Methodological Perspectives on Second Language Prosody

Papers from ML2P 2012

edited by
Maria Grazia Busà, Antonio Stella
CAN A GUIDED RHYTHMIC APPROACH CONTRIBUTE TO THE ORAL PERFORMANCE OF LEARNERS OF L2 ENGLISH?

A CASE STUDY

Francis Picavet,1,2 Véronique Auberge,2 Solange Rossato2

1Institut Polytechnique de Grenoble, 2Université Stendhal Grenoble 3
Francis.Picavet@ensae3.grenoble-inp.fr, Veronique.Auberge@gmail.com
solange.rossato@u-grenoble3.fr

ABSTRACT
The stress-timed vs. syllable-timed dichotomy between English and other languages (e.g. French), is at the origin of a major challenge for teachers and students of English L2 alike. Lexical stress—not a free parameter in English words—plays a fundamental role in the intelligibility of the message. As the rhythmic, basic beat-bearer of the utterance, its absence or misplacement can dramatically impair interaction with a native speaker. The experimentation described in this paper is based on an original pedagogical approach targeted at the acquisition of rhythm, using a metronome frequently during oral production tasks in class. An identical test was used, before and after the work with the metronome, enabling the measurement and analysis of progress-related data. The productions of 75 students were recorded during the test, so as to build a corpus. A preliminary single case study is related here. Using PRAAT software, measures were made and the collected data analyzed in order to determine whether there was progress. First results show definite progress in some parameters (vowel duration, pitch) and stagnation in others (intensity, vowel quality). This preliminary study paves the way to a future broader study based on part, or all, of the recorded corpus.

Keywords: prosody, lexical stress, metronome, linguistic rhythm, oral performance

1. INTRODUCTION

Speech rhythm plays a fundamental role in any language. Yet, unless they undertake linguistic studies of their mother tongue (L1), native speakers (NSs) are usually not aware of the rhythmic patterns that they apply to their utterances, so natural it is to them. When the learning of a foreign language (L2) is involved, the difference in speech rhythms between one’s L1 and L2 is most often at the origin of a poor performance, at best impairing communicative efficiency; at worst leading to incomprehension, or even confusion—and this regardless of the phonetic or structural correction.

Throughout history, French has been one the greatest contributor to the English lexicon. It is estimated that around thirty per cent of the English lexicon is borrowed from French, and another thirty per cent is borrowed from Latin, often through French [4, 18]. In spite of this theoretical “advantage”, most francophones find it difficult to express themselves efficiently in L2 spoken English: this is mainly due to the stress-timed vs. syllable-timed dichotomy [1, 5, 6, 15]. Through this, easily recognizable words, with a similar spelling, will undergo a dramatically different phonological treatment by NSs, notably through the use of lexical stress. Then, the “advantage” usually turns out to be a handicap.

The stress-timing phenomenon is not limited solely to the words of French or Latin origin, of course: lexical stress applies to all vocabulary items at various degrees, all the time, either by its presence or its absence. In spite of its essential role, the rhythm of the English language has too often come second—if at all—to the teaching of lexical, structural and phonetic correction. Yet, rhythm is key to prosody, through which so much non-verbal fundamental information is conveyed; this is especially true for English [3, 4, 7, 9, 10, 11, 17, 19]. Prosody is the first perceptive competence of the newborn child, the earliest appearance of the infant’s oral production [6]; nonetheless, prosody (and, thus, rhythm) is ever so often the last acquired L2 competence and the eventual competence taught. In this article, a pedagogical method in which linguistic rhythm is given its foremost place is described and the results of a preliminary case study are given and discussed.

Under this approach, the tasks of rhythm acquisition and phonetic correction, though
considered complementary, are separated. Students are invited to take part in explicit rhythmical practice reading tasks under the guidance of a metronome. The basic idea is that the pendular beat of the metronome gives the expression “stress-timing” its true sense, with each main-stressed syllable in the utterance falling on a beat. It is believed that using the metronome as a time-keeping reference will lower the “cognitive cost” for the student and, thus, enhance performance. The underlying hypothesis is that prosodic production is partially independent from the phonemic-lexical-syntactic chain.

This introduction exposed the object and background of the study. Thereafter, the experimental methodology and the studied parameters are presented, followed by an analysis, interpretation and discussion of the early results of the case study. Conclusions will be drawn and the perspectives for future research, highlighted.

2. EXPERIMENTAL METHODOLOGY AND STUDIED PARAMETERS

In this second part, the parameters that will be analyzed in the third part are presented. Then, the ecological conditions in which learning tasks were implemented will be described. Finally, the testing method, from which the relevant data were gathered, will be exposed.

2.1. The parameters of lexical stress

Lexical stress parameters have been rather well defined in scientific literature, although authors do not always agree about which of these parameters is the most prominent \([4, 11, 19]\). Moreover, there can be much intra- and inter-speaker variability as to the way and degree to which these parameters are used, through the application of focal accents, notably. The parameters that have been retained in this study are listed below.

- Pitch, fundamental frequency \(F_0\) (in Hertz [Hz]). A stressed syllable usually means higher pitch.
- Intensity \(I\), in sound emission; loudness, in perception (in decibels [dB]). A stressed syllable usually means a more intense sound
- Duration \(d\), or length, or “quantity” of vowels, consonants, syllables (in milliseconds [ms]). A stressed syllable usually means either a longer or richer sound (“long” vowels \([V_{\text{long}}]\) or diphthongs \([V_{\text{diph}}]\), e.g. beat, cake), or a clearer, shorter, snappy sound in the case of “short” vowels \([V_{\text{short}}]\), e.g. bit, cell.
- Sound quality: a stressed syllable usually means well-defined vowel sounds; weak vowels \(/\acute{a}/, /\circ/ \text{ et } /\acute{u}/ \text{ [V}_{\text{weak}}\] will not appear in them, normally.

Finally, stress, not a free parameter in English, plays a fundamental role in differentiating minimal pairs, e.g. the word \(\text{contrast}\) can either be a noun or a verb; stress differentiates one from the other, as \(/k\text{kn træst}/\text{ and } /k\text{kn 'træst}/\).

In this work, for the sake of clarity, it was chosen to allocate 3 levels of stress to syllables: \(\text{St1}\), \(\text{St2}\) and \(\text{St3}\). \(\text{St1}\) refers to all the syllables by default, unless stressed otherwise; an \(\text{St1}\) syllable bears no mark and is commonly referred to as “unstressed”. \(\text{St3}\) refers to the main stress (‘), such as it appears in the International Phonetic Alphabet (IPA). \(\text{St2}\) refers to secondary stress (\(\prime\)).

2.2. Ecological conditions: the metronome and its use

The implemented and tested teaching method is based on the use of a metronome during reading tasks. The method was tested with volunteer students for ten weekly sessions of about thirty minutes each, between January and May 2012.

A typical session would consist in the elaboration of a short text on a word processor by the students, on the subject of the day (four to ten sentences). Once the text was typed, students would highlight all the \(\text{St3}\) syllables in the sentences, with possible recourse to pronunciation dictionaries according to needs. Then students’ readings were recorded in a digital-audio language laboratory a number of times, for practice and critical listening, by both student and teacher (high-quality sampling at 32 kHz). Each \(\text{St3}\) syllable had to fall on the metronome beat. A final take was sometimes made for assessment. Students were also invited to say sentences without reading them, if they knew them by heart, as well as recording them in a final take without the metronome, so as to go back to a more natural, less rigid speech delivery, but keeping the beat in mind.

The tempo of the metronome was made to vary between 60 and 80 beats-per-minute (bpm), so as to make students aware of the influence of tempo on speech delivery. Students knew, and were reminded, that NSs do not speak with such rigid metronome-like rhythm. However, one of the main interests of the method is to emphasize the acceleration that must take place when several \(\text{St1}\) or \(\text{St2}\) syllables have to fit in between two \(\text{St3}\) ones, which is a natural practice of NSs. The direct consequences of this acceleration are the NSs’ observed phenomena of vowel reduction, vowel centralization and, sometimes, vowel elision, which were also expected in the case of our
students. Digital online metronomes were used, for convenience and user-friendliness, enabling individual listening through the students’ headsets without disturbing the neighbouring students’ rhythms.

A single pronunciation model was given as a reference to the students, for the sole purpose of our study: British Standard Received Pronunciation (RP), as defined in [4], [11] or [19]. Of course, students are sensitized to other models outside of the current study (notably: General American [GA]), and variation from the models, including the modern evolution of English as a worldwide language [14], is frequently illustrated.

2.3. Construction of a corpus

A total of seventy-five students from the first author’s institute of engineering (Institut Polytechnique de Grenoble, France) volunteered to take part in the experiment. They were tested before \( T_{\text{init}} \) (early January 2012) and after \( T_{\text{term}} \) (late May 2012) the work with the metronome, as described in section 2.2, with the test detailed in section 2.4 below. The gathered results of these two instances of the same test are the basis of this study.

All students belonged to the first and second years of the institute (third and fourth year at university). The vast majority were around 20 or 21 years old. About 40% were female students and around 15% were non-francophones. Roughly 60% studied another foreign language and 20%, a third foreign language. A reference group of an extra 15 students, who did not follow the metronome method, was tested in the same conditions.

2.4. Testing the method

The same test reading-test was used initially \( (T_{\text{init}}) \) and after \( (T_{\text{term}}) \) the work with the metronome, in order to determine the contribution of the latter to the potential progress made by the students, in terms of stress, rhythm and phonetic rendering. Firstly, a subjective comparative analysis was made by their teacher (the first author); then a detailed acoustic analysis was carried out, both to highlight differences and, possibly, progress between \( T_{\text{init}} \) and \( T_{\text{term}} \).

The test consisted in the reading of 12 phonemically-balanced sentences, which included the following criteria:

- six occurrences of four \( V_{\text{long}} /\text{iw}/-\text{aw}/-\text{aw}/-\text{aw}/ \) in a pseudo-random distribution
- variation of the lexical-morphological-syntactic place of the stressed syllables
- variation of the length of the sentences

- The 12 sentences are presented in Table 1 below.

**Table 1.** The 12 sentences used in the \( T \) test. Words bearing one of the four \( V_{\text{long}} \) have been shaded. \( S_n \) stands for “sentence”.

<table>
<thead>
<tr>
<th>( S_n )</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sn01</td>
<td>A reform in higher education took place this autumn and the situation has improved at last.</td>
</tr>
<tr>
<td>Sn02</td>
<td>A recent survey has shown that more and more female teenagers opt for scientific studies.</td>
</tr>
<tr>
<td>Sn03</td>
<td>Engineering studies involve a lot of work on projects.</td>
</tr>
<tr>
<td>Sn04</td>
<td>Each new class of students is surveyed in terms of academic achievements.</td>
</tr>
<tr>
<td>Sn05</td>
<td>The record of employment of graduates coming through is very good.</td>
</tr>
<tr>
<td>Sn06</td>
<td>University authorities make concerted efforts on student well-being.</td>
</tr>
<tr>
<td>Sn07</td>
<td>Each spring, those who have a good musical level give a concert.</td>
</tr>
<tr>
<td>Sn08</td>
<td>At the start of the year, the directors project an information slide-show.</td>
</tr>
<tr>
<td>Sn09</td>
<td>This graph shows how our students’ origin and age are related, in one piece of information.</td>
</tr>
<tr>
<td>Sn10</td>
<td>The year before had been quite ordinary, but a drop is recorded this month, and it’s sharp!</td>
</tr>
<tr>
<td>Sn11</td>
<td>On average, we lose three per cent of boys per annum — that’s true!</td>
</tr>
<tr>
<td>Sn12</td>
<td>On that chart, we can see how they choose their future place of study too.</td>
</tr>
</tbody>
</table>

2.5. The case study

Analysing all the data of whole corpus in detail could take a considerable amount of time, especially if the analysis is done manually, by only one or two researchers. Before any such task is undertaken, it was considered wise to engage in a preliminary case study in order to validate the feasibility of the observation method. All the segmentation was done manually, using PRAAT software [2] and appropriate scripts to gather data.

One subject was selected among the seventy-five, as being “in the average”, both in his personal and student statuses. This subject was given the pseudonym of Jean (/ʒæn/), here a French male first name. Jean is a French student of 22 years old, who:

- has studied L2 English for 10 years, essentially at school
- visited Britain for 2 weeks during a holiday
- studies L3 Spanish.

It was decided to opt for a male student, as results due to a normally lower \( F_0 \) frequency than for girls were expected to be easier to analyze and provide clearer graphs.
3. ANALYSIS, INTERPRETATION AND DISCUSSION OF EARLY RESULTS

3.1. Jean’s teacher’s perceptive analysis

The graph of figure 1 shows stress distribution in the first sentence of the test, read by Jean in the $T_{\text{init}}$ and $T_{\text{term}}$ versions, as it was perceived by Jean’s teacher.

**Figure 5.** Stress distribution in Jean’s production of first sentence Sn01, in tests $T_{\text{init}}$ and $T_{\text{term}}$ compared to NSs’ expected stress patterns (Jean’s teacher’s perceived data).

The Stress_Norm set of bars indicates the expected Received Pronunciation (RP) stresses for the three stress levels. The dominance of the St1 type appears clearly, as it should, this being the most common type. Out of the 26 syllables, a total of seven main-stress St3-type syllables in the sentence is found, and only four St2.

In the initial Stress_Init set of bars, the situation is entirely reversed, with an overwhelming dominance of St2 stresses, patently characteristic of the regular beat of a syllable-timed French utterance, almost devoid of dynamic contrast. The Stress_Term bars show dramatic change, potentially indicating improvement, provided the stresses fall on the “correct” syllables.

3.2. Jean’s phonemic performances

This part focuses on the spectral distribution of the four selected $V_{\text{long}}$ (/iː/-/ɑː/-/ɔː/-/uː/), as illustrated in table 1 above (NB: /ɔː/ because of its neutral sound and central position in the vowel diagram was not selected). It was mainly though the classical analysis of the first two formants ($F_1$ and $F_2$, harmonic frequencies of $F_0$) that Jean’s production was assessed. $F_1$, mainly an indicator of the movement of the lower jaw, against $F_2$, an indicator of the anterior-posterior movement of the tongue, enable the positioning of vowels on such figures as the classical IPA vowel diagram [11] [17] [19].

**Figure 6.** Vowel diagrams of Jean’s production in tests $T_{\text{init}}$ and $T_{\text{term}}$ (at bottom) compared to various NSs’ reference studies, based on the 12 sentences of the test.

Figure 2 shows Jean’s isolated vowel diagrams (in the hatched thick line: $T_{\text{init}}$; in the continuous thick line: $T_{\text{term}}$ — averaged from the data of the 12 sentences) against NSs’ ones, as found in three previous example RP studies [8] [12] [13] (thin lines), and against Jean’s teacher’s own production (thick dashed line). What must be noticed here is that not only Jean’s phonemic performance leaves much room for progress, distant as it is from all of the references, but also that his $T_{\text{term}}$ production is in regression compared to the $T_{\text{init}}$ one.

3.3. Duration, pitch and intensity: practicing for contrast

In the study of Jean’s case, raw vowel and syllable durations, calculated consonant duration, vowel and syllable elasticity, instantaneous speech delivery speed and syllable acceleration, were all studied and analyzed in detail; as well as pitch and intensity. The main factor that was sought was contrast. Indeed, in order to have progress in the field of oral expression, the stress-induced contrast has to increase in some or, preferably, all of the parameters.

Figure 3 shows the example of the vowel length distribution for 23 vowels of sentence Sn10. On this graph the general contrast in the results has almost systematically increased in $T_{\text{term}}$, both upward and downward. Moreover, some improvement can also be noticed in the matching of the stress level variation and the $T_{\text{term}}$ curve.

Figure 4 enables the comparison between the mean values of $F_0$ for the two tests, for sentence Sn01. A sharp increase in the contrast is to be

---

1 Please note that the scope of the case study and the amount of analyzed data is too considerable to be entirely presented within the limits of this article. Thus, only a few examples to illustrate the approach and results are given here.
noticed here, especially in the direction of lower F₀ values in the middle part of the graph.

**Figure 3.** Vowel length distribution in Jean’s production of sentence Sn10, in tests T_init and T_term, compared to expected Stress_Norm.

![Figure 3](image)

**Figure 4.** Mean F₀ distribution in Jean’s production of sentence Sn01, in tests T_init and T_term, compared to expected Stress_Norm.

![Figure 4](image)

4. CONCLUSION AND PERSPECTIVES

The main conclusions that were drawn from the detailed study of Jean’s learning strategy were that a) the dynamics of his rhythm production was reorganized around the lexical stress, for more efficiency through a higher contrast; b) a majority of the V_long had been lengthened and the V_shor and V_weak had been shortened; c) he implemented an original intonation strategy that meant a higher contrast due to lowering F₀ more than increasing the frequency; d) the intensity contrast is in direct relation with the intrinsic value of the phonemes and shows less “prosodic freedom” than duration or pitch, thereby undergoing less fluctuation.

Through this preliminary case study, the observation methodology can be validated, as much for the experimental approach as for the analysis of data. General progress in rhythm and intonation was noticed. Still, some improvement can be made in terms of graphic representation, for example by using tones instead of the Hertz unit for F₀ values. Some trials have been made with the aim of representing stress on 3D graphs, featuring duration, pitch and intensity all at once, but they are not fully satisfactory yet. Phonetic correction, however, is rather on the decrease; this can be explained by the concentration required by the tasks. It is believed that this will be temporary and that once basic rhythmic language reflexes are acquired, this aspect will improve too.

5. REFERENCES