Frequency as (dis)advantage to word stress acquisition

Post Da Silveira, A.

Note: To cite this publication please use the final published version (if applicable).

FREQUENCY AS (DIS)ADVANTAGE TO WORD STRESS ACQUISITION

Amanda Post da Silveira

Leiden University Center for Linguistics, the Netherlands
a.post.da.silveira@hum.leidenuniv.nl

ABSTRACT

This is an initial study on the acquisition of English word stress by Brazilian Portuguese learners. The results show that incorrect productions typically had penultimate stress. This phenomenon is the focus of our present investigation. Two major factors are likely to be influencing these inaccurate productions: (i) the high frequency of penultimate stress in the first-language lexicon and (ii) category assimilation of phonetic features that characterize English word stress.

Keywords: prosody, word stress, frequency, phonetic assimilation, second language acquisition

1. INTRODUCTION

Based on the description and analysis of a corpus of English (L2) spoken language by Brazilian Portuguese (BP or L1) learners, we, in a prior study, observed that penultimate stress is correctly produced in approximately 80% of the cases. However, when an analysis of their errors was undertaken, penultimate stress was the most common type of erroneous stress production, ca. 50% in BP English.

Our hypothesis is that the high frequency of the penultimate stress pattern in the L1 is triggering what was called by Trubetzkoy [8], the L1 filter, so that the phonetic features that characterize English word stress are not perceived and/or produced. Or even if the L2 stress is correctly perceived, it seems that in some cases learners do not reproduce such information in their L2 productions. The consequence of the erroneous or mismatched perception/production is that phonetic features that are distinct in both languages are assimilated according to the patterns that learners were more exposed to, such as L1 and/or L2 predominant stress patterns and lexical frequency in both L1 and L2 [2].

With the present study, we intend to give an account of the role played by word stress pattern frequencies in L1-Brazilian Portuguese and BP-English second language acquisition (SLA). This study serves as a basis for a future study on the investigation of the assimilation of phonetic categories triggered by frequency.

2. LEXICAL FREQUENCY AND SLA

Language comprehension is partly determined by the learners' great amount of statistical information concerning how lexical items behave in their own language. The investigation of pattern frequencies in the languages implicated in SLA is a starting point for the discovery of how constructions emerge and the ways in which the token frequency affects the productivity of a pattern. The phonetic-phonological encoding is based on the retrieval of phonetic-phonological distributional properties of words in the input [5]. Thus, the degree of accurate perception and production in the target language would be dependent on the correct weighting range of association for each element of the language input.

Experimental and theoretical linguists have approached the question of how learners store and process phonetic categories in the L2. Most of the existing studies investigate segmental phenomena. while a more limited number of studies have approached the effects of frequency in the acquisition of L2 prosody [1, 6].

The present experimental study intends to add evidence, by means of frequency counting in corpora (comprising multiple databases) and based on a production experiment, to test 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.

2.1. Word stress in English and in BP

The frequency distribution data on English word stress patterns is based on Clopper's study [3]. Firstly, Clopper examined the combination of syllable number and stress location. The measures of frequency were calculated based on the lexical

frequency data provided by the Hoosier Mental Lexicon (HML), an online version of Webster's Pocket Dictionary that includes 20,000 words and the Brown University database. The token frequency is shown in table 1:

Table 1: Sum frequency relating word size and stress position in English databases.

Words size/ stress position	Sum frequency (absolute and relative numbers)	
2 syllables	**87,574**	**(59.06%)**
Final	19,881	(22.71%)
Penultimate	67,693	(77.29%)
3 syllables	**41,234**	**(27.80%)**
Final	1,398	(3.40%)
Penultimate	15,278	(37.05%)
Antepenult.	24,558	(59.55%)
4 syllables	**19,491**	**(13.14%)**
Final	3,549	(18.20%)
Penultimate	6,831	(35.05%)
Antepenult.	9,014	(49.25%)
Pre-antepenult.	97	(0,50%)
Total	**148,299**	

From table 1, we can firstly observe that there is a tendency for shorter words (2- and 3-syllables) to be more frequent than longer words (4-syllables) in the lexicon. In relation to words stress, final syllable stress is the least frequent overall. Considering the sum frequency, we can also infer that penultimate stress is the most assigned word stress pattern in English, since 2-syllable penultimate stress is the most frequent overall, while for 3-syllable and 4-syllable words, the antepenultimate stress is the most frequent pattern.

In BP, a study on word stress frequency was carried out based on ASPA (a statistical database for sound patterns of BP). The corpus consists of 607.392 words (types) in a total of 228.766.402 tokens [4]. The analysis shows that there is a highly significant difference in type frequency among the stress patterns in BP, $\chi^2(2)=5642,69$, $p<0,0001$, the second-syllable stress pattern being the predominant one (both on verbs and non-verbs), as demonstrated in table 2:

Table 2: Type frequency of word stress assignment in verbs and non-verbs in Brazilian Portuguese.

Stress position	Verbs		Non-verbs	
Final	10,617	(27.3%)	14,200	(18.2%)
Penultimate	**27,730**	**(71.4%)**	**52,271**	**(67.1%)**
Antepenult	482	(1.2%)	11,389	(14.6%)

Clearly, the percentage of penultimate syllable stress assignment is very high, approximately 70%. Thus, as we are taking into account token frequencies as evidences for word stress acquisition, our hypothesis is that the penultimate word stress pattern robust frequency in both English and Brazilian Portuguese lexica will be transferred to the BP learner's productions in English. Such transfer may implicate a facilitation to English word stress acquisition as well as a disadvantage, since other least frequent patterns in the target language may be assimilated.

3. THE EXPERIMENT

In order to observe speech patterns emergent in English word stress acquisition by Brazilian learners, we ran a production experiment. The stimulus materials included 135 words with stress on either the first, second or third syllable. Word length varied from two to four syllables. The stimuli contained monomorphemic and polymorphemic words. The polymorphemes were divided into three groups according to the behavior of the suffixes in relation to the primary stress of the root words: (i) stress neutral (no change in the position of root word primary stress), (ii) stress bearing (shift of the primary stress from the root onto the suffix) and (iii) stress attracting (shift of the stress to a position one or two syllables before the suffix [10]). Words were read in isolation as well as in carrier sentences. The resulting corpus contained 7,560 tokens.

3.1. Informants

The informants were sixteen undergraduate students of English Language at a Brazilian Southern Federal university whose first instruction in spoken English began in the undergraduate course. All of them were between 18 and 24 years old. Informants were from mixed semesters, yielding different amounts of experience with the target language, and were blocked in groups of four per semester. We expected to find progression in their L2 speech accuracy, so that students in the eighth semester would perform better than the students in the second semester.

3.2. Stimuli and procedures

The informants read aloud lists of isolated words printed on cards. Afterwards, they read the same words in carrier sentences. Each informant performed the experiment twice within an interval

of 6 months. Recordings were made in a quiet room on a digital recorder. Each recording lasted approximately 12 minutes and they were transcribed by ear by an expert based on IPA. The transcriptions were reviewed by two experts, non-native speakers of English. Later, the data were transcribed in Praat software. Stress position was indicated by the author, on the basis of auditory information only.

3.3. Results

Firstly, there were no significant differences in the BP-English production of word primary stress from the first in relation to the second data collection, $\chi^2(1)=2.06$, p=0.15, n.s.. On both occasions, the word stress was produced on the correct syllable in 80% of the cases. Also, the data showed no significant influence of the context in which the words were read, i.e., whether in isolation or in carrier sentences, $\chi^2(1)=3.18$, p=0.075, n.s.. Surprisingly, there were even no significant differences in the performance of the four distinct semester groups. The percentage of errors is described as follows: 2nd semester (26.5%), 4th semester (26.4%), 6th semester (20.7%) and 8th semester (26.3%).

The percentages presented above and the observation that there was no difference in the performance of students in six months indicate that longer exposure to the L2 alone has not caused a significant enhancement in students' performance, at least not in terms of word stress. Accordingly, the tokens collected from different groups of students and/or at different moments in their training will be considered as a single, homogeneous language sample.

We observe that there is no significant difference in stress assignment concerning penultimate and antepenultimate syllable stress. There is a larger difference between penultimate and last syllable stress assignment, although not very prominent. The percentages of correct production of English word stress by BP native speakers are: final syllable (70.0%), penultimate (81.4%) and antepenultimate (79%).

From these results we could infer that the subjects are proficient, since the percentages of correct productions are within 70 to 80% approximately. Also, the percentages are very similar between the antepenultimate and penultimate stress patterns. The final stress pattern

is the least correctly assigned, 70%, but it is also in a relatively high range of correctness.

However, looking into the incorrect tokens, we observe a considerable difference among the stress patterns erroneously assigned, as the following percentages show: last syllable (18%), penultimate (53%), antepenultimate (24%) and pre-antepenultimate (5%). The most frequent error type was to stress the penultimate syllable. Comparing the percentages from correct and incorrect outputs one can state that the penultimate stress assignment is high in correct productions, but it can be highlighted as significantly higher in comparison to the other patterns when we look at the incorrect productions. It can be an indication that the system which accounts for the learners' productions in this case is actually that of the L1, not of the L2. Also, the high frequency in which the penultimate stress pattern is assigned in the lexicon of each of the languages involved, strongly motivates its use both in correct and in incorrect productions in the target language. But this account by itself does not answer the question whether the observed pattern is actually acquired on the basis of exposure to the L2 or whether it is merely the result of transfer from the L1.

Additional evidence that the high frequency of penultimate stress in the L1 actually plays a role in the L2 system, can be found in the percentages of incorrect production of morphologically simple and complex words with different suffix types, as shown in figure 1:

Figure 1: Percentages of word stress type erroneously produced in relation to the word suffixation type.
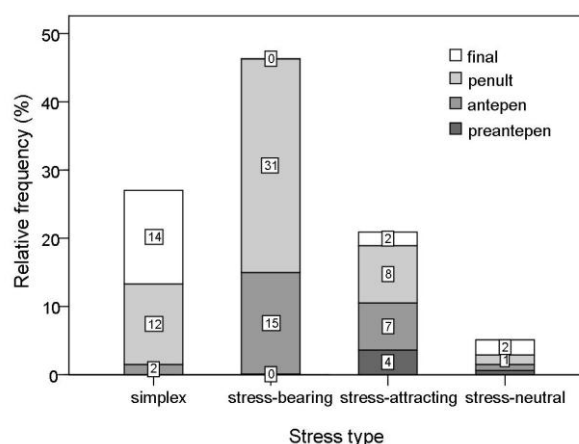
The penultimate stress is clearly more often assigned to the stress-bearing suffix. In previous studies, complex words containing stress-bearing suffixes have shown to be a challenge to the acquisition of English word stress by BP native

speakers, signaled by its random production, ca. 50% [7]. In this case, it seems that learners recur to the pattern that is more frequent in their system, the penultimate stress.

There is also another effect of frequency to be considered specifically in this group of suffixed words. Data from two databases (one formal and another informal) show that most of the complex words containing stress-bearing suffixes are a lot less frequent than their corresponding root words. Such root words also bear penultimate syllable stress in 80% of the cases, as in *mountain/ mountaineer* and *question/questionnaire*. Data description can be observed in table 3:

Table 3: Token frequency for roots and suffixed words extracted from IPhOD oral database [9] and Google.

Simplex/suffixed word	IPhOD		Google x 10^6	
Mountain/-eer	33	0	236	6
Employ/-ee	12	26	36	91
Refuge/-ee	7	7.0	22	15
Kitchen/-ette	90	3.0	293	12
Picture/-sque	168	9.0	705	16
Question/-naire	259	37	699	27
Vision/-aire	58.0	0	222	0

In sum, here we can infer that at least three factors are conspiring to the assignment of penultimate stress: (i) the high frequency of root words in the English lexicon that favor a more robust representation of such words regarding stress, (ii) the fact that the highly frequent root words also bear stress on the penultimate syllable and (iii) the strong representation of penultimate stress in the BP lexicon.

4. CONCLUSIONS

This study intends to give an account of the role frequency plays in the acquisition of second-language prosody.

Firstly, we 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.

In general, BP native speakers showed high percentages of correct word stress productions of English suffixed and non-suffixed words. However, we identified that stress assignment in incorrect productions bias towards the second syllable. Evidence has shown that it is due to the transfer of the robust representation of the penultimate stress pattern in both L1 and L2 lexicons, which may be causing the assimilation of specific categories that characterize word stress in the L2. A fact that corroborates this hypothesis is that, in less frequent word stress patterns, as complex words with stress-bearing suffixes, learners seem to opt for the "safest choice", assigning penultimate stress.

In future studies we will test how frequency distribution affects the perception/production so that acoustic-articulatory categories that are distinct in L1 and L2 are assimilated.

5. REFERENCES

[1] Altmann, H., Vogel, I. 2002. *L2 Acquisition of Stress: the Role of L1*. Paper presented at the *DGfS Annual Meeting "Multilingualism Today"* Mannheim, Germany.

[2] Best, C.T., McRoberts, G.W., Goodell, E. 2001. American listeners' perception of nonnative consonant contrasts varying in perceptual assimilation to English phonology. *Journal of the Acoustical Society of America* 109, 775-794.

[3] Clopper, C.G. 2002. Frequency of stress patterns in English: A computational analysis. *Indiana University Linguistics Club Working Papers* 2.

[4] Cristófaro-Silva, T., de Almeida, L.S., Fraga, T. 2005. ASPA: A Formulação de um Banco de Dados de Referência da Estrutura Sonora do Português Contemporâneo. *XXV Congress of Brazilian Society of Computing Science* São Leopoldo, RS, Brazil.

[5] Ellis, N.C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24, 143-188.

[6] Peperkamp, S., Dupoux, E. 2002. A typological study of stress 'deafness'. In Gussenhoven, C., Warner, N. (eds.), *Papers in Laboratory Phonology* Berlin: Mouton de Gruyter, 7, 203-240.

[7] Post da Silveira, A. 2010. *Estrategias de Reparo na Aquisicao do Acento Primario do Ingles por Falantes Nativos de PB*. Master Thesis, UFSM, Santa Maria. *http://cascavel.cpd.ufsm.br/tede/tde_busca/arquivo.php? codArquivo=3192*

[8] Trubetzkoy, N. 1939. Grundzüge der Phonologie. *Travaux du cercle linguistique de Prague* 7.

[9] Vaden, K.I., Halpin, H.R., Hickok, G.S. 2009. Irvine Phonotactic Online Dictionary, Version 2.0. [Data file]. *http://www.iphod.com*

[10] Zamma, H. 2003. Suffixes and Stress/Accent Assignment in English and Japanese: A survey. In Honma, T., et al. (eds.), *A New Century of Phonology and Phonological Theory*. Tokyo: Kaitakusha, 456-469.