LEVELS OF STRESS IN ENGLISH: THEIR PHONETIC REALITY AND THEORETICAL STATUS

Piotr Gąsiorowski[1], Michał Jankowski[1], Krzysztof Sawala[1]

[1]Adam Mickiewicz University, Poznań (Poland)
gpiotr@wa.amu.edu.pl

ABSTRACT

We argue that in order to present a physically and psychologically realistic description of lexical and phrasal stress in English, it is necessary (and sufficient) to distinguish three levels of syllable prominence, while a fourth, unstressed type of syllable may only contain a "weak" vowel. We test our hypothesis empirically by analysing a rich selection of units (polysyllabic words, compounds, and fixed phrases displaying complex prosodic patterns) extracted from a large corpus of recorded samples of British and American English. We aim to to develop a parsimonious model of English stress consistent with our observations. We also propose a principled system of stress marking for multifarious applications, including pronouncing dictionaries.

Keywords: stress, levels of prominence, English, prosody, phonetics.

1. INTRODUCTION

1.1. Traditional approaches: stress as a multi-valued feature

Early approaches to stress disregarded its complex prosodic nature and reduced it to a purely physical attribute (respiratory activity producing varying degrees of loudness). Stress was typically equalled with "the comparative force with which the separate syllables of a sound-group are pronounced" [14] or with "intensity or loudness" consisting in "greater amplitude of sound-waves" and resulting from more "energetic" respiratory and articulatory activity [2]. Trager and Smith [15] separated stress from intonation, claiming that accent equals loudness while intonation equals pitch and that in a phonological analysis, an utterance may be stripped of pitch phonemes and junctures, leaving all the stresses intact. Like Chomsky and Halle (*SPE*, [3]) almost two decades later, they suggested that multiple levels of stress must be distinguished for an accurate analysis. In the *SPE* model, stress was a multi-valued feature assigned to vowels in abstract representations by phonological rules; the physical realisation of stress levels was of secondary importance; a speaker who had internalised the rules in question would recover an abstract stress pattern even on the basis of imperfect phonetic cues. Such approaches met with criticism from those reluctant to posit fine distinctions along a dimension that could be studied "using no analytical technique other than the linguist's listening to himself saying various test utterances" [1].

1.2. Binary representations of stress

For practical purposes, a "binary" representation of stress often seems sufficient. It is favoured by many phoneticians [10] and by some of the major pronouncing dictionaries [8, 18]. Only one or two degrees of stress are treated as relevant. In transcribing them, marks for primary (tonic) stress and secondary stress are used, and the following assumptions are made:

- a syllable is either stressed or unstressed,
- lexical items may be multiple-stressed, in which case one and only one of the stressed syllables receives a primary stress (ˈ), while other stressed syllables (if any) receive secondary stresses (ˌ),
- primary stress is an accentual overlay upon secondary stress,
- syllables that are not reduced and contain full vowels may nevertheless remain unstressed.

Thus, for example, according to the conventions used in [18] we have stress patterns like the following:

(1) ˌanaesˌthesiˈology
(2) conˌciliˈation
(3) ˌreconciliˈation

Note the inconsistencies: the secondary stress of *conciliation* (itself reflecting the primary stress of *conciliate*) is no longer marked when the prefix *re-* is added, though (1) shows that there may be two or even more secondary stressed preceding the primary one. The /ɪ/ in the fourth syllable of *reconciliation* exhibits no evidence of phonetic reduction. It is hard to see why a transcriber should treat (1) and (3) differently. Also the current IPA chart [IPA] provides only two stress marks, as described above.

1.3. Metrical phonology: stress as a strength relation between prosodic units

Early analyses of stress were based on observations of acoustic phenomena – usually non-instrumental and impressionistic. In the mid-1970s, studies of

stress relied more on formal models than phonetic observables. Metrical phonology developed the formalism of metrical trees [11, 6, 7], abandoning the "linear" analysis of stress as a multi-valued feature assigned to prominence-bearing segments (mostly vowels). Stress was redefined as a contrast between "strong" and "weak" prosodic units in a hierarchy of inclusion (syllable ⊂ metrical foot ⊂ phonological word ⊂ intonational phrase). The grid notation [11, 13, 5] represents levels of prominence as different heights of grid columns, but this is an accidental feature of the theoretical model; grid height has no inherent phonetic realisation. The fully mature versions of the metrical theory of stress, as in [7, 16], have produced representations employed in lexical phonology, optimality theory, and other areas of phonological research. Metrical phonology has developed into a sophisticated theory which captures important generalizations about the diverse stress systems of the world's languages, makes it possible to build a complete typology of such systems, and characterizes them in terms of a parametric model.

Nevertheless, some aspects of the notion of stress remain controversial. In particular, the units routinely employed in metrical phonology are often defined with little regard to direct empirical evidence. Different phonologists have mutually incompatible views on fundamental issues like syllable structure and foot boundaries, and even on the reality of some of the units involved. Predictions concerning the phonetic realisation of prosodic prominence (i.e. the acoustic correlates of stress and of its discrete levels) do not follow naturally from the formal models.

1.4 The aim of the present study

The objectives of our research include the empirical investigation of the acoustic properties of stress with a view to resolving the issue if there is empirical support for the existence of different levels of stress and if their number is constrained. In a way, we want to step back from models overburdened with theoretical apparatus to fundamental questions. What we are seeking is a phonological representation driven primarily by observable and measurable empirical evidence. We hope that the impact of the findings will not be restricted to studies of English stress.

2. CORPUS AND METHODS

We have compiled an electronic corpus of approximately 140,000 lexical items comparable in size and structure to the sets of entries featured in such existing corpora as CELEX, Corpus of Contemporary American English, British National Corpus, Longman Pronouncing Dictionary, Cambridge Pronouncing Dictionary, CMU Pronouncing Dictionary, etc. Every item of the corpus is assigned a phonetic transcription according to the principles worked out by the authors of the project, separately for Standard British English and General American English. Apart from single-word entries, there are additionally about 8,000 complex compounds and phrases. What we are especially interested in are long words and multi-word units that may be expected to show maximally complicated stress patterns.

2.1 Acoustic data and its analysis

Items of interest are searched for with the use of search engines available at the audio archives of Internet radio sites such as the National Public Radio (US) and the BBC (UK). The use of corpora of spontaneously produced live English for an acoustic analysis guarantees that the material is "ecologically valid" in comparison with laboratory recordings. The quality of MP3 files available from the Internet sources in question (with a compression of 128 kbps at the frequency of 44,100 Hz) is more than sufficient to carry out acoustic analysis with the help of Praat. Similar quality files have been used in various acoustic researches of stress-related phenomena. As we are interested in relative changes of pitch, amplitude, duration etc. within individual items in order to detect statistically significant differences between degrees of stress (if any), speaker-to-speakers differences should not have any impact on the credibility of the results.

The items for which recordings are available are subjected to detailed acoustic analysis. Its primary goal is to identify systematically occurring correlates of distinct degrees of non-primary stress in order to establish whether these correlates are sufficiently robust to justify positing multiple levels of stress in naturally spoken English.

2.2 Preliminary findings

Our preliminary results suggest that four distinguishable levels of prominence are both necessary and sufficient for a realistic description of our data. We use the following numerical representation of prominence levels: **1** – primary stress; **2** – secondary stress; **3** – tertiary stress; **4** – no stress. The corresponding stress marks are respectively as follows: [ˈ], [ˈ], [ˌ], none, as in *documentary evidence* [ˈdɒk.jəˌmen.tᵊr.i ˊev.ɨ.dənᵗs] (UK transcription), the stress contour being **24344144** in our notation. It should be noted that level **2** as phonetically distinct from level **3** occurs only pretonically (before the primary stress of a phrase). Levels **1** and **2** are characterised primarily by their tonal prominence and clearly connected with the structure of intonational phrases rather than

individual words (see below). Level **3** is distinguished from level **4** by its strong vocalism (stable formant structure) and greater duration (the latter, however, is neutralised word-finally).

We provisionally use the numerical representation of stress contours, as well as the corresponding transcription, for all the items in our database. This allows us to sort our items according to the type of contour they represent and to collect exhaustive sets of examples illustrating any possible stress pattern.

3. CRITICAL STRESS PATTERNS

We focus on selected phenomena that throw light on the nature of prominence levels. Here we present two of them.

3.1 Iambic reversal

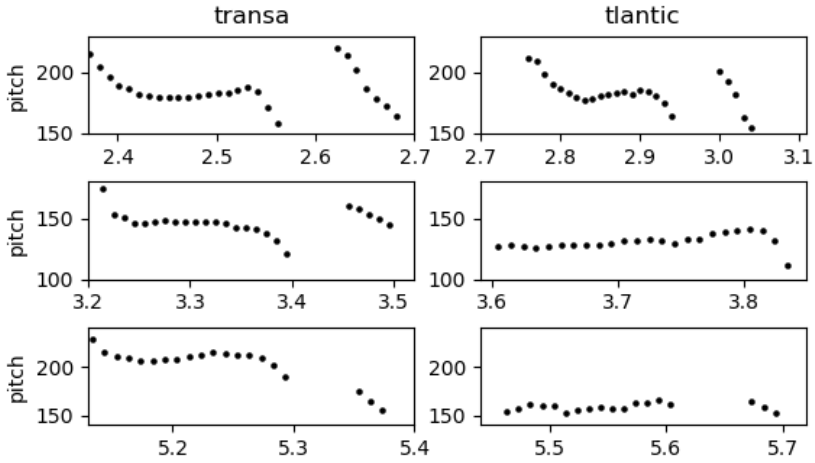
Iambic reversal in English is defined as a leftward shift of the primary stress in a word followed, within the same prosodic unit, by another, more strongly stressed word. The primary stress may only retract to a syllable carrying a secondary stress. In the usual notation, *ˌthirˈteen* becomes *ˈthirˌteen* in *ˈthirˌteen ˈmen* to avoid a "stress clash". Metrical phonology defines stress clash as non-eurhythmic adjacency of "strong" units on the same tier of metrical structure. Iambic reversal is thus modelled as a rule of rhythmic repair reversing the strength values of sister nodes (at foot level) in a metrical tree: (w s) → (s w). Alternatively, the top grid mark of a column is shifted to another column on the left. In either case, there is an actual movement of an originally assigned primary stress during the derivation of a phrasal stress pattern.

There is, however, no observable difference between the prominence contours of *thirteen men* (if pronounced in isolation) and *chimpanzee*. Since in the latter example the reinforcement of the initial syllable merely demarcates the initial boundary of a unit, there is no reason to treat the first syllable of *thirteen men* differently.

In order to reduce the possible interference of vowel-quality contrasts with the acoustic correlates of stress, we focus on polysyllabic words with at least one non-primary stress, in which the stressed syllables contain the same vowel phoneme, e.g. /æ/ in *acrobatic* or *charismatic*. Figure 1 shows the pitch contours for three out of ten occurrences of *transatlantic* in phrases such as *transatlantic flight* in out database.

The first example is exceptional in that both prosodic feet display nearly identical pitch trajectories. We interpret it as a marked, emphatic pronunciation with extra prominence given to all stressed syllables. The other two examples illustrate what we consistently observe in the vast majority of cases: a more dynamic pitch movement and a greater peak–valley distance in the first foot. It is remarkable that the domain over which the pitch movement is realised is the whole foot, not just the stressed syllable, and that what counts is not the height of the pitch but its dynamics across the foot domain (a phenomenon still incompletely understood, see [17]). If the first syllable carries a low rising head tone, the rise peaks on the second, unstressed syllable, and may even drop sharply at the right boundary of the foot.

Figure 1. Pitch trajectories in three realisations of *transatlantic* in the environment of "iambic reversal".



We conclude that no real "reversal" takes place here. In a normal intonational phrase, no primary stress is realised phonetically in words that would display it when pronounced in isolation, except in the word aligned with the nuclear tone of the phrase. All other stresses are in principle equal (level **3**), but since English intonational phrases with multiple stresses involve head tones, such a tone docks onto a stressed syllable near the left edge of the phrase, raising its prominence to level **2**. In most cases it is the first stressed syllable of a word, though pretonic slopes of prominence decreasing over a span of three syllables (**234**) are avoided in English. Therefore, the head tone may be exceptionally realised on the second of two consecutive stressed syllable in such cases, as in *ambiguous remarks* (**3244 41**). In traditional parlance, iambic reversal fails in a phrase like this.

3.2 Posttonic Stress

There is much inconsistency in the marking of posttonic (post-primary) stresses by phoneticians. Note the following examples, taken from [18] (the conventions of [8] are also somewhat inconsistent, but at the same time slightly different from those of [18]):

(4) ˈsteamˌroll
(5) ˈson-in-ˌlaw
(6) ˈcompound
(7) ˈhesitate
(8) ˈbusinessman
(9) ˈspelling pronunciˌation
(10) ˈword count
(11) ˈwordˌorder
(12) ˈalligator
(13) ˈarchˌangel
(14) ˈcucumber

The general rule in is that a final syllable with a strong vowel is not regarded as stressed (6, 7, 8, 10), but transparent compounds (4, 5) may be arbitrarily exempted. A posttonic strong-vowel syllable followed by a weak one is generally marked as stressed in compounds (11) and prefixations (13) but not elsewhere (12, 14). Some cases (9) require special handling.

We were interested in finding any systematic difference between compounds and simplex words in terms of pitch dynamics, duration or amplitude. No difference was found. The final or non-final position of a strong-vowel syllable turned out to be irrelevant in "minimal pairs" like *eyelash* : *eyelashes*. We therefore assign level **3** prominence to any strong-vowel syllables following a primary stress. Thus, *cucumber*, *word order* and *archangel* all represent a **134** stress contour, while *steamroll*, *compound* and *word count* are all **13**.

4. DISCUSSION

Our analyses show that the derivation of stress contours from metrical trees or grids is superfluous. In English, the presence of stress manifests itself as strong vocalism. A lexical primary stress is realised as a characteristic pitch trajectory only when aligned with the nucleus of an intonational phrase. Although its location depends to some extent on the segmental structure of the word (the count of syllables and their weight) and on morpholexical factors (lexical class membership, derivational history), idiosyncratic exceptions and conflicting tendencies make English stress unpredictable (or at least not fully predictable), therefore "phonemic".

The same is true of non-primary stresses. In many cases their location can be predicted, being redundantly assigned by a default rule of iterative foot construction, but (especially in rare words and in many derivational morphemes) strong vowels appear in positions not favoured by rhythmic preferences. Cyclic effects (strong vocalism reflecting a stressed vowel of the derivational base) may also clash with the rhythm, and the conflict may be resolved in different ways, as in *e'lectric* → *eˌlec'tricity* (cyclic) or *ˈe (ˌ)lec'tricity* (rhythmic).

Level **3** prominence is largely encoded in lexical representations as a "strong" vowel whose distinctive features must be fully articulated. One vowel per word is lexically designated as a potential docking site for the nuclear tone of an intonational phrase, manifesting itself as level **1** prominence if the word is pronounced in isolation and therefore constitutes an utterance on its own. Level **2** prominence may result from the attachment of a boundary tone to strong syllables near the left edge of an intonational phrase. When a word is pronounced in isolation, this "secondary" stress is realised phonetically. We emphasise the fact that it does not result from any transformations of an abstract metrical structure. All surface prominence above level **3** in English results from the interplay between full vocalism and intonational factors. "Word stress" is in fact phrasal stress reduced to a restricted domain – an utterance consisting of a single word.

In connection with postulating different levels of prominence in English, we have observed a phenomenon overlooked in other studies taking a similar approach (see the analysis of German prominence and intonation in [9]): the important role of the foot as the domain across which contrastive tonal features are distributed. Weak syllables play a crucial role in the realisation and recognition of salient prominence cues spreading from a preceding stressed syllable. At any rate, we find it to be true of English but do not rule out the possiblity that in other languages the relevant tonal cues may be more localised, i.e. confined to stressed syllables.

5. CONCLUSIONS

In principle, since level **2** prominence is not "phonemic" and does not have to be lexically specified, it could be left undistinguished from level **3**. On the other hand, it differs acoustically from both level **1** and level **3**, and it is not always immediately obvious when it can be assigned to a pretonic strong syllable. Furthermore, there is a difference between levels **2** and **3** with regard to phonological processes such as optional destressing (as in the middle syllable of *BBC*). For these reasons, we recommend that level **2** be marked in pronouncing dictionaries, if only to help foreign learners to develop a native-like prosodic competence. It is often claimed that the failure to "apply iambic reversal" – or, in our view, to place the head tone in its correct position – results in an unnatural pronunciation. Thus, in our database, both a word such as *chimpanzee* [ˈʧɪmˌpænˈzi] and a phrase like *Chinese room* have a **231** stress contour, All other levels of prominence are contrastive and must be duly marked.

Acknowledgements
This research was supported by a project grant from the Polish National Science Centre (NCN), OPUS11 2016/21/B/HS2/00616.

6. REFERENCES

[1] Beckman, M. E. 1986. *Stress and non-stress accent*. Dordrecht: Foris Publications.

[2] Bloomfield, L.1933. *Language*. New York: Holt.

[3] Chomsky, N., Halle, M. 1968. *The sound pattern of English*. New York, Evanston, London: Harper and Row.

[4] Full IPA chart https://www.internationalphoneticassociation.org/content/full-ipa-chart.

[5] Halle, M., Vergnaud, J.-R.1987. *An essay on stress*. Cambridge, MA: MIT Press.

[6] Hayes, B. 1981. *A metrical theory of stress rules*. Bloomington, IN: Indiana University Linguistics Club.

[7] Hayes, B. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

[8] Jones, D. (ed.). 2011. *Cambridge English pronouncing dictionary* (18th edition, revised by Roach, P., Setter, J., Esling, J.). Cambridge: CUP.

[9] Kügler, Frank et al. 2015. DIMA – Annotation guidelines for German intonation. Proceedings of the 18th International Congress of Phonetic Sciences (https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0317.pdf).

[10] Ladefoged, P. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Oxford: Blackwell.

[11] Liberman, M., Prince, A. S. 1977. On stress and linguistic rhythm. *Linguistic Inquir*y 8: 249–336.

[12] Prince, A. S. 1983. Relating to the grid. *Linguistic Inquiry* 14, 19–100.

[13] Selkirk, E. O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

[14] Sweet, H. 1877. *A handbook of phonetics: Including a popular exposition of the principles of spelling reform*. Oxford: Clarendon.

[15] Trager, G. L. and Smith, Jr., H. L. 1951. *An outline of English structure*. Studies in Linguistics, Occasional Papers 3. Norman, OK: University of Oklahoma Press.

[16] van der Hulst, H. G. 2009. Brackets and grid marks or theories of primary accent and rhythm. In Eric Raimy and Charles Cairns (eds.), *Contemporary views on architecture and representations in phonological theory*, Cambridge MA: MIT Press, 225–245.

[17] Wang, D. and Narayanan, S. 2007. An acoustic measure for word prominence in spontaneous speech. *IEEE Transactions on Audio, Speech, and Language Processing*. 15(2): 690–701.

[18] Wells, J. C. (ed.). 2008. *Longman pronouncing dictionary* (3rd edition). Harlow: Pearson Education.