# Stress Patterns of Compounds and MWEs in English and Bulgarian

**Bistra Popovska**
Department for English Studies
Faculty of Philology
The Paisii Hilendarski University of Plovdiv
`bistra63@abv.bg`

**Rositsa Dekova**
Department for English Studies
Faculty of Philology
The Paisii Hilendarski University of Plovdiv
`rosdek@gmail.com`

## Abstract

The paper presents an ongoing research on the stress patterns of compounds and MWEs of the type ADJ+N and their corresponding free NPs in English and Bulgarian. The research focuses on the identification and the formal representation of the possible stress patterns of compounds and MWEs and free NPs. During our research so far, we have compiled a corpus of over 2000 compounds and MWEs, approx. 1000 for each language – English and Bulgarian. Our theoretical framework includes elements from different theories, i.e. the Generative Phonology Theory, the Metrical Theory, and the Theory of Primary accent first which all define the stress as a prosodic element. Our main goals are to specify the prosodic region where the stress is defined in English and Bulgarian MWEs and noun phrases and to define the main features of the stress in MWEs and free NPs in English and Bulgarian. The results of our research can serve for implementation into NLP modules for spoken language processing and generation.

## 1. Introduction

The opportunities of modern language technologies resulting in the accumulation of huge data bases and large-scale theoretical generalizations relate to a rapid development of phonological and cognitive-linguistic prosodic studies and particularly those focused on stress (see Patseva, 2011, among others). One phenomenon which has received little attention in previous linguistic work on speech prosody is the use of contrastive stress patterns to distinguish meanings at the suprasegmental levels. The types of stress contrast can be exemplified by minimal pairs, such as the compound *gréenhouse* vs. the phrase *green hóuse* in English and червеношийка (*tchervenoshéeka* 'robin') and червена шийка (*tchervéna sheeka* 'red neck') in Bulgarian. Previous studies using behavioural methods (e.g. Cutler and Otake 1999, Cutler and Van Donselaar 2001) and electrophysiology (e.g. Friedrich et al. 2001) suggest that listeners use lexical stress information during spoken word identification. However, the distinction between compound and phrasal stress and the role it plays in online comprehension remain relatively unexplored, and represent the focus of the present study.

Our corpus includes more than 1000 MWEs for each of the languages English and Bulgarian. The English compounds and phrases have been compiled by hand, while the Bulgarian MWEs were extracted from *The Bulgarian Dictionary of MWEs* (Koeva et al., 2016; Stoyanova and Todorova, 2014).

## 2. Stress as prosodic element

In linguistic terms speech prosody studies stress and intonation. In the present research we focus on the stress patterns of compounds and MWEs of the type ADJ+N and their corresponding free NPs in English and Bulgarian.

Firstly, we have examined the main prosodic concepts relating to stress, starting with the prosodic hierarchy, phonological aspect of the syllable and the main phonetic and phonological phenomena of rhythm as discussed by Dimitrova (1998, 1999), Giegerich (2005, 2011), Halle and Vergnaud (1987), Prince and Smolensky (2002), Savitska and Boyadzhiev (1988), Tilkov (1982), among others.

Secondly, we have outlined the prosodic area in which stress is defined using the terminology apparatus of the contemporary approaches to prosody. We have studied the phonetic and phonological aspects of stress, following the logic of scientific research, which has led us to the paradigm of the linguistic tendencies of the last decades. For the purposes of our research, we have considered different theoretical approaches, i.e. the Generative Phonology Theory, the Metrical Theory, and the Theory of Primary accent first.

## 3.    Characteristics, functions and position of stress in English and Bulgarian

In our study, we follow traditional ways of defining and describing the characteristics and functions of Bulgarian stress – phonetic, positional and phonological (word stress and phrasal stress) as discussed by Kurlova (1997), Misheva (1991), and Tilkov and Boyadzhiev (1978), among others. However, we also introduce an analysis from the point of view of the Primary Accent First theory of Van der Hulst (2002, 2009).

A far as the stress in English is concerned, we have described it according to the four variable indicators: intensity, pitch, vowel quality and vowel duration, following Collins and Mees (2003).

The study investigates firstly word stress: the degrees of stress, lexically designated stress in English and then switch stress. We have formulated some word stress guidelines concerning words consisting of two or three syllables (in most of the cases stress falls on the first syllable, e.g. *súmmer*) , longer words having four or more syllables (in most cases stress falls on the antepenultimate syllable, e.g. *solubílity*), prefix words (in shorter words the main stress generally falls on the syllable after the prefix, e.g. *repláy*), word endings (on ending itself, e.g. *himsélf,* on syllable preceding ending, e.g. *defícient*).

Secondly, we have investigated stress in compounds and MWEs (Initial Element Stress, e.g. *Rússian class*. Final Element stress, e.g *Russian roulétte*). Collins and Mees (2003) have formulated some stress guidelines for compounds, defining word shape: The Manufactures Rule, according to which if the compound contains the material from which the item is manufactured, the stress falls on the final element, e.g. *apricot brándy*; and the Location Rule which dictates that if the compound contains location, the stress usually falls on the final element, e.g. *London Eye* (Collins and Mees 2003). They have formulated some further stress pattern guides related to the above (for example food labels generally have stress on the final element, e.g. *scrambled eggs*).

Finally, we have also investigated the phonetic and phonological characteristics of English stress according to Giegerich (2005) from the point of view of the Metrical phonology.

## 4.    Stress patterns of compounds, MWEs and Phrases in English and Bulgarian

### 4.1.    The importance of the contrast

Contrastive word stress plays an important role in the differentiation between compounds and phrases. The current research investigates the development of compound and phrasal stress in both production and perception, an area considerably neglected by previous studies on linguistic stress. By comparing directly production and perception we aim to prove that compounds are generally not mistaken for phrases while phrases are often mistaken for compounds.

According to some authors (Chomsky and Halle 1968:15) compound and phrasal stress can be assigned in English by the Compound Stress Rule according to which stress is placed on the first segment of the compound. In contrast, the Nuclear Stress Rule dictates that phrasal stress is assigned to the rightmost phrase segment. The abovementioned difference in the placement of stress generally allows listeners to discriminate between compounds and phrases with identical constituents.

According to others (e.g. Bloomfield 1933:228, Giegerich 2006), however, the stress criterion in English, commonly invoked in attempts to draw the compound-phrase distinction, is getting less reliable: it fails to correlate with other (semantic, syntactic) criteria; it draws on incomplete and flawed generalizations regarding stress in compounds and phrases. A fictitious category distinction arises for pairs of semantically very similar constructions such as *Christmas púdding* and *Chrístmas cake. Ice cream* has a variable stress pattern – for some speakers it is a compound and for others it is a phrase. Then that distinction needs to be revisited.

### 4.2.  Stress patterns of compounds and MWEs in English and Bulgarian

The current study describes the types of noun-centred (e.g. *wíndmill*) and verb-centred (e.g. *táxi-driver*) compound nouns in English. It aims to define the labels for stress in compound nouns and collocations. In addition, we investigate the single stress pattern and the double stress pattern in compounds and collocations together with the stress pattern in three-word compound nouns.

We define the types of compound nouns in Bulgarian with or without a linking element (e.g. *злоб**о**днéвка* (*zlobodnévka* 'burning topic') and *свръхпроизвóдство* (*svruhproizvódstvo* 'overproduction'), respectively); with or without suffixation (e.g. *първоклас**ник*** (*purvoklásnik* 'first-grader') vs. *буквояд* (*bukvoyád* 'verbalist')). As far as the stress patterns of compound nouns are concerned, we point out the reasons for their formation: extralinguistic and linguistic (semantic, syntactic and morphological).

### 4.3.  Stress patterns of free phrases in English and Bulgarian

The research investigates the types of free phrases in English and Bulgarian together with their stress patterns. The terms *theme* and *rheme* are introduced together with the way they relate to stress. We aim to describe the factors, determining the relation between *rheme* and stress. For example, the reason can be syntactic, as the noun modifier can bear stronger stress, e.g *бял кон* (*byál kon* 'white horse').

### 5.  Conclusions and future work

Following our theoretical approach, we have planned some experimental research focusing on compound/phrasal distinction. The experiment is designed to satisfy two experimental task types: production and perception, using as stimuli minimal pairs of segmentally identical but prosodically distinct phrases and compounds such as *bláckboard* and *black bóard*. A statistical analysis of the results of the experiments will provide empirical evidence to support our theoretical model.

Finally, we believe that a thorough investigation and a proper formal representation of stress patterns of compounds, and especially MWEs as opposed to free phrases will contribute greatly to the tasks of spoken language processing and generation.

## References

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Chomsky, N. and Halle, M. (1968). *The Sound pattern of English*, NY: Harper & Row.

Collins, B. and Mees M.I. (2003). *Practical Phonetics and Phonology, A Resource Book for Students*. Routledge, pages 109–116.

Cutler, A. and Van Donselaar, W. (2001). Voornaam is Not a Homophone: Lexical Prosody and Lexical Access in Dutch. *Language and Speech*, 44:171–195.

Cutler, A. and Otake, T. (1999). Pitch Accent in Spoken-word Recognition in Japanese. *Journal of the Acoustical Society of America*, 105:1877–1888.

Dimitrova, S. (1998). Bulgarian Speech Rhythm: Stress-Timed or Syllable-Timed, *Journal of International Phonetic Association*, 27 (1-2):27–33.

Dimitrova, S. (1999). *Ritmichna organizatsiya na angliyskata i balgarskata rech: sapostavitelno fonetichno izsledvane*. PhD Dissertation Abstract, Sofia.

Friedrich, C.K., Alter K. and Kotz S.A. (2001). An Electrophysiological Response to Different Pitch Contours in Words. *Neuroreport*. Oct 29; 12(15):3189–91.

Giegerich, H. (2005). *Metric Phonology,* Cambridge Studies in Linguistics, Cambridge University Press.

Giegerich, H. (2006). Attribution in English and the Distinction between Phrases and Compounds, In: Petr Rosel (ed.) *Englisch in Zeit und Raum – English in Time and Space: Frschungsbericht fur Klau Faiss*. Trier: Wissenschaftlicher Verlag Trier.

Giegerich, H. (2011). *Compounding and Lexicalism, Handbook of Compounding.* Oxford: Oxford University Press.

Halle, M. and Vergnaud J.-R. (1987). An Essay on Stress. *Current Studies in Linguistics 15*). Cambridge, Mass.: MIT Press.

Hulst, H.G. van der (2002). Stress and Accent. In: Nadel, L. (ed.). *Encyclopedia of Cognitive Science.* vol. 4, 4. London: Nature Publishing Group, pages 246–254.

Hulst, H.G. van der (2009). *Brackets and Grid Marks or Theories of Primary Accent and Rhythm.* In: E. Raimy and C. Cairns (eds.). *Contemporary views on architecture and representations in phonological theory.* MIT press.

Koeva, S., Stoyanova, I., Leseva, S. and Todorova, M. Semi-automatic Compilation of the Dictionary of Bulgarian Multiword Expressions. In *Proceedings of the GLOBALEX Workshop*, LREC, 2016, ELRA.

Kurlova, R. (1997). *Fonetika i fonologia*, Blagoevgrad.

Misheva, A. (1991). *Intonatsionna sistema na balgarskiya ezik*. Sofia, BAS.

Patseva, M. (2011). Dinamika na konotatitvnoto znachenie. *Littera et Lingua.*

Patseva, M. (2007). Aktsentni modei pri nyakoi sashtestvitelni imena ot mazhki I zhenski rod. *Godishnik na Fakulteta po slavyanski filologii na SU*, vol. 96–98.

Prince, A. and P. Smolensky 2002 (1993). *Optimality Theory. Constraint Interaction in Generative Grammar*, ROA 537 0802, Rutgers

Savitska, I. and Boyadzhiev, T. (1988). *Bulgaro-polska supostavitelna gramatika*, vol. 1. Bulgarian Academy of Sciences.

Tilkov, D. (1982). Srichka. In *Gramatika na balgarskiya knizhoven ezik. T. 1. Fonetika.* Sofia, BAS, pages 154–158

Tilkov, D. and Boyadzhiev, T. (1978). *Udarenieto v balgarskiya knizhoven ezik.* Sofia, Narodna prosveta Publishing House.