

DOI: <https://doi.org/10.18485/philologia.2020.18.18.1>

UDC: 811.111'373

■ LEXICAL STRESS PATTERNS IN HIGH-FREQUENCY WORDS OF SPOKEN ENGLISH

VALENTINA RAPAJIĆ¹University of Defence
Military Academy
Belgrade, Serbia

Obrasci leksičkog naglaska u najfrekventnijim rečima engleskog jezika značajni su kako za nastavu i učenje, tako i za istraživanja u oblasti usvajanja engleskog jezika, iako im je poklanjano mnogo manje pažnje nego efektima učestalosti u drugim segmentima jezičke strukture. U ovom radu su obrasci leksičkog naglaska opisani i rangirani po zastupljenosti u najfrekventnijoj engleskoj leksici u opštem i akademskom registru. Obrasci identifikovani u korpusu koji čine reči od 2 do 6 slogova u frekvencijskoj listi *Longman Communication 3000* (koja pruža podatke o najfrekventnijim rečima u opštem registru engleskog jezika) upoređeni su sa podacima koje su prethodni istraživači sakupili u korpusima zasnovanim na bazi podataka *Hoosier Mental Lexicon* (koja pruža podatke o vrednovanju poznatosti i vremenu odziva za visokofrekventne reči kod izvornih govornika engleskog jezika) i listi *Academic Word List* (koju čine najfrekventnije reči u akademskom diskursu). Iako su ova tri korpusa različiti po veličini i domenu, kod dva korpusa iz oblasti opšteg engleskog jezika zapažaju se izrazite podudarnosti kod dvosložnih i trosložnih reči, a kod četvorosložnih podudaranja postoje kod sva tri korpusa. Time su jasno potvrđeni dominantni obrasci leksičkog naglaska – a to znači oni kojima su učenici najviše izloženi. Uvid u zastupljenost obrazaca leksičkog naglaska olakšava izbor stavki za leksička vežbanja u nastavi engleskog jezika (usmerena ne samo na vežbanje izgovora nego i na razvoj leksikona kod učenika), kao i izbor stimulusa u eksperimentima kojima se proučavaju L1 i L2 uticaji u razvoju njihovog međujezika.

Ključne reči: engleski jezik, leksički naglasak, obrasci leksičkog naglaska, visokofrekventna opšta engleska leksika, visokofrekventna engleska leksika u akademskom diskursu, nastava i učenje, usvajanje stranog jezika.

1 Kontakt podaci (Email): valentinrapajic@yahoo.co.uk

1. INTRODUCTION

In this day and age when successful (enough) communication is often taken as the only (realistic) goal of language learning and the only measure of its success, it should be noted that lexical-stress-related issues still have a role in intelligibility. Misplaced word stress has been shown to hinder communication (Benrabah 1997: 161), especially when coupled with omission of vowel reduction in unstressed syllables (Lepage and Busà 2013; Lepage 2015). Zielinski (2008) showed that native speakers rely on word stress and the number of syllables to determine unintelligible or near-unintelligible words (containing pronunciation errors) in L2 English recorded speech.

Whereas the importance of lexical item frequency for language learning has been well established (Liu & Nation 1985; Laufer & Nation 1995; Nation 2006; Cobb 2007; Ellis 2012, vol. 1), frequency of a suprasegmental feature such as the lexical stress pattern has not attracted as much attention (Murphy 2004; Baker & Murphy 2011). However, lexical stress patterns (henceforth LSPs) are important for vocabulary development and consolidation in language learning, but also in research on second language acquisition, in terms of L1 vs. L2 lexical stress patterns frequency.

The focus of this paper is the lexical stress pattern (henceforth LSP), a suprasegmental feature that combines the number of syllables in a word and placement of primary and secondary stress, if the latter is present. In this research, the LSPs used in the comparison of the corpora based on frequency lists are described in terms of primary stress only (and, of course, the number of syllables), disregarding the presence or absence of secondary stress. When secondary stress data are included (see Table 2), primary-stress defined LSPs are subcategorized into more detailed patterns. For reasons of clarity, the full investigation of percentages and rankings of LSPs across the corpora was carried out only on the primary-stress defined LSPs. Subcategorised LSPs are exemplified in Table 2.

It has been known since Zipf (1935, 1949) that a small number of words occur in texts many times, while the majority of words occur rarely, or only once. In other words, a small number of high-frequency words are likely to make up most of any text, whereas a great number of low-frequency words have only a small share. Later research² confirmed this finding, and led to the compilation of high-frequency lexis lists, which proved to be significant in EFL materials design and in the selection of vocabulary for dictionary definitions.

Insight into the LSPs of the most frequent English words is relevant for both second language teaching and learning (sensitising the learners to lexical prosody, increased exposure to English lexis) and SLA research (experiment design and selection of stimuli). These implications will be addressed in more detail in the conclusion.

2 For further information on vocabulary frequency lists and their history see Nation and Waring (1997) and Přecechtěl (2016).

2. PREVIOUS RESEARCH

2.1 CLOPPER (2002)

Clopper used the *Hoosier Mental Lexicon* (Nusbaum, Pisoni & Davis 1984) as a resource for research into the most frequent syllable-stress patterns³ in English. The *Hoosier Mental Lexicon* (henceforth HML) provides familiarity ratings and response times for 20,000 words in the *Merriam-Webster Pocket Dictionary*. She compiled a corpus of disyllabic, trisyllabic, and tetrasyllabic words (11,999 in total), and identified nine different syllable-stress patterns (or LSPs in our terminology). Noting “the inherent imbalance in the distribution of stress over syllables in multisyllabic words in English”, Clopper sums up the main findings of her study: a) “two- and three-syllable words are more likely to have primary stress on the first syllable than on any other syllable”, and b) “[f]our syllable words, however, are more likely to have primary stress on the second or third syllable than on the initial or final syllable” (Clopper 2002: 8).

It should be noted that while Clopper was interested in identifying the most frequent LSPs in English, the goal of this research is more modest: identifying the LSPs of the most frequent English words.

While Clopper used various statistical procedures other than percentages, since the total number of words in her corpus is available as well as the count for each LSP within it, it was possible to calculate the share of each LSP in the corpus, express it in percentages, rank the LSPs accordingly, and compare the data with the AWL- and LC3000-based corpora.

2.2 POST DA SILVEIRA (2011)

In a study of English word stress acquisition by speakers of Brazilian Portuguese, Post da Silveira tested “the hypothesis that the regularities which emerge in L2 learners’ productions, specifically related to word stress, are due to the robust representation triggered by L1 and L2 stress and lexical frequency” (2011: 1634). Although the stimuli included monomorphemic and polymorphemic disyllabic, trisyllabic, and tetrasyllabic words, the author reports correct and incorrect production (in percentages) only in terms of the syllable stressed (ultimate, penultimate, or antepenultimate), but not in terms of the lexical stress patterns.

The study found that “BP native speakers showed high percentages of correct word stress productions of English suffixed and non-suffixed words” (Post da Silveira 2011: 1637), but it also found preference for stress placement on the penultimate syllable; although penultimate stress was correctly assigned in most items, the most mistakes also concerned erroneous assignment of penultimate stress. The author attributes this to “the transfer of the robust representation of the penultimate stress pattern in both L1 and L2 lexicons” (ibid.).

3 Clopper defines “syllable-stress patterns” as “various combinations of syllable number and primary stress location” (Clopper 2002: 2); in this research they are termed “lexical stress patterns”.

Noteworthy is a reference to the informants' learning experience. "[W]e observed that amount of implicit exposure to word stress frequency is not triggering enhancement in learners' performance, since there were no significant differences in production within the distinct semesters of instruction in the L2" (Post da Silveira 2011: 1637). This observation seems to signal a need for seasoned lexical stress-related vocabulary practice, that could bring about progress in lexical stress production competence.

Even in terms of absolute stress, the claim that penultimate stress is predominant in English cannot be accepted (cf. Cutler and Carter 1987). To verify this, we revisited Clopper's data, which Post da Silveira cites as her baseline research, in order to calculate and compare the percentages of penultimate and initial stress, respectively, in disyllabic, trisyllabic, and tetrasyllabic words of the HML, taken together. While penultimate stress is present in 31.89% of these words, word-initial stress is present in 34.95%. It should be pointed out that there is some overlap, of course, because disyllabic words with penultimate stress and disyllabic words with initial stress are simply two ways of describing the same LSP. Generalisability of these findings depends on how representative the HML is of the English language as a whole. Also worth noting is the fact that Clopper's data apply to disyllabic, trisyllabic, and tetrasyllabic words only, but not to polysyllabic words beyond 4-syllables, whose share is not insignificant.

2.3 MURPHY AND KANDIL (2004)

Murphy and Kandil analysed the Academic Word List (henceforth AWL) developed by Coxhead (2000) in order to "identify patterns of word-level stress and their frequency of occurrence" (Murphy and Kandil 2004: 64). Acknowledging that the main purpose of the AWL is vocabulary and reading instruction, they add that "it could also serve as a useful resource for teachers who prepare EAP learners as more competent speakers and listeners in academic settings" (Murphy and Kandil 2004: 65). The objective of their analysis is to gain insight into "how to sequence the introduction of word-stress patterns in EAP courses" (ibid.) by providing a complementary source of information in syllabus and lesson planning, all in an effort to enhance both speaking and listening training outcomes, since lexical stress influences intelligibility of speech and provides important auditory signals that facilitate listening comprehension. The authors emphasize that teachers should also draw learners' attention to "the predictable ways in which suffixation alters word-level stress patterns in English" and that EAP learners should be introduced to "such predictable patterns in conjunction with the frequency data reported in the study".

Murphy and Kandil compiled a corpus of all polysyllabic words (ranging from 2 to 7 syllables) in AWL: 2979 out of 3109 words. The authors identified 39 LSPs, calculated the share of each LSP in their corpus, expressed it in percentages, and ranked the LSPs. They discovered that the highest ranking 14 patterns account for over 90% of the polysyllabic words in AWL, with 25 other patterns making up the rest of their AWL-based corpus.

Critiquing Murphy and Kandil's study, Torrie (2016) remarks that an effort to memorize a lexical stress pattern and a corresponding notation for each word would

put an additional unnecessary burden on learners. However, Murphy and Kandil never suggest that lexical stress notations or lexical stress patterns should be memorized as additional items. "Rather than taking the time to provide instruction in what may appear to learners to be an overwhelming inventory of 39 possible stress patterns in academic words, our analysis revealed that just 14 patterns encompass 90% of the polysyllabic items in the combined AWL headwords and sublists inventories" (Murphy and Kandil 2004: 71). The notations are simply description conventions, for ease of reference in the survey of the multitude of patterns.

To our knowledge, this is the first study to document that a relatively small number of stress patterns are reflected in the pronunciations of a large number of academic words. This finding signals that the process of gaining control over word-level stress patterns of academic vocabulary may be more manageable than previously believed, and it should be taken as encouraging news by EAP learners and their teachers. It would be helpful, for example, for EAP learners and teachers to prioritize their efforts by highlighting attention to words already included in course materials that fit the relatively higher-frequency patterns. [...] At 90% coverage, we consider an inventory of the fourteen most frequent patterns to be a manageable number to feature in either courses or lesson segments of EAP speech-intelligibility instruction. [...] [T]he remaining ten percent of AWL items are tied to 25 additional patterns, each of which is of relatively low frequency of occurrence (...) Such information illuminates where learners with beginning to high-intermediate levels of proficiency in English can most profitably spend their time and energy when it comes to learning to recognize, identify, and use stress patterns of polysyllabic words. (Murphy and Kandil 2004: 71-73)

Our position is that the LSP, together with its prominence or its infrequency, is an important descriptor, an element of auditory mental representation, alongside the sequence of phonemes, when storing words in the mental lexicon⁴. The benefit of the fact that various LSPs are not equally conspicuous is that they can help sensitise learners to a variety of patterns through contrasting the more familiar LSPs with the less frequent ones.

4 Without going into detail regarding the mental lexicon debate (whether L1 and L2 mental lexicons are separate or integrated, whether there is an interlanguage system that is neither L1 or L2, etc.), suffice it to say that there is evidence that some components of word sound structure are imprinted more than others, e.g. initial syllables more than final ones, and both more than median ones; the rhythmic structure (lexical stress patterns, V.R.) of a word is also an important contributor to the auditory presentation in the mental lexicon (Aitchison 2004). Meara (1983) claimed not only that L2 mental lexicon is organized phonologically rather than semantically, but that it is in fact largely phonological in nature.

3. PRESENT RESEARCH

3.1 THE AIM OF THE STUDY

The aim of the present study is to establish whether the words of the same number of syllables in the three corpora exhibit the same tendencies in terms of LSPs; in other words, whether LSPs have the similar shares across the corpora markedly different in size⁵. If the distribution of LSPs is the same in all three corpora, and furthermore skewed in favour of the same LSPs in each corpus, the distinction between typical and atypical LSPs can be supported by frequency data. Such data, as previously mentioned, are significant for both SLA research and English teaching practice.

3.2 MATERIALS AND METHODS

The corpus for this research was the *Longman Communication 3000* frequency list (henceforth LC3000) that actually contains 3593 most frequent words in spoken and written English, based on statistical analysis of the Longman Corpus Network. Two sets of words were excluded from this list: 1324 monosyllabic words, and 592 words that are highly frequent in writing only. This yielded a total of 1677 highest-frequency words which are very common in either both speech and writing or in speech only. The selection ensures that the lexical patterns involved are those that are heard most often; words encountered in reading only may or may not have an LSP representation in the mind of the learner.

The *American Heritage Dictionary* was consulted for syllabification and stress assignment, to make the present research findings comparable to those in Murphy and Kandil (2004). Following Murphy and Kandil's approach, for entries with multiple transcriptions, only the first transcription was included in the analysis.

Clopper (2002) and Murphy and Kandil (2004) studied lexical stress patterns to different depths and based on different corpora. Clopper did not study words beyond 4 syllables in length in the HML, whereas Murphy and Kandil studied all polysyllabic words in AWL. Clopper was concerned with primary stress only, whereas Murphy and Kandil took both primary and secondary stress into consideration. Clopper expressed the frequency of each LSP as the sum of frequencies of all words in the HML-based corpus that exhibit the pattern. Murphy and Kandil expressed the frequency of an LSP as the percentage of the AWL-based corpus taken up by the words exhibiting it. However, the fact that both studies provide data on the number of words exhibiting particular LSPs and the total number of words examined, makes them comparable to each other as well as to the present study.

The notation for LSPs is adapted from the *Carnegie Mellon University Pronouncing Dictionary* (CMU PD). The number of digits equals the number of syllables. The syllable marked 1 carries primary stress. Whereas in CMU unstressed syllables are marked 0 and those carrying secondary stress are marked 2, in the notation applied here the syllable marked 0 is either unstressed or carrying secondary stress.

5 LC3000 to HML ratio is 1 : 7.15. Although LC3000 to AWL ratio is not that high (1 : 1.78), it still means that AWL is almost twice the size of LC3000.

Shown in the section below are the percentages of words in the three corpora according to the number of syllables, the percentages of LSPs in each set defined by the number of syllables, and the LSPs rankings.

3.3 RESULTS AND DISCUSSION

Three sets comprised of words from 2 to 5 syllables are compared in Figure 1 according to their participation in the LC3000-, HML-, and AWL-based corpora. For purposes of simplicity and ease of reading, these corpora will be termed LC3000, HML, and AWL, respectively.

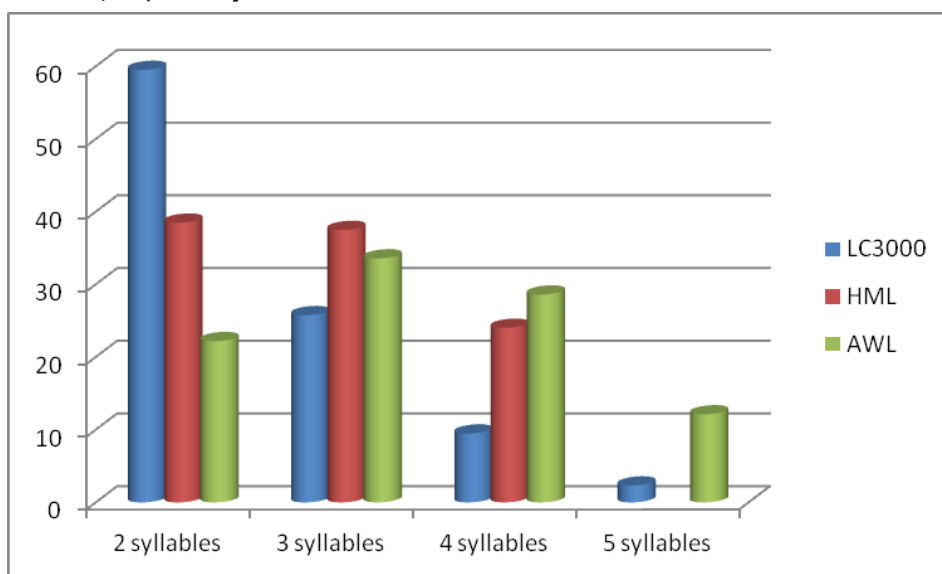


Figure 1. Percentages of words in the three corpora according to the number of syllables

While disyllabic words overwhelmingly dominate other polysyllabic words in LC3000, it is not the case with AWL, where disyllabic words are clearly outnumbered by trisyllabic and tetrasyllabic words, the former dominating the set. In HML (words ranging from 2 to 4 syllables only), disyllabic words are negligibly more represented than trisyllabic words, with tetrasyllabic words being outnumbered by both. The sharply descending pattern in LC3000 is not replicated by the sets in either of the other two corpora. With dissimilarities abundantly evident, the LSPs emerging as prevailing in the sets defined by the number of syllables in at least two of the three corpora can be validated as typical LSPs in English.

The subsets defined by the number of syllables are broken down according to the position of primary stress. The results are shown in Table 1. It should be noted that 6- and 7-syllable words are absent from this overview, because in LC3000 only one hexasyllabic word was found, and no septasyllabic words.

LPSS	LC3000 1,677 words			HML 11,999 words			AWL 2,979 words		
	Count	%	Rank	Count	%	Rank	Count	%	Rank
10	713	71.44	1	3624	78.46	1	321	48.56	2
01	285	28.56	2	995	21.54	2	340	51.44	1
100	254	58.8	1	2619	58.22	1	368	36.8	2
010	166	38.42	2	1510	33.57	2	557	55.7	1
001	12	2.78	3	369	8.2	3	75	7.5	3
1000	43	27.04	2	497	17.24	3	270	31.69	2
0100	75	47.17	1	1331	46.18	1	345	40.49	1
0010	40	25.16	3	1017	35.29	2	236	27.7	3
0001	1	0.63	4	37	1.28	4	1	0.12	4
10000	0	0	/				21	5.82	4
01000	13	33.34	2				164	45.43	1
00100	12	30.77	3				112	31.02	2
00010	14	35.9	1				64	17.73	3
00001	0	0	/				0	0	/

Table 1. LSPs in LC3000, HML, and AWL, defined by the number of syllables and primary stress: count, percentage, and rank. The matching rankings are highlighted.

Once again, it should be noted that the adaptation of CMU PD notation for this purpose does not take secondary stress into account; therefore, the opposition *1: 0* means *primary stress : secondary stress/absence of stress*.

The data from Table 1 are visualised in Figures 2 and 3, where correspondences as well as stark differences can be easily observed.

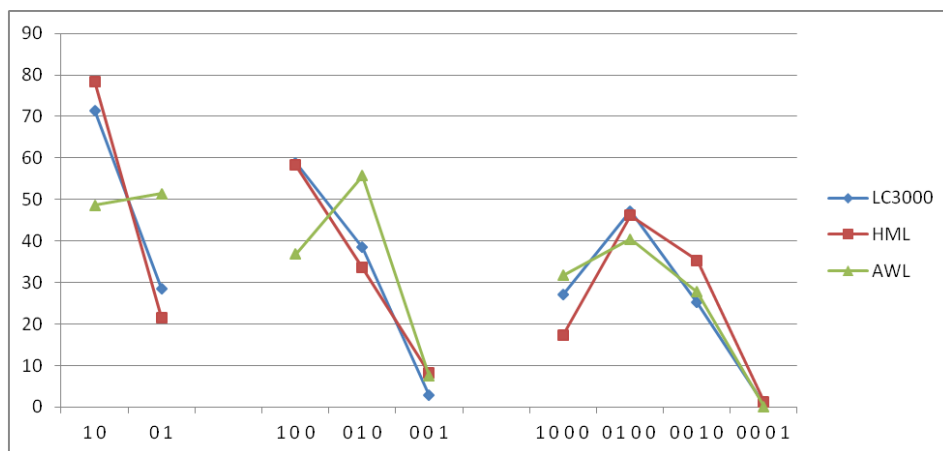


Figure 2. The percentages of LSPs in 2-, 3-, and 4-syllable words in HML, AWL, and LC3000

Regarding disyllabic words, those in LC3000 and HML exhibit the same ranking and similar percentages: there are approximately three times more words with stress on the first syllable, than those stressed on the second. The word-initial LSP is clearly dominant. In AWL, however, the situation is not only reverse (it is the word-final LSP that is more represented), but the difference in percentages between the more represented LSP and the less represented one is very slight.

With trisyllabic words there are again marked correspondences between LC3000 and HML. Initial-stress words are not only ranked first, but the difference in their percentages is negligible. The second-ranking LSPs (medial stress) are also similar in percentages. The third-ranking LSPs (final stress) are markedly fewer; still, in HML they are slightly more represented than in LC3000. AWL is different again: word-initial and word-final LSPs have the reverse rankings and percentages compared to both LC3000 and HML.

Regarding tetrasyllabic words, the situation visibly changes. With both disyllabic and trisyllabic words the correspondences between LC3000 and HML are strong. Compared to them, distributions of LSPs in AWL are reverse in disyllabic words, and markedly different in trisyllabic ones. But with tetrasyllabic words, it is AWL that corresponds with LC3000 in both rankings and percentages of LSPs, whereas HML differs slightly from both.

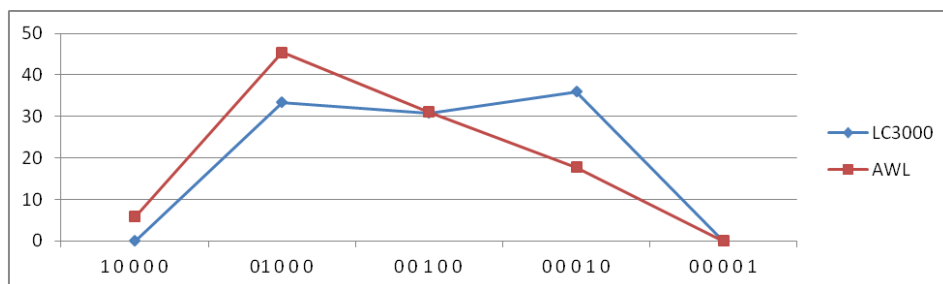


Figure 3. The percentages of LSPs in 5-syllable words in AWL and LC3000

Data on LSPs beyond four syllables are absent for HML, and those for LC3000 and AWL are therefore displayed separately. No pentasyllabic words with initial stress are recorded in LC3000, and none with word-final stress in either LC3000 or HML. The only intersection concerns words with antepenultimate stress (ranked third in LC3000 and second in HML).

Finally, words exemplifying subcategorised LSPs (those where both primary and secondary stress are included in the notation) are shown in Table 2.

HML	AWL	LC3000	Number of syllables	Examples
2syl-1pri	2-1	10	2	<i>versions, winter</i>
	2-1-2	12		<i>networks, software</i>
2syl-2pri	2-2	01		<i>approach, police</i>
3syl-1pri	3-1	100	3	<i>analyst, excellent</i>
	3-1-2	120		<i>formatted, somebody</i>
	3-1-3	102		<i>institute, restaurant</i>
3syl-2pri	3-2	010		<i>commitment, remember</i>
	3-2-3	012		<i>adulthood, beforehand</i>
	/	210		<i>unlikely</i>
3syl-3pri	3-3	/	<i>unaware</i>	
	3-3-1	201	<i>guarantee, afternoon</i>	
4syl-1pri	4-1	1000	4	<i>variable, comfortable</i>
	4-1-3	1020		<i>qualitative, supermarket</i>
	4-1-4	1002		<i>visualize, literature</i>
4syl-2pri	4-2	0100		<i>complexity, community</i>
	4-2-1	/		<i>practitioner</i>
	4-2-4	0102		<i>facilitate, relationship</i>
4syl-3pri	4-3-1	2010		<i>economic, conversation</i>
	/	0010		<i>electronic</i>
4syl-4pri	4-4-1	2001		<i>nevertheless, nevertheless</i>

HML	AWL	LC3000	Number of syllables	Examples
	5-1	/	5	<i>culturally</i>
	5-1-3	/		<i>specifiable</i>
	5-1-4	/		<i>regulatory</i>
	5-2	01000		<i>considerable, occasionally</i>
	5-2-4	01020		<i>discriminating, refrigerator</i>
	5-2-5	/		<i>contextualize</i>
	5-2-1-4	/		<i>incorporated</i>
	5-3-1	20100		<i>methodologies, international</i>
	/	00100		<i>electricity</i>
	5-4-1	20010		<i>implementation, organisation</i>
	5-4-2	02010		<i>environmental, accommodation</i>
	6-1-3		6	<i>justifiably</i>
	6-2			<i>inevitably</i>
	6-2-4			<i>administratively</i>
	6-2-5			<i>conceptualizing</i>
	6-3-1			<i>philosophically</i>
	6-4-1			<i>variability</i>
	6-4-2	020100		<i>predictability, responsibility</i>
	6-5-1	/		<i>re-evaluation</i>
	6-5-2	/		<i>intensification</i>
	6-5-3-1	/		<i>differentiation</i>

Table 2. Lexical stress patterns with both primary and secondary stress included in the notation, exemplified by words from AWL and LC3000

Original LSP notations from two previous studies are retained in the table. As previously mentioned, an adapted CMU PD notation is applied in the present study.

The AWL examples from Murphy and Kandil (2004) are listed first, and the LC3000 examples are listed next (except for those LSPs for which examples were found in one study only). In some cases examples from the former are inflexions.

4. CONCLUSION

Lexical stress patterns were compared in terms of percentage and rank (the latter within the sets defined by the number of syllables). Despite marked differences in the size of the corpora in this investigation as well as the variation in prevalence (disyllabic words in LC3000 and HML, trisyllabic words in AWL), notable correspondences were found between LC3000 and HML regarding both disyllabic and trisyllabic words, whereas with tetrasyllabic words correspondences were found in all LSPs in LC3000 and AWL, as well as regarding the most and the least represented LSPs in HML.

Certain generalisations about typical LSPs as well as the differences between the LSPs regarding frequent lexis in general English and frequent lexis in academic discourse in English can now be made. In general English word-initial stress is clearly dominant in disyllabic words, whereas in the most frequent academic lexis slightly more disyllabic words carry stress on the second syllable than on the first. In the most frequent general English lexis most trisyllabic words carry stress on the initial syllable, but in the most frequent academic lexis the medial syllable is most often stressed. In both general English and academic English, as far as the most frequent lexis is concerned, most words carry antepenultimate stress, and the words with ultimate stress are the fewest.

The fact that there are differences between LSPs in general and academic English, at least as far as the most frequent disyllabic and trisyllabic words are concerned, actually represents an excellent learning and teaching opportunity: academic lexis can be practiced together with everyday lexis with the same LSPs (everyday lexis presumably being more familiar), especially since there are LSPs that are frequent in one domain but less so in the other.

LSP frequency data can facilitate the selection of lexis for pronunciation practice tasks in general, regardless of type of discourse: contrasting the words with frequent and infrequent LSPs can sensitise learners to lexical prosody. Such tasks can provide both vocabulary practice and additional exposure to English lexis.

In SLA research, insight into the most represented LSPs regarding the most frequent English words and those LSPs that are relatively rare can facilitate the selection of stimuli for stress-placement experiments involving foreign speakers of English. For instance, if informants who do not share the same L1 are asked to assign stress a) to frequent (and presumably familiar) words with common and uncommon LSPs, and b) to rare (and presumably unfamiliar) words with common and uncommon LSPs, a positive correlation of LSPs frequency and correct stress assignment would suggest L1-based generalisation rather than L2 influence. On the other hand, if significant similarities in stress misplacement are found in the informants with the same L1, this would point to transfer from L1. Such research, concerning the dynamics of L2-based generalisations and L1 transfer, can shed more light on the lexical prosody of interlanguage.

REFERENCES

- Aitchison, J. 2004. *Words in the Mind: An Introduction to the Mental Lexicon*. 3rd Edition. Oxford: Basil Blackwell Publishers.
- Baker, A. & J. Murphy. 2011. Knowledge base of pronunciation teaching: staking out the territory. *TESL Canada Journal* 28(2), 29–50.
- Benrabah, M. 1997. Word-stress – a source of unintelligibility in English. *International Review of Applied Linguistics* 35(3), 145–165.
- Clopper, C. G. 2002. Frequency of stress patterns in English: A computational analysis. *IULC Working Papers* 2(1), 1–9.
- Cobb, T. 2007. Computing the vocabulary demands of L2 reading. *Language Learning & Technology* 11(3), 38–63.
- Coxhead, A. 2000. A new academic word list. *TESOL Quarterly* 34(2), 213–238.
- Cutler, A. & D. Carter. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* 2, 133–142.
- Ellis, N. 2012. What can we count in language, and what counts in language acquisition, cognition, and use? In S. Gries & D. Divjak (eds.) *Frequency Effects in Language Learning and Processing*. 2 vols. Berlin: De Gruyter Mouton.
- Meara, P. 1983. Word associations in a foreign language. *Nottingham Linguistics Circular* 11(2), 29–38.
- Murphy, J. 2004. Attending to word stress while learning new vocabulary. *English for Specific Purposes* 23, 67–83.
- Murphy, J. & M. Kandil. 2004. Word-level stress patterns in the academic word list. *System* 32(1), 61–74.
- Lepage, A. & M. Busà. 2014. Intelligibility of English L2: The effects of incorrect word stress placement and incorrect vowel reduction in the speech of French and Italian learners of English. *Concordia Working Papers in Applied Linguistics* 5, 387–400.
- Lepage, A. 2015. *The Contribution of Word Stress and Vowel Reduction to the Intelligibility of the Speech of Canadian French Second Language Learners of English*. Laval University. Unpublished doctoral dissertation.
- Liu, N. & I. S. P. Nation. 1985. Factors affecting guessing vocabulary in context. *RELC Journal* 16(1), 33–42.
- Longman Communication 3000. [Internet]. Available at: https://www.lex tutor.ca/freq/lists_download/longman_3000_list.pdf [22. 11. 2020].
- Nation, P. & R. Waring. 1997. Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (eds.) *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press, 6–19.
- Nusbaum, H., D. Pisoni, & C. Davis. 1984. Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception, Progress Report 10*. Bloomington, IN: Speech Research Laboratory, Indiana University, 357–376.
- Post da Silveira, A. 2011. Frequency as (dis)advantage to word stress acquisition. *Proceedings of the ICPHS XVII*, 1634–1637.
- Přecechtěl, V. 2016. *English Vocabulary Frequency and Its Use at Lower Secondary Schools*. Pedagogická Fakulta Univerzita Palackého. Unpublished master's thesis.
- The American Heritage Dictionary of the English Language*. [Internet]. Available at:

- <https://ahdictionary.com/> [22. 11. 2020].
The Carnegie Mellon University Pronunciation Dictionary. [Internet]. Available at: <http://www.speech.cs.cmu.edu/cgi-bin/cmudict> [22. 11. 2020].
- Torrie, H. 2016. Examining word-level stress patterns: Comment on Murphy & Kandil (2004). [Internet]. Available at: <http://www.heathertorrie.com/2016/05/examining-word-level-stress-patterns.html> [22. 11. 2020].
- Zielinski, B. W. 2008. The listener: No longer the silent partner in reduced intelligibility. *System* 36, 69–84.
- Zipf, G. K. 1935. *The Psycho-Biology of Language*. Boston: Houghton Mifflin.
- Zipf, G. K. 1949. *Human Behaviour and the Principle of Least Effort. An Introduction to Human Ecology*. Cambridge, MA: Addison-Wesley.

SUMMARY

LEXICAL STRESS PATTERNS IN HIGH-FREQUENCY WORDS OF SPOKEN ENGLISH

Lexical stress patterns exhibited by the most frequent English words are significant for teaching practice as well as for SLA research, although they have received much less attention than frequency effects in other segments of language structure. This paper describes and ranks lexical stress patterns according to their share in the most frequent lexis in general and academic registers. The patterns identified in the corpus consisting of 2- to 6- syllable words in the *Longman Communication 3000* frequency list (that provides data on the most frequent words in general English) are compared to previous researchers' data obtained from the corpus based on the *Hoosier Mental Lexicon* (that provides data on native speakers' familiarity ratings and response time for high-frequency words) and the corpus based on the *Academic Word List* (consisting of the most frequent words in academic discourse). Although the three corpora vary in size and domain, in the two general English corpora there are strong correspondences regarding 2- and 3-syllable words; with 4-syllable words correspondences are noted in all the three corpora. This validates dominant lexical stress patterns, to which the learners are most often exposed. Insight into the representation of lexical stress patterns in high-frequency lexis facilitates the selection of items for vocabulary exercises in language learning (intended not only for pronunciation practice but also for learners' vocabulary development), as well as the selection of stimuli for experiments regarding L1 and L2 effects in interlanguage development.

KEYWORDS: English, lexical stress, lexical stress patterns, high-frequency general English lexis, high-frequency academic English lexis, teaching and learning, SLA.

ARTICLE INFO:

Original research article
 Received: February 19, 2020
 Revised: November 23, 2020
 Accepted: December 7, 2020