Linguistica Silesiana 39, 2018
ISSN 0208-4228

MARCIN FORTUNA
University of Gdańsk
marcin.fortuna@ug.edu.pl

PROSODIC TYPOLOGY IN LIGHT OF LICENSING SCALES

The objective of the paper is to use the model of Complexity Scales and Licensing (Cyran 2003, 2010) to account for the existence of two prosodic types: 'syllable' and 'word' languages (Auer 1993, Szczepaniak 2007), which roughly correspond to syllable-timed and stress-timed languages. We will postulate that these categories are not primitive and that many of their phonological characteristics can be derived from simpler mechanisms of licensing. It will be also argued that the phenomenon of contrast plays an important role in prosodic typology and may influence syllable structure. Languages use more marked syllabic configurations in order to optimise contrast expression. We will carry out an analysis on a simple hypothetical language in order to demonstrate the interdependence of syllabic complexity and the contrastive potential of a syllabic unit.

Keywords: *phonology*, *CVCV*, *syllable structure*, *prosody*, *typology*

1. Introduction

The paper focuses on the approach to prosodic typology proposed by Auer (1993), Szczepaniak (2007), Reina & Szczepaniak (2014), according to which languages choose either the syllable or the phonological word as the primary prosodic unit and organize most of the phonological system around it (i.e. they 'profile' it). Yoruba is a canonical 'syllable language', whereas English is a prime example of a 'word language'.

In its original formulation, this perspective is inherently based upon the traditional units of phonological analysis: syllables and phonological words. However, there are some approaches to phonology which do not recognize traditional units of Prosodic Hierarchy. For example, CVCV phonology (Scheer 2004, 2012) assumes that syllable structure is expressed via lateral forces and that higher-order units (like phonological words) are only diacritics storing morphosyntactic information, and thus they are redundant.

The aim of this paper is to make an attempt to capture the observations made by the advocates of the 'syllable-' vs. 'word-language' typology from the point of view of the CVCV model (especially in the implementation of Cyran 2003, 2010) without reference to prosodic units like syllables and phonological words. It will be argued that this model allows us to express generalisations in a comparably efficient manner and to shed some new light on prosodic typology.

2. Syllable and word languages

The division into syllable and word languages is not an entirely new approach. It essentially builds upon older work on syllable-timing and stress-timing (Donegan and Stampe 1983, Dauer 1983), but unlike its predecessor, it denies that the typological division should be based solely on rhythm. The focus is on a wider variety of phonological factors.

Both prototypical 'syllable languages' and prototypical 'word languages' come with a bundle of characteristic phonological correlates. One of the fundamental differences between syllable and word languages lies in the complexity of syllable structure. Syllable languages have a strong preference for simple syllables, in most cases consisting of a single consonant followed by a single vowel and no coda (CV syllables). At the same time, usually they do not display discrepancies with regard to the distribution of syllables of various types within a word. All syllables permissible in a language can occur in every conceivable position, regardless of such factors as stress. In contrast, word languages typically display complex onsets and codas, with huge differences between the stressed and unstressed position. Syllable languages are much less likely to violate the Sonority Sequencing Principle, whereas word languages are much more likely to have vowel reduction. Syllable languages often simplify extant consonant clusters and epenthesise vowels, whereas word languages drop vowels and insert consonants. Resyllabification across morphosyntactic boundaries can be expected in a syllable language, and to a much smaller extent in a word language. Finally, syllable languages make use of such devices as vowel harmony and tone, which are hardly ever found in word languages.

The following table (based on Auer 1993, Szczepaniak 2007) summarises these typological correlations for ease of reference.[1]

It needs to be kept in mind that there are no languages which can be unambiguously assigned to only one of these types (Auer 2014: 3). One should not expect to find a language which would not display any features of the opposing type whatsoever. The typology is rather viewed as a continuum, with each language occupying a given position on the scale, but features characteristic of one type will be usually dominant.

[1] Only the features relevant to this paper were included.

Table 1. Features of syllable and word languages

Syllable languages	Word languages
small syllable complexity	big syllable complexity
little dependence between syllable complexity and stress	differences in syllable complexity between stressed and unstressed syllables; complex consonantal structures on word edges
sonority scale observed	sonority scale violated
no vowel reduction	phonological or phonetic vowel reduction
vowel epenthesis	vowel elision
consonant cluster simplification	consonant epenthesis
resyllabification across morphological and syntactic boundaries	no resyllabification across morphological and syntactic boundaries
vowel harmony	no vowel harmony
tone	no tone
length opposition possible in all positions	length opposition possible only in stressed syllables

As can be inferred from Table 1, in a canonical ‘syllable language’ most phonological rules aim at optimising the syllable, i.e. bringing it as close as possible to the unmarked CV structure. In contrast, ‘word languages’ mostly profile the phonological word, by emphasising the stressed syllable and word edges.

2.1. A UG parameter?

From the perspective of I-language, one could be tempted to state that the ‘syllable-/word-language’ dichotomy is encoded in Universal Grammar in the form of a binary parameter. This is not a very likely option, especially due to the continuum nature of the distinction. In the case of a parameter we would expect a more rigorous, and much less selective, concentration of the features of one type in each language.

However, regardless of the existence of many languages falling in the middle of the continuum, the commonly attested typological correlations between all of the phenomena ascribed to syllable or word language are quite striking and still demand an explanation. The explanation can be to some extent dependent on UG-related factors. It will be argued in Section 4 that a UG parameter is actually involved, but it will be given a much more specific and parsimonious formulation.

3. CVCV approaches to syllable structure

Our explanation of the prosodic typology will be formulated within an unconventional phonological model, which explicitly rejects traditional units of prosodic organization. CVCV (Lowenstamm 1996, Szigetvári 1999, Scheer 2004, 2012) is defined by Scheer (2004) as a *lateral* or *syntagmatic* theory of phonology. According to its underlying assumptions, the phonological constituent structure is entirely flat: there are no arboreal units like syllables or rhymes. Instead, the smallest prosodic constituent is a CV unit, i.e. a sequence of an onset and a nucleus. It is impossible in CVCV to have a 'C' or a 'V' unit in isolation. However, not all of the CV units need to be associated with melody (phonetic substance) – some of them may be left unexpressed. The following figure contains a CVCV representation of the English word *guarantee*, which will illustrate the constituent structure advocated by the model.

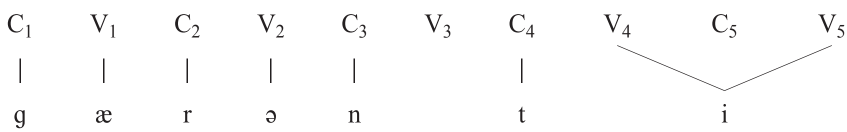

Figure 1. Guarantee in CVCV

As we can see, a consonant cluster (in this case /nt/) contains an empty nucleus position, whereas a long vowel (in this case /iː/) straddles an empty onset.

The effects and regularities normally ascribed to syllabic arborescence (such as phonotactics, or lenition and fortition) are accounted for by means of the so-called lateral forces, relations between segments in a flat string. What kind of lateral forces we should recognize is a matter of dispute: it differs from implementation to implementation. Due to space restrictions, we will review only the lateral forces recognized in the model relevant to this paper: Cyran's (2003, 2010) *Complexity Scales and Licensing* theory.[2]

3.1. Licensing scales

The *Complexity Scales and Licensing* (CSL) model uses interonset relations to account for the existence of consonant clusters. Since there are no constituents like 'rhyme', and the onset cannot branch, each consonant cluster needs to be assumed to contain an empty nucleus. Whether a given consonant sequence is

[2] For a fuller discussion of lateral forces in other implementations the interested reader may consult Szigetvári (1999), Ségéral and Scheer (2001), Scheer (2004), Fortuna (2015).

parsed as a 'branching onset' or a 'coda-onset cluster' can be typically read off from the sonority profile of the participating consonants.

The lateral force responsible for the creation of consonant clusters is referred to as Interonset Government, which the stronger (less sonorous) member of the cluster exerts upon the weaker (more sonorous) consonant. Coda-onset clusters, in which the right-hand member is typically stronger, are created by LIO (Leftward InterOnset Government), and branching onsets by RIO (Rightward InterOnset Government).

Another crucial lateral force in the CSL toolbox is licensing. Licensing is a force which legitimises the existence of consonantal structures. Every V-position in a CV-slot licenses its C-position. Also empty V-positions license their onsets – this is how words ending in a consonant can exist, e.g. *bat*.

Licensing also plays the fundamental role in the process of cluster formation, since Interonset relations need to be licensed too. To be able to contract a relation, each governor needs to be licensed by the following nucleus to govern. This relation is called Government-Licensing.

Figure 2 illustrates all of the lateral forces we have described so far:

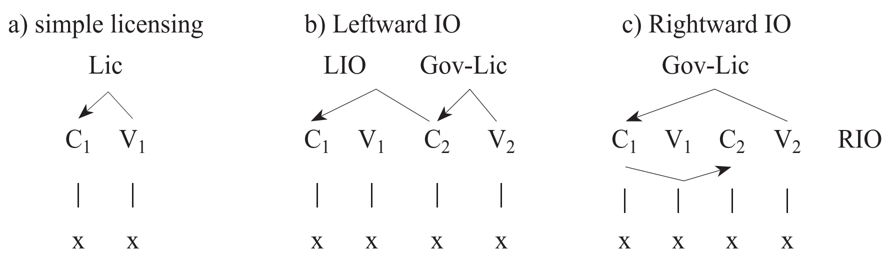

Figure 2. Interonset relations in Cyran (2010)[3]

An observation which is quite easy to make on the basis of Figure 2 is that there are differences in the difficulty of licensing respective phonological structures. The situation in 2a is the most trivial one: a single consonant is licensed by the following nucleus. In 2b the nucleus needs to license the whole coda-onset cluster, being thus responsible for *two* preceding consonants. Finally, in 2c licensing is not local, with the nucleus forced to reach the second preceding consonant. So single consonants are easier to license than coda-onset clusters, and both these structures are easier to license than branching onsets. Interestingly, the scale of licensing difficulty nicely coincides with the scale of markedness of consonantal structures proposed by Kaye and Lowenstamm (1981), according to which branching onsets are more marked than coda-onset clusters, which are consequently more marked than single consonants.

---

[3] Cyran (2010) uses a notation with 'O' for onset and 'N' for nucleus, however, we use a uniform notation with 'C' and 'V' throughout the paper.

This implicational relation is inbuilt in the mechanism of CSL. If a language has nuclei which are strong enough to license LIO, it follows naturally that they can also license single consonants. If nuclei can license RIO, they must also be able to license the two simpler structures. The licensing strength of nuclei is assumed to be a phonological parameter. The parameter values form a three-step scale:

	Structure	**Example**	**example language**
I	*Ca*	ba**by**	Desano
II	*R.Ta*	wi**nte**r	Japanese
III	*T.Ra*	**tra**p	English

Figure 3. Licensing strength of nuclei in Cyran (2010: 93)

Languages which pick value ‘I’ lack consonant clusters. Languages of the ‘II’ type have only coda-onset clusters, but not branching onsets. Finally, languages picking ‘III’ display all three syllabic structures.

The three levels of complexity are juxtaposed with another variable: the types of nuclei carrying the burden of licensing. The stronger the nucleus, the bigger its relative licensing power. A full vowel will always be able to license more (or at least *not less*) than a schwa (or another reduced vowel), and a schwa will always be able to license more (or at least *not less*) than an empty nucleus. These relationships are depicted in Figure 4:

		[a]		[ə]		[Ø]
I	C_	Ca	⊂	Cə	⊂	CØ
		∩		∩		∩
II	RT_	RTa	⊂	RTə	⊂	RTØ
		∩		∩		∩
III	TR_	TRa	⊂	TRə	⊂	TRØ

Figure 4. The full net of syllable structure relations in CSL

It implies that even if a language possesses all syllabic configurations, it does not follow that all of them will be found in all positions within the word. To exemplify this, let us take a cursory look at the right edge of English words. As it is evident from Figure 3, English allows all three types of syllabic structure, but at the end of the word only single consonants and clusters of falling sonority can be found. Clusters of rising sonority are absent from this position.[4] This is the evidence for the smaller licensing potential of the word-final empty nucleus. It can license level I and II, but not III. Exactly this kind of behaviour would we expect according to CSL assumptions – an empty nucleus has a smaller licensing strength than full and reduced nuclei.

4. A CVCV approach to prosodic typology

Let us now attempt to combine the two concepts introduced in the previous sections. The theory of prosodic profiling makes several useful generalisations, using at the same time traditional units of the Prosodic Hierarchy. It is to a large extent successful at characterizing cross-linguistic phonological patterns, even though it does not achieve full empirical coverage.

We also argued in section 2.1 that when approached from the I-language angle, the two poles postulated by proponents of the theory cannot be considered to be two settings of a phonological parameter. Now, having introduced CVCV and CSL, we will propose that some aspects of the typological division can be dependent on UG-related factors and can be ascribed to a binary parameter, with some non-grammatical aspects also playing a role.

How can CVCV contribute to the problem? There are no syllables, so the syllable cannot be profiled. Similarly, the Phonological Word is not always recognized by CVCV phonologists (see Scheer 2012 for a well-articulated criticism of the Prosodic Hierarchy). There is no way to express the generalisation directly. However, this is not a problem, since CVCV still allows us to refer to individual rules governing formation of syllable structure and other phonological phenomena.

We assume that out of all fundamental differences between prototypical 'syllable' languages and prototypical 'word' languages, as enumerated in Table 1, only one is parametrically encoded. This difference, listed in all descriptions of the 'profiling' approach to prosodic typology, is the presence of reduced vowels. Syllable languages do not have vowel reduction and reduced nuclei, whereas word languages display phonological and/or phonetic vowel reduction in unstressed positions.

Note that in Cyran's CSL approach reduced nuclei are a distinct type of licensor. Languages may use schwa (or another reduced vowel) or not – it is a matter of parametric choice. We assume this parametric choice to lie at the heart of the basic prosodic division among the world's languages. Syllable languages use mainly full nuclei as licensors in both stressed and unstressed

---

[4] Of course, structures with a syllabic consonant, like *bottle*, are disregarded in this context. They are assumed to host a filled nucleus, and thus they are not real clusters.

syllables. Word languages typically use full vowels in stressed positions and reduced vowels in unstressed positions. Thus, the surface distinction may boil down to the underlying two-way vs. three-way distribution of licensors. This is the contribution of UG to the surface pattern.

What about the other characteristic features of each type? Why do they pattern together so often? We will attempt to answer these questions in the following subsections.

4.1. Syllable complexity

It is indeed the case that prototypical 'syllable' languages tend not to display complex consonant structures. It may come as a surprise from the point of view of CSL, since these languages abound in full vowels, which are the best licensors. Of course, a language may pick any value out of the three provided by UG, and no choice is inherently 'better' or 'worse' than any other. The choice of the parametric value for full vowels should also be independent of the presence or absence of other licensors in a language (as long as the golden rule of CSL is preserved and other licensors cannot license more than full vowels do). Nevertheless, the empirical record appears to prove otherwise – there is an implicational tendency for languages with only full vowels to display simple syllabic structures. Interestingly, another logically possible type of language, in which all three licensors are present, but they all pick the value of I (i.e. have only simple consonants), also seems to be rare (or non-existent?). The mainstream for languages is to be positioned somewhere between the top left and bottom right corner of the diagram of syllabic configurations introduced in Figure 4. It is less usual, however, to pick one of the values in the left bottom or right top corner, without extending right or down respectively. The following figure illustrates the tendency:

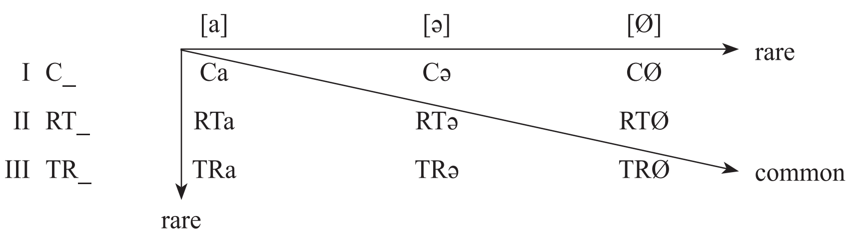

Figure 5. The statistical tendency in the distribution of parametric settings on syllable structure

A potential explanation of this tendency needs to transcend beyond the discussion of parameters. It is widely agreed upon that parameters do not have 'preferences'. Each setting is as good as any other. If there is a statistical bias towards one setting, it needs to be attributed to (intra- or extragrammatical) factors independent of the parameter inventory itself. In our attempt to solve this problem, we will employ a concept crucial to many (but not all) approaches to phonology: contrast.[5]

We assume that a licensor and its licensing domain (i.e. the preceding consonant or cluster) is a relevant contrastive unit of language. Every language needs to be able to generate as many contrastive units as it needs to cover the necessary lexical contrasts. This is true both at the segmental and the suprasegmental level. It is clear that various licensors will be successful to a varying degree at the ability to generate contrastive units. This is due to two reasons. First, the differing licensing potential. Second, the contrastive value of the licensor itself. There is always more than one full vowel, and each of them contrasts with the remaining ones. The inventory of reduced vowels is usually much smaller (often limited to only one vowel).

In canonical 'syllable languages', which primarily make use of full nuclei as licensors in both stressed and unstressed positions, many contrastive CV units can be generated. When concatenated, they are able to cover the necessary contrasts without the need to employ more complex syllabic structures. All syllables hold the same degree of responsibility for the duty of expressing lexical contrasts. To fulfil it, the system often does not need a more marked value for the parameter of syllabic complexity than just 'I'. Even if it does, the inventory of clusters is normally relatively small. The permissible coda-onset clusters are often limited to the best syllable contacts, i.e. sonorant-obstruent, and the permissible branching onsets are often just plosive-approximant sequences. There is no need for any more sophisticated clusters.

In 'word languages', the responsibility for contrast is unevenly distributed between stressed and unstressed syllables. The nuclei of unstressed syllables usually host reduced vowels, so unstressed syllables are much less useful, since the contrastive potential of the whole licensing domain projected by an unstressed nucleus is diminished. The lexical contrastive burden needs to be taken over by the stressed nucleus of the word, which hosts a full vowel. Consequently, in order to be able to cover all the necessary contrasts, it often has to license more complex consonant structures, including more marked clusters. Hence the concentration of consonants around the accented peak.

4.1.1. Illustration on a hypothetical language

To illustrate the general idea, let us analyse a hypothetical, cooked-up language, which has the following inventory at its disposal: 5 obstruents, 5 sonorants, 5 vowels. We will then run this language through all logically

[5] Contrast, which was the fundamental concept in structuralist linguistics, conspicuously fell out of grace with the rise of generative phonology. However, within modern generative linguistics there are some approaches to phonology centered upon the notion of contrast, most notably the Toronto school: Dresher (2009), Hall (2007).

available parametric settings for syllabic complexity. At the settings which involve reduced vowels, the language will be assumed to also have /ə/ in its inventory. For the sake of simplicity and transparency, we will eliminate as many other potential points of variation as possible. Let us universally assume that an RT cluster can be made of any sonorant followed by any obstruent, whereas a TR cluster is formed by any obstruent followed by any sonorant. All words consist of two licensing domains, with stress always falling on the first one. We also take for granted in our simulation that a reduced or empty nucleus can never occur in a stressed syllable, and that in languages with reduced vowels, a full vowel cannot occur in an unstressed syllable. These relations mimic the situation which frequently takes place in natural languages: if a language differentiates full and reduced vowels, it often reserves the former for stressed syllables and the latter for unstressed ones. Finally, we disregard the possibility of a language limiting the left-edge phonotactics to just TR clusters (which means that if full vowels in a language license level II, we include #RTV sequences in our computations).

The number of logically possible contrastive licensing domains in such a system is presented in Table 2.

Table 2. Possible contrastive domains in a hypothetical system

Structure / Licensor	**_a**	**_ə**	**_Ø**
C_	10 x 5 = 50	10 x 1 = 10	10 x 1 = 10
RT_	5 x 5 x 5 = 125	5 x 5 x 1 = 25	5 x 5 x 1 = 25
TR_	5 x 5 x 5 = 125	5 x 5 x 1 = 25	5 x 5 x 1 = 25

In the first row, we multiply the number of possible consonants (10) by the number of possible licensors (5 full vowels, 1 reduced vowel and 1 empty nucleus). In the second and third row we multiply the number of available sonorants and obstruents by the number of available nuclei. We can clearly see that full nuclei are able to generate many more contrastive units than any other type of licensor.

Let us now run our hypothetical language through all logically possible parametric settings and try to find out how many two-domain-long words can be possibly generated at each setting.

1)		[a]	[ə]	[Ø]
I	C_	✓		
II	RT_			
III	TR_			

2,500 combinations

2)		[a]	[ə]	[Ø]
I	C_	✓		
II	RT_	✓		
III	TR_			

30,625 combinations

3)		[a]	[ə]	[Ø]
I	C_	✓		
II	RT_	✓		
III	TR_	✓		

90,000 combinations

The configurations (1-3) are peculiar since languages with these settings only make use of full vowels as licensors. A natural consequence is that in these three cases both licensing domains in our hypothetical language would need to contain full vowels. This fact leads to significant proliferation of possible two-domain-long words. For instance, in a language with setting (3) there are 300 possible contrastive units (50 with a single consonant, 125 LIO domains and another 125 RIO domains). After concatenation we receive the astounding 90,000 possible words.

In all of the remaining cases only the first contrastive unit will host a full vowel in our calculations, whereas the second one will be headed by a reduced or empty vowel.

Parametric configurations (4-5) are in a sense the opposite of the settings in (2-3). Such languages would choose to expand their inventory of syllabic structures by accepting new licensor types in the system, keeping at the same time the licensing abilities of all licensors at the lowest level.

4)		[a]	[ə]	[Ø]
I	C_	✓	✓	
II	RT_			
III	TR_			

500 combinations

5)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_			
III	TR_			

1,000 combinations

We clearly see that such systems have a much sparser generative potential. In system (4) the number of contrastive units is 500 (50 units in the first domain multiplied by 10 in the second one), rising to 1,000 (50 x 20) in system (5).

6)		[a]	[ə]	[Ø]
I	C_	✓	✓	
II	RT_	✓		
III	TR_			

1,750 combinations

7)		[a]	[ə]	[Ø]
I	C_	✓	✓	
II	RT_	✓		
III	TR_	✓		

3,000 combinations

8)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓		
III	TR_			

3,500 combinations

9)		[a]	[ə]	[Ø]
I	C_	✓	✓	
II	RT_	✓	✓	
III	TR_			

6,125 combinations

10)		[a]	[ə]	[Ø]
I	C_	✓	✓	
II	RT_	✓	✓	
III	TR_	✓		

10,500 combinations

11)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	
III	TR_			

7,875 combinations

12)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓		
III	TR_	✓		

6,000 combinations

13)		[a]	[ə]	[Ø]
I	C_	✓	✓	
II	RT_	✓	✓	
III	TR_	✓	✓	

18,000 combinations

14)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	✓
III	TR_			

12,250 combinations

15)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	
III	TR_	✓		

13,500 combinations

16)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	
III	TR_	✓	✓	

21,000 combinations

17)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	✓
III	TR_	✓		

21,000 combinations

18)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	✓
III	TR_	✓	✓	

28,500 combinations

19)		[a]	[ə]	[Ø]
I	C_	✓	✓	✓
II	RT_	✓	✓	✓
III	TR_	✓	✓	✓

36,000 combinations

Systems (6-19) follow a different logic. We increment the inventory of permitted structures by progressing along both axes at the same time. Adding each tick raises the generative potential of the language, retaining at the same time some balance between the available inventories of licensors and levels of complexity.[6]

Languages may be assumed to aim at effective utilisation of segments and their combinations to express lexical contrasts. The parametric settings which are effective in this regard form a skew line from the left top corner to the right bottom corner. At the same time, languages which pick either the value in the left bottom corner or the one in the right top corner are assumed to be ineffective and therefore dispreferred. The former would generate many more contrastive units than necessary, the latter could generate too few. As we can see,

[6] An inquisitive reader could notice that we disregarded systems which omit the middle step in the inventory of licensors and make use of only full vowels and empty nuclei (Polish is a prime example of such a system). Including separate tables for such systems is needless, since the number of combinations will be equivalent with the systems using only full and reduced vowels, i.e. (4, 6, 7, 9, 10, 13).

incorporation of the concept of contrast in the explanation predicts a preference for precisely the kind of distribution which is presented in Figure 5.

Of course, this demonstration could be viewed by some as too simplistic. But please note that it follows the mode of argumentation used in generative linguistics for decades – in order to discover the fundamental properties of human language faculty, we are obliged to disregard all secondary, distracting, and irrelevant factors. Only then can we single out basic mechanisms. Nobody denies that other factors are also at play in prosodic typology. For instance, the number of vowels and consonants in a language heavily bears on the number of possible contrastive units and for some languages it can mask the tendency presented above. However, the drive to effectively use the segmental resources for generating contrastive units can still be a property of human language. It may just not manifest itself in some cases.

From this perspective, the syllable vs. word language dichotomy can be also viewed as a dichotomy on the mode of unveiling contrastive lexical information. In syllable languages the information is unveiled in a stable and continuous way, i.e. every subpart of the phonetic string contributes a similar amount of information. In word languages the lexical information is concentrated around stressed nuclei, which are interspersed with unstressed nuclei. Thus, the process of unveiling lexical information takes place in a wave form.

4.2. Other features

Let us briefly address the remaining dissimilarities between syllable and word languages, as listed in Table 1. It is assumed that all of them are explicable without explicit reference to the conventional prosodic notions and that the groupings of features depicted there can be derived from the theoretical apparatus of CVCV and CSL.

4.2.1. Vowel/consonant insertion vs. deletion

A tendency of syllable languages is to epenthesise vowels or delete consonants in order to get rid of consonant clusters and enforce the CV syllable structure (or a structure sufficiently close to it). This phenomenon can be easily accounted for by the relatively poor licensing potential of nuclei, resulting from a given parameter setting. If a nucleus comes across a consonantal sequence which it can parse as neither a coda-onset cluster nor a branching onset, the system simplifies the sequence in one of two available ways: it either vocalizes the V-slot available in the cluster, or it eliminates one of the consonants. The desired effect is the same, with a consonant-vowel sequence in the output.

Canonical word languages have an opposite tendency. They feature vowel deletion and consonant epenthesis. Diachronically, vowel deletion may be seen as a natural continuation of the process of vowel reduction. In the process of transition from a syllable language to a word language the duration of unstressed vowels gets increasingly smaller, until complete disappearance in some cases. Nonetheless, the proliferation of reduced and empty nuclei in a language is accompanied by boosting their licensing power. Equipped with many good licensors, a typical word language may lose some of its vowels without any consequences, as long as the resulting clusters are within the licensing scope of the nuclei.[7]

Consonant epenthesis can be interpreted as a reflex of the strength of nuclei too. Consider Middle High German *saf* 'juice', which acquired a /t/ at the right edge in Early Modern High German (Szczepaniak 2007: 251). The empty nucleus following /f/ ranked up on the licensing scale and became capable of licensing consonant clusters with even more ease than before. In some cases this excessive licensing potential remains unaccommodated, in some other cases paragoge takes place and a LIO domain is built.

4.2.2. Resyllabification across morphosyntactic boundaries

Resyllabification across boundaries is typologically much more common in syllable languages than in word languages. Usually it takes place in sequences like VC#V, in which the last consonant of the first word is incorporated into the second word as its onset. If it occurs, it is seen as syllable structure optimisation and enforcement of the CV structure regardless of the intervening boundary. If it does not occur, its lack is interpreted as a symptom of a language protecting word edges and profiling thus the phonological word.

The CVCV/CSL perspective on this problem can derive the phenomenon from the properties of empty nuclei, present at the end of every consonant-final word. Final empty nuclei in syllable languages, if they exist at all, have very poor licensing skills, and in order to optimise the licensing relationships, the system aims at maximising the role of full nuclei. Hence, if there is a possibility to employ a full nucleus in the licensor role, the system enforces spreading from the uncovered V-position which was made available postlexically.

In contrast, word languages feature much stronger and more independent empty nuclei at the right edge. They are so strong that they may not need any external help. Even if resyllabification is possible, it does not take place, since there is no motivation for such a process at this particular parameter setting.

[7] From the synchronic point of view, there is one more relevant aspect of vowel deletion, namely vowel-zero alternations. A vowel is sometimes syncopated only when another vowel follows in the next syllable, and it surfaces when the next V-slot is empty, e.g. German *samm<u>e</u>l<u>Ø</u>n* [ˈzaməln] 'collect, inf.' vs. *samm<u>Ø</u>l<u>e</u>* [ˈzamlə] 'collect, 1p.sg.'. For more information on how vowel-zero alternations are analysed in Government Phonology and CVCV, see Kaye et al. (1990), Rowicka (1999), Scheer (2004) (among others).

4.2.3. Vowel harmony, tone, and long segments in the unstressed position

The frequent absence of such phenomena as vowel harmony, tone, or length contrast in the unstressed position also receives a direct explanation within the CVCV framework. The crux of the problem lies in the frequent restrictions on the content of unstressed nuclei in these languages and can be easily accounted for with reference to autosegmental licensing (e.g. Goldsmith 1990).

It can be stated that an unstressed nucleus can host only a limited subset of distinctive features. One of the ramifications of this fact is the aforementioned limited contrastive potential entertained by such nuclei. It is not the only ramification, though. If an unstressed nucleus is unable to sustain a contrast between, for example, /e/ and /o/, then it naturally follows that it is not capable of accepting features spreading from other nuclei or of hosting tone features. In this way, 'word' languages, amply using reduced vowels, are at the same time unlikely to feature vowel harmony or to use tone (except in stressed syllables).

5. Consequences

The proposed approach attempts to account for the attested prosodic typology of languages by combining strictly formal concepts (licensing scales as the UG basis) with the extragrammatical concept of the necessity of expressing lexical contrast. This perspective has a huge explanatory potential and sheds new light on both the theory of 'prosodic profiling' and on licensing scales.

Its main contribution to the former is the reduction of the syllable/word language distinction to much more basic devices. What lies at the heart of the phenomenon is the fundamental mechanisms of cluster licensing and autosegmental licensing, from which the remaining phenomena can be derived. It is not the case that the linguistic system selects one of the extant prosodic categories and tries to 'optimise' it. Or, at least, not at the I-language level. The collection of features associated with each type, as presented in Table 1, is principally an E-language-based generalisation. However, the reason for their clustering with each other is a little unclear. It does not follow from anything. With the CVCV/CSL approach, we are much closer to providing a unified explanation, since all generalisations are expressible in terms of nuclei and their parametrically determined abilities and limitations.

Furthermore, relegation of most of the type-specific phenomena beyond the binary parameter explains the continuum nature of the distinction. Even though a language either uses the reduced nucleus as a licensor or not (there is nothing in between), languages may differ to a large extent with the degree to which they implement other features characteristic of each language type. It becomes visible especially from a diachronic perspective. Szczepaniak (2007) meticulously documents a drift of German from a canonical syllable language to a canonical word language. (Early) Old High German had virtually all syllable language features, such as simple syllable structure, vowel harmony (in the form of umlaut), and vowel quantity contrast in unstressed positions (at least in some dialects). It also mostly lacked reduced vowels.[8] In the course of transition to Middle High German, one of the most crucial changes was full-blown vowel reduction, which struck all unstressed syllables. From our point of view, that was the time at which a change at the I-language level took place. German introduced a new type of licensor: schwa. However, note that other syllable language features did not disappear overnight. German needed several centuries to reach the state in which it is now. Before, it belonged to a mixed type for a long time. Such mixed types are cross-linguistically quite common, and languages often drift in one or the other direction. Importantly, their existence does not disprove any of our claims.

And what can our approach contribute to the Complexity Scales and Licensing theory? It adds an extragrammatical dimension to it, which explains the actually attested tendencies in phonological typology. The idea of the optimisation of the use of contrastive units explains why there are few languages in which, for instance, full vowels in all positions of the word coexist with complex syllabic structures (which would amount to setting III-1 in Table 4 and Table 5). We propose that languages climb up the licensing scales, or markedness scale in general, only when there is such pressure from other parts of the system or even from beyond the system.[9] UG is an important, or even the most fundamental, ingredient of human language, but it does not exist in splendid isolation. Parameter settings can be influenced by other factors too.

6. Final remarks

Note that this proposal does not aim to constitute a full explanation of prosodic typology from the perspective of CSL. All the time we are just talking about *tendencies*. The proposal adds just one factor (contrast) to a very complex problem. Undoubtedly, prosodic typology is a big topic, and most certainly it will never be able to be fully reduced to a binary parameter, or to be fully generalised with just a small set of descriptive categories. Nonetheless, it is hoped that this paper has provided a fresh perspective, which can be a point of departure for further, more detailed, research.

---

[8] As a matter of fact, the interpretation of some graphemes is ambiguous. Some studies of OHG vocalism include /ə/ in the inventory (e.g. in Williram, see Valentin 1978: 379). Nevertheless, it needs to be kept in mind that there is no one-to-one relationship between the ‘reduced’ status and any particular phonetic value. A language can have a mid-central /ə/ phoneme, which does not display any properties of a reduced vowel and has the same distribution as all other vowels. Beyond doubt, Old High German unstressed syllables can host many different vocalic qualities, and there is no reason to assume that the system made use of the ‘reduced’ type of licensor.

[9] Note that this conclusion is (completely unintentionally) compatible with Optimality Theory (Prince and Smolensky 1993), which is inherently based upon the interaction of markedness constraints with other constraints.

References

Auer, P. 1993. Is a rhythm-based typology possible? A study of the role of prosody in phonological typology. *KontRI Working Paper 21*. Universität Konstanz.

Auer, P. 2014. Preface. In J.C. Reina and R. Szczepaniak (eds.), *Syllable and word languages*. Berlin/Boston: Mouton de Gruyter.

Cyran, E. 2003. *Complexity scales and licensing strength in phonology*. Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego.

Cyran, E. 2010. *Complexity scales and licensing in phonology*. Berlin/New York: Mouton de Gruyter.

Dauer, R. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.

Donegan, P.J., and D. Stampe 1983. Rhythm and holistic organization of language structure. In J.F. Richardson, M. Marks and A. Chukerman (eds.), *Papers from the parasession on the interplay of phonology, morphology and syntax, Chicago, 22-23 April 1983*, 337-353. Chicago: Chicago Linguistic Society.

Dresher, E.B. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Fortuna, M. 2015. *Double licensing phonology*. München: Lincom.

Goldsmith, J. 1990. *Autosegmental and metrical phonology*. Oxford: Basil Blackwell, Ltd.

Hall, D.C. 2007. *The role and representation of contrast in phonological theory*. Doctoral Dissertation, University of Toronto.

Kaye, J., and J. Lowenstamm 1981. Syllable structure and markedness theory. In A. Belletti, L. Brandi and L. Rizzi (eds.), *Theory of markedness in generative grammar. Proceedings of the 1979 GLOW Conference*. 287-315. Pisa: Scuola normale superiore.

Kaye, J., J. Lowenstamm and J.R. Vergnaud 1990. Constituent structure and government in phonology. *Phonology* 7: 193-231.

Lowenstamm, J. 1996. CV as the only syllable type. In J. Durand and B. Laks (eds.), *Current trends in phonology. Models and methods*, 419-441. Salford, Manchester: European Studies Research Institute, University of Salford.

Prince, A., and P. Smolensky 1993. *Optimality Theory: Constraint interaction in generative grammar*. Rutgers University Center for Cognitive Science Technical Report 2.

Reina, J.C., and R. Szczepaniak (eds.) 2014. *Syllable and word languages*. Berlin/Boston: Mouton de Gruyter.

Rowicka, G. 1999. *On ghost vowels. A Strict CV approach*. The Hague: Holland Academic Graphics.

Scheer, T. 2004. *A lateral theory of phonology. What is CVCV and why should it be?* Berlin/New York: Mouton de Gruyter.

Scheer, T. 2012. *Direct interface and one-channel translation.* Berlin/New York: Mouton de Gruyter.

Ségéral, P., and T. Scheer 2001. La Coda-Miroir. *Bulletin de la Société de Linguistique de Paris* 96, 107-152.

Szczepaniak, R. 2007. *Der phonologisch-typologische Wandel des Deutschen von einer Silbensprache zu einer Wortsprache.* Berlin/New York: Mouton de Gruyter.

Szigetvári, P. 1999. *VC phonology: a theory of consonant lenition and phonotactics.* Doctoral dissertation, Eötvös Loránd University.

Valentin, P. 1978. The simplification of the unstressed vowel systems in Old High German. In J. Fisiak (ed.), *Recent developments in historical phonology*, 373-389, Den Haag: Mouton de Gruyter.