Czech in Formal Grammar

Mojmír Dočekal & Markéta Ziková

2009
LINCOM EUROPA

Published by LINCOM GmbH 2009.

LINCOM GmbH
Gmunderstr. 35
D-81379 Muenchen

LINCOM.EUROPA@t-online.de
www.lincom-europa.com

webshop: lincom.eu

Printed in E.C.
Printed on chlorine-free paper

Die Deutsche Bibliothek - CIP Cataloguing-in-Publication-Data

A catalogue record for this publication is available from Die Deutsche Bibliothek (http://www.ddb.de)

Preface

This book is comprised of papers from the conference *Czech in Formal Grammar,* whose first meeting was held on February 12-14, 2009, at Masaryk University in Brno.

The programme of the conference was constituted on the basis of anonymously reviewed abstracts. It consisted of 23 papers and 3 invited talks focused on topics from phonology (Tobias Scheer), syntax (Ludmila Veselovská and Petr Karlík) and semantics (Hana Filip). 18 papers of those presented at the conference appear in their revised and edited versions in this volume.

We would like to thank our colleagues who reviewed either abstracts for the conference or papers for this book. We are also grateful to the head of the Department of Czech Language, Jana Pleskalová, and the head of the Department of Linguistics and Baltic Languages, Ondřej Šefčík, for financial support.

Markéta Ziková (Department of Czech Language)
Mojmír Dočekal (Department of Linguistics and Baltic Languages)

Faculty of Arts
Masaryk University in Brno

Contents

Problems and Possibilities of the Annotation of the Interpropositional Discourse Relations in PDT 2.0*

Zuzanna Bedřichová
Charles University, Prague

The following paper results from the preparation of the annotation of interpropositional discourse relations that is currently worked on at the Institute of Formal and Applied Linguistics at the Charles University at Prague by Eva Hajičová, Zuzanna Bedřichová, Šárka Zikánová and Lucie Mladová. The goal of this report is not to present the pilot annotation in a complex way but to point out the problems and possibilities of such annotation. As the annotation is based on the Prague Dependency Treebank version 2.0 (PDT 2.0) corpus, especially its tectogrammatical layer (underlying syntax and semantics), the problems we will deal with will concern primarily the relation between the existing description of the tectogrammatical layer and the possibility to annotate the new discourse relations. After a brief introduction of the PDT 2.0 and of the interpropositional relations already marked in some way in the PDT 2.0 we will deal with the main problematic areas and the solutions we adopted. The subject of the first area is the problem of proposition and its relation to the tectogrammatical tree (dependency graph) representing a surface sentence. The second area deals with the diversity of the interpropositional discourse relations (hereafter referred to as IDR). At the end of this paper we will try to decide whether we can, despite the diversity of these relations, consider an independent layer of discourse or not.

1. PDT 2.0

The Prague Dependency Treebank 2.0 is a manually annotated corpus of 2 million tokens (for the morphological layer) that aims to describe a natural inflected language – Czech. The language description is divided into several hierarchic layers: the morphological, the analytical (syntactic) and the underlying tectogrammatical (underlying) layer. A hierarchy exists between these layers: the superior layer constitutes the formal representation of the inferior layer:

(1) a. *Obamův módní styl je obdivován všemi Američany bez ohledu na*
Obama's chic fashion is admired all-Ins.Pl Americans-Ins.Pl without regard
barvu pleti.
colour skin.
'Obama's clothing fashion is admired by all Americans disregarding the color of their skin.'

b. morphological layer: Americans – noun
analytical layer: Americans – indirect object
tectogrammatical layer: Americans – actor

This description of language starts from the units of superior layer that constitute the units of inferior layer: word is formed by morphological units; phrase is formed by lexemes. In our annotation we will proceed primarily from the representation of language on the

* This paper was written with the support of GAUK, n° of the grant 424/2009.

Representational and Procedural Sandhi Killers: Diagnostics, Distribution, Behaviour

Tobias Scheer
Université de Nice, CNRS 6039

1. Introduction: a representational and a procedural sandhi-blocker

Phonological processes usually apply across morpheme-, but not across word boundaries.[1] This is the situation that phonologists consider to be default, and on which phonological theories are built. Cases where phonology applies across word boundaries attract specific attention: traditionally they are called (external) sandhi (or connected speech), and generative phonology has devised an entire theory in order to handle them: the subject matter of Prosodic Phonology (Selkirk 1981, 1986, Nespor & Vogel 1986) and of its instrument, the *Prosodic Hierarchy*, is to formalize the intervention of syntactic conditions (i.e. the relationship between words) in phonological matters.

Phonologists agree that external sandhi is a situation where all barriers between words have been removed and phonology "does not see" the morpho-syntactic boundary, just like it may or may not "see" morphological boundaries: while the boundary of class 1 affixes such as *-al* is invisible (*parént-al* is computed as a single stress domain and therefore receives regular penultimate stress), the boundary of class 2 affixes like *-hood* is taken into account by the stress assigning mechanism (the root is a stress domain by itself and receives penultimate stress: *párent-hood*). In other words, the default is that specific action is taken in order to prevent phonology from applying across word boundaries: barriers are established that provoke the anti-sandhi situation, i.e. when phonological processes are blocked by word boundaries. It is the exceptional absence of these barriers that produces the sandhi situation.

This article inquires on the exact nature of the barriers at hand. A priori there are two possible reasons for the phonological incommunication across words: one is procedural, the other representational. The latter is phonology-internal in the sense that the barrier is some phonological structure: a representational unit that carries morpho-syntactic information and was inserted into the phonological representation inhibits cross-word communication. In SPE, a phonological process was blocked by the hash mark # if not otherwise specified in the rule. Prosodic Phonology autosegmentalized hash marks by representing relevant morpho-syntactic information in terms of an autosegmental arboreal structure, the Prosodic Hierarchy. In this environment, processes do not apply across word boundaries because they are made sensitive to a condition that specifies their domain of application, for example "process X applies within the Prosodic Word", and the Prosodic Word coincides with the morpho-syntactic word.

By contrast, the procedural reason for anti-sandhi behaviour is extra-phonological in the sense that phonological computation is blocked without any participation of phonology itself. Since Chomsky et al. (1956:75), morpho-syntax and phonology communicate through cyclic derivation, a mechanism that has known various implementations (as the Transformational Cycle in SPE, as the *Phonological Cycle* in Mascaró 1976) and is called derivation by phase in current minimalist syntax, of which it is the spine (Chomsky 2000, 2001 and following). The idea is that phonological (and semantic) interpretation follows morpho-syntactic structure from inside out, i.e. from the most to the least embedded item. A no look-back proviso that is

[1] Here and below, "word" is shorthand for "some morpho-syntactically relevant chunk at about the word size"; it does not imply any take on what counts as a "word" in which language (clitics etc.).

called the *Phase Impenetrability Condition* (PIC) today[2] inhibits cross-phase communication: previously interpreted material is "frozen" and inaccessible for further computation. Therefore, if two words [[A] [B]] sit in different phases, both A and B will be interpreted by themselves and hence frozen by Phase Impenetrability. When phonology applies to the outer cycle [A B], the shape of A and B that was achieved on the previous cycle cannot be modified. This makes sure that a given phonological process applies word-internally, but not across words (in case A and B are words).

The typical division of labour that is practised in phonology since the early 80s is that a procedural solution (i.e. Lexical Phonology[3]) is applied to internal anti-sandhi (i.e. processes that cannot cross morpheme boundaries), while the representational tools of the Prosodic Hierarchy deal with external anti-sandhi (i.e. processes that cannot apply across word boundaries). This custom notwithstanding, the two kinds of sandhi-blockers may in principle compete.[4] On very rare occasions they actually do in the literature: nasal assimilation in English for example has been analyzed along both lines.

The nasal of *in-* does, but the nasal of *un-* does not assimilate to a following stop: compare *un-predictable* with *im-possible* (cf. *in-offensive*). The procedural solution is based on a contrasting cyclic structure (e.g. Kaye 1995): [[un] [predictable]] vs. [in possible]. *Un-* being spelt out in isolation,[5] its nasal is frozen by the PIC and cannot assimilate to the following stop on the outer cycle. The competing representational solution builds on the contrasting definition of prosodic tree structure (phonological phrasing) (Rubach & Booij 1984, Rubach 1984:221ff, Vogel 1991): while *un-* is a Prosodic Word (PrW) in its own right, *in-* is part of the PrW of the root, i.e. $[[un]_{\omega}\ [predictable]_{\omega}]$ vs. $[in\ possible]_{\omega}$. Nasal assimilation then is specified for applying only within PrWs.

2. Representational communication with phonology

2.1. Carriers of morpho-syntactic information reduce to syllabic space

It was mentioned that the default assumption among phonologists is that the word boundary causes phonological incommunication: the Prosodic Hierarchy enters the scene only when phonology jumps over word boundaries (or below the word level in matters related to prefixes and compounds). Below I look at the distribution of the two types of sandhi-killers in greater

[2] The history and various incarnations of this device are further discussed in Scheer (forth a,b).

[3] Today OTed versions thereof, i.e. Stratal OT (Kiparsky 2000, Bermúdez-Otero forth) or DOT (Rubach 1997 and following). The same is true for Prosodic Phonology, which today has also been carried over into the constraint-based environment of OT. Nothing in the discussion below depends on eventual differences between the original theories of the 80s and modern OTed versions thereof.

[4] Voices that call this division of labour into question are very rare. Based on the fact that the area below the word level may in principle be covered by representational means (the Prosodic Hierarchy), while the area at and above the word level may not be handled with the tools of Lexical Phonology, Selkirk (1984) and Inkelas (1990) argue that the latter are redundant and need to be done away with. The literature indeed draws an asymmetrical picture regarding phonological traces of cyclic spell-out: sequences of morphemes and sequences of words are sent to phonology in several chunks of growing size, but this piecemeal fire leaves only traces in the interpretation of the former. That is, no cases seem to be on record where phonology records traces of the cyclic spell-out of words. This is encoded in the architecture of Lexical Phonology, where lexical phonology is, but postlexical phonology is not cyclic. I call the absence of phonological traces of the cyclic spell-out of words the word-spell-out mystery (Scheer 2009a, forth a) because there is no reason why word sequences should not react on the same conditions that produce cyclic effects on morpheme sequences. Whether phonological effects of the spell-out of words are really absent from the record is an empirical question that deserves attention.

[5] Newell & Scheer (2007) and Newell (2008:178ff) discuss the syntactic aspects of this analysis (late insertion, a-cyclic merger).

detail and from the perspective of a specific assumption regarding the representational device. Just like hash marks #, the Prosodic Hierarchy is a diacritic and therefore does not qualify for the representation of morpho-syntactic information in phonology. Like other modules, phonology understands only its own vocabulary, to which Prosodic Words, Prosodic Phrases and the like do not belong. Truly phonological vocabulary is the one that is used when phonological processes apply in absence of any extra-phonological conditioning (e.g. when a palatalization turns a [k] into [t͡ʃ] before front vowels). Unlike melodic primes (frontness, occlusion etc.) and bottom-up constructed units such as syllable structure and feet, items of the Prosodic Hierarchy are top-down constructions (they are unpredictable from lower units) and intervene only when morpho-syntactic information needs to be imported. They are therefore alien and just as diacritic as #s (Scheer 2008a, forth a).

On the other hand, melodic items are not possible carriers of morpho-syntactic information either: while a segmental/melodic effect can of course be the *result* of morpho-syntactic conditioning, nobody has ever seen a phenomenon, let alone proposed an analysis, whereby a melodic prime, say [+high], is inserted into the phonological string in place of a #, or of the construction of some unit of the Prosodic Hierarchy. I therefore conclude in Scheer (forth a) that the only representational items that qualify as carriers of morpho-syntactic information in phonology are syllabic units: they are non-melodic and at the same time piece and parcel of the truly phonological vocabulary.

2.2. Direct Interface and CVCV: CV units are the carriers of boundary information

The idea that only those units qualify as carriers of morpho-syntactic information that are used in phonological processes which that are unconditioned by extra-phonological factors is called Direct Interface in Scheer (2008a, forth a). "Direct" refers to the fact that all diacritic mediation between morpho-syntactic structure and phonological interpretation is done away with: there is no buffer such as hash marks or the Prosodic Hierarchy whose only function is to store morpho-syntactic information. Rather, morpho-syntactic information is directly translated into truly and active phonological vocabulary. The discussion below is based on these premises.

If thus for the reasons mentioned syllabic space alone qualifies for the representation of morpho-syntactic information in phonology, what kind of item is exactly inserted in the linear string of course depends on the specific theory of syllable structure that is assumed. This consequence is most welcome since it allows phonological theories to be refereed not only on purely phonological grounds, but also according to their behaviour at the interface. That is, different theories of syllable structure provide different units as carriers of morpho-syntactic information and hence make different predictions. Therefore the comparative merit of competing phonological theories may also be assessed at the interface.

Note that this contrasts with buffer-based theories of the interface: whatever the contrast of phonological theories, it is levelled out at the interface since the representational means that are used at the interface – hash marks or the Prosodic Hierarchy – are the same for all phonological theories. Hence there is no theory-specific behaviour at the interface, and the interface will never be able to contribute an arbitral award to the competition of different phonological theories.

The evaluation of different phonological theories according to their behaviour at the interface must be done elsewhere. In the discussion below, the predictions of a specific theory of syllable structure, so-called CVCV (Lowenstamm 1996, Scheer 2004, Cyran 2003, Szigetvári 1999, Szigetvári & Scheer 2005), are made explicit. The developments suppose some familiarity with this framework.

In CVCV, the inventory of autonomous syllabic items reduces to an onset followed by a nucleus, i.e. a CV unit. Empty CV units are thus the only representational items that may carry boundary-information. Lowenstamm (1999) has proposed that the contrast between those languages where only word-initial TR-clusters are allowed (TR-only languages) and those where any cluster occurs (anything-goes languages) is expressed by a CV unit, the socalled initial CV (see Scheer 2004:§83 for an overview).[6]

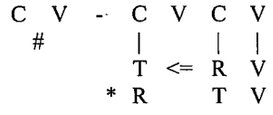
While the (word-)initial occurrence of CV units that carry morpho-syntactic information is best studied (e.g. Ségéral & Scheer 2001, 2005, 2008, Szigetvári 1999, Seigneur-Froli 2003, 2006), boundary information has also been found to be carried by CV units elsewhere: examples are the negative marker in Kabyle Berber (Bendjaballah 2001), a verbal marker in Chleuh Berber (Lahrouchi 2001), a tense marker in German strong verbs (Bendjaballah & Haiden 2003a,b), the so-called derivational syllable in Guerssel & Lowenstamm's (1990) analysis of Classical Arabic and the syllabic support for intrusive consonants in French that occur between the root and certain suffixes (e.g. *stabilo-t-er* 'to mark with a Stabilo pen', Pagliano 2003).

2.3. Three for the price of one: the initial CV and its predictions

Let us now concentrate on the word-initial CV. Table (1) below shows the situation of the beginning of the word in presence and in absence of the initial CV.

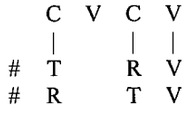
(1) presence vs. absence of the initial CV (effect one)

a. initial clusters: initial CV present



*#RT: two empty nuclei in a row

b. initial clusters: initial CV absent



#RT: ok

Under (1) there are two empty nuclei to be covered. In case of an initial #TR cluster, the sonorant and the obstruent can establish a relationship (<=) which circumscribes the intervening empty nucleus, and the first filled nucleus of the word can govern the empty nucleus of the initial CV. The structure is therefore well-formed. The same is not true for initial #RT, #TT and #RR clusters, though, since the consonants at hand cannot establish any relationship (why this is so is explained at greater length in Scheer 2004, Ségéral & Scheer 2008). There are thus two empty nuclei in a row that require government, and the structure is ill-formed. By contrast under (1) where the initial CV is absent, there is only one empty nucleus to be governed whatever the word-initial cluster, and this can be done by the first filled nucleus of the word. Hence both #TR and #RT clusters are well-formed in this language.[7]

Beyond the regulation of word-initial clusters, the initial CV impacts phonology yet in two other ways (Scheer 2004:§87, forth a, Ségéral & Scheer 2008): its presence makes word-initial consonants strong (while they are weak in its absence) and enforces the presence

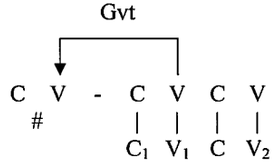
[6] T is shorthand for any obstruent, R for any sonorant.

[7] Unless otherwise specified, #RT is shorthand for all initial non-TR clusters, i.e. #RT, #RR and #TT.

of the first vowel of the word (while this vowel may be absent in absence of the initial CV). This is illustrated under (2) below.

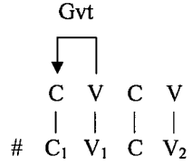
(2) presence vs. absence of the initial CV (effects two and three)

a. initial simplex C: initial CV present



#C strong: it escapes Gvt
V_1 cannot be absent: two empty nuclei in a row

b. initial simplex C: initial CV absent



#C is governed (= intervocalic)
V_1 can be absent: only one empty nucleus

Under (2), the first filled nucleus of the word must govern the empty nucleus of the initial CV. Therefore the word-initial consonant escapes government, whose effect is to inhibit the segmental expression of its target. The ungoverned C_1 under (2) is thus strong. By contrast under (2), the first nucleus of the word has no governing duty and therefore targets C_1, which is weak. The alliance of the effects on clusters and word-initial consonants is tested by Seigneur-Froli (2003, 2006) on the grounds of the Greek pattern: while French for example is TR-only and has word-initial consonants that are strong in diachronic evolution, Greek has also non-TR initial clusters (*#pt, #kt, #mn*) and concomitantly weak word-initial consonants.

Regarding the third effect, the absence of the first vowel of the word produces an ill-formed structure under (2) since it creates a sequence of two empty nuclei. Under (2) where the initial CV is absent on the other hand, nothing withstands the first vowel of a word to alternate with zero: it may always been governed by the following nucleus, which has no other governing duties. Czech for example illustrates this pattern: *pes* 'dog Nsg' appears as *ps-a* in Gsg, and this concords with the fact that Czech is an anything-goes language (i.e. has words with non-TR-initial clusters: *lžíce*, *rtuť* 'spoon, quicksilver' etc.).

In sum, thus, the parameterisation of the initial CV has (at least) three empirical consequences, which the theory predicts to co-occur in the way shown under (3) below.

(3) typological predictions made by the parameterisation of the initial CV

	in a language where the initial CV is present	in a language where the initial CV is absent
a.	word-initial consonants are strong	word-initial consonants are non-strong
b.	the first vowel of words may not alternate with zero	the first vowel of words may alternate with zero
c.	initial clusters are restricted to #TR	there are no restrictions: #TR, #RT, #TT and #RR clusters may occur

Note that these predictions are anything but trivial: they chain together three empirical situations that otherwise seem to be unrelated. Also, they are empirically explicit and may be easily falsified: any language that displays one of the three properties of the righthand or the lefthand column under (3) must also instantiate the two other properties of the same column.

Space restrictions preclude further discussion of the empirical record (see Scheer 2004:§87, Ségéral & Scheer 2008 for greater detail).

2.4. Syllabic consonants: a fourth effect of the initial CV?

A diachronic process in Czech seems to suggest that another parameter could depend on the presence vs. absence of the initial CV. Old Czech (OCz) features both syllabic and so-called trapped consonants: while the former continue Common Slavic (CS) *tьrt*, the latter appear in place of CS *trьt*. Even though both origins produce CRC clusters in OCz, these CRC were clearly distinct according to their diachronic identity. This is witnessed by OCz verse where CRC < *tьrt* counts as a syllabic peak, while CRC < *trьt* does not. The existence of a minimal pair has led Trubetzkoy (1939:199) to establish a "correlation of syllabicity" for OCz: syllabic *dr̥žěti* 'to hold' (< CS *dьržati*, cf. Po *dzierżyć*, MCz *držet*) vs. trapped *držěti* 'to tremble' (< CS *drъžati*, cf. Po *drżeć*).[8]

The CS contrast between tьrt and trьt that was preserved in OCz in the syllabic vs. trapped coat was abandoned in further evolution: in Modern Czech all trapped consonants become syllabic (e.g. Trávníček 1935:57f,111ff,226ff, Lehr-Spławiński & Stieber 1957:97ff, Komárek 1969:60f, 82, 97ff, 127ff, Liewehr 1933:93f,162f). This evolution is observed across the board, with two exceptions: palatalized trapped *ř* remains trapped to date (in all positions: initial *řvát*, medial *křtít,* final *pepř* 'to shout, to baptize, pepper' are monosyllables), and so do word-initial trapped liquids. Hence OCz trapped *trvati*, *bratr* > bisyllabic MCz syllabic *trvat*, *bratr* 'to last, brother', against OCz trapped *rdieti*, *lháti* > monosyllabic MCz trapped *rdít se*, *lhát* 'to go red, to lie'.[9]

There is good reason why ř cannot become syllabic: it is an obstruent (like the uvular "r" of French or German, it has a voiced and a voiceless version). In Czech (and many other languages), however, only sonorants can be syllabic. On the other hand, the positional bias of liquids is intriguing: it looks like the word-initial context prevents trapped consonants from becoming syllabic. Why should that be so?

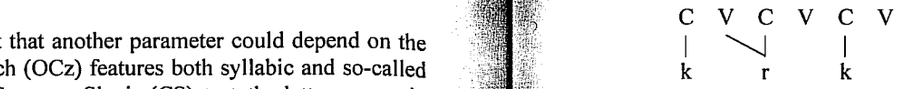
In order to answer this question, let us look at what kind of animal syllabic consonants are. Syllabic consonants have been understood as consonants that branch on a neighbouring nucleus. This accounts for their hermaphrodite identity: their body is consonantal, but their behaviour is vocalic (see Scheer 2008b for an overview). Two options have been entertained: the consonant may either branch on the preceding or on the following nucleus. Both are depicted in table (4) below, which also indicates followers of either solution.

[8] This is regular textbook evidence reported for example by Komárek (1969:82) and Liewehr (1933:94). For instance, syllabic *dr̥žěti* 'to hold' (< CS *dьržati*) counts for 3 syllables in typical 8-peak Old Czech Alexandrine verse (*to jmě drzal takým kmenem*, Kat. verse 24), while trapped *držěti* 'to tremble' (< CS *drъžati*) weighs only 2 syllables (*všecko pohanstvo drzezalo*, Kat. verse 2803). The philological basis of this distinction is discussed at greater length in Scheer (2004:§277).

[9] The empirical situation as well as criteria that allow us to distinguish between syllabic and trapped consonants are exposed at greater length in Scheer (2004:§277, 2008b). Consequences of the evolution from trapped to syllabic consonants on vowel length are discussed in Ziková & Karlík (2009).

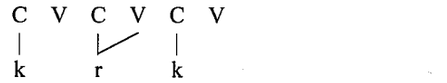
(4) phonological identity of syllabic consonants

a. solution 1: left-branching



Harris (1994:224f), Hall (1992:35f), Wiese (1996), Szigetvári (1999:117ff), Toft (2002), Scheer (2004, 2008b)

b. solution 2: right-branching

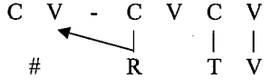


Rowicka (1999:261ff, 2003), Blaho (2001, 2004), Rennison (1999b:333ff), Ziková (2008), Scheer (2009b)

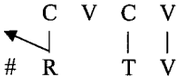
Czech is an anything-goes language (it features #RT-initial items, e.g. *rdít se*, *lhát* 'to go red, to lie') and allows the first vowel of words to alternate with zero (*pes - psa* 'dog Nsg, Gsg'). It therefore lacks the initial CV. Given this situation, there is an obvious reason why word-initial trapped liquids cannot become syllabic if syllabic consonants are left-branching: they have nothing to branch on.

(5) word-initial syllabic consonants and the initial CV

a. TR-only language: CV present



b. anything-goes language: CV absent



This analysis predicts that the existence of word-initial syllabic consonants is a fourth effect of the parametric distribution of the initial CV: they can occur in TR-only languages, but are banned from anything-goes languages. Unfortunately, this prediction turns out to be wrong. Serbo-Croatian for example has #TT and #RR clusters (e.g. *bdjeti* 'to keep guard', *ptica* 'bird', *tkati* 'to weave', *mrak* 'darkness', *mlad* 'young'), but also word-initial syllabic consonants: *rvati se* (trisyllabic) 'to fight', *rt* 'foothills' (monosyllabic), *rđa* (bisyllabic) 'rost', *rčak* (bisyllabic) 'hamster'.

It must therefore be concluded that there is no correlation between the presence/absence of word-initial syllabic consonants and the presence/absence of the initial CV. That is, the reason why Czech has no word-initial syllabic consonants is not the absence of the initial CV. In Scheer (2009b), I argue that syllabic consonants are right-branching as under (4). There are no word-initial syllabic consonants in Czech because this position is the only one where Old Czech trapped sonorants that are candidates for becoming syllabic are not preceded by other consonants, i.e. do not engage into a branching onset-type relationship: compare the final (...TR#: *bratr* 'brother') and medial (...TRT: *trvat* 'to last') with the initial (#RT...: *lžíce* 'spoon') position. On this analysis, becoming syllabic is a reaction (a repair) on the absence of a vowel to the right of *branching onsets*: TRø is not viable, but #Rø is.

3. Procedural communication with phonology

3.1. Process-specific PIC

Let us now look at the other potential sandhi-blocker, *Phase Impenetrability*. A classical and cross-linguistically pervasive observation is that cross-word phonology is process-specific. It

is not true that languages set a binary parameter that defines whether phonological processes do or do not apply across word boundaries. Typically some processes do, while others are blocked. Hence if the PIC is responsible for sandhi-blocking, it must apply to processes *à la carte*.

The process-specificity of the PIC may be illustrated in English: stress assignment is strictly limited to the word, but there is a lot of (external) sandhi. As indicated by its name, word stress is strictly bound by the limits of the word. It was already mentioned that affix classes bear on stress placement: in the traditional SPE-terminology, affixes may be stress-shifting (*párent – parént-al*) or stress-neutral (*párent – párent-hood*). Stress placement is always calculated with respect to the right edge of the word: in our example it is always penultimate (*párent – parént-al*), except when the morpheme boundary of class 2 affixes interferes, in which case the root is treated as a stress-relevant domain (i.e. a word) in its own right (*párent-hood*). The domain of stress-assignment is thus either the word or a subset thereof, the root – but never any chunk larger than the word. That is, there are no cases where word-stress would be calculated for a unit of, say, two words: the stress of *parént-al* will not shift when another word is added. *Paréntal tásks* bears two independent word stresses, and there is no global calculus that would produce something like **parentál tasks*, i.e. where penultimate stress is assigned to the domain [parental tasks] that is made of both words.[10]

Word stress thus indicates that there is some barrier that prevents phonology from applying to chunks that are larger than the word: we are in presence of a sandhi killer. This does not mean, however, that there are no phonological processes in English that apply across word boundaries. One was already mentioned in note 10: stress clash. Another is t-flapping that occurs in certain (American) varieties (Kahn 1976 and much subsequent literature): according to standard descriptions (e.g. Nespor & Vogel 1986:46f, 224ff), flapping applies in whatever syntactic environment, provided that the /t/ is word-final and intervocalic. Hence [ɾ] appears word-internally in *city* and *atom* (word-internally, the /t/ must also be post-tonic), but also word-finally in *at issue*, *a white owl*, *invite Olivia*, *at eleven* or *just the other night a racoon was spotted in our neighbourhood.*

A hard fact about sandhi phonology is thus that it is process-specific. Phonologists have tried to accommodate this situation in various ways. One is to have a more fine-grained (i.e. weaker) definition of Phase Impenetrability: previously interpreted strings are not frozen altogether; rather, only those phonological properties that are due to previous phonological computation are frozen, i.e. cannot be undone. This is Kaye's (1992, 1995) solution. In this perspective, further stress shift after the word level is blocked because stress was assigned by previous computation. By contrast, flapping across word boundaries can go into effect because the /-t/ was not modified by previous phonological action.

Another way of expressing roughly the same idea is the distinction between structure-building and structure-changing processes that was introduced in the 80s in order to rescue the *Strict Cycle Condition* (SCC) which stipulates that rules apply only to derived environments (Kiparsky 1982a:46ff, 1982b:160ff, see Scheer forth a,b). Structure-building operations such as stress assignment (on the assumption that stress incarnates as metrical structure, i.e. grids or feet) were allowed to apply to non-derived items (Halle & Vergnaud 1987:84ff), while structure-changing operations such as segmental changes (i.e. which modify pre-existing structure) – flapping in our case – can only apply to derived environments. An environment was derived either when the string was heteromorphemic, or when the target of the modification was the result of the application of a previous phonological process. The latter condition describes the same situation as Kaye's version of the PIC: items cannot be further modified if they are the result of the application of a phonological process.[11]

Another approach to the process-specificity of sandhi-blocking is – cast in modern vocabulary – to make Phase Impenetrability phase-specific. In the vocabulary of the 80s, Mohanan & Mohanan (1984) and Halle & Mohanan (1985:95ff) argue for the stratum-specificity of the SCC: in English, stratum 1 is, but stratum 2 is not cyclic (that is, stratum 1 does, but stratum 2 does not respect the SCC/the PIC).

A more direct implementation of process-specificity is the core idea of Lexical Phonology according to which word- and sentence phonology, i.e. the phonology of strings of morphemes and the phonology of strings of words, are two distinct computational systems, lexical and postlexical. Individual rules are assigned to either, or to both. In our example, the word stress rule is part of the lexical, but not of the post-lexical phonology. By contrast, flapping is present in both computational systems.

Finally, the alternative that I consider in Scheer (2009a, forth a) is also a faithful transcription of the observational fact: a process-specific PIC. In this perspective, it is specified for each process whether its application is subject to the PIC or not. In our example, stress assignment is, flapping is not. The difference with respect to the lexical vs. postlexical distinction of Lexical Phonology is that there is only one computational system, i.e. only one phonology (one set of rules, one constraint ranking). Process-sensitive PIC has also been proposed in syntax (Bošković 2007), and is implied by Marvin's (2002) analysis of English stress (where secondary, but not primary stress is marshalled by the PIC).

From all this it may be concluded that it is certainly necessary and useful to determine the phase structure of a language – knowing its precise contours, however, does not reveal much about the phonological consequences of phases since Phase Impenetrability (i.e. sandhi-blocking) is not an automatic consequence of a phase.

3.2. The initial CV is out of business for process-specific patterns

Process-specific sandhi blocking may also be (and commonly is) analyzed in terms of the Prosodic Hierarchy, i.e. representationally. In the perspective of Prosodic Phonology, this is indeed straightforward since as in Lexical Phonology, individual phonological processes are associated to a specific domain of application: we have seen in section 1 that on one possible analysis, the application of nasal assimilation in English is restricted to the *Prosodic Word*. Unlike in Lexical Phonology, though, the domain-specificity of phonological processes is done in the frame of one single computational system.

In principle, this option is not available when the representational carrier of morpho-syntactic information in phonology is a CV unit. The fundamental difference between the arboreal structure of the Prosodic Hierarchy and a CV unit is the fact that the latter is a *linear object*: it is inserted into the linear string and hence precisely located as an item that follows morpheme X and precedes morpheme Y. There is no way to talk about the linearized

[10] It is true that word stress may be subject to further modification in external sandhi: the contrast between *thirtéen* and *thírteen mén* follows the well-known pattern of so-called stress clash that modifies word stress when two word stresses are adjacent (Liberman & Prince 1977 and subsequent literature). Obviously, however, this is an independent phonological process that has got nothing to do with word stress assignment: it further modifies the *result* of word stress and obeys triggering conditions that are completely different from those that define word stress assignment. We are thus in presence of two distinct phonological processes. The relevant generalization for the present argument is simply that the phonological mechanism which is responsible for word stress is strictly bound by the limits of the word.

[11] Flapping in monomorphemic roots (*city*, *atom*) should be ruled out even when Kiparsky's distinction between structure-building and structure-changing operations is applied. The way out of this is to assume that the flap in monomorphemic items is lexically recorded: there is no evidence for an underlying /t/ (except spelling and its diachronic identity). But this discussion is idle anyway since Kiparsky (1993) has declared the bankruptcy of the entire derived environment programme.

CV in terms of a domain: like SPE's hash mark, it defines a specific *point* in the linear string. Conversely, there is no way to talk about the domains that are defined by the Prosodic Hierarchy in punctual terms, or in terms of linear precedence: a unit of the arboreal structure, say, the Prosodic Word, is not located between any two morphemes – it spans a number of them (Scheer forth a discusses this contrast at greater length).

Therefore individual processes cannot be associated with a specification for any domain of application: the CV unit does not define any such domain. Rather, it is piece and parcel of the string that is submitted to the phonological module for computation. Since modules cannot backtrack the origin of the items that they work on, the CV unit that represents morpho-syntactic information and any other CV unit that belongs to a specific morpheme are indistinguishable. As a result, once a CV unit is inserted into the input string to phonological computation, it cannot disappear or be selectively "seen" by individual phonological processes.

A la carte-visibility of the CV unit is precisely what Balogné-Bérces (2004, 2005) proposes: each phonological process is specified for ignoring (flapping in our case) or not ignoring (word stress) CV units that carry morpho-syntactic information and are present anyway. For the reasons mentioned, this does not appear to be an option: phonology is unable to tell "morpho-syntactic" and "truly phonological" CV units apart. Also note the contrast with respect to the perspective of Prosodic Phonology where nothing has to be "switched off" of made "invisible" (arboreal structure is appealed to or not by a rule).

Finally, another reason that makes à la carte-visibility incompatible with Direct Interface is that it revives the functioning of diacritic SPE-type hash marks. These were inserted into the linear string, but had no effect unless they were appealed to by a rule. CV units that are only "switched on" for certain processes are the same kind of "sleeper". Being a "sleeper", though, is the trademark of diacritics: the goal of Direct Interface is precisely to do away with representational items that are inserted into the phonological string but do not have any effect (see section 2.2).

4. The initial CV and connected speech

4.1. Introduction

Let us now look at a concrete case of external sandhi that may illustrate the general pattern. The goal is to evaluate the consequences of the initial CV in this context. We will see that it cannot be present when phonology applies across words since it would block external sandhi. The question, then, is how this relates to the fact that TR-only languages, which are supposed to possess the initial CV, may also show external sandhi. The language chosen, Belarusian, is precisely of this kind: TR-only *and* accommodating connected speech.

Before looking at the data, a disclaimer is in order: up to this point, connected speech was presented as a binary phenomenon, i.e. which either does or does not apply across word boundaries. However, many phonological processes apply across certain word boundaries, but not across others, depending on the syntactic relationship. Much of the Prosodic Phonology literature is concerned with determining the exact syntactic conditions of connected speech. The example of English flapping was chosen on purpose because it is usually described as a process that applies across word boundaries no matter what their syntactic nature (e.g. Nespor & Vogel 1986:225). This complete insensitivity for syntactic conditioning is also found in the description of other processes such as gorgia toscana (Marotta 2008), and it is assumed for the Belarusian case below. The data discussed come from my own work with an informant, but I have not worked enough on the precise syntactic conditions of the process in order to assert that no syntactic boundary can inhibit it at all. The demonstration is not impacted, though,

should there be additional syntactic sandhi-blockers: the purpose is merely to evaluate the influence of the initial CV in case phonology applies across word boundaries.

4.2. Word-initial epenthesis in Belarusian

Belarusian i-prothesis concerns the typical Slavic CVC roots which bear a vowel that alternates with zero. In Czech for example, the word for "lion" is *lev* in Nsg, but appears as *lv-a* in Gsg. All (vowel-initial) case markers provoke the absence of the root vowel.

The Belarusian situation is interesting because in addition of the regular vowel-zero alternation, a prothetic [i] appears when the root is in zero grade and either utterance-initial or preceded by a consonant-final word. By contrast, no prothesis is observed after vowel-final words. Table (6) below provides illustration. The example used is *lev*, but note that all CVC roots of the language whose vowel alternates with zero behave in the same way (e.g. *l'on* 'linen', *lob* 'forehead' etc.).

(6) i-prothesis before CVC roots that occur in zero grade

	context	example	gloss
a.	## __CVC	lev	lion NOMsg
	## __CøC-V	i-lva	lion GENsg
b.	...C # __CøC-V	brat i-lv-a	the brother of the lion
c.	...V # __CøC-V	śastra lv-a	the sister of the lion
d.	...C # __CVC	tam joś lev	there is a lion
e.	...V # __CVC	malady lev	young lion

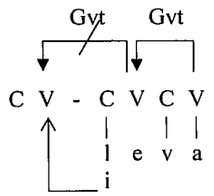
Under (6), the Gsg of the Nsg *lev* is *i-lv-a* when occurring in utterance-initial position, or when quoted in isolation. The prothesis also appears when the word comes to stand after a consonant-final word as under (6): *brat i-lv-a*. By contrast, the prothetic vowel is absent in case the preceding word ends in a vowel as under (6): *śastra lv-a*. The prothesis is also absent when the root vowel is phonetically present, and in this case irrespectively of the lefthand context: after C-final words under (6) as much as after V-final words under (6).

The prothesis is thus sensitive both to properties of its own word and to properties of the preceding word (if any). It establishes an equivalence between two consonants in utterance-initial position (/##lva/ → [ilva]) and a sequence of three consonants that belong to different words (/brat lva/ → [brat ilva]). These contexts contrast with the situation where the same clusters are preceded by a vowel-final word (/śastra lva/ → [ɕastra lva]).

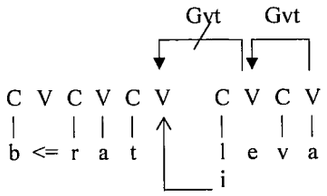
Let us start by looking at the prothesis-triggering contexts. We know what the situation looks like for *brat ilva*: the epenthetic vowel is inserted into the word-final empty nucleus of *brat*. This is shown under (7): the root vowel is silenced by the action of the case marker *-a*, and since phonology applies across word boundaries a sequence of two empty nuclei in a row is created. This configuration is ill-formed, and epenthesis is the response to the problem.

(7) epenthesis into the leftmost of two empty nuclei in a row

a. epenthesis into the (utterance-) initial CV



b. epenthesis into the final empty nucleus of the preceding word



Given this situation, the fact that epenthesis also occurs utterance-initially means that *lva* must be preceded by an empty nucleus in this position. This is shown under (7) and means that the initial CV occurs utterance-initially, but not word-initially in utterance-internal position. Indeed, its presence under (7) would create a sequence of three empty nuclei in a row, an ill-formed configuration that cannot be rescued by the epenthesis observed.

The absence of the initial CV from the beginning of the word is also critical for the configuration where no epenthesis occurs: were a CV unit inserted before the noun *lva* in *śastra lva*, prothesis into this word-initial CV would have to occur since two empty nuclei in a row would be created. As a matter of fact, though, no prothesis is observed: there is no reason for epenthesis to occur because the word-final nucleus that is empty under (7) is filled in *śastra lva*, and hence no ungoverned empty nucleus is encountered.

The conditions of epenthesis are thus very simple: all and only those empty nuclei that remain ungoverned receive melodic content in order to make the structure well-formed. A necessary ingredient of this analysis is that the initial CV occurs in utterance-, but not in word-initial position in Belarusian. In other words, in order for connected speech to go into effect, all sandhi-killers, representational as much as procedural, need to be eliminated: no empty CV may inhibit communication between words, and the words at hand must belong to the same phase (or alternatively, the process must not be subjected to Phase Impenetrability).

5. The distribution of the initial CV

Let us now take stock of the diagnostics that may be used in order to determine the (non-)distribution of the initial CV. We have seen that the initial CV must not be present when phonology applies across word boundaries. This diagnostic suffers no exception at least for processes that are conditioned by syllabic factors (i.e. involving constituent structure and lateral relations). That is, purely melody-driven processes such as, say, palatalization, may well go into effect across word boundaries even with an intervening CV unit: the spreading of melodic items may be insensitive to extra space that needs to be crossed.

Another unmistakable diagnostic for the absence of the initial CV is the presence of RT-initial morphemes in a language. The occurrence of an initial CV at the left edge of words in anything-goes languages is out of the question since the structure created, /CV-RøTV…/, is ill-formed: two empty nuclei occur in a row, and their computation in the same phase is also guaranteed.

On the other hand, a strictly TR-only lexicon does not allow for any conclusion regarding the distribution of the initial CV. Belarusian is a strict TR-only language, and we have seen that this is not an obstacle for connected speech that supposes the absence of the initial CV at the left edge of words. Put differently, there is no correlation between the type of

morpheme-initial clusters that the lexical inventory of a language allows for and external sandhi.

The Belarusian case thus shows that while the presence of RT-initial morphemes enforces the absence of the initial CV in word-initial position, the absence of the initial CV in this position does not "force" a language to develop RT-initial words. Lexical properties impact grammar, but the reverse is not true – at least when all lexical items can "survive" computation without problem, which is the case in Belarusian. Lexicon optimization describes a situation where lexical items are shaped according to the conditions that they experience during computation, but this concerns only those lexical items that will turn out to be ill-formed upon phonological computation for one reason or another ("prefer inputs that are well-formed outputs", e.g. Hale 1973:420, Yip 1996, Prince & Smolensky 1993:§9.3).

While there is no obligation for Belarusian to develop RT-initial words, the absence of the initial CV in word-initial position creates the *possibility* for such words to exist. They could happily survive after vowel-final words (only one empty nucleus: …V#RøTV…), but would produce two empty nuclei in a row utterance-initially (CV-RøTV…) and after consonant-final words (…Cø#RøTV…). However, Belarusian is equipped with a repair mechanism that inserts a vowel into the leftmost empty nucleus when a sequence of two empty nuclei is encountered (*brat i-lv-a* 'the brother of the lion'). Therefore RT-initial words could happily enter the language tomorrow.

A repair mechanism could also be imagined in anything-goes languages: if a repair effect were systematically observed word-initially with cluster-initial morphemes (but not with morphemes that begin with a single consonant), this could be ascribed to the presence of a word-initial CV.

Alongside these diagnostics for the absence of the initial CV, there are also diagnostics for its presence. This is when the two online effects of the initial CV are observed: if the initial consonant of words is strong no matter what (i.e. independently of the properties of the preceding word), or if the first vowel of words cannot alternate with zero no matter what (i.e. independently of the properties of the preceding word), words must be preceded by an empty CV unit.

It is useful indeed to distinguish online from lexical effects: table (3) identifies three correlates of the presence/absence of the initial CV. One is lexical and hence does not depend on phonological computation: we have seen that even though Belarusian *could* in principle host RT-initial words, it is strictly TR-only. By contrast, the two other effects of the initial CV depend on online computation: vowel-zero alternations and the strength of consonants are the result of phonological operations. Hence if word-initial consonants are systematically strong irrespectively of the shape of the preceding word, or if first vowels of words are unable to alternate with zero in the same conditions, this must be due to the presence of an empty CV to their left.

In sum, thus, if the management of process-specific connected speech can only be procedural, the management of the two online effects of the initial CV can only be representational.

6. Phonological and syntactic evidence for phase boundaries

6.1. What the initial CV is initial of are phases

In the above discussion we have come across two patterns: the initial CV may occur either word- or utterance-initially. These two chunk sizes were defined on the basis of positive evidence: languages show the two online effects of the initial CV at their left edge. Inside these chunks phonological computation is continuous: no edge-effects are observed within

words (there is nothing like morpheme-initial strength) in those languages where the initial CV is word-initial, and edge-effects are unheard of within utterances in the other type of language (precisely because phonology applies across word boundaries).

It may thus be concluded that the portions of the string that we are talking about are computational domains, i.e. phases in modern vocabulary. This means that what the initial CV is actually initial of are phases (in the sense of current syntactic theory). The initial CV may therefore be regarded as a marker of phase boundaries. We have come across phonological evidence for two chunk sizes that have this status, but nothing indicates a priori that there is no language that shows phonological effects of phase boundaries for some intermediate chunk size, i.e. between the word and the utterance. At the same time, it is obvious that these two specific chunk sizes are critical and recurrent "barriers" for phonological processes across languages.

6.2. Phases may or may not leave phonological traces

Given our ability to detect phase boundaries on phonological grounds, an interesting question is of course how the phonological definition of phases correlates with morpho-syntactic phase structure. The mapping is certainly not trivial. 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 provides an overview): the DP track is followed, and also DP-internal phases are argued for (Matushansky 2005). TP is also 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 being 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, but even on the most conservative count, i.e. Chomsky's initial vP and CP, there is a "syntactic" phase between the word and the utterance: vP. Less conservative perspectives place many more phase boundaries in this area, none of which seems to leave phonological traces.

It is hard to believe that this is due to insufficient analysis, or to the lack of cross-linguistic study of phonological traces of phase boundaries. That is, it is hard to imagine a language where word-initial consonants are strong, and first vowels of the word stable, but only in words that happen to be vP-initial (or TP-initial etc.). Also, we have seen a language, Belarusian, where there is definitely no phonological trace of the spell-out of chunk sizes that range between the word and the utterance: (at least) vP will be a phase in Belarusian as well, but its spell-out does not leave any phonological trace.

All this does not really come as a surprise: we know that procedural sandhi-killers are process-specific (see section 3.1): should the PIC be responsible for blocking word-stress computation across word boundaries in English, it applies *à la carte* to stress, but not to flapping. This leads to the same conclusion as for the distribution of the initial CV: chunks (i.e. phases) must be determined in morpho-syntax and exist independently of whether they are "mobilized" in phonology, i.e. of whether they leave a phonological trace. Phonological traces are produced by either making a (specific) phonological computation subject to the PIC

at a specific chunk size, or by inserting a CV unit at the left edge of a phase. That is, phase boundaries may, but do not need to be accompanied by phonological effects.

Given our current understanding of syntactic phase structure and the (incomplete) inventory of chunk sizes that produce phonological traces, the only good match between "syntactic" and "phonological" phases is the CP, which corresponds to what was called the utterance above. The other chunk size that produces phonological effects (and massively so), the word, is not anything that is known to be a syntactic phase.

Table (8) below recapitulates the bumpy match between syntactic phases and chunk sizes that produce phonological effects.

(8) bumpy match between syntactic and phonological evidence for phases

phases (syntactic evidence)	phases (phonological evidence)	
CP	utterance	good match
vP	–	no phonological trace
TP	–	no phonological trace
DP	–	no phonological trace
…	–	no phonological trace
–	word	no syntactic trace

The bottom line is thus that somewhere a decision is made regarding the distribution of the initial CV over phases: this phase gets one, but that phase does not. A subset of the match between phases and initial CVs may be universal: some phases, say, TP (if the phasehood of TP is confirmed) may be disqualified altogether for carrying the initial CV. This may explain the lack of phonological response for most of the lines under (8).

In any event, though, there must be some space for parameterisation since we know that the word may or may not be elected as a CV-carrying phase (connected speech is the result of its non-election).

7. Conclusion

This article has proposed a number of criteria that allow us to determine whether the "barriers" that prevent processes from applying across word boundaries are of representational or procedural nature. The diagnostics are established on the grounds of assumptions regarding representational and procedural interface management: on the latter side, current syntactic phase theory, i.e. including Phase Impenetrability, is assumed; on the former, CVCV in general and Lowenstamm's (1999) initial CV in particular are the basis of the discussion. Developing some tenets of Direct Interface (Scheer 2008a, forth a), it is argued that the insertion of CV units into the linear string at morpheme/word boundaries is the only way for morpho-syntax to bear on phonology by representational means.

Given these premises, process-specific sandhi-blocking, a hard fact of sandhi phonology, cannot be due to representational intervention: it must be a consequence of a PIC condition, which is or is not associated to individual phonological processes. That is, it is important and useful to know what the "syntactic" phase structure in a language looks like, but this does not tell us anything about its phonological consequences: phases may or may not enforce Phase Impenetrability on the phonological side.

The same generalisation is made on the representational side. Parametric variation was studied regarding the question what the initial CV can be initial of. Two cases have been identified: the initial CV heads words in some languages, utterances in others (and there may well be other chunk sizes that are able to receive a CV unit). In the former case it acts as a sandhi-blocker, while phonology freely applies across word-boundaries in the latter. Or rather, to be precise, external sandhi supposes the absence of both sandhi blockers, representational and procedural alike: in order for phonology to apply across word boundaries, no PIC must be associated to the chunk that is defined by the boundary in question, and no CV unit must stand in the way.

Given that the initial CV always heads computational domains, it turns out to be phase-initial in all cases. Its distribution, though, is just like the distribution of the PIC: phases may or may not be headed by CV units on a language-specific basis. That is, the absence of the initial CV in word-initial position in Belarusian does not mean that words are not phases in this language. It just means that Belarusian does not distribute a CV unit with this phase. Or, in other words, that units that the initial CV can be initial of are only phases, but that not every phase is headed by an initial CV.

In sum, the skeleton of cross-word phonology is phase structure: every phase boundary may be armed with a PIC and an initial CV. On the zero hypothesis that the distribution of both is independent, a four-way parametric space is opened for each phase. Table (9) below shows the situation for the phase that is associated to the word.

(9) sandhi-killers: parametric variation at the word level

language	symptoms			the word-level phase is	
	preceding word taken into account	word-initial C is	first vowel of the word alternates with zero without repair	preceded by an initial CV	is armed with a PIC
a. French, English etc.	–	strong	no	yes	yes
b. ?	–	strong	no	yes	no
c. Greek	–	weak	yes	no	yes
d. Belarusian	✓	weak	yes	no	no

It was mentioned that phonological processes can apply across word boundaries only when neither sandhi-killer is active as under (9)d. The presence of only one barrier creates two different patterns. In one configuration, (9)c, cross-word phonology that regards the strength of word-initial consonants and the alternation of the first vowel of words is inhibited by the PIC, but the initial CV is absent. This means that the preceding word is cut off and will not be taken into account for the calculation of word-internal matters. The absence of the initial CV, however, makes word-initial consonants weak, and the first vowel of the word may freely alternate with zero (without this provoking any repair). This is the pattern found in Greek (see Seigneur-Froli 2003, 2006).

The reverse configuration (9)b is a situation where the PIC does not inhibit cross-word phonology, but where a CV unit separates words. This pattern will act as a sandhi-killer for syllable-based processes: word-initial consonants will be strong, and first vowels of words will not be able to alternate with zero (without repair). Purely melody-driven processes such as palatalizations, however, will be able to cross the word boundary (in case they are unimpacted by the PIC). I am not aware of any language that instantiates this pattern, which however does not appear to be outlandish.

Finally, the "regular" Indo-European case is (9)a where all communication is cut off between words: the preceding word is never taken into account for the calculus of the strength of word-initial consonants and the ability of the first vowel of words to alternate with zero (there is a PIC effect on these processes), no melody-driven processes cross word boundaries (there is a PIC effect on these processes as well), word-initial consonants are strong and the first vowel of words does not alternate with zero (effects of the initial CV).

References

Aelbrecht, Lobke. 2008. Licensing ellipsis as Agree. Paper presented at the *27th West Coast Conference on Formal Linguistics*, UCLA, May 16–18 2008.

Balogné-Bérces, Katalin. 2004. Connected speech phenomena in Strict CV phonology. *The Even Yearbook* 6: 1–10.

Balogné-Bérces, Katalin. 2005. Strict CV Phonology and the English Cross-Word Puzzle. Doctoral dissertaton, Eötvös Loránd University, Budapest.

Baltin, Mark. 2007. Deletion versus Pro-forms: a false dichotomy? Ms., New York University.

Bendjaballah, Sabrina. 2001. The negative preterite in Kabyle Berber. *Folia Linguistica* 34: 185–223.

Bendjaballah, Sabrina & Martin Haiden. 2003a. Meaningful vowels. In P. Bottari (ed.) *Proceedings of the XXVIII Incontro di Grammatica Generativa*. Lecce: Congedo Editore.

Bendjaballah, Sabrina & Martin Haiden 2003b. Templatic Architecture. *Recherches Linguistiques de Vincennes* 32: 157–168.

Bermúdez-Otero, Ricardo. forth. Stratal Optimality Theory. Oxford: Oxford University Press.

Blaho, Sylvia. 2001. The representation of Slovak syllabic consonants in strict CV. *The Odd Yearbook* 6: 3–24.

Blaho, Sylvia. 2004. Syllabic consonants in strict CV. MA thesis, Pazmany Peter University.

Bošković, Željko. 2007. Agree, Phases, and Intervention Effects. *Linguistic Analysis* 33: 54–96.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels & J. Uriagereka (eds.) *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: The MIT Press. 89–155.

Chomsky, Noam. 2001. Derivation by Phase. In M. Kenstowicz (ed.) *Ken Hale: A Life in Language*. Cambridge, Mass.: The MIT Press. 1–52.

Chomsky, Noam. 2004. Beyond explanatory adequacy. Structures and Beyond. In A. Belletti (ed.) *The Cartography of Syntactic Structures. Volume 3.* Oxford: Oxford University Press. 104–131.

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

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

Cyran, Eugeniusz. 2003. *Complexity Scales and Licensing Strength in Phonology*. Lublin: KUL.

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

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

Epstein, Samuel & T. Daniel Seely. 2002. Rule Applications as Cycles in Level-Free Syntax. In S. Epstein & T. D. Seely (eds.) *Derivation and Explanation in the Minimalist*

Program. Oxford: Blackwell. 65–89.
Epstein, Samuel & T. Daniel Seely. 2006. *Derivation in Minimalism*. Cambridge: Cambridge University Press.
Guerssel, Mohand & Jean Lowenstamm. 1990. The derivational morphology of the Classical Arabic Verb. Ms., Université Paris 7 & UQAM.
Hale, Ken. 1973. Deep-surpface canonical disparities in relation to analysis and change: an Australian case. In T. A. Seboek (ed.) *Diachronic, areal, and typological linguistics*. The Hague: Mouton. 401–458
Hall, Tracy. 1992. *Syllable Structure and Syllable-Related Processes in German*. Tübingen: Niemeyer.
Halle, Morris & Karuvannur Mohanan. 1985. Segmental Phonology of Modern English. *Linguistic Inquiry* 16: 57–116.
Halle, Morris & Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, Mass.: The MIT Press.
Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
Hinterhölzl, Roland. 2006. The Phase Condition and cyclic Spell-out: Evidence from VP-topicalization. In M. Frascarelli (ed.) *Phases of Interpretation*. Berlin: Mouton de Gruyter. 237–259.
Inkelas, Sharon. 1990. *Prosodic Constituency in the Lexicon*. New York: Garland.
Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. Doctoral dissertation, MIT.
Kaye, Jonathan. 1992. On the interaction of theories of Lexical Phonology and theories of phonological phenomena. In U. Dressler, H. Luschützky, O. Pfeiffer & J. Rennison (eds.) *Phonologica 1988*. Cambridge: Cambridge University Press. 141–155.
Kaye, Jonathan. 1995. Derivations and interfaces. In J. Durand & F. Katamba (eds.) *Frontiers of phonology*. London & New York: Longman. 289–332.
Kiparsky, Paul. 1982a. Lexical morphology and phonology. In In-Seok Yang (ed.) *Linguistics in the morning calm*. Seoul: Hanshin. 3–91
Kiparsky, Paul. 1982b. From Cyclic Phonology to Lexical Phonology. In H. van der Hulst & N. Smith (eds.) *The structure of phonological representations I*. Dordrecht: Foris. 131–175.
Kiparsky, Paul. 1993. Blocking in nonderived environments. In S. Hargus & E. Kaisse (eds.) *Studies in Lexical Phonology*. San Diego: Academic Press. 277–313.
Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17: 351–365.
Komárek, Miroslav. 1969. *Historická mluvnice česká. Hláskosloví*. Praha: SPN.
Lahrouchi, Mohamed. 2001. Aspects morpho-phonologiques de la dérivation verbale en berbère (parler chleuh d'Agadir). Doctoral dissertation, Université Paris 7.
Lehr-Spławiński, Tadeusz & Zdzisław Stieber. 1957. *Gramatyka historyczna języka czeskiego*. Warszawa: PWN.
Liberman, Mark & Alan Prince. 1977. On Stress and Linguistic Rhythm. *Linguistic Inquiry* 8: 249–336.
Liewehr, Ferdinand. 1933. *Einführung in die historische Grammatik der tschechischen Sprache. 1. Teil: Lautlehre, Erste Lieferung*. Brünn: Rohrer.
Lowenstamm, Jean. 1996. CV as the only syllable type. In J. Durand & B. Laks (eds.) *Current trends in Phonology. Models and Methods*. Salford, Manchester: ESRI. 419–441.
Lowenstamm, Jean. 1999. The beginning of the word. In J. Rennison & K. Kühnhammer (eds.) *Phonologica 1996*. La Hague: Holland Academic Graphics. 153–166
Marotta, Giovanna. 2008. Lenition in Tuscan Italian (Gorgia Toscana). In J. Brandão de Carvalho, T. Scheer & P. Ségéral (eds.) *Lenition and Fortition*. Berlin: Mouton Mouton de Gruyter. 235–271.

Marvin, Tatjana. 2002. Topics in the Stress and Syntax of Words. Doctoral dissertation, MIT.
Mascaró, Joan. 1976. Catalan Phonology and the Phonological Cycle. Doctoral dissertation. MIT.
Matushansky, Ora. 2005. Going through a phase. *MIT Working Papers in Linguistics* 49: 157–181.
Mohanan, Karuvannur & Tara Mohanan. 1984. Lexical Phonology of the consonant system in Malayalam. *Linguistic Inquiry* 15:575–602.
Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
Newell, Heather. 2008. Aspects of the morphology and phonology of phases. Doctoral dissertation, McGill University, Montreal.
Newell, Heather & Tobias Scheer. 2007. Procedural First. Paper presented at the *38th Poznań Linguistic Meeting*, Poznań, September 13–16 2007.
Pagliano, Claudine. 2003. L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie. Doctoral dissertation, University of Nice.
Prince, Alan & Paul Smolensky. 1993. Optimality Theory. Constraint Interaction in Generative Grammar. Ms, Rutgers University, University of Colorado.
Rennison, John. 1999. Syllables in Western Koromfe. In H. van der Hulst & N. Ritter (eds.) *The syllable, Views and Facts*. Berlin, New York: Mouton de Gruyter. 311–347
Rowicka, Grażyna. 1999. On Ghost vowels. A Strict CV Approach. Doctoral dissertation, Leiden University.
Rowicka, Grażyna. 2003. /r/ syllabicity: Polish *versus* Bulgarian and Serbo-Croatian. In S. Ploch (ed.) *Living on the Edge. 28 papers in honour of Jonathan Kaye*. Berlin, New York: Mouton de Gruyter. 511–526.
Rubach, Jerzy. 1984. *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris.
Rubach, Jerzy. 1997. Extrasyllabic consonants in Polish: Derivational Optimality Theory. In I. Rocca (ed.) *Derivations and Constraints in Phonology*. Oxford: Clarendon. 551–581.
Rubach, Jerzy & Geert Booij. 1984. Morphological and prosodic domains in Lexical Phonology. *Phonology Yearbook* 1: 1–27.
Scheer, Tobias. 2004. *A Lateral Theory of Phonology. Vol.1: What is CVCV, and why should it be?* Berlin: Mouton Mouton de Gruyter.
Scheer, Tobias. 2008a. Why the Prosodic Hierarchy is a diacritic and why the Interface must be Direct. In J. Hartmann, V. Hegedüs & H. van Riemsdijk (eds.) *Sounds of Silence: Empty Elements in Syntax and Phonology*. Amsterdam: Elsevier. 145–192.
Scheer, Tobias. 2008b. Syllabic and Trapped Consonants in (Western) Slavic: the Same but yet Different. In G. Zybatow, L. Szucsich, U. Junghanns & R. Meyer (eds.) *Formal Description of Slavic Languages: The Fifth Conference, Leipzig 2003*. Frankfurt am Main: Peter Lang. 149–167.
Scheer, Tobias. 2009a. Intermodular Argumentation and the Word-Spell-Out-Mystery. In K. Grohmann (ed.) *Explorations of Phase Theory: Interpretation at the Interfaces*. Berlin: Mouton de Gruyter. 23–65.
Scheer, Tobias. 2009b. Syllabic and trapped consonants in the light of branching onsets and licensing scales. In G. Zybatow, U. Junghanns, D. Lenertová & P. Biskup (eds.) *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure. Proceedings of FDSL 7, Leipzig 2007*. Frankfurt am Main: Peter Lang.
Scheer, Tobias. forth a. *How morpho-syntax talks to phonology. A survey of extra-phonological information in phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton Mouton de Gruyter.
Scheer, Tobias. forth b. Intermodular argumentation: morpheme-specific phonologies are out

of business in a phase-based architecture. In N. Shir & L. Rochman (eds.) *The Sound Pattern of Syntax*. Oxford: OUP.
Ségéral, Philippe & Tobias Scheer. 2001. La Coda-Miroir. *Bulletin de la Société de Linguistique de Paris 96*:107–152.
Ségéral, Philippe & Tobias Scheer. 2005. What lenition and fortition tells us about Gallo-Romance. In T. Geerts, I. van Ginneken & H. Jacobs (eds.) *Muta cum Liquida*. In *Romance Languages and Linguistic Theory 2003*. Amsterdam: Benjamins. 235–267.
Ségéral, Philippe & Tobias Scheer. 2008. The Coda Mirror, stress and positional parameters. In J. Brandão de Carvalho, T. Scheer & P. Ségéral (eds.) *Lenition and Fortition*. Berlin: Mouton Mouton de Gruyter. 483–518.
Seigneur-Froli, Delphine. 2003. Diachronic consonant lenition & exotic word-initial clusters in Greek: a unified account. In M. Stavrou-Sifaki & A. Fliatouras (eds.) *Studies in Greek Linguistics. Proceedings of the 23rd annual meeting of the department of linguistics*. Thessaloniki: University of Thessaloniki. 345–357.
Seigneur-Froli, Delphine. 2006. Le Statut phonologique du début de mot grec. Lénitions consonantiques et libertés phonotactiques initiales dans la diachronie de la langue commune et dans le dialecte de Lesbos. Doctoral dissertation, University of Nice.
Selkirk, Elisabeth. 1981 [1978]. On prosodic structure and its relation to syntactic structure. In T. Fretheim (ed.) *Nordic Prosody II*. Trondheim: TAPIR. 111–140.
Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: The MIT Press.
Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology* 3:371–405.
Szigetvári, Péter. 1999. VC Phonology: a theory of consonant lenition and phonotactics. Doctoral dissertation, Eötvös Loránd University, Budapest.
Szigetvári, Péter & Tobias Scheer. 2005. Unified representations for the syllable and stress. *Phonology* 22:37–75.
Toft, Zoë. 2002. The phonetics and phonology of some syllabic consonants in Southern British English. *ZAS Papers in Linguistics* 28:111–144.
Trávníček, František. 1935. *Historická mluvnice československá*. Praha: Melantrich.
Trubetzkoy, Nikolai S. 1939. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
Vogel, Irene. 1991. Level ordering in Italian Lexical Phonology? In P. M. Bertinetto, M. Kenstowicz & M. Loporcaro (eds.) *Certamen Phonologicum II*. Torino: Sellier & Rosenberg. 81–101.
Wiese, Richard. 1996. *The Phonology of German*. Oxford: Oxford University Press.
Yip, Moira 1996. Lexicon Optimization in languages without alternations. In J. Durand & B. Laks (eds.) *Current trends in Phonology. Models and Methods*. Salford, Manchester: ESRI. 759–790.
Ziková, Markéta. 2008. Alternace e-nula v současné češtině – laterální autosegmentální analýza. Doctoral dissertation, Masaryk University in Brno.
Ziková, Markéta & Petr Karlík. 2009. Syntax a fonologie staročeského supina. Paper presented at the conference *Užívání a prožívání jazyka*, Prague, April 15–17 2009.

Paradigmatic Bridge – Accentual and Quantitative Paradigms of Czech and Slovak Feminine *a*-stems

Roman Sukač
Silesian University in Opava

1. Problem

Observing the case forms of Czech disyllabic *a*-stems we notice that the Nsg and Gpl differ in quantity: *kráva – krav, hlína – hlín, slina – slin* although all those words had acute intonation in Proto-Slavic: **kőrva – kőrvъ, *glĩna – glĩnъ, slĩna – slĩnъ*. The similar quantitative dispersion can be seen in Moravian dialects (*krava – krav/krava – kráf*), in Silesian dialects *krova – kruv* and also in Slovak *krava – kráv, hlina – hlín, slina – slín* but *sláva – sláv*. To explain those phenomena, we must come back to the original accentual distribution of nouns.

Since the revolutionary works of Stang (1957), the modern Balto-Slavic accentology operates with accentual paradigms (*AP*). All Proto-Slavic derivates and non-derivates are distributed among the three accentual paradigms. *APa* had a constant root stress and acute intonation[1], *APb* contained oxytona[2] created by a progressive stress shift due to the Dybo's Law (Dybo 1962). Mobilia with circumflex intonation in stem stressed forms and ending stress in other case forms were distributed in *APc*.[3] Pre-Stang classical accentology thought that acute in disyllabic forms continues as length in Czech and that elsewhere in the West Slavic territory the acute length was shortened. Acute length can be seen only in Czech dialects (*kráva*) and Upper Sorbian (*kruwa*). According to Kortlandt (1975), the length in *kráva* type is due to the lengthening in the open syllable.[4] Combining this claim with Bethin's opinion that Slavic has a tendency to create disyllabic domains (Bethin 1998), the length in *kráva* is quite understandable. *APb* is reflected in Czech either as a pretonic length, e.g. *brázda* or as a length from the so called compensatory lengthening, e.g. *stůl*. Otherwise it is short, e.g. *skot*. Nouns from *APc* are quantitatively short, e.g. *brod,* or are lengthened due to the compensatory lengthening, e.g. *hnůj*. Details and complexity of compensatory lengthening in Slavic are not dealt here.

Synchronically, the quantity in Czech disyllabic substantives is distributed as follows: first, we have a derivative length as in *[rybář] – ry[bařík]* type where the quantity alternation operates in disyllabic maximally threemoraic domain.[5] Derivates posses the tendency to rhythmicity. This rhythmicity is either regular or less regular – some derivates have a normal rhythmic law (as *-nie* deverbatives in Old Czech[6], hypocoristics and prefixed deverbatives in Modern Czech[7]), others have only rests of rhythmicity and the distribution of length is asymetric (Modern Czech *-ař/-ář* substantives).[8] Derivative length is stable in Czech paradigms, the length does not change.

[1] Nsg **kőrva*, Gsg **kőrvy*.
[2] Nsg **kosà*, Gsg **kosý*.
[3] Nsg. *golvà*, Asg. **gôlvǫ*.
[4] "A short rising vowel in an open first syllable of disyllabic words was lengthened in early Czech unless the second syllable contained a long vowel." (Kortlandt 1975:19)
[5] Brackets show the domain. In disyllabic (*-ař/-ář*) substantives the quantity cannot exceed three moras. Adding a long suffix (*-ík*) to a disyllabic (*-ář*) substantive, the domain shifts to the left, so there is no ***rybářík*.
[6] e.g. *dávati > dávanie, dělati > dělánie*
[7] *Kateřina >Káťa/Katka, nalévat > nálevka*, see Bethin (2003a,b) for details.
[8] Asymetry means that in some derivates containing the disyllabic maximally three-moraic domain we cannot say why the length is distributed in such a way. We know that the domain is active but the quantitative behavior