsky, P. 1982. “From Cyclic Phonology to Lexical Phonology”. In: van der Hulst, H. and N. Smith (eds.), *The structure of phonological representations*. Dordrecht: Foris. 131–175.
Lowenstamm, J. and J. Kaye. 1986. “Compensatory lengthening in Tiberian Hebrew”. In: Wetzels, L. and E. Sezer (eds.), *Studies in compensatory lengthening*. Dordrecht: Foris. 97–132.
Lowenstamm, J. 1996. “CV as the only syllable type”. In: Durand, J. and B. Laks (eds.), *Current trends in phonology*. Manchester: ESRI. 419–441.
Lowenstamm, J. 2003a. “À propos des gabarits”. *Recherches Linguistiques de Vincennes* 32. 7–30.
Lowenstamm, J. 2003b. “Remarks on mutae cum liquida and branching onsets”. In: Ploch, S. (ed.), *Living on the edge: 28 Papers in Honor of Jonathan Kaye*. Berlin: Mouton de Gruyter. 339–363.

Manzini, M.R. and K. Wexler. 1987. "Parameters, binding theory, and learnability". *Linguistic Inquiry* 18. 413–444.
McCarthy, J. 1979. *Formal problems in Semitic phonology and morphology*. New York: Garland.
Nespor, M. and I. Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
Phillips, C. 2003. "Linear order and constituency". *Linguistic Inquiry* 34. 37–90.
Pléh, C., A. Lukacs and M. Racsmany. 2003. "Morphological patterns in Hungarian children with Williams syndrome and the rule debates". *Brain and Language* 86. 377–383.
Ségéral, P. 1995. Une théorie généralisée de l'apophonie. [Unpublished PhD dissertation, Université Paris 7.]
Ségéral, P. and T. Scheer. 1998. "A generalized theory of Ablaut: The case of Modern German strong verbs". In: Parodi, T. (ed.), *Models of inflection*. Tübingen: Niemeyer. 28–59.
Vergnaud, J.-R. 2003. "On a certain notion of 'occurrence': The source of metrical structure, and of much more". In: Ploch, S. (ed.), *Living on the edge: 28 Papers in Honor of Jonathan Kaye*. Berlin: Mouton de Gruyter. 599–632.
Wexler, K. 1998. "Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage". *Lingua* 106. 23–79.