

A semi-automatic assessment of lexical stress patterns in non-native English speech

Évelyne Cauvin¹, Laure Pairet²

¹CLILLAC-ARP (EA 3967) - Sorbonne Paris Cité - Université Paris Diderot, France

²CURAPP (UMR 7319), Université de Picardie Jules Verne, France

<https://doi.org/10.36505/ExLing-2018/09/0007/000340>

Abstract

This study investigates the acquisition and assessment of English lexical stress placement by French native speakers. Previous research on the evaluation of stress patterns has been centred on selections of syllables in written words, but little is known about the learners' realisations in a spoken context. We aim to shape methodological criteria to semi-automatically assess auditory stress pattern realisations based on perception tests in read speech. By refining the methodology and running statistical tests, we have been able to find out more consistent data to create a more adequate variable to assess L2 learners' lexical stress proficiency.

Keywords: semi-automatic assessment, French learners of English, lexical stress models, evaluation grid, multidimensional analysis

Introduction

This study on lexical stress patterns in oral speech addresses intonation in the wider meaning of the suprasegmental field since *It also involves the study of the rhythm of speech, and (in English, at any rate) the study of how the interplay of accented, stressed and unstressed syllables functions as a framework onto which the intonation patterns are attached* (Wells [2006] 2009).

It is basically intended to be a response to the increasing pressure of finding ways to grade the proficiency of an international workforce in spoken English, which is in line with the *CEFRL* (2001). It paves the way to finding *criteria features* (Hawkins, Buttery 2010) in order to assess non-native spoken English at the suprasegmental level.

Extensive research has been done in lexical stress assignment and patterns, focusing on words out of context and based on data collected from dictionaries (Wells 2008, Roach *et al.* 2011). Descriptive findings have led to the formulation of rules for both native and non-native speakers of English (Guierre [1984] 1992, Duchet & Fryd 1998, Fournier 2010). Assessing lexical primary stress assignment is general practice to many teachers of English on an auditory basis and in written tests, yet disconnected from context and connected speech. Horgues (2010) used arrows to distinguish between right or left shift in learner speech. Cauvin (2017) devised a methodology to assess and semi-automatically grade

lexical stress assignment proficiency in non-native speech production in context. Four specialists listened to the recordings of 15 French learners of English, representative of 155 learners reading a text (*Longdale-Charlipbonia corpus*¹), and assessed 6 polysyllable realisations according to native productions found in reference pronunciation dictionaries (see above) as well as by listening to two native English speakers reading the same text as models. *Excel* spreadsheet formulas are the basis of the creation of a variable to assess lexical stress pattern realisations. Coining each word with right or wrong stress pattern assignment seems unsatisfactory due to a lack of adequacy with other prosodic variables. Nevertheless, it was a first step towards creating a comprehensive evaluative suprasegmental methodology.

Methodology and analysis

This research is based on Cauvin (2017)'s methodology and it analyses her data per learner, group and word more thoroughly by using statistical tests.

Factors at work in L2 lexical stress acquisition

As written above, 15 learners representative of a 155 learner corpus were selected according to their reading speed performance — three of them were selected in the first and ninth deciles, the first and third quartiles, and the median. Two more variables define the learners such as their age and time spent in an English-speaking country.

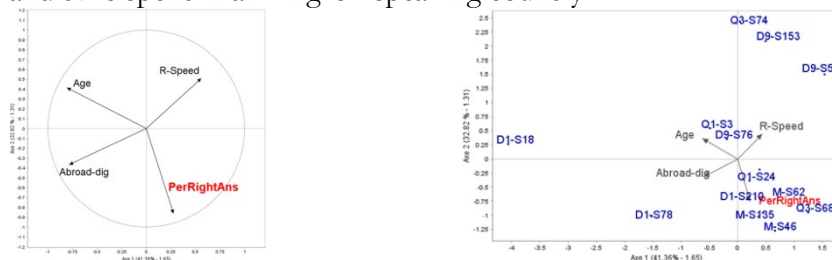


Figure 1. Representation of quantitative variables on the 1st plane of the PCA correlation circle and representation of the individuals on the first plane of the component space.

In Figure 1, the correlation circle shows that the percentage of correct answers (PerRightAns arrow) is orthogonal to the learners' reading speed (R-Speed arrow) and the time spent in an English speaking (Abroad-dig arrow). Neither the reading speed nor the time spent speaking English influences lexical stress mastery, which is acquired via specific teaching.

The most proficient students are in the bottom part of the PCA factor map (Escofier, Pagès 2008).

Polysyllabic word analysis

In the reading of the *Longdale-Charlipbonia* text, Cauvin (2017) selected six polysyllabic words in context in Figure 2 introduced with their stress pattern² (Stress Pattern), their syllable number (nSyll), their stress number (nStress), the potential presence of a prefix (Prefix), and a strong or weak suffix (Suffix).

Table 1. Polysyllabic word selection and their characteristics (their stress pattern, their syllable number, their stress number, the potential presence of a prefix, and either a strong or weak prefix).

Words	Stress Pattern	nSyll	nStress	Prefix	Suffix
Conversation	/2010/	4	2	pre	strong-suf
Limited	/100/	3	1	no-pre	weak-suf
Determined	/010/	3	1	pre	weak-suf
Trespassers	/100/	3	1	no-pre	weak-suf
Prosecuted	/1000/	4	1	no-pre	weak-suf
Musicians	/010/	3	1	no-pre	strong-suf

We devised a more complex grading scale with 5 for a correct answer, 4 for ignoring a secondary stress, 3 for creating or inverting primary and secondary stresses, 2 for stressing a syllable before the assigned placement, 1 for doing so after it and 0 for stressing the word at the end or when the stress realisation is unclear. Applying this refined assessment grid to Cauvin (2017)’s polysyllabic data resulted in a new hierarchy of word stress difficulty (Figure 3).

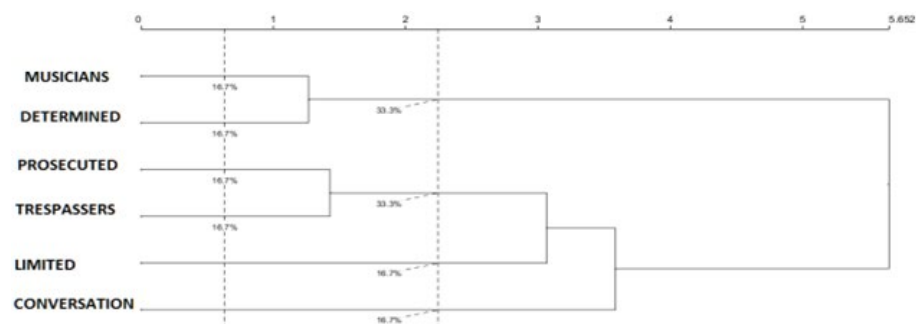


Figure 2 . Ascendant hierarchical clustering of the polysyllabic words with a dendrogram using 3 variables.

According to three assessment variables (Cauvin (2017)'s true/false grading scale, the current grading variable with six marking degrees from 0 to 5, and the sum of the learners' answers marked differently from 0), this hierarchical clustering analysis of the words shows that *musicians* and *determined* present a similar degree of difficulty, with the lowest success scores. At the bottom end of the dendrogram, *conversation* remains apart since it gathers the best results. *Limited* is next, and *trespassers* and *prosecuted* are grouped together at a higher difficulty level.

Discussion and conclusion

Contrarily to the 2017 true/false grading scale (Figure 4a), the results of the new one (Figure 4b) follow a normal distribution, similar to that found in most histograms depicting other prosodic variables, which shows a greater methodological consistency in assessing lexical stress. Finding more specific data in lexical stress assignment refines the semi-automatic assessment model.

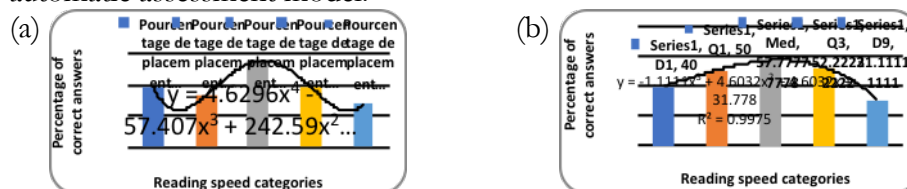


Figure 3. Assessing quality realisations per reading speed category (in percentage) with a) a 0-1 grading scale and b) a 0-5 grading scale

Notes

1. *Longdale-Charlipbonia*: [<http://www.chillac-arp.univ-paris-diderot.fr/projets/>]
2. Guierre ([1984] 1992)'s methodology: "1" for primary stress, "2" for secondary stress and "0" for an absence of stress.

References

- Cauvin, É. 2017. Élaboration de critères prosodiques pour une évaluation semi-automatique des apprenants francophones de l'anglais. Paris 7.
- Duchet, J.-L., Fryd, M. 1998. Manuel d'anglais oral pour les Concours. CNED: Didier Érudition.
- Fournier, J.-M. 2010. Manuel d'anglais oral. Paris: Ophrys.
- Guierre, L. [1984] 1992. Drills in English Stress-Patterns. Paris: Armand Colin-Longman.
- Hawkins, J.A., Buttery, P. 2010. Criterial Features in Learner Corpora: Theory and Illustrations. *English Profile Journal*, 1, e5. (40-62).
- Horgues, C. 2010. Prosodie de l'accent français en anglais et perception par les auditeurs anglophones. Université Paris Diderot - Paris 7.
- Wells, J. C. [2006] 2009. *English Intonation: An Introduction*. CUP.