Chunk definition in phonology: prosodic constituency vs. phase structure

Tobias Scheer

1. Introduction[1]

Everybody agrees that the linear string is cut into a number of chunks that are phonologically relevant in the sense that they delineate the application of phonological processes. Everybody also agrees that chunk definition is done outside of the phonology: the existence of chunks is a prerequisite for phonological computation to run. The question that is addressed on the pages below is whether chunks are defined derivationally, i.e. by cyclic derivation which today is called phase theory, or representationally, i.e. by the constituents of the Prosodic Hierarchy.

The two ways of defining chunks are competing for about three decades, but the competition was systematically downplayed, and the theories that are affiliated to each option since the 80s, Lexical Phonology on the derivational and Prosodic Phonology on the representational side, were officially declared to peacefully coexist. That they are in conflict, however, is obvious: Occam commands that there be only one way in which chunks are defined. While peaceful coexistence was practised until recently and both sides tried not to raise the issue (with two notable but short-lived exceptions in the 80s and one modern attempt to predict the assignment of interface phenomena to either means of chunk definition), the conflict has come to the fore as phase theory developed on the syntactic side.

Below the historical situation is recalled, and the recent debate is described: the origin of the latter are phase-based interface theories such as Distributed Morphology, which argue that phase theory is needed for independent (syntactic) reasons, while there is no syntactic or other independent motivation for prosodic constituency. Therefore the Prosodic Hierarchy is redundant and needs to be done away with (Pak 2008, Samuels 2009). A related question is whether translation (i.e. the conversion of morpho-syntactic into phonological vocabulary) and hence modularity (Indirect Reference) is abandoned together with the Prosodic Hierarchy, and whether we therefore move back to Direct Syntax approaches of the 80s (where phonology may directly refer to morpho-syntactic categories). It is shown that there is no necessity or urge to go that way: the fate of the Prosodic Hierarchy is entirely independent of the question regarding the modular architecture of grammar. It is argued that the Prosodic Hierarchy indeed needs to be abandoned, but yet for another reason: because it is a diacritic and hence anyway ruled out on modular grounds. But this does not mean that phonology can make direct reference to syntax: if grammar is modular, it cannot.

The overall result regarding chunk definition, then, is that chunks are only delineated procedurally: the input to phonological computation are pieces of increasing size, and these pieces are defined by phase theory. There is no representational way of identifying phonologically relevant pieces. Finally, the consequences of this position are examined: since syntactically and phonologically relevant chunks only marginally coincide (the only good match is the CP), phase theory needs to be adapted to phonological needs.

---

[1]   The present article is a reorganised piece of Scheer (forth), and makes constant reference to Scheer (2011) where relevant topics from (the history of) interface theory are discussed at greater length. The author is affiliated to Université de Nice - Sophia Antipolis and CNRS 6039 Bases, Corpus, Langage (scheer@unice.fr).

2. Two options for defining phonologically relevant chunks

The two competing ways of defining phonologically relevant chunks, as well as the fact that they are around for quite some time, may be illustrated by nasal assimilation in English. In this language, the nasal of *in-* does, but the nasal of *un-* does not assimilate to a following stop: compare **un***-predictable* with **im***-possible* (and **in***-audible*). The contrast must be due to an idiosyncratic property of the prefixes whereby *in-* is in some way "closer" to the root than *un-* and therefore behaves like monomorphemic NC clusters, which are always homorganic (e.g. *dump, rent, sink*). The classical derivational solution of Lexical Phonology proposes that this more or less intimate relationship be expressed by the fact that *in-* is a class 1 affix, while *un-* belongs to class 2. Nasal assimilation is a level 1 rule, and hence absent from level 2 where *un-* is fed into the derivation. Since *un-predictable* is absent from level 1, it remains unaffected by the rule. By contrast *in-*, being a class 1 affix, concatenates with the root at level 1 and is then subject to nasal assimilation, which is active at this level.

Another derivational implementation is found in Kaye (1995), whose theory of the interface is along the lines of what today is known as phase theory (Uriagereka 1999, Chomsky 2000 and following), even though it predates the syntactic version for a couple of years (see Scheer 2011:§287 for the historical background). Kaye holds that the contrast between the two prefixes is that they induce different domain structure (i.e. phase structure in modern terminology): [[un] [predictable]] vs. [in possible]. *Un-* being spelt out in isolation,[2] its nasal is "frozen" by the PIC (Phase Impenetrability Condition) and cannot assimilate to the following stop on the outer phase. Kaye's phonological version the PIC is a no look-back device that prohibits further modification of material that was already subject to computation on a previous cycle.[3] By contrast, *in-* and the root cohabitate in the same phase and are therefore computed together: the nasal assimilates. Note that unlike in Lexical Phonology, there is only one computational system (rather than two sets of rules, i.e. level 1 and level 2); therefore nasal assimilation applies to every phase. In Kaye's system, the more or less intimate relationship of the two prefixes with the root is thus directly expressed by phase structure.

The competing representational solution builds on contrasting prosodic phrasing: while *un-* is a Prosodic Word (PrW) in its own right, *in-* is part of the PrW of the root, as shown under (1) below. Nasal assimilation is then specified for applying only within PrWs (Rubach & Booij 1984, Rubach 1984:221ff, Vogel 1991)..

(1)   representational solution: separate PrW for *un-*
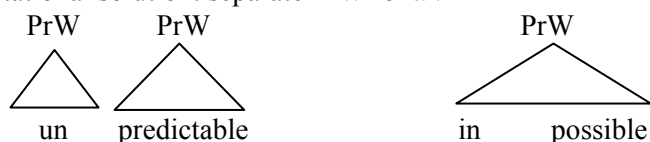
The difference between the derivational and the representational solution is the way in which it is decided that *un-* has a less intimae relationship with the root than *in-*. Prosodic constituency is the result of mapping. Following modular principles, morpho-syntactic structure cannot be accessed directly by the phonology (more on this below). It is therefore translated into vocabulary that can be interpreted by the phonology, prosodic constituency, to which phonological instructions make reference.

By contrast, nothing is translated on the phase-based account: phase structure is created by the spell-out mechanism, which is sensitive to lexically (or universally) defined phase heads: vP and CP on Chomsky's original count, but many more and lower functional projections have been granted phase head status in the literature since then (see the summary in section 8.3).

---

[2]   Newell & Scheer (2007) and Newell (2008:178ff) discuss the syntactic aspects of this analysis.
[3]   Scheer (2011:§302) discusses further detail and provides comparison with the syntactic PIC.

3. The landscape in the 80s: peaceful coexistence that is rarely challenged
3.1. Direct Syntax: no modules, hence no need for translation

In order to describe how chunk definition worked in the 80s, it is useful to begin with so-called Direct Syntax. Direct Syntax was a movement in the late 70s and early 80s that grew out of the feeling that something was wrong with SPE-type boundary theory, which seemed to lead nowhere (Rotenberg 1978, Clements 1978, Pyle 1972, Hyman 1978:459, Kenstowicz & Kisseberth 1977, Scheer 2011:§131 provides an overview). The alternative was to do away with translation altogether: instead of storing morpho-syntactic information in a sponge – hash marks at that time – and then having phonological processes make reference to that buffer, phonology could shortcut and *directly* mention morpho-syntactic information in its instructions. Hence instead of "the phonological event X happens in presence of #" and # representing the beginning of an adjunct, the phonological instruction would directly say "the phonological event X happens at the beginning of an adjunct."

That is, there are no hash marks or any other carriers of morpho-syntactic information anymore: phonology makes direct reference to any morpho-syntactic information that is required.

The founding battle of Prosodic Phonology was against Direct Syntax. The central claim was *Indirect* Reference, i.e. the assertion that SPE (and all other interface theories since the 19th century) were right in providing for carriers of morpho-syntactic information in phonology (hash marks in SPE): translation (of morpho-syntactic into phonological vocabulary) exists, and this is the only way that morpho-syntactic information can reach phonology. The direct mention of morpho-syntactic categories in phonological rules is therefore prohibited.

Representatives of Direct Syntax at that time were for example Kaisse (1983, 1985) and Odden (1987, 1990). The conflict was decided in favour of Indirect Reference within a couple of years, and Prosodic Phonology stands (almost) unchallenged since then.[4]

An issue that for some reason did not play any role in the debate between Direct Syntax and Indirect Reference is modularity (Scheer 2011:§414). The former approach violates modularity, while the latter applies it to language. Translation is a necessary consequence of modularity, i.e. the idea that the mind and grammar are organised in a number of distinct computational units each of which works with a domain-specific vocabulary. Hence there is no way in which phonological computation could understand, parse or process any morpho-syntactic vocabulary such as "adjunct". In Cognitive Science, the fact that each computational system works with a proprietary vocabulary (and hence does not understand the vocabulary of other modules) is called domain specificity (e.g. Gerrans 2002, Fodor 2000:58ff).

Modularity thus defines a major front line among interface theories: those that practice translation (e.g. Prosodic Phonology) are instantiations of modularity, while those which believe that translation does not exist (e.g. Direct Syntax) are incompatible with the modular architecture of grammar that defines the generative enterprise since its inception in the 50s (the inverted T is made explicit in Chomsky 1965:15ff, see also Gardner 1985).

3.2. Peaceful coexistence I: only prosodic constituents define chunks at the word level and above

Since the establishment of representational and hence translation-created chunk definition in the early 80s by virtue of Indirect Reference and in terms of prosodic constituents (Selkirk 1981 [1978], 1984), the question arose – or rather should have arisen – in which relationship the new domains stand with the regular chunks that are defined by cyclic derivation. Derivationally defined cycles were known under various labels (the transformational cycle in SPE, the phonological cycle in Mascaró 1976, see Scheer 2011:§100) since Chomsky et al. (1956:75).

The reason why this obvious issue was hardly debated, and why the "official" solution was one of peaceful coexistence, is the non-cyclic nature of postlexical phonology. A headstone of Lexical Phonology is that there are two distinct computational systems that assess strings of morphemes (i.e.

---

4   This episode of the history of the interface is discussed at greater length in Scheer (2011:§407).

chunks up to the size of words) and strings of words (i.e. chunks of word size and larger). The former set of rules defines lexical, the latter postlexical phonology.

Associated to this distinction is the idea that lexical phonology is cyclic, while postlexical phonology is not. This fundamental claim was never motivated or explained as far as I can see (Scheer 2011:§791), but is supported by the pervasive absence of no look-back phenomena above the word level. That is, there does not appear to be any phonological process among words that is influenced by the fact that a more embedded item was already spelled out earlier (while phenomena of that kind below the word level are well known: affix class-based effects). Today one would say that for some strange reason there are no PIC effects observed when words and larger chunks are spelled out, while they are frequent below the word level. This was called the word spell-out mystery in Scheer (2009b) (see also Scheer 2011:§786).

Whether the empirical generalisation is true or not (i.e. whether there are really no PIC-effects in external sandhi phonology), and if so whatever the mysterious reason for that, the fact is that postlexical phonology is non-cyclic according to the standards of Lexical Phonology. This means that there are no cyclicity-defined chunks above the word level. Hence there is no competition between Lexical Phonology and Prosodic Phonology above the word: the only way that chunks can be defined for words and bigger pieces is by the prosodic constituents of the Prosodic Hierarchy.

In the ideal world of peaceful coexistence, then, chunk definition is both derivational and representational, and the two means of identifying phonologically relevant stretches of the linear string are in complementary distribution according to the size of the items at hand: the grouping of small pieces (morphemes) is defined by cyclic derivation, while the grouping of bigger pieces (words) is done by translation-born domains.

3.3. Peaceful coexistence II: the demarcation line between prosodic and cyclic phenomena

Peaceful coexistence in fact comes in a number of flavours, and the question is how it can be determined whether a given interface phenomenon is driven by a procedural (i.e. cyclic, phase-based) or representational (prosodic constituency) mechanism. The base line answer was reported in the previous section: smaller chunks (morphemes) are the matter of cyclic management, while bigger chunks (words) are dealt with by prosodic categories. This is Hayes' (1989 [1984]) line of attack, which has the merit of drawing a clear demarcation line.

Another approach is Nespor & Vogel's (1986:18f, 27ff): these authors try to divide processes according to whether they make "reference to morphological structure and/or specific morphological elements" (Nespor & Vogel 1986:18). It is shown in Scheer (2011:§435) that this distinction is inoperative: Nespor & Vogel (1986) do not mention that on the count of Prosodic Phonology, mapping rules decide whether or not (and how) morpho-syntactic structure is transformed into prosodic structure. Prosodic constituency encodes "morphological structure and/or specific morphological elements" just as much as cyclic structure does.

Alongside with Nespor & Vogel's soft version of peaceful coexistence, there is a radical version advocated by Rubach & Booij (1984) (among others, see Scheer 2011:§440). On this count, below the word level phonological rules make random reference to either morphological (cyclic) structure or prosodic constituency, and there is no principle that defines which phenomena use one or the other strategy. Lexical Phonology and Prosodic Phonology are thus both candidates for managing everything that is going on below the word level; whether one or the other comes to bear is random and unimportant.

This is what the landscape looked like in the 80s, and there were no innovations for about 30 years. In the recent Stratal OT literature that adapts the Lexical Phonology heritage to the modern constraint-based environment (see Scheer 2011:§483), though, serious effort is made to develop a consistent synthesis of the cyclic and the prosodic constituency-based approach to the interface. In order to make both compatible, the old question regarding the division of labour is addressed anew by Bermúdez-Otero (2011:2025ff) and Bermúdez-Otero & Luís (2009).

These authors develop three criteria in order to tell cyclic from prosodic effects.[5]

(2)   division of labour according to Bermúdez-Otero and Luís
       a.   gradience
           prosodic categories can trigger gradient (nondiscrete) phonetic effects; cyclic domains cannot.
       b.   cyclic locality
           phonology cannot access the internal morphosyntactic structure of cyclic domains, but it can access all levels of prosodic structure.
       c.   coextensiveness
           every cyclic domain is exactly coextensive withy some morphosyntactic constituent; a prosodic category need not be.

This is an interesting catalogue, which needs to be run against a larger empirical record. In any event, as far as I can see it is the first attempt to tell two classes of interface phenomena apart on the grounds of pre-theoretical criteria. These two classes are then equated with the two means by which morpho-syntax can bear on phonology, procedural and representational.

3.4. Elimination of cycles: prosodic domains below the word level

The relationship between procedural and representational devices as conceived in the 80s, though, is asymmetrical: while there is no way to make postlexical phonology cyclic (its non-cyclicity is one of the founding statements of Lexical Phonology), nothing in principle withstands prosodic constituents to extend below the word level, i.e. to also group sequences of morphemes.

This is precisely the line of attack of Selkirk (1984:412ff) and Inkelas (1990): the title of the latter work is "Prosodic Constituency in the Lexicon". Both argue that there must not be several devices that do the same job, i.e. chunk definition. Either all pieces are grouped by prosodic domains, or they are grouped by cyclic spell-out.

Inkelas (1990) therefore argues that the technology which works for all piece types and sizes should be generalised, and the other abandoned. In order to do so, she introduces a new set of prosodic constituents which, just like their higher level peers, are created by mapping rules (Inkelas 1990:33f). Inkelas holds that there are as many layers below the word as there are lexical levels (strata) in Lexical Phonology.[6]

4. Constraint-based instead of rule-based mapping
4.1. Mapping is done in modular no man's land.

In order to approach the modern situation of prosodic constituency, let us have a closer look at the evolution of the central operation of translation, which is a direct consequence of the aforementioned Indirect Reference, and hence of modularity. In a modular perspective, mapping cannot be located in either morpho-syntax or phonology. In the original rule-based environment of the 80s, mapping rules do the job of translation, and have therefore access to both morpho-syntactic and phonological structure: they can read the former, and create part of the latter (the Prosodic Hierarchy). Since modules do not understand the language of other modules, the interface is necessarily located in modular no man's land, i.e. somewhere after morpho-syntax and before phonology. This location of translation was consensual (e.g. Nespor & Vogel 1986).

---

[5]   There is also a fourth criterion based on sociolinguistic variation, which however is theory-dependent.
[6]   Selkirk's and Inkelas' economy-based argument in favour of generalised prosodic domains is discussed at greater length in Scheer (2011:§429).

4.2. Constraint-based mapping: ALIGN

In the new constraint-based environment of the 90s, Prosodic Phonology was adapted to the general OT architecture where phonological well-formedness is defined in relative, rather than in absolute terms: the interaction of a set of ranked constraints defines the least contravening solution among a set of candidates. Applied to the interface, the new perspective may be introduced like this: the Prosodic Hierarchy is a device that represents the (mis)match of morpho-syntactic constituency and phonologically relevant chunks of the linear string. The operation that relates both, mapping, may thus be described as the (mis)alignment of the edges of morpho-syntactic constituents with the edges of prosodic constituents.

This perspective on mapping was introduced by Selkirk (1984:52ff), who builds prosodic structure by alignment rules (e.g. Basic Beat Rules, Demi-Beat Alignment DBA, see Scheer 2011:§426). Constituent margins then become the centre of interest two years later in Selkirk's (1986) edge-based mapping (Scheer 2011:§386). Finally, McCarthy & Prince (1993) generalise alignment to more empirical situations and make it the central tool for the interface with morpho-syntax (see Itô & Mester 1999, McCarthy & Prince 2001:vii, Peperkamp 1995:227ff for a historical overview).

Today ALIGN is a constraint family with a uniform template: the left or right edge of a given unit coincides with the left or right edge of another unit. The units in question may be phonological, morphological or syntactic, and both units involved in an alignment constraint may belong to the same area (phonology, morphology, syntax, see for example Yip 1998).

4.3. Parallel mapping: translation and reference to prosodic constituency are conflated

The following example from German provides illustration of how alignment works (Kleinhenz 1998:39f). The word *auf-essen* "to eat up" is made of a stem (*essen*) and a prefix (*auf*). It is realised with an epenthetic glottal stop *auf-ʔessen*. The glottal stop is non-contrastive: it is always filled in word-initially in case the word begins with a vowel underlyingly.

According to Kleinhenz' analysis, the pronunciation at hand is the result of the ranking ONSET >> ALIGN (Stem, L, PrWd, L) >> DEP. The alignment constraint demands that the beginning of the stem coincides with a prosodic word; regular mapping would express the same requirement by the fact of assigning two separate prosodic words to the prefix and the stem.

Now ONSET demands the presence of a consonant in the first syllable of *essen* – hence the epenthetic glottal stop. The competing solution that syllabifies the prefix-final consonant into the onset of the stem fails because of ALIGN: in *au.fessen* the stem and the prosodic word would be misaligned, and this is fatal since ALIGN outranks the epenthesis-hostile DEP. In a language with the reverse ranking, *au.fessen* without glottal stop would be optimal.

Regular rule-based mapping would describe this variation by a parameter on the assignment of the prosodic word, which will or will not encompass prefixes. That is, the domain of glottal stop insertion is the prosodic word: glottal stops are inserted at their left edge. In German, the prefix makes a prosodic word of its own; the prefix-final consonant cannot join *essen* because there is no syllabification across prosodic word boundaries (in German). *Essen* will therefore be subject to glottal stop insertion. In the hypothetical language where *au.fessen* is encountered, the prefix and the stem cohabitate within a single phonological word: this allows the prefix-final consonant to join *essen*, and no glottal stop will be inserted.

As may be seen, OT produces the winner without mentioning any particular process and thus without any process (or rule) making reference to the Prosodic Hierarchy. The effect is achieved by the interspersing of alignment with regular phonological constraints that are not involved in issues related to the interface (ONSET and DEP).

This system significantly contrasts with rule-based mapping: it conflates a two-step procedure into one single constraint. On the regular account, prosodic constituency is *first* created by mapping; phonological rules *then* make reference to the prosodic structure. These steps are procedurally ordered,

and phonological rules refer to the carrier of morpho-syntactic information in the description of their environment, as much as they may refer to truly phonological conditions.

Neither of these mechanisms survives in constraint-based mapping: there is no procedural difference between the translation of morpho-syntactic into prosodic constituency, and there is no reference made by phonology to this structure. Also, the instructions responsible for phonological computation, constraints instead of rules, do not mention prosodic constituency anymore: ONSET, DEP or other constraints do not make reference to the Prosodic Hierarchy. Prosodic structure appears only in one constraint family, ALIGN, which is specialised in importing morpho-syntactic information into phonology.

4.4. Parallel mapping transfers the translation *into* the phonology and thereby gives up on modularity

On the face of it, the constraint-based architecture looks like an interesting simplification of the mapping mechanism: a two-step procedure whereby mapping precedes reference to prosodic constituency is incompatible with anti-derivationalism and therefore replaced by an entirely parallel one-step mechanism.

The extension of the parallel philosophy to intermodular relations, however, gives up on modularity altogether. Mapping rules have access to both morpho-syntactic and phonological structure since they convert one into the other. Modules are domain specific and hence cannot understand the vocabulary of other modules, or parse their categories. Therefore, translation cannot be done in either morpho-syntax or phonology: it must be done in modular no man's land.

The effect of parallel mapping, however, is that translation is done *in* the phonology: ALIGN constraints are interspersed with regular phonological constraints in the same constraint ranking. This situation is incompatible with a modular architecture. Also note that ALIGN is deeply rooted in OT: together with syllable-related constraints (ONSET, NOCODA), it is one of the rare constraint families that are accepted and practised in all OT quarters. Doing OT without, say, LAZY or NOSTRUCTURE is one thing and certainly workable (there are many versions of OT from which these constraints are absent); doing it without ALIGN is an entirely different job, and it is hard to see today how this could work.

5. Prosodic constituency vs. phases
5.1. Phase-based mapping: phases are "prosodic islands"

More recently, constraint-based mapping was impacted by and adapted to syntactic phase theory. Kratzer & Selkirk (2007) introduce the idea that chunks which are defined by the syntax as phases also grossly correspond to constituents of the Prosodic Hierarchy. Specifically, they propose that "the highest phrase within the spellout domain is *spelled out as a prosodic major phrase*" (Kratzer & Selkirk 2007:106, emphasis in original). We are thus still in the environment that was defined by Selkirk (1986) where three prosodic constituents are recognised: the major phrase (called prosodic phrase then, and corresponding to X" in an X-bar structure), the prosodic word (corresponding to X°), and an intermediate item called minor phrase (small phonological phrase then, corresponding to X'). Hence note that what Kratzer & Selkirk call phrases are regular constituents of the Prosodic Hierarchy (in Selkirk's 1986 version of prosodic constituency), while a phase is a procedurally (and syntactically) defined chunk (also called spellout domain in the quote).

On Kratzer & Selkirk's (conservative: see the discussion in section 8.3) assumption that only CP and vP are phases, the idea is thus that spell-out domains, i.e. CPs and vPs, correspond to major phrases on the phonological side. This is supposed to be a universal equivalence and the result of mapping. Language-specific variation in prosodic phrasing is then achieved not by the syntax-phonology mapping as before (see section 4.3), but purely phonologically by "prosodic markedness constraints, which operate to produce surface prosodic structures that are more nearly phonologically ideal" (Kratzer & Selkirk 2007:126). This is a significant departure from a Prosodic Phonology

essential: mapping becomes universal and phase-driven, while the great amount of language-specific variation in prosodic phrasing is achieved *in* the phonology by purely phonological mechanisms.

Familiar mapping is thus broken into two pieces: one that is done by phases and hence outside of the phonological computation, another that (as before in constraint-based mapping) is located *in* the phonology. Kratzer & Selkirk's mapping thus continues to violate modularity, which commands that translation be done in modular no-man's land.[7]

Other work along the idea that phases form "prosodic islands" includes Piggott & Newell (2006), Dobashi (2003), Ishihara (2007) and Kahnemuyipour (2009) (Elordieta 2008:274ff offers an informed survey).

5.2. Equating prosodic constituents and phases is dangerous for the former: another round of Direct Syntax is lurking

The idea that phases – which did not exist in the 80s when Prosodic Phonology was developed – and constituents of the Prosodic Hierarchy are identical is obviously appealing: both delineate chunks of the linear string that serve as domains for the application of phonological processes, and this is what prosodic constituency is all about. The Occam-based argument that having two independent means of chunk definition is not an option was already made on several occasions. Hence the modern idea in a phase-based environment that when Prosodic Phonology talked about prosodic constituents in the 80s, what was really meant were phases.

This move is dangerous for Prosodic Phonology, though: the question arises what prosodic constituency is good for in the first place if the job can be done by phase theory, which is needed anyway since phases are a general architectural property of grammar that is also central in current minimalist syntax. If having both is redundant and one has to go, no doubt the loser is the phonology-specific prosodic constituency, which can be reduced to phases (but the reverse it not true).

Therefore, quite unsurprisingly the equation of prosodic constituents with phases has led to another round of Direct Syntax: the redundancy-based objection against the Prosodic Hierarchy was made in the mid-80s by defenders of this approach (e.g. Kaisse 1985:156 note 1, 110 note 1, 1990:128f, see Scheer 2011:§409). Prosodic Phonology has burdened phonology with extra arboreal structure plus extra computation, the mapping rules (or today alignment constraints). These devices have no other purpose than making morpho-syntactic structure available inside the phonological module. By contrast, the alternative solution that makes direct reference to morpho-syntactic categories achieves the same result without any additional structure or computation.

This is exactly the line of attack of Pak (2008:42ff, in a Distributed Morphology environment) and Samuels (2009:284ff), who are dissatisfied with prosodic constituency and make the obvious point that if all information necessary is provided by phase structure, the Prosodic Hierarchy is redundant and needs to be done away with (see also Seidl 2001).

In an additional move that does *not* follow from the absence of the Prosodic Hierarchy, Samuels and Pak throw out translation and hence leave modular grounds. This, however, is inconsistent with their own framework, since phase theory is the central device of *intermodular* communication and therefore supposes a modular environment. How could one argue for modularity plus phase theory on the one hand, and violate modularity on the other by eliminating translation and hence domain specificity?

---

[7] Another issue is the fact that there are more layers of the Prosodic Hierarchy than just the major phrase: what happens with them if mapping reduces to the equivalence "phase = major phrase"? In a note, Kratzer & Selkirk (2007:125) hint at the possibility that the "prosodic word could be understood as the spellout of lexical (not functional) heads, while intonational phrase could be the spellout of 'comma phrase.'"

5.3. Reaction by orthodox Prosodic Phonology: prosodic constituents ≠ phases

In the view of some proponents of regular Prosodic Phonology, Selkirk's move to equate prosodic constituents with phases locates her in the camp that strives to abolish prosodic constituency by dissolving it into phases. While this is the reverse movement with respect to Inkelas' (1990) elimination of cyclic structure through the extension of prosodic constituency below the word (see section 3.4), it shows that the peaceful coexistence of cyclic and prosodic chunk definition that was alleged in the 80s has no standing anymore. The conflict now clearly appears, and this is all to the good: the two options are clearly characterised, and under Occam's rule one will have to go if they do the same labour.

The reaction against the equation of prosodic constituents with phases is led by Lisa Cheng and Laura Downing: one of their talks is called "Prosodic domains do not match spell-out domains" (Cheng & Downing 2011a, see also Downing 2010 and Cheng & Downing 2011b).

The alternative that Cheng & Downing (2007, 2009) propose works with alignment constraints that take phases as an argument: ALIGN-L(PHASE, INTP) "align the left edge of a phase with the left edge of an Intonational Phrase." On this view, phase theory also impacts chunk definition, albeit only indirectly: the independence of prosodic constituency through its genesis via ALIGN-based mapping is preserved.

6. The Prosodic Hierarchy is a diacritic
6.1. Interim summary

Up to this point, the historical development of the two chunk-defining mechanisms was reviewed. Thus far the conclusion is that
– prosodic constituency and phase structure are in competition. Both cannot be correct, and one will have to go.
– all actors agree that phase theory is a valid and necessary ingredient of grammar: it mediates between morpho-syntax and phonology.
– therefore, implicitly, they agree that grammar is modular, and hence that translation exists.

What could an alternative look like if prosodic constituency is done away with, but translation maintained? In other words: what will the output of translation look like? This is the question that we turn to now, and which provides an independent argument in the debate between the two chunk-defining mechanisms: the Prosodic Hierarchy does not qualify because it is a diacritic. Hence phonologically relevant chunks of the linear string are defined by phase structure, and what we need is a non-diacritic output of translation.

The argument made is thus conceptual. Of course the empirical situation needs to be carefully considered independently, but this will have to be done elsewhere. The approach to the representational side of the interface (i.e. non-diacritic translation) that is lined out below is called Direct Interface (because there is no buffer between morpho-syntactic and phonological objects, see Scheer 2009a,c, forth).

6.2. Prosodic Phonology lays claim to boundaries: they are the old buffer, prosodic domains are the modern buffer

In an overview article that anchors the legitimacy of Prosodic Phonology (with respect to Direct Syntax), Vogel & Kenesei (1990:344) review the arguments in favour of Indirect Reference. One point they make is historical: all interface theories have been indirect thus far, so there is probably something to this approach. They single out SPE as a forerunner of Indirect Reference.

"Working within the SPE framework, Selkirk [1972] modifies the original proposal by showing that at least in certain types of phonological phenomena, interaction between the two components is only indirect. Word boundaries (#'s) inserted into a string on the basis of syntactic structure determine where external sandhi rules apply. Phonological rules thus do not directly 'see' syntactic structure, but rather access only strings of segments and boundaries." Vogel & Kenesei (1990:344)

Representatives of Prosodic Phonology thus lay claim to the equivalence of hash marks (#s) and prosodic constituency. The same line of reasoning is found in another overview article by Inkelas & Zec (1995). The authors call p-structure the level of representation that mediates between morpho-syntax and phonology; they explicitly identify boundaries as the ancestor of this mediating structure, whose more recent incarnation is the Prosodic Hierarchy.

"An early version of p-structure was proposed in SPE and developed in subsequent work (Selkirk, 1972, 1974; Rotenberg, 1978). According to this view, domains of phonological rules are expressed in terms of phonological boundary symbols, generated by rules. […] Far more constrained is the 'prosodic' view of p-structure. Under this view, p-structure occupies a level with its own hierarchical organization and a high degree of autonomy." Inkelas & Zec (1995:537f)

If thus prosodic constituency is but a more advanced version of boundaries that presents a number of advantages, it must have the same formal properties as its predecessor. The two quotes clearly show that prosodic constituency, just as hash marks, is a diacritic: it serves no other purpose than replicating phonologically relevant morpho-syntactic information in phonology. This is the essence of diacritic translation, which is based on a buffer (or a sponge): phonologically relevant information is stored into a diacritic, which is transported into phonology where its load is released. We have seen that this is true for the original implementation of the Prosodic Hierarchy where translation was rule-based, as much as for the OTed version thereof, constraint-based mapping.

Although proponents of Prosodic Phonology are prompt to point out that SPE-style boundaries are odd because they are diacritic (Scheer 2011:§373), nobody seems to wonder whether the Prosodic Hierarchy could be a diacritic as well. Nobody seems to worry either when the Prosodic Hierarchy is advertised as the direct surrogate of arbitrary and diacritic boundaries (see also Scheer 2008).

6.3. Definition of the term "diacritic"
6.3.1. A diacritic is an alien

A formal definition of what exactly counts as a diacritic must rely on the alien status of the object in question in the environment where it evolves. A workable definition appears under (3) below.

(3)   definition of the term "diacritic"
        a diacritic is a non-native object in module X: it is only used when information from outside of X is processed. It is absent from events that do not appeal to extra-Xal information.

Hash marks and omegas ω (i.e. prosodic words) meet these conditions: they are non-phonological intruders in the phonological world which are injected for the *exclusive* purpose of storing extra-phonological information.

Also, they are systematically absent from phonological processes that do not use extra-phonological information. For example, an ordinary palatalisation that turns k into t͡ʃ before front vowels involves consonants, vowels, velarity, palatality, occlusion, affrication and the like, i.e. all pieces of the domain-specific vocabulary that is used and managed in phonological computation. Such a process does not appeal to any extra-phonological information: this would only be the case if the

description were, say, "k turns into t͡ʃ before front vowels, but only in case there is a morpheme boundary between the trigger and the target."

Hence there is an objective and pre-theoretical means to tell processes apart that do or do not use extra-phonological information. Therefore, we can be sure that *only* domain-specific vocabulary, i.e. the one that is used in the computational system which carries out phonological computation, occurs in the latter. By contrast in the former, the information that is processed by phonological computation is blended: the specifically phonological vocabulary cohabitates with carriers of extra-phonological information.

If some item, then, never occurs in the "pure" processes, i.e. those that only use specifically phonological vocabulary, we can safely conclude that it is an alien. Obviously, hash marks as much as omegas (prosodic words), phis (prosodic phrases) or any other prosodic constituent from the prosodic word upwards,[8] are never found to participate in processes that do not use morpho-syntactic information. For example, there is no palatalisation of the pure kind "k turns into t͡ʃ before front vowels" where a hash mark, an omega, a phi or anything of that kind is needed. Therefore all of these items are diacritics.

6.3.2. Apples and bananas in phonology, but not in syntax

Another obvious issue is that like boundaries, the units of the Prosodic Hierarchy are arbitrarily chosen and named: "ω" (the phonological word), "φ" (the phonological phrase) etc. are not any less arbitrary than "+" or "#". Calling a unit whose exclusive purpose is to store and release some information a hash mark, an omega, a banana or an apple is indifferent: any name will do the job.

Saying that an omega is only shorthand for a real linguistic object, the phonological word, does not help: the same may be said about + and #, only that a regular scientific-sounding terminology has never been introduced for these objects.

Finally, pointing out that omegas and phis represent certain stretches of the linear string which coarsely correlate with morpho-syntactic divisions does not make them less arbitrary. Everybody knows that the linear string is chunked into stretches that define the domain of application of phonological processes, and that these stretches more or less closely follow morpho-syntactic structure. The issue is not the coarse equivalence between morpho-syntactic structure and phonologically relevant stretches – it is the nature of the items that are supposed to be inserted into the phonology in order to carry this information.

Finally, it is interesting to observe that only phonologists seem to be happy to live with apples and bananas in their theory: there is no equivalent in morphology, syntax or semantics. No representative of these disciplines would accept, say, an ωP (omega phrase), a #P (hash mark phrase) or a 🍌P (banana phrase). Nodes in morpho-syntactic structure project something, and this something is recorded in the lexicon: items only qualify if they belong to the domain-specific vocabulary of the morpho-syntactic computational system (number, gender, person etc.).

6.3.3. The output of translation has always been a diacritic

The phonological literature stigmatises some diacritics as arbitrary diacritics, while others are advertised as "truly phonological objects". Nespor & Vogel (1986:3) for example call boundaries "pseudo-phonological terms" and argue that phonology should only be able to refer to truly

---

8   The heterogeneous character of the Prosodic Hierarchy, i.e. the cohabitation of bottom-up (syllables, feet, morae) and top-down constructions (prosodic words and higher) is admitted and made explicit in the literature (e.g. Nespor & Vogel 1986:109, Nespor 1999:119, see also Scheer 2011:§401). While syllables, feet and morae are projections of (phonological) terminals, higher units of the Prosodic Hierarchy are not: they dominate terminals, but are not their projection in any way – a unique situation in linguistic structure (more on that in Scheer forth).

phonological objects (just as syntax can only make reference to truly syntactic objects). Selkirk (1984:32, 409f) and Nespor & Vogel (1986:27ff, 110ff) are along the same lines.

I have only come across one voice that clearly identifies the Prosodic Hierarchy for what it is. Without surprise, this voice comes from Direct Syntax quarters where of course no buffer is needed: Kaisse (1990:128) calls attention to the redundant and diacritic character of prosodic constituency. She points out that the direct syntax option "does not require the postulation of constituents that are needed only to describe the sandhi phenomena in question."

In a broader historical perspective, then, the output of translation has always been diacritic, with various degrees of camouflage, though. Structuralism has tried to make the carrier of morpho-syntactic information a truly phonological object that has got nothing to do with morphology. Juncture phonemes were the result of the descriptivism-rooted requirement of Level Independence: the bottom-up discovery procedure did not allow phonology to contain any morpho-syntactic information. Much effort was put into the camouflage of the extra-phonological identity of morpho-syntactic information: as indicated by their name, juncture phonemes were supposed to be phonemes, that is truly phonological objects (Scheer 2011:§§59, 61).

SPE basically does the same thing, only that the phonological currency has changed: phonology is now made of segments (rather than of phonemes), which means that boundaries are [-segment] segments. The generative camouflage is not really less outlandish than its structuralist predecessor: # is not any more a segment than it is a phoneme. The unwarranted consequences of its alleged segmental status were made explicit early on, namely by Pyle (1972) (see Scheer 2011:§136). This time, however, the masquerade was entirely transparent right from the start: unlike in structuralist theory where the phonemic status was given real credit, hardly anybody took the [-segment] camouflage seriously. The naked # was taken for what it really is in all phonological quarters: a unit whose only purpose is to store and release morpho-syntactic information.

Finally, in the early 80s when all areas of phonology were autosegmentalised, linear diacritics (boundaries) were replaced by autosegmental diacritics (the Prosodic Hierarchy). As we have seen, the argument which was used in order to promote this move was precisely the idea that diacritics do not qualify (Scheer 2011:§373).

In sum, since structuralist times and up to the present day, the output of translation has always been a diacritic. The choice of the diacritic was according to whatever was the basic representational currency of the time: phonemes in structuralism, segments in SPE, autosegmental domains when phonology became autosegmental. Protagonists of the respective theories were aware of the diacritic character of carriers of morpho-syntactic information to various degrees: high for SPE-type boundaries, choked but sensible for juncture phonemes, not at all for prosodic constituency.

7. Non-diacritic translation and the Direct Effect
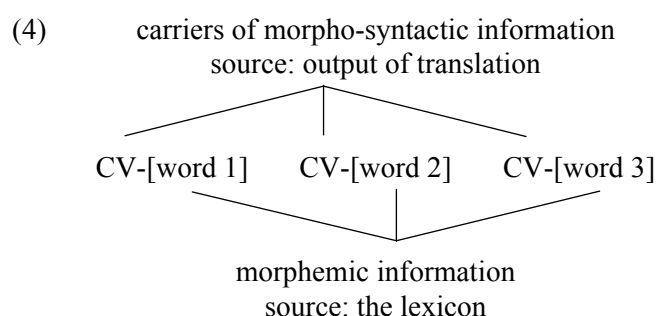7.1. An output of translation that qualifies: syllabic space

If the Prosodic Hierarchy is a diacritic and if diacritics do not qualify for the output of translation, what is the alternative? What would non-diacritic translation look like? We have seen that a non-diacritic is an object that exists in phonology anyway, even in absence of any appeal to extra-phonological information. Lass' (1971) proposal that a boundary materialises as the feature [-voice] in phonology therefore satisfies the non-diacritic requirement: [-voice] exists in phonological processes that have got nothing to do with extra-phonological information. However, we know that melodic primes are no good candidates for the output of translation because melody and morpho-syntax are incommunicado altogether. Nobody (except Lass) and no interface theory has ever tried to make features the carrier of morpho-syntactic information: the fact that morpho-syntax and the area below the skeleton are incommunicado is a well-established (if tacit) fact (see Scheer 2011:§660).

Lowenstamm (1999) has introduced another type of carrier: syllabic space. He argues that the beginning of the word materialises as an empty onset-nucleus pair, the so-called initial CV. Syllabic space is certainly not a diacritic since it is a necessary ingredient of phonology even in absence of extra-phonological factors. The idea that carriers of morpho-syntactic information in phonology

identify as syllabic space can be implemented in any (phonological) theory. In the particular theory of CVCV (Lowenstamm 1996, Scheer 2004, Cyran 2010, among others), it identifies as the initial CV because of the internal logic of this framework where the minimal (and actually maximal) building block of constituent structure is a CV unit. In other theories, syllabic space will take on other forms: x-slots, morae, regular syllabic constituents such as onsets etc.[9]

Another interesting property of syllabic space is that it is linear (local), rather than autosegmental (more on this distinction in Scheer 2009a,b, forth): just like SPE-type boundaries, syllabic space is necessarily inserted into the linear string at morpho-syntactic divisions. An individual piece of the linear string *belongs* to an autosegmental domain (such as the prosodic word), but it cannot *belong* to a boundary. On the other hand, a boundary (or some piece of syllabic space) has a linear location: it follows some piece, and precedes some other piece. It does not make sense to talk about domains that intervene between two pieces: domains are *made* of pieces, but they are not defined by a linear precedence relation with the items that they dominate.

Table (4) below provides a schematic representation of how syllabic space (here the initial CV), representing non-morphemic information, concatenates with regular morphemic material.

(4)   carriers of morpho-syntactic information
              source: output of translation

CV-[word 1]   CV-[word 2]   CV-[word 3]

                    morphemic information
                    source: the lexicon

Carriers of morpho-syntactic information can thus be both local and non-diacritic. Counter to what early Prosodic Phonology claimed, nothing withstands the existence of *non-diacritic boundaries*.

7.2. Diacritic sleepers vs. phonologically meaningful objects

An intrinsic property of diacritics such as hash marks or omegas is to be unable to make predictions. This is because they are empty shells: they do not carry any information except the morpho-syntactic load that they are designed for. Phonology does not react on the simple presence of a hash mark or a prosodic word – these items can only bear on phonology if the analyst has made a phonological process sensitive to them.

Hash marks, omegas and the like may thus be described as passive sleepers. They merely sit in phonological representations without producing any effect by themselves: nothing happens unless a phonological instruction makes reference to them. For example, there is no way to know whether they rather favour or disfavour consonant clusters in their vicinity. In principle they can trigger (or inhibit) any phonological process and its reverse.

By contrast, if non-diacritic objects, i.e. phonologically meaningful vocabulary, carry morpho-syntactic information, phonological computation will react on their bare presence. This is what I call the Direct Effect (see Scheer 2009a,c, forth): the simple presence of phonologically meaningful objects makes predictions. Let us consider the following example in order to flesh out the difference between diacritic sleepers and phonologically meaningful items.

---

9   This is all to the good because different predictions will be made by the different vocabulary chosen, and hence competing phonological theories may be assessed according to their behaviour at the interface (see Scheer 2009a,c, forth).

(5)   equally probable rules?
    a. $V \rightarrow \emptyset$ / #C__CV
    b. $\emptyset \rightarrow V$ / #C__CV

Both rules under (5) are equally probable and equally natural from the point of view of a theory that uses diacritic boundaries: no property of the theory favours or disfavours the epenthesis into an initial cluster, or the deletion of a vowel in this context. Every phonologist knows, however, that (5b) is an attested phonological process, while (5a) is not on record. That is, there is no "masochistic" language that would delete vowels in initial clusters (and only in this context).[10]

Therefore theories that cannot discriminate between (5a) and (5b) have a problem, and the reason why they are in trouble is that the critical information, i.e. word-initiality, is conveyed by a diacritic hash mark. The result is the same in case the prosodic word or some other prosodic constituent carries this information: anything and its reverse may happen at the left edge of a prosodic constituent.

A look at a non-diacritic alternative shows that the two rules at hand are discriminated as soon as the extra-phonological information comes as a real phonological object that impacts phonology directly and does not need to be explicitly mentioned in rules (or constraints) in order to produce an effect. Suppose that an empty CV unit carries the information regarding word initiality along the lines of Lowenstamm (1999) that were discussed above, and in the syllabic environment of Government Phonology.

(6)   deletion vs. insertion of the first vowel in a word in CVCV
    a. deletion: ill-formed   b. insertion: structure saved
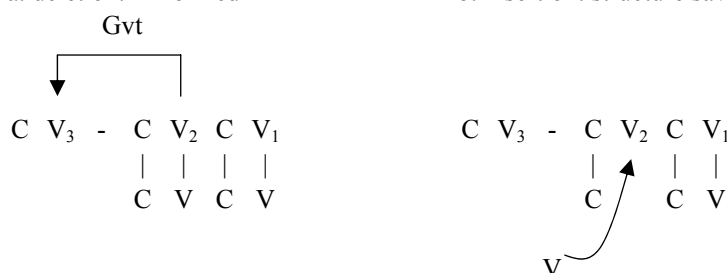
The leftmost CV unit under (6a,b) is the representative of the beginning of the word in phonology. Being a truly phonological object, it is parsable by phonological computation. Under (6b) where the first nucleus of the root is empty, its presence creates a sequence of two empty nuclei. Since V_1 can only govern V_2, V_3 will remain ungoverned, which means that the structure is ill-formed. Therefore an epenthesis into V_2 will rescue the word.

On the other hand, the structure under (6a) is well-formed: V_2 governs V_3, and no empty nucleus remains orphan. The deletion of the content of V_2, however, creates a sequence of two empty nuclei and therefore makes the structure ill-formed: this is the "masochistic" option.

It is thus predicted that the deletion rule (5a) is impossible, while the epenthesis rule (5b) is regular – exactly what we find across languages. Hence there is a clear difference between non-predictive diacritics which allow anything and its reverse to happen in their vicinity, and truly phonological objects that have a predictable effect on the well-formedness of phonological structure.

7.3. Stable effects of the beginning of the word

The beginning of the word is well suited to illustrate that morpho-syntactic information (at least sometimes) has predictable and cross-linguistically stable, rather than arbitrary effects. Table (7) below

---

[10]   Note that rule (5a) says that vowels are deleted *only* when they occur in word-initial clusters. Of course there are languages where vowels are deleted in this context (e.g. Czech *pes - ps-a* "dog Nsg, Gsg), but they will then also be deleted elsewhere (Czech *loket - lokt-e* "elbow Nsg, Gsg").

shows the three effects that the left edge of the word produces language after language (Scheer 2004:§87, 2009a,c, forth, Ségéral & Scheer 2008).

(7) stable effects of the beginning of the word across languages
   a. word-initial clusters
       in some languages initial clusters are restricted to #TR. In others they have the same distribution as internal clusters. But there is no language where they are restricted to non-#TR clusters (i.e. to #RT, #TT and #RR).[11]
   b. strength of word-initial consonants
       in some languages word-initial consonants are especially strong. In others, they do not have any peculiar behaviour regarding strength. But there is no language where they are especially weak.
   c. deletion of the first vowel of the word
       in some languages the first vowel of words is unable to alternate with zero. In others it does not show any peculiar behaviour when compared to other vowels. But there is no language where non-initial vowels are unable to alternate with zero, while initial vowels do.

Languages seem to make a binary choice: either the beginning of the word is in no way peculiar in comparison to what happens morpheme-internally, or it is outstanding and allows only for a characteristic subset of the options that are well-formed elsewhere.

The conclusion is that if the effect of the beginning of the word is not arbitrary, the representational identity of the object by which it is represented must not be arbitrary either. Rather, we are looking for one single object that produces the three effects at hand: three for the price of one. In any event, objects such as the hash mark or some prosodic constituent that do not produce any effect at all (or rather, that tolerate any effect and its reverse under the appropriate rule or constraint) do not qualify. If the effect is predictable, the identity of its trigger cannot be arbitrary.

8. Phase theory adapted to the demands of phonology
8.1. Desiderata for phase-based chunk definition

Let us now have a closer look at what phase theory needs to be able to do when the perspective is seriously entertained that phase structure is the only chunk-defining mechanism. In short, phase structure must delineate chunks that are small enough so that every phonologically relevant stretch is a phase. Phase heads and hence phase structure, however, are only determined on the basis of morpho-syntactic evidence in current phase theory.

Recall from sections 3.2 and 3.3 that the issue does not arise below the word level, i.e. for sequences of morphemes. For this chunk size, the definition of phonologically relevant pieces has always been done derivationally. The issue here is how to adapt the traditional mechanisms of Lexical Phonology where the PIC plays no role (see the analysis of English nasal assimilation in section 2) by the phase-based analysis where visible traces of phase structure are PIC effects – but this is a different question (see Scheer 2011:§647 for discussion).

At and above the word level, phases need to take over the function of prosodic constituency. Since the 80s, the basic argument against a perspective where morpho-syntactic divisions alone define phonologically relevant chunks is so-called non-isomorphism. The following section discusses this point.

In order to be able to define the domains of application of phonological processes, phases need to be able to block their application across relevant boundaries. The phase-based technology for this task is the PIC, which "freezes" previously interpreted strings. How exactly the adaptation of the PIC to the blocking of processes across chunk boundaries is to be operated needs to be worked out.

---

[11] T is shorthand for any obstruent, R for any sonorant.

Finally, an issue is the fact that by and large morpho-syntactic evidence for phases does not coincide with phonological evidence. This is workable as long as there is an inclusive relationship, i.e. as long as all "phonological phases" are a proper subset of "morpho-syntactic phases". This also depends on the evolution, or rather, atomisation of phase theory in current minimalist syntax. All that is discussed in section 8.3 below.

8.2. Non-isomorphism evaporates if looked at through the prism of local boundaries

Since Chomsky & Halle (1968:371f), what will be called non-isomorphism in the 80s is illustrated by an example that is about a cat, a rat and cheese. In Prosodic Phonology, the argument was made in Selkirk's (1981 [1978]:138) founding article and runs through the entire literature up to the present day (e.g. Nespor & Vogel 1986:37ff, Vogel 1999:117ff).

(8)   a.   This is [the cat that caught [the rat that stole [the cheese]]]
        b.   [This is the cat] [that caught the rat] [that stole the cheese]

The major syntactic divisions of the sentence under (8a) do not coincide with its intonational structure under (8b). Hence, goes the argument since SPE, whatever drives phonology to decide that intonation is as under (8b), it is not the output of the syntactic module: there is no node in the syntactic tree that uniquely dominates every intonational span of the sentence. This discrepancy between (8a) and (8b) is what Prosodic Phonology (and the generative mainstream since then) calls non-isomorphism: syntactic structure and the structure that is necessary in order to properly run phonology do not always coincide.

Upon closer inspection, however, it appears that non-isomorphism is a fact about the prism that Prosodic Phonology uses in order to look at the evidence, rather that about the linguistic fact itself. When we look at the same evidence through the glasses of local boundaries (instead of the domain-based prism), non-isomorphism evaporates: intonational chunks simply begin with every CP. Hence it is enough to say that the intonation-building mechanism starts a new unit every time it hits a CP-boundary upon spell-out. Intonational and syntactic structure are thus perfectly isomorphic, and no specific autosegmental constituency on the phonological side is needed in order to express the relevant generalisation (see Scheer 2011:§419 for further discussion).

8.3. Phonological and morpho-syntactic evidence for phases

Let us now consider how morpho-syntactic phase structure correlates with the phonological demand for chunk definition in case this definition is also done by phases. The field that we are talking about is the one from the word size upwards, and phase theory needs to be flexible enough to define all chunk sizes that are needed in order to take over the chunk-defining function from prosodic constituency. That is, there must be a phase for every domain in which some phonological process applies (at whose boundaries it cannot cross).

In order to evaluate the situation, let us first have a brief look at the kind of phases that are defined for purely morpho-syntactic purposes in current syntactic discussion. Chomsky's (2000) original take on phasehood identifies CP and vP, maybe DP (Chomsky 2005:17f.), as phase heads. Since then there is a constant trend to grant phasehood to smaller and smaller chunks (den Dikken 2007:33 and Grohmann 2007 provide an overview, see also Scheer 2011:§773): the DP track is followed, but DP-internal phases are also argued for (Matushansky 2005). TP is another item under debate: while Chomsky (e.g. 2000:106, 2004:124) is explicit on the fact that TP does not qualify as a phase head (because it is not propositional), den Dikken (2007) points out that according to Chomsky's own criteria, this conclusion is far from obvious. TP is indeed assumed to act as a phase head in a growing body of literature, and nodes below TP such as Voice[0] (Baltin 2007, Aelbrecht 2008) and AspP (Hinterhölzl 2006) are also granted phasehood. The vanishing point of the atomization of

phasehood is a situation where all nodes trigger interpretation; or, in other words, where interpretation occurs upon every application of Merge. This radical position – Spell-out-as-you-Merge – is defended by Samuel Epstein and colleagues: Epstein *et al.* (1998), Epstein & Seely (2002, 2006).

The field is in steady movement, and the atomisation of chunks that are granted phase status continues. At some point, all chunks that are needed for phonological purposes may also found to be morpho-syntactically motivated phases. In case there is a phonological demand for a given chunk to be a phase, but no morpho-syntactic response, this may be taken to be genuinely phonological evidence for phasehood. In current practice, phase heads are detected only on morpho-syntactic grounds, but of course there is no reason for this monoculture. If phases transport chunks between morpho-syntax and phonology, it is to be expected that they leave traces in the latter that betray their existence.

In any event, we know that external sandhi is process-specific (Scheer 2011:§823). If the PIC is responsible for blocking word stress computation across word boundaries in English, it applies to stress, but not to external sandhi phenomena such as t-flapping (Scheer 2009c). Therefore it is not true that a given boundary, the word in our case, is a barrier for all processes. In terms of phase theory, phases do or do not come with a PIC: the word is a phase (in English), but only stress sees the barrier – t-flapping does not.

In other words, the PIC applies à la carte: there is a decision made somewhere regarding the armament of phases and processes with a PIC: some phases will be made visible in the phonology by the association of a PIC effect for some processes, others will not. These choices produce the parametric variation that is known as variation on prosodic phrasing in Prosodic Phonology. In all cases, the non-intervention of morpho-syntax, i.e. the absence of a PIC on a given phase, produces a situation where phonology applies across word boundaries, i.e. where external sandhi occurs. Note, however, that in addition to derivational sandhi-blockers (the PIC) there are also representational sandhi-blockers (the aforementioned syllabic space that may be inserted at morpheme breaks, see Scheer 2009c).

8.4. Spell-out is symmetric, but allows for free rides

If the perspective that was lined out is on the right track, phase boundaries may, but do not need to be accompanied by phonological effects. This is interesting in the context of the debate regarding asymmetric spell-out, i.e. the independent access of LF and PF. A basic (if often tacit) assumption of phase theory is that LF and PF phases are always concomitant: when a given node is spelled out, its content is sent to and interpreted at both LF and PF. It is obvious that phase theory would be significantly weakened if it turned out that a given node could be independently spelled out at LF and PF. Chomsky (2004) is explicit on this.

(9)   "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. Assume that to be true. Then Φ [the phonological component] and Σ [the semantic component] apply to units constructed by NS [narrow syntax], and the three components of the derivation of <PHON, SEM> proceed cyclically in parallel. L [language] contains operations that transfer each unit to Φ and Σ. In the best case, these apply at the same stage of the cycle. […] In this conception there is no LF: rather, the computation maps LA [lexical array] to <PHON, SEM> piece-by-piece cyclically." Chomsky (2004:107)

Responding to empirical pressure from various sides, though, independent LF and PF spell-out is proposed or considered by, among others, Marušič (2005), Marušič & Žaucer (2006), Felser (2004), Matushansky (2005), den Dikken (2007), Megerdoomian (2003) and Caha & Scheer (2008).

There is no doubt that simultaneous phases are to be preferred. The question is whether they resist empirical pressure. The parametric scenario of the above discussion maintains simultaneous spell-out, but in fact allows for interpretation to be vacuous: the phase structure of a sentence is uniquely defined at the morpho-syntactic level. Every time a phase head is hit upon spell-out, its

content is sent to both LF and PF. Every phase is thus processed by both interpretational modules, but this does not mean that an effect is systematically encountered: there are "free rides". That is, at least for PF, a phonological footprint is only left behind if either a PIC condition or a representational unit (such as syllabic space) is associated to a phase. Whether this is the case or not is a matter of a parametric choice. In other words, all phases are treated in the same way by morpho-syntax, but not by PF and LF.

9. Conclusion

On the preceding pages it was argued that of the two currently entertained mechanisms which compete for the definition of phonologically relevant chunks, prosodic constituency (representational) and phase structure (procedural), only one can survive: the same labour must not be done twice. After about three decades of peaceful coexistence, bringing to the fore the competition is a virtue by itself. Leaving the empirical assessment for another venue, two conceptual arguments are made in disfavour of prosodic constituency: for one thing, phase theory is needed independently in current minimalist syntax, and for the communication between morpho-syntax and phonology, while prosodic constituency has no motivation outside of phonology. But also, prosodic constituency is disqualified on modular grounds by the fact that it is a diacritic. Modularity-violating Direct Syntax is not an option either.

The article then inquires on two issues that beg the question when this reasoning is followed: 1) what, if any, is the output of translation, i.e. which representational objects that carry morpho-syntactic information are inserted into phonological representations?, and 2) how does phase theory need to be adapted in order to be able to serve its new phonological function? The answer to the former question is syllabic space, to the latter that there is a phase skeleton common to morpho-syntax and phonology, but that not every phase necessarily leaves a trace in the phonology. Whether or not a phonological footprint is left behind depends on parametric choices that concern the armament of phases à la carte with either a PIC (derivational) or a unit of syllabic space (representational).

What is gained at the end of the day is this: 1) an interface theory where the same labour is not done twice, 2) a somewhat weakened, but unified phase theory that does labour both in morpho-syntax and phonology, and 3) a phase theory where evidence for phasehood may be both phonological and morpho-syntactic (while thus far only the latter domain is phase-defining).

References
Items followed by "WEB" are available at www.unice.fr/scheer/.

Aelbrecht, Lobke 2008. Licensing ellipsis as Agree. Paper presented at the 27th West Coast Conference on Formal Linguistics, UCLA 16-18 May.
Baltin, M. 2007. Deletion versus Pro-forms: a false dichotomy? Ms., New York University.
Bermúdez-Otero, Ricardo 2011. Cyclicity. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell Companion to Phonology*, 2019-2048. Malden, MA: Wiley-Blackwell.
Bermúdez-Otero, Ricardo and Ana Luís 2009. Cyclic domains and prosodic spans in the phonology of European Portuguese functional morphs. Paper presented at OCP 6, Edinburg 24 January. Handout available at http://www.bermudez-otero.com/bermudez-otero&luis.pdf.
Caha, Pavel and Tobias Scheer 2008. The Syntax and Phonology of Czech Templatic Morphology. In Andrei Antoneko, John Bailyn and Christina Bethin (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics. The Stony Brook Meeting 2007*, 68-83. Ann Arbor: Michigan Slavic Publications. WEB.
Cheng, Lisa and Laura Downing 2007. The prosody and syntax of Zulu relative clauses. *SOAS Working Papers in Linguistics and Phonetics* 15: 51-63.
Cheng, Lisa and Laura Downing 2009. Where's the topic in Zulu? *The Linguistic Review* 26: 207-238.

Cheng, Lisa and Laura Downing 2011a. Prosodic domains do not match spell-out domains. Paper presented at Phonology in the 21st century: In Honour of Glyne Piggott, Montreal 7-9 May.

Cheng, Lisa and Laura Downing 2011b. To phrase or not to phrase: on the (non-)congruence of focus and prosody. Paper presented at GLOW 34, Vienna 27 April - 1st May.

Chomsky, Noam 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond. The cartography of syntactic structures, Volume 3*, 104-131. Oxford: Oxford University Press.

Chomsky, Noam 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.

Chomsky, Noam and Morris Halle 1968. *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.

Chomsky, Noam, Morris Halle and Fred Lukoff 1956. On Accent and Juncture in English. In Morris Halle, Horace Lunt, Hugh McLean and Cornelis van Schooneveld (eds.), *For Roman Jakobson. Essays on the occasion of his sixtieth birthday*, 65-80. The Hague: Mouton.

Clements, George 1978. Tone and syntax in Ewe. In Donna Jo Napoli (ed.), *Elements of stress, tone and intonation*, 21-99. Washington, D.C.: Georgetown University Press.

Cyran, Eugeniusz 2010. *Complexity scales and licensing in phonology*. Berlin: Mouton de Gruyter.

den Dikken, Marcel 2007. Phase Extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33: 1-41.

Dobashi, Yoshihito 2003. Phonological Phrasing and syntactic derivation. Ph.D thesis, Cornell University.

Downing, Laura 2010. An edge-based approach to the alignment of syntactic phases and prosodic phrases. *Transactions of the Philological Society* 108: 352-369.

Elordieta, Gorka 2008. An overview of theories of the syntax-phonology interface. *Journal of Basque Linguistics and Philology* 42: 209-286.

Epstein, Samuel, E.M. Groat, R. Kawashima and H. Kitahara 1998. *A Derivational Approach to Syntactic Relations*. Oxford: OUP.

Epstein, Samuel and T. Daniel Seely 2006. *Derivation in Minimalism*. Cambridge: CUP.

Epstein, Samuel and T.Daniel Seely 2002. Rule Applications as Cycles in Level-Free Syntax. In Samuel Epstein and T.Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 65-89. Oxford: Blackwell.

Felser, Claudia 2004. Wh-copying, phases, and successive cyclicity. *Lingua* 114: 543-574.

Fodor, Jerry 2000. *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, Mass.: MIT Press.

Gardner, Howard 1985. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Gerrans, Philip 2002. Modularity reconsidered. *Language and Communication* 22: 259-268.

Grohmann, Kleanthes 2007. Spelling out Dynamic Interfaces. *Linguistic Analysis* 33: 197-208.

Hayes, Bruce 1989 [1984]. The Prosodic Hierarchy in Meter. In Paul Kiparsky and G. Youmans (eds.), *Manuscript circulated since 1984, published 1989 in Rhythm and Meter*, 201-260. Orlando, Florida: Academic Press.

Hinterhölzl, Roland 2006. The Phase Condition and cyclic Spell-out: Evidence from VP-topicalization. In Mara Frascarelli (ed.), *Phases of Interpretation*, 237-259. Berlin: de Gruyter.

Hyman, Larry 1978. Word Demarcation. In Joseph Greenberg (ed.), *Universals of Human Language, Vol 2*, 443-470. Stanford: Stanford University Press.

Inkelas, Sharon 1990. *Prosodic Constituency in the Lexicon*. New York: Garland.

Inkelas, Sharon and Draga Zec 1995. Syntax-phonology Interface. In John Goldsmith (ed.), *The Handbook of Phonological Theory*, 535-549. Oxford: Blackwell.

Ishihara, Shinichiro 2007. Major phrase, focus intonation, multiple spell-out. *The Linguistic Review* 24: 137-167.

Itô, Junko and Armin Mester 1999. Realignment. In René Kager, Harry van der Hulst and Wim

Zonneveld (eds.), *The prosody-morphology interface*, 188-217. Cambridge: Cambridge University Press.

Kahnemuyipour, Arsalan 2009. *The Syntax of Sentential Stress*. Oxford: OUP.

Kaisse, Ellen 1983. The syntax of auxiliary reduction in English. *Language* 59: 93-122.

Kaisse, Ellen 1985. *Connected Speech. The interaction of Syntax and Phonology*. London, New York: Academic Press.

Kaisse, Ellen 1990. Toward a Typology of Postlexical Rules. In Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 127-143. Chicago: Chicago University Press.

Kaye, Jonathan 1995. Derivations and Interfaces. In Jacques Durand and Francis Katamba (eds.), *Frontiers of Phonology*, 289-332. London & New York: Longman. Also in SOAS Working Papers in Linguistics and Phonetics 3, 1993, 90-126. WEB.

Kenstowicz, Michael and Charles Kisseberth 1977. *Topics in Phonological Theory*. New York: Academic Press.

Kleinhenz, Ursula 1998. On Words and Phrases in Phonology. A comparative study with focus on German. Ph.D thesis, University of Tübingen.

Kratzer, Angelika and Elisabeth Selkirk 2007. Phase theory and prosodic spellout: the case of verbs. *The Linguistic Review* 24: 93-135.

Lass, Roger 1971. Boundaries as obstruents: Old English voicing assimilation and universal strength hierarchies. *Journal of Linguistics* 7: 15-30.

Lowenstamm, Jean 1996. CV as the only syllable type. In Jacques Durand and Bernard Laks (eds.), *Current trends in Phonology. Models and Methods*, 419-441. Salford, Manchester: ESRI. WEB.

Lowenstamm, Jean 1999. The beginning of the word. In John Rennison and Klaus Kühnhammer (eds.), *Phonologica 1996*, 153-166. La Hague: Holland Academic Graphics. WEB.

Marušič, Franc 2005. On non-simultaneous phases. Ph.D. dissertation, SUNY, Stony Brook.

Marušič, Franc and Rok Žaucer 2006. On the intensional feel-like construction in Slovenian. *Natural Language and Linguistic Theory* 24: 1093-1159.

Mascaró, Joan 1976. Catalan Phonology and the Phonological Cycle. Ph.D. dissertation.

Matushansky, Ora 2005. Going through a phase. *MIT Working Papers in Linguistics* 49: 157-181.

McCarthy, John and Alan Prince 1993. Generalized Alignment. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1993*, 79-153. Dordrecht: Kluwer. Abridged version in McCarthy, John (ed.) 2004. Optimality Theory in Phonology, 451-463. Oxford: Blackwell.

McCarthy, John and Alan Prince 2001. Prosodic Morphology. Constraint Interaction and Satisfaction. Ms, ROA #482.

Megerdoomian, Karine 2003. Asymmetries in Form and Meaning: Surface Realization and the Interface Conditions. Paper presented at Approaching Asymmetry at the Interfaces, UQAM, Montreal.

Nespor, Marina 1999. Stress Domains. In Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 117-159. Berlin: Mouton de Gruyter.

Nespor, Marina and Irene Vogel 1986. *Prosodic Phonology*. Dordrecht: Foris.

Newell, Heather 2008. Aspects of the morphology and phonology of phases. Ph.D dissertation, McGill University Montreal.

Newell, Heather and Tobias Scheer 2007. Procedural First. Paper presented at the 38th Poznań Linguistic Meeting, Poznań 13-16 September. WEB.

Odden, David 1987. Kimatuumbi phrasal phonology. *Phonology* 4: 13-26.

Odden, David 1990. Syntax, lexical rules and postlexical rules in Kimatuumbi. In Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 259-277. Chicago: University of Chicago Press.

Pak, Majorie 2008. The postsyntactic derivation and its phonological reflexes. Ph.D dissertation, University of Pennsylvania.

Peperkamp, Sharon 1995. Prosodic constraints in the derivational morphology of Italian. *Yearbook of Morphology 1994*: 207-244.

Piggott, Glyne and Heather Newell 2006. Syllabification and the spell-out of phases in Ojibwa words.

McGill Working Papers in Linguistics 20: 39-64.

Pyle, Charles 1972. On Eliminating BM's. In Paul Peranteau, Judith Levi and Gloria Phares (eds.), *Papers from the eighth regional meeting of the Chicago Linguistic Society*, 516-532. Chicago: Chicago Linguistic Society.

Rotenberg, Joel 1978. The Syntax of Phonology. Ph.D dissertation, MIT.

Rubach, Jerzy 1984. *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris.

Rubach, Jerzy and Geert Booij 1984. Morphological and prosodic domains in Lexical Phonology. *Phonology Yearbook* 1: 1-27.

Samuels, Bridget 2009. The structure of phonological theory. Ph.D dissertation, Harvard University.

Scheer, Tobias 2004. *A Lateral Theory of Phonology. Vol.1: What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.

Scheer, Tobias 2008. Why the Prosodic Hierarchy is a diacritic and why the Interface must be Direct. In Jutta Hartmann, Veronika Hegedüs and Henk van Riemsdijk (eds.), *Sounds of Silence: Empty Elements in Syntax and Phonology*, 145-192. Amsterdam: Elsevier. WEB.

Scheer, Tobias 2009a. External sandhi: what the initial CV is initial of. *Studi e Saggi Linguistici* 47: 43-82. WEB.

Scheer, Tobias 2009b. Intermodular Argumentation and the Word-Spell-Out-Mystery. In Kleanthes Grohmann (ed.), *Explorations of Phase Theory: Interpretation at the Interfaces*, 23-65. Berlin: Mouton de Gruyter. WEB.

Scheer, Tobias 2009c. Representational and procedural sandhi killers: diagnostics, distribution, behaviour. In Mojmír Dočekal and Markéta Ziková (eds.), *Czech in Formal Grammar*, 155-174. München: Lincom. WEB.

Scheer, Tobias 2011. *A Guide to Morphosyntax-Phonology Interface Theories. How Extra-Phonological Information is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.

Scheer, Tobias forth. *Direct Interface and One-Channel Translation*. Berlin: Mouton de Gruyter.

Ségéral, Philippe and Tobias Scheer 2008. The Coda Mirror, stress and positional parameters. In Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral (eds.), *Lenition and Fortition*, 483-518. Berlin: Mouton de Gruyter. WEB.

Seidl, Amanda 2001. *Minimal Indirect Reference: a theory of the syntax-phonology interface*. London: Routledge.

Selkirk, Elisabeth 1972. *The phrase phonology of English and French*. Ph.D. dissertation MIT, published 1980 by Garland Press.

Selkirk, Elisabeth 1974. French liaison and the X-bar notation. *Linguistic Inquiry* 5: 573-590.

Selkirk, Elisabeth 1981 [1978]. On prosodic structure and its relation to syntactic structure. In Thorstein Fretheim (ed.), *Nordic Prosody II*, 111-140. Trondheim: TAPIR.

Selkirk, Elisabeth 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: MIT Press.

Selkirk, Elisabeth 1986. On derived domains in sentence phonology. *Phonology* 3: 371-405.

Uriagereka, Juan 1999. Multiple spell-out. In Samuel Epstein and Norbert Hornstein (eds.), *Working Minimalism*, 251-282. Cambridge, Mass.: MIT Press.

Vogel, Irene 1991. Level ordering in Italian Lexical Phonology? In Pier Marco Bertinetto, Michael Kenstowicz and Michele Loporcaro (eds.), *Certamen Phonologicum II*, 81-101. Torino: Sellier & Rosenberg.

Vogel, Irene 1999. Subminimal constituents in prosodic phonology. In Stephen Hannahs and Mike Davenport (eds.), *Issues in phonological structure*, 249-267. Amsterdam: Benjamins.

Vogel, Irene and István Kenesei 1990. Syntax and semantics in phonology. In Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 339-363. Chicago: University of Chicago Press.

Yip, Moira 1998. Identity Avoidance in Phonology and Morphology. In Steven Lapointe, Diane Brentari and Patrick Farrell (eds.), *Morphology and its Relation to Phonology and Syntax*, 216-246. Stanford: CSLI Publications.