Proceedings of the
Phonetics Teaching and Learning Conference
UCL, London, 5–7 August 2015

Editors
Joanna Przedlacka
John Maidment
Michael Ashby
PTLC2015

Proceedings of the Phonetics Teaching and Learning Conference,
London, 5–7 August 2015

edited by

Joanna Przedlacka, John Maidment, Michael Ashby
CONTENTS
(Papers are in alphabetical order of first author’s surname)

About PTLC .................................................................................................................. 7
Scientific Committee ..................................................................................................... 9

L2 pronunciation stability: how the pronunciation of advanced EFL students is affected by a lack of explicit instruction
Gwen Brekelmans ......................................................................................................... 11

Assessment of acoustic phonetics exam questions in an EFL environment
Mercedes Cabrera-Abreu and Francisco Vizcaíno-Ortega ........................................... 17

Teaching the Prosodic Configuration of Lectures: Thematic Structure, Intonation, and L1 Interference
Marina Noelia Cantarutti .............................................................................................. 23

Self-learning Catalan phonetics. From description to application
Josefina Carrera-Sabaté, Imma Creus Bellet and Clàudia Pons Moll ............................. 29

Making your phonetics and phonology lessons interesting
Alice Yin Wa Chan ......................................................................................................... 35

The learning of English intonation by Spanish speakers in a distance education environment
Eva Estebas Vilaplana .................................................................................................... 39

Techniques to improve the perception of the English fallrise tone
Smiljana Komar ............................................................................................................ 45

Learning English phonetics: preferences and attitudes
Pekka Lintunen and Aleksi Mäkilähde ........................................................................ 51

Laboratory activities for large and online phonetics classes
Timothy Mills, Karen Pollock and Benjamin V. Tucker ............................................... 57

Teaching a phonetics course with a high student enrolment using a blended-learning format
Shawn Nissen ................................................................................................................ 63

Assessing the phonetic level in L2 vowels production with the Visuvo software: case of Czech learners acquiring contrasts between French mid vowels e/ε, ø/œ and o/ɔ
Nikola Paillereau ........................................................................................................ 67

Different designs of university courses based on phonetics in Thailand
Rungpat Roengpitya ..................................................................................................... 71
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using flipped learning to support the development of broad phonetic transcription skills among L1 and L2 English-speaking students</td>
<td>Jane Setter</td>
</tr>
<tr>
<td>Teachability and learnability of English tonicity for Japanese junior high school students</td>
<td>Yusuke Shibata, Masaki Taniguchi and Tamikazu Date</td>
</tr>
<tr>
<td>Perception of four General British vowels by Slovenian university students of English as a Foreign Language</td>
<td>Andrej Stopar</td>
</tr>
<tr>
<td>An Analysis of Japanese Junior High School Textbooks as Pronunciation Teaching Materials</td>
<td>Junko Sugimoto and Yoko Uchida</td>
</tr>
<tr>
<td>Cross-linguistic perception of English /p/-/b/ in noise by two levels of Japanese learners of English</td>
<td>Shinichi Tokuma and Won Tokuma</td>
</tr>
<tr>
<td>Phonetics instruction in L2 French: Contributions of segments, prosody and fluency to speech ratings</td>
<td>Pavel Trofimovich, Sara Kennedy and Josée Blanchet</td>
</tr>
<tr>
<td>The effects of high-variability phonetic training on Cantonese ESL learners' production of English /ɪ/-/i:/ contrast - an acoustic analysis</td>
<td>Janice Wing Sze Wong</td>
</tr>
</tbody>
</table>
About PTLC

The Phonetics Teaching and Learning Conference was set up as one of the deliverables of a project funded by the Higher Education Funding Council for England and the Department of Education of Northern Ireland. The project, which ran from 1997 to 2000, had UCL as its lead institution and partner institutions at the universities of Cambridge, Central England, Newcastle, Ulster, Westminster and York.

The first meeting of PTLC was at UCL in 1999, and inaugurated a biennial series of conferences which built upon and expanded the original successful formula. Over the first 14 years researchers from a wide range of countries presented no fewer than 138 papers, all of which were double-blind reviewed by members of a scientific committee before acceptance and publication.

These present proceedings of PTLC2015 add a further 19 papers from 12 countries. To re-emphasise the broad remit of PTLC, the conference programme included 2 invited talks on diverse topics. Eleanor Lawson (University of Glasgow) spoke on ‘Seeing Speech and Dynamic dialects – building vocal tract imaging resources for use in Phonetics teaching’, while David Deterding (University of Brunei Darussalam) gave a talk with the title ‘Teaching Pronunciation to Avoid Misunderstandings’.

PTLC clearly met—and continues to meet—an internationally-felt need. It remains the only regular forum dedicated to research in the teaching, learning and assessment of phonetics.

JOHN MAIDMENT (CHAIR)
Scientific committee PTLC2015:

Mercedes Cabrera Abreu: Universidad de Las Palmas de Gran Canaria
Michael Ashby: UCL
Patricia Ashby: University of Westminster
Joan Carles Mora Bonilla: Universitat de Barcelona
Eva Estebas: Universidad Nacional de Educación a Distancia
Ulrike Gut: University of Münster
Rachael-Anne Knight: City University
Smiljana Komar: University of Ljubljana
María Luisa García Lecumberri: University of the Basque Country
John Maidment: UCL (Chair)
Joanna Przedlacka: UCL and University of Oxford
Alex Rotatori, Università degli Studi di Tor Vergata, and Collegio IPASVI, Rome
Jane Setter: University of Reading
Hiroko Saito: Tokyo University of Foreign Studies
Gladys Saunders: University of Virginia
Masaki Taniguchi: Kochi University
J. C. Wells: UCL
L2 pronunciation stability: how the pronunciation of advanced EFL students is affected by a lack of explicit instruction

Gwen Brekelmans
Radboud University
gwen.brekel@gmail.com

ABSTRACT
Pronunciation teaching tends to take a back seat in second language teaching. Even advanced students of English still have some trouble in maintaining the advanced level of their pronunciation. This paper aims to show how stable the pronunciation of advanced students of English is when they no longer have any explicit pronunciation teaching, and when their exposure to the desired accent is minimal. The influence of a stay abroad is also investigated.

It turns out that overall, the student's pronunciation tends to remain fairly stable with only a marginal deterioration occurring after pronunciation instruction was stopped, but there are specific features that show distinct patterns of deterioration or improvement. Additionally, the pronunciation of students who had spent time abroad improved instead of stabilised.

Keywords: pronunciation, EFL, pronunciation stability, pronunciation instruction.

1. INTRODUCTION
In attaining proficiency in an L2, pronunciation is often seen as the last hurdle learners need to overcome. However, even very advanced learners of English seem to have trouble retaining an advanced level in their pronunciation. Many anecdotes were found of L2 learners indicating that they feel their pronunciation deteriorates after some time of not having had explicit pronunciation teaching, which is not an unlikely change. Both [6] and [10] have shown that even relatively limited explicit phonetic instruction can be beneficial to learner's pronunciation. Another factor in the learners' deteriorating pronunciation may be the increased use of the L1: [5] showed that the degree to which the L1 is used can be a negative influence on the accentedness with which a learner pronounces the L2.

What this paper aims to show is how the pronunciation of advanced Dutch learners of English is affected by a lack of explicit pronunciation instruction, to see whether their accent improves, deteriorates, or remains stable over time. This is examined through means of a phonetic analysis of both spontaneous and read speech produced by a group of advanced Dutch university students of English. Their speech was recorded at different moments during their studies, at which the amount of explicit pronunciation instruction they received varied.

In addition to the lack of explicit instruction possibly negatively influencing the learners' pronunciation, this research will include subquestions regarding several independent variables that might be of influence. One factor that might be of influence is exposure to the pronunciation of the people they interact with most, or the variety of English they are most exposed to; speech accommodation theory poses that speakers adapt their pronunciation to that which they are exposed to [see 4 for an overview of the theory]. Another factor that might be influential is whether the student spent time abroad in an English-speaking country, and in accordance with this, how many English-taught courses they had in the academic year they were recorded.

Taking these variables into account, this means there are essentially four ways in which the pronunciation of the advanced Dutch students might develop when explicit pronunciation instruction is stopped: it can remain stable and on the same level as it was when they still received instruction, it can improve and come to sound more native-like, it can deteriorate and become more Dutch, or it can change towards a variety of English to which the students are more exposed, i.e. American English. This paper hypothesises that overall, the students' pronunciation will become more Dutch. It is expected that going abroad and having more English-taught courses will be beneficial to having a native-like pronunciation.

2. METHOD
In order to investigate this hypothesis, the pronunciation of advanced Dutch university students of English was recorded. The students' current pronunciation was compared to their own pronunciation that had been previously recorded at pronunciation exams, at which times they did have pronunciation classes. Their overall pronunciation was analysed on the basis of a list of features as shown in section 2.3.1.
2.1 Participants

For this study, 12 native Dutch third-year students of English from Radboud University participated in the recordings (9 women and 3 men, mean age 22.1), all of whom had chosen the British (RP) pronunciation track. Four of the students had spent a semester abroad in an English-speaking country (either in England or in the Republic of Ireland). In the academic year in which they were recorded, the students did not have any pronunciation instruction or any classes focusing on oral communication skills. Up to the time of recording, the average number of English-taught courses that the students had taken in that academic year, including any taken abroad, was five (out of a total of eleven courses they had to take). This is barely half of the English-taught courses they had in previous years, when obligatory courses ranged from eight to eleven individual courses.

2.2 Materials

The materials the participants were asked to record consisted of three parts, representing three different types of speech: one text that was familiar to them through previous exams (the phonetic text ‘Arthur the Rat’), a set of 10 sentences used in the grading of the pronunciation exams focusing on particular pronunciation difficulties for Dutch students (e.g. The work in the brewery has ruined his shoes for the vowel combinations), and three open questions in which they were asked to elaborate on topics relevant to their pronunciation (their stay abroad, their classes, and their opinion on their own pronunciation). These materials were compared to the recordings of the students’ own pronunciation made as part of the pronunciation exams at the end of their first and second year at university. The structure of these exams was identical to that of the current recording, with the only difference being in the content of the tasks.

2.3 Procedure

The recordings took place in the language lab at the Radboud University. This lab is equipped with 22 PCs, and one PC on which the recording program operates. For each session, a maximum of five students were recorded at a time, to ensure the participants were not distracted by having others reading out the same text right beside them. The participants were seated at a PC running on Windows 7, which was equipped with a standard-issue headset used in the language lab. ReLANpro [11] was used to record the students, which was controlled from the experimenter’s desktop. The optimality of the recording procedure was confirmed in a pilot study in which three graduate students participated.

Beforehand, the participants were not told the goal of the study in which they were participating, but since the structure of the material was similar to their own pronunciation exams, the study’s focus on pronunciation will have been obvious to them. It was made clear to the participants that they would not be graded for anything that they did during the session, and that their data would only be used for research purposes.

The setting and procedure used in the recordings were identical to that used by the pronunciation teacher in the pronunciation exams which was when the data from the participant’s first and second year was collected. The only difference between those exams and the current recordings was the students receiving a grade for their pronunciation exams.

2.3.1 Features

The features that were used to investigate the learner’s pronunciation were chosen based on the list of features that is used to grade the student’s pronunciation exams. This list, as shown in Table 1 below, focusses on those features of RP English that have shown over the years to be the most difficult for advanced Dutch learners and to be the best indicators for students’ overall pronunciation proficiency (Kamara, personal communication).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Specific attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>vowel quality</td>
<td>/æ vs. e, o, ʌ, ɔ, əʊ, ɜː/</td>
</tr>
<tr>
<td>word and sentence stress</td>
<td>compounds</td>
</tr>
<tr>
<td>duration of voiced portions</td>
<td>fortis vs. lenis endings</td>
</tr>
<tr>
<td>consonants</td>
<td>/θ, ð/; /s,f/ vs. /z,v/</td>
</tr>
<tr>
<td>voiceless stops</td>
<td>aspiration and preglottalisation</td>
</tr>
<tr>
<td>non-rhoticity</td>
<td>US English influence</td>
</tr>
<tr>
<td>undesirable assimilations</td>
<td>regressive voicing /b,d/</td>
</tr>
<tr>
<td></td>
<td>progressive voicing /v,z/</td>
</tr>
<tr>
<td></td>
<td>intervocalic voicing</td>
</tr>
<tr>
<td>strong vs. weak forms</td>
<td>overuse of strong forms</td>
</tr>
<tr>
<td>liaison</td>
<td>pre-vowel glottal closure</td>
</tr>
<tr>
<td>intonation</td>
<td>flat intonation, ‘uptalk’</td>
</tr>
</tbody>
</table>

Table 1: Features of interest under investigation in the student’s pronunciation.
Importantly, all of these features are dealt with extensively in the pronunciation classes at the Radboud University [see 8 for the book that is used during this course]. The second column of Table 1 indicates to which issues particular attention will be paid for each of the features, as they are known to occur in Dutch learner's pronunciation. These issues will be discussed in more detail below [for an overview of the differences between the phonetics of Dutch and English, see 2].

The mentioned set of vowels is under closer investigation since they do not occur in Dutch and are not always mapped onto a correct vowel representation as a result of this; this is particularly the case for the distinction between /ɛː/ and /eː/, as both vowels tend to be mapped onto Dutch /ɛ/ [2]. The duration of the voiced portions is a known troublesome area, with Dutch learners not always being able to distinguish between word-final voiced and voiceless consonants as Dutch has final devoicing [9].

Additionally, for the consonants, those that do not occur in Dutch can cause problems (/θ, ð/), and the distinction between word-initial voiced and voiceless fricatives might result in problems depending on the native region of the speakers [7]. Since Dutch, unlike English, does not aspirate its word-initial voiceless stops nor preglottalises the word-final ones [8], these areas can be particularly troublesome. Finally, the assimilations that are used in English differ from those that occur in Dutch [3], so undesirable transfer of regressive voicing, progressive devoicing, and intervocalic voicing are things to look out for.

2.3.2 Data analysis

The participants' recordings were analysed on the basis of the feature list as presented in section 2.3.1. For each feature, the RP-like quality of the participant's current pronunciation was determined in both quality and quantity, which was expressed through means of a percentage. This percentage was then compared to their pronunciation of that feature in previous years. In the end, this percentage was converted to a grade for the student's current overall pronunciation as well as for the separate features, and this grade was compared to their previous grades to see if the general tendency was an improvement or a deterioration.

Praat [1] was used to analyse the average vowel quality and the quality and precision of the consonants, to measure the duration of the voiced portions, to check for the presence of aspiration or affrication and preglottalisation, and for undesirable voice assimilations, as well as to determine the intonation pattern that had been used. The other features were phonetically transcribed and checked for accuracy.

3. RESULTS AND DISCUSSION

Overall, as Figure 1 shows, the participants' pronunciation was slightly less good in the third year than it was in their second year, but it was better than in their first year of studying English, despite the fact that they did not receive any explicit pronunciation instruction or communication classes anymore (mean year 1: 77.1%; mean year 2 = 81.9%, mean year 3 = 81.5%). Both the difference between the first and the second year, t(863) = -28.4, p < .001 d = 0.2, and the difference between the second and the third year, t(863) = 12.7, p < .001 d = -0.02, were significant and represented small effects.

Figure 1: Average pronunciation grade for all participants, indicated per year.

However, a closer analysis paints quite a different picture for the separate features. Figure 2 shows that the individual patterns of change for all features differed quite a bit. Interestingly, the change in the features can to be categorised in three distinct patterns: the first pattern only keep increasing from year 1 to year 3 (though not necessarily massively so), the second remains relatively stable overall, and the third pattern increases from year 1 to year 2, but then decreases in year 3. Features that fall into the first category are word and sentence stress, use of strong and weak forms, liaison, and voiceless stops. The second category contains consonants and the duration of voiced portions. Finally, the third category houses most features: rhoticity, undesirable assimilations, intonation, and vowel quality.

There does not seem to be a very clear reason for a particular pattern occurring for a specific feature. Strikingly, the features that show an increase between the first and the second year but decrease in the third year, were those that started out lower than the other features; all but the rhoticity
percentage started out at or below 80%, while in the first year the other features had an RP-like percentage of 85% or more.

**Figure 2:** The overall RP-like percentage per year for each feature.

Overall, it could be concluded that the biggest change for the participant's pronunciation from the first to the second year was to be found in the quality of the vowels: in the second year, these were 74% RP-like, while in the first year these were much more consistently replaced with the closest Dutch equivalent with an RP-like quality of 62%. From the second to the third year, the RP-like quality of the vowels dropped slightly to 71%, though still remaining well above the percentage at which they started. It seemed as if the participants were trying to fine-tune their vowels but not quite succeeding. This is likely due to them receiving more exposure to different varieties of native English as well as to other Dutch students speaking accented English, and adapting their own pronunciation accordingly.

**Figure 3:** The average overall pronunciation grade for participants who had spent time abroad and those who had not.

Whether the participant had spent time abroad in an English-speaking country during their studies turned out to be of influence on their overall pronunciation grade. As can be seen in Figure 3, the average pronunciation grade for those students who did not spend time abroad went down slightly when compared to the previous year, but the average pronunciation of those students who had spent time abroad, usually in the first part of the third year, increased quite steadily. The mere fact of having to speak only English for a term seems to have been enough to enable students to improve their pronunciation.

A final note, though unrelated to the pronunciation of the students, is that the vocabulary of the students contained more Americanisms than it did in the previous pronunciation exams. This seems to indicate that they were exposed more to US English than they had been before, and consequently incorporated features of it in their own speech.

4. CONCLUSION

This paper aimed to show the effect of a lack of explicit pronunciation instruction on the pronunciation of advanced learners of L2 English. Without continued pronunciation instruction, the students' pronunciation did improve overall. Certain features turned out to be more robust than others: consonants and the duration of voiced portions remained very similar through all three years, while word and sentence stress, use of strong and weak forms, liaison, and voiceless stops steadily increased and rhoticity, undesirable assimilations, intonation, and vowel quality showed a more variable pattern of initial increase but a slight subsequent deterioration. This means that the learners' pronunciation has not stabilised in such a way that all features remain at a certain level of RP-like quality, but that the individual features seem to develop separately.

The value of pronunciation teaching turns out to be very high, even for advanced students. Additionally, not being exposed to varieties of British English as much as they were before might have been a big influence on the pronunciation of the participants: since they were exposed more to mainly American English through the media, and Dutch accented English from other Dutch natives, there was less input that matched their pronunciation model than before.

Overall, the students' pronunciation seems to not have stabilised entirely after their third year, with only those students who spent time abroad improving still, while the other students marginally deteriorated. This has confirmed the original hypothesis: it was expected that unless they had spent a term abroad, the students' pronunciation would actually deteriorate.
5. REFERENCES


ASSESSMENT OF ACOUSTIC PHONETICS EXAM QUESTIONS IN AN EFL ENVIRONMENT

Mercedes Cabrera-Abreu, Francisco Vizcaíno-Ortega

Instituto Universitario de Análisis y Aplicaciones Textuales (IATEXT), Universidad de Las Palmas de Gran Canaria
mercedes.cabrera@ulpgc.es francisco.vizcaino@ulpgc.es

ABSTRACT

This paper offers an assessment of the efficiency of three exam questions on acoustic phonetics in an EFL context (Spanish-English). The goal of these questions is twofold: (i) to find out whether students are able to establish a correlation between the articulatory description of English sounds and their acoustic manifestation both in spectrograms and in waveforms; and (ii) to confirm their capacity to conduct a contrastive acoustic analysis between the two languages. Results indicate that those questions which concentrate on English sounds only are well handled by students, whilst the only question concerning contrasts between both languages turns out to be poorly answered.

Keywords: acoustic cues; articulatory correlates; contrastive analysis; exam questions.

1. INTRODUCTION

Teaching and learning English phonetics and phonology in an EFL environment seems to be changing in the past few years. A quick look at the degrees in English studies in Spain evidences that in some universities the content courses devoted to this discipline exhibit a gradual shift towards the combination of both acoustic phonetics and articulatory phonetics, rather than the traditional articulatory description of sounds that has characterised these courses until recently, and which is still most prevalent in the majority of universities in the country.

While it is true that knowledge of the organs of speech, or of how or where sounds are produced in the oral cavity, is a fundamental part in the description of the sounds of a language, no study of speech sounds is complete without an analysis of the acoustic signals formed by the movement of the vocal organs of speakers in the production of speech. Aspects like friction, retroflexion or rounding, among others, clearly call for acoustic descriptions if they are to be properly defined. Other connected speech phenomena, like nasal assimilation, have recently been investigated resorting to acoustic information [1]. We strongly believe that the study of acoustic cues observed in both spectrograms and waveforms helps students in English phonetics and phonology to enhance their understanding of the pronunciation of the language and to correct their own. Among the benefits of acoustic analysis in practical sessions, [2] mentions that “the unnecessary and damaging division between research and practice is broken down”, and claims that the acoustic analysis performed in class, with both students’ and lecturers’ recordings subjected to immediate measurement, serves as a stimulus for students to view the renditions of their utterances objectively, which offers them the possibility to improve their production. In addition, free-of-charge download of acoustic analysis software encourages students to use it also as an out-of-class learning experience.

In English Phonetics and Phonology, a second-year obligatory course in the Degree of Modern Languages at the University of Las Palmas de Gran Canaria, the discipline-specific academic competences regarding acoustic analysis are the following: i) ability to perform both articulatory and acoustic analyses of English speech sounds; ii) ability to self-diagnose and improve one’s pronunciation of English; and iii) use of speech analysis software. The purpose of the present paper is to assess how effective our exam questions about acoustic phonetics are within the acoustic approach we pursue in the teaching / learning process. For this, in §3 we analyse the wrong answers to three questions about acoustics included in the final written exam paper of the obligatory course aforementioned.

2. COURSE METHODOLOGY

In this course, priority is given both to the acoustic description of English sounds, and to the allophonic variants of consonants only. Articulatory descriptions are also included, and care is taken that an immediate mapping is established between the observed acoustic cues and their articulatory correlates.

One of the activities which supports this learning process is the following: students choose and record
their own production of a short utterance; following the spectrographic display on the classroom screen, they are asked to describe the acoustic cues observed (for example, the ones for voice), and to provide their articulatory correlates. This activity prepares them in advance for one of the exam questions (see question 1 in §2.1).

During the course, students are also instructed about the most common pronunciation errors made by Spanish speakers of English [3]; furthermore, in weekly lab-sessions, students are invited to record, diagnose and correct their own pronunciation, and that of others.

In order to provide further assistance to the learning process through this contrastive analysis, specific video recordings were designed [4]. In one of these videos, a narrator’s voice advises the student to closely follow a set of specific instructions for the completion of question number 2 described in §2.2, and then turns to solve the question by performing the requested acoustic analysis of a particular word.

Another activity included in the learning process consists in the drawing of a stylized spectrogram for a short utterance like, for example, a beach. Here, students’ understanding of the acoustic cues and their temporal alignment with a phonetic transcription is practised (see §2.2). In the following sections, we describe and illustrate the 3 questions about acoustics included in the written exam paper, which also incorporates a phonemic transcription of English, a find the odd-one-out exercise, images of spectrograms and waveforms, as can be seen in figures 1 and 2, and finally an exercise in which students draw a spectrogram as described in §2.2. A pass in this paper constitutes an achievement of 50 out of 100 marks.

2.1. Question 1: identification of acoustic cues and brief reference to their articulatory correlates

In this question students find both the waveform and the spectrogram of a short utterance like, for instance, a belch. A set of numbered acoustic cues are highlighted in the spectrogram (figure 1), and in the waveform. The student is first required to transcribe the utterance phonetically and to align it with the spectrogram, then to label the acoustic cues, and finally to specify the articulatory correlate for each acoustic cue observed in the spectrogram. As far as the waveform is concerned, he/she is asked to align the phonetic transcription with it, and also to label the numbered acoustic cues. 13 out of 26 marks constitute a pass for this question.

Figure 1: Spectrogram for the identification of acoustic cues.

2.2. Questions 2 and 3: contrastive acoustic analysis and spectrogram drawing

In question 2 (figure 2), students are faced with two spectrograms, one that corresponds to the production of an English word by a proficient non-native speaker of English (any of the two authors, who are trained phoneticians and native speakers of Spanish), and another that illustrates the incorrect pronunciation of the same word by the same speaker, who reproduces a broad Spanish accent, only that he/she is extremely careful to incorporate a single error per word while an effort is made to keep the production of the rest of the segments unchanged. A clue pointing at the location of the error in the given word is also provided. There is no audio available, so that students are informed that the production on the left hand side of the figure belongs to an English speaker, while the one on the right belongs to a Spanish speaker with a strong Spanish accent who is trying to imitate the English speaker. Students are expected to identify and describe acoustically the error produced; they are also asked how the error should be corrected.

Figure 2: Spectrogram for a contrastive acoustic analysis. The correct production is displayed on the left, while the incorrect production is located on the right.

Figure 2 exemplifies one of the most common errors produced by Spanish speakers of English: the use of an epenthetic /ə/ at the beginning of words starting with s- for ease of pronunciation, given the phonotactic differences between the two languages.
(thus, in figure 2 step /∀στεπ/ results in /ςστεπ/). Three more typical errors are included in this question; for instance, the use of a voiceless alveolar fricative instead of a palato-alveolar fricative in the word ocean (ˊ/∀ς\v ις\v ις/ instead of /∀ς\v ις\v ις\v ις/); absence of devoicing of the voiced labio-velar approximant /w/ in the word tweet (ˊ/[∀τωι;τ]/ instead of [∀τωι;τ]); and the use of a voiced velar approximant instead of a velar plosive in the word eager ([∀t;fς] instead of [∀t;γς]). As in question 1, 50% of the marks for questions 2 and 3 constitutes a pass in each case.

Finally, in question 3 students are requested to draw a stylized spectrogram for a short utterance of the type the keys. In order to exert some control over the resulting drawing, a checklist is included that guides students to draw the acoustic manifestation of voice, place and manner for each segment. Here, what we test is students’ knowledge about the acoustic cues together with their temporal alignment with a phonetic transcription.

3. RESULTS, ANALYSIS AND DISCUSSION

In table 1 we show the numbers and percentages of pass and fail obtained by our students in these three questions.

Table 1: Students’ performance in questions 1, 2 & 3.

<table>
<thead>
<tr>
<th>Question</th>
<th>Pass</th>
<th>Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>78 (72%)</td>
<td>30 (28%)</td>
</tr>
<tr>
<td>Question 2</td>
<td>25 (23%)</td>
<td>83 (77%)</td>
</tr>
<tr>
<td>Question 3</td>
<td>86 (80%)</td>
<td>22 (20%)</td>
</tr>
</tbody>
</table>

3.1 Question 1

Here, there is a high score of pass, 72% of students. Notice that all students who failed (28%), also failed question 2, but not question 3 (14 out of 30 passed question 3). The results indicate that question 1 is well tackled by students. Furthermore, by examining in more detail the responses of students who failed this question, we notice that they perform better in the identification of cues for consonants than for vowels. As far as their articulatory correlates are concerned, however, these are either left unanswered, or are rather inaccurately referred to. This is, for instance, the case of the articulatory correlates of vowel transitions flanking a voiced plosive. Instead of a response of the type ‘this is the approach phase from the vowel to the consonant, and the release phase from the consonant into the vowel’ what we find is ‘the air is stopped in the oral cavity, and then, it goes out’, thus, missing any reference to the approach phase, and not using an adequate terminology. In relation to the analysis of the waveform, its cues are, for the most part, correctly identified, and only occasionally some students resort to terminology which is specific to the spectrogram instead of the waveform (for example, ‘vertical striations’ instead of ‘periodicity’).

3.2 Question 2

As far as question 2 is concerned, the results are reversed: there is a low percentage of students (23%) who passed, while there is a high score of fails (77%). These results exhibit the worst students’ performance.

The low percentage of pass for question 2 seems to suggest that it is rather complex for students despite the clear, detailed instructions supplied. In the following sub-sections, we describe the recurrent errors in the answers given by those students who failed this exam question.

3.2.1 Wrong acoustic cue

It is frequently the case that students identify an acoustic cue which is irrelevant for the case at stake. For instance, in the case of the word ocean the following answer is attested: ‘The consonant /Σ/ is more intense than in the case of the English speaker, because the strong random energy is darker’. The expected answer should make reference to two different frequency peaks of strong random energy corresponding to /ς/ and /Σ/, rather than to a difference in intensity.

3.2.2 Relevant but incomplete acoustic information

A token in which the acoustic information is incomplete is found in the analysis of the word eager: ‘the Spanish speaker isn’t pronouncing a plosive; that’s why we cannot see silence or the explosion bar. To correct this, he should make a stop, and decrease the amplitude and make a silence’. Although two relevant and correct acoustic cues are mentioned here, the student fails to refer to the presence of voice bar, striations and formants for the incorrect voiced velar approximant consonant that appears in the Spanish case. In addition, he/she does not suggest what the Spanish speaker may be pronouncing.

There are other exam responses in which some relevant acoustic information is indeed included, but the issue at stake is either missed altogether, or inaccurately described.
3.2.3 Labels for articulatory information only

Rather frequently, we have encountered answers in which students only offer the labels of sounds, and even though they identify the phenomenon correctly, no clear acoustic detail is included. A case in point is the following analysis of the word ocean: ‘the Spanish speaker is probably using the fricative sound /s/ when he should use the sound /Ω/, as the frequency should be lower than it is’.

On other occasions, students may add some broad articulatory information (together with sound labels), as in the following description of the word tweet: ‘the Spanish speaker doesn’t produce aspiration. He/she should let the air escape after the release of the voiceless plosive and before the vowel is pronounced’. Notice, furthermore, that this student’s response reflects a misunderstanding of the acoustic difference between aspiration and devoicing, since he/she makes explicit reference to the former when this is a case of the latter.

3.2.4 Nonsense or unanswereds

Among the 118 exam papers marked, we have found cases in which either the complete question or the analysis of a single word is left unanswered. The analysis of tweet is most frequently left out, possibly because of the following various reasons: i) the complexity of the phenomenon itself; ii) the fact that the devoicing of approximants is the first phenomenon which is introduced from the acoustic perspective in the syllabus, and students are still becoming familiar with acoustic phonetics; or iii) since it is often confused with aspiration, students may hesitate before responding, and decide not to answer at all.

Finally, overall nonsense answers are also attested, as in the following student’s description of step as /σοντεπι/ instead of /στεπι/: ‘at the beginning, the /s/ sound was not aspirated, so we can find voice bar, which means that there is vocal fold activity’. As can be seen, the student lacks basic knowledge to perform the contrastive analysis, as he/she refers to an unattested phenomenon in either language, and misses the issue at stake altogether.

In view of the low percentage of pass and of the different types of incorrect answers, we can assert that this question is indeed complex. It includes too many distinct acoustic phenomena, and presents the student with cases for which several aspects of contrastive analysis have to be borne in mind: i) identification and description of presence/absence of acoustic cues (the case of epenthesis); ii) awareness of the dissimilar acoustic cues concerning differences in the phonemic inventory within one language (the case of /σ/ vs. /Ω/ in English); and iii) awareness of allophonic variation within a single language (the case of [o] vs. [o̞] in English, and the case of [γ] vs.[f] in Spanish).

3.3 Question 3

The results for question 3, which exhibit the highest percentage of pass (80%), indicate that it is also well handled by students. Drawing a spectrogram seems to be a straightforward means for students to show their knowledge about acoustic cues. A possible reason for this high percentage of pass may be that students have resorted to the drawing of spectrograms as a technique for revision.

Despite this, however, we still encounter errors of the following type: (i) acoustic cues are misaligned with the phonetic transcription, or they are incorrectly added; for instance, a voiceless plosive is aligned well after the explosion bar, or a voice bar is included for such plosives. (ii) The expected distance between formants is not represented, or their transitions point towards the wrong direction in the frequency axis; (iii) no difference is shown between distinct types of energies: rather than representing vertical striations as vertical lines, and random energy as irregular or chaotic pencil strokes, only the latter are used for both cases. Lastly, (iv) there is a mismatch between the phonetic transcription and the spectrogram. For instance, a voiceless unaspirated velar plosive is present in the phonetic transcription, but weak random energy corresponding to +VOT is included in the spectrogram.

4. CONCLUSIONS

After examining students’ performance in the three questions, we can conclude that exam questions 1 and 3 have proved to be highly effective, for they show a high percentage of pass. The low percentage of pass for question 2 may indicate any of the following options which future research will discern: (i) this question may be too demanding, calls both for a revision of its design in terms of a reduction of cases to be analysed, and for a lesser degree of complexity; (ii) the marking system may need to be revised in terms of, for instance, assigning a higher mark to this question; and/or (iii) the teaching method may require some adjustments in order to target contrastive analysis more effectively. Despite all this, it might also be the case that this question should remain unaltered, since it also helps to discriminate between poor and excellent students, while questions 1 and 3 do not.
5. REFERENCES


1 We are grateful to two anonymous reviewers for their comments which have helped us to improve the present paper.
TEACHING THE PROSODIC CONFIGURATION OF LECTURES: THEMATIC STRUCTURE, INTONATION AND L1 INTERFERENCE

Cantarutti, Marina Noelia

ISP Joaquín V González – ENSLV Sofía B de Spangenberg – Profesorado del Consudec
marina.cantarutti@bue.edu.ar

ABSTRACT
The lecture is the chosen speech genre in the academic world for the distribution of knowledge. Among other defining characteristics, lectures are organised along a particular thematic structure that signals to the audience how the information status and its progression is to be interpreted; and the instructions that this structuring provides are either accompanied or disambiguated through prosodic choices in the systems of tonality, tonicity and tone. This paper reviews some basic research on the prosodic configuration of lectures in General British and Riverplate Spanish, and reports a pilot experience in training Spanish-speaking teacher trainees in the production of typical prosodic patterns in the lecture genre, with a special focus on thematic structure, and with a secondary aim of reducing patterns of interference from L1 through the training of metaphonological awareness and self-regulatory skills.

Keywords: prosody, interference, metaphonological awareness, SFL, Discourse Intonation

1. INTRODUCTION

“Teacher talk” is made up of different speech genres distinguishable not only by a particular schematic structure of stages and conversational routines, but also through intonation patterns making up a particular prosodic configuration. One of the speech genres closely related to teaching is the lecture, and its textual and prosodic configuration as a rule could be said to mirror its “hybrid” status between the oral and written medium, as Biber and Conrad [1] state. The lecture genre could be said to be clearly organized through the staged presentation of “what is already known or predictable, and what is new or unpredictable” [12], which are always in tension. This interaction of the New and the Given generally builds a pattern of “periodicity” [18] of information waves embedded into a particular thematic structure. This organisation consists in the marking of “points of departure” that frame the message and prepare the audience to apprehend its contents. The role of prosody in this textual function of informational organisation is essential, and for future teachers of English from a different L1 background, an appropriate handling of the intonation system that makes the perception of “periodicity” possible is an important asset, as it is through prosody that the speakers aid the audience’s understanding of what is “foreground”, and what, “background” information [13], what information “drives discourse forward” and what acts like a “consolidatory loop” in the development of the text [3]. The presentation of information distribution relies mostly in the system of tonality, and the signaling of information status and relevance is effected through tonicity and tone.

The present paper will review some of the characteristics of the prosodic configuration of thematic structure in lectures in English and in Riverplate Spanish, from the contributions of Systemic Functional Linguistics (SFL) and Discourse Intonation (DI), to predict possible cases of L1 interference. In addition, a report will be made on the preliminary results of a pilot experience in the teaching of prosody and thematic structure in lectures to a group of native Spanish-speaking teacher trainees in Buenos Aires, Argentina. It is believed that the training of metaphonological awareness [10, 23] of the connection between patterns of thematic structure and prosody could help non-native speakers of English self-regulate [22] and fine-tune their L2 intonational choices in their delivery of the lecture genre.

2. THE TEXTUAL AND PROSODIC CONFIGURATION OF LECTURES

From a Systemic Functional Linguistic perspective, genre is “a staged, goal-oriented, purposeful activity in which speakers engage as members of our culture” (Martin, 1984 as cited in [9]) and it has a particular set of manifestations in the context of situation. Socially speaking, the lecture is one of the institutional means through which knowledge is distributed, subject to the on-line planning restrictions of the spoken medium and the lexical density of spoken mode, thus its denomination of “hybrid” in [1]. Plus, the lecture event promotes a generally assymetrical relationship, with a discourse dominant lecturer who is in a position to claim discourse dominance and control. Thus, lecturers can provide the audience with procedural instructions to
allow their interlocutors to “surf smoothly on the
peaks, and slide through the troughs on the
information flow” (Bowcher, 2003 in [24]).

2.1. Thematic Structure

One of the forms in which linguistic choices can create “texture” [11] is what is known as thematic structure, that is, the sequencing of initial “points of departure” [9] followed by their rhemes, elements that may act as message cores. Lectures tend to display a clear and patterned structure of clause simplexes and complexes made up of topical themes that present a starting point from which the rest of the section of the message is to be interpreted. The presence of topical themes of unmarked types facilitates the introduction of the participants -topics or people generally marked by pronouns or noun phrases-, whereas marked types introduce circumstantial information -place, time, instruments, conditions, presented by phrases or clauses. These thematic choices are generally associated, though not necessarily, with Given information, and it is prosody that will help establish their degree of salience and status. Even though themes can also present other types of information and be of equative, predicative, interpersonal and textual kinds, in this paper, our concern is with topical themes, given their connection to the New-Given patterns of periodicity.

2.2. Prosodic Configuration of Lectures in General British and Riverplate Spanish

These characteristics of the lecture genre invite a staged presentation of the content to an audience with different degrees of previous knowledge. From the point of view of prosody, the distribution of content into “units of information as the speaker perceives them” [19] is accomplished through the intonation system of tonality, a “phonological way of representing pieces of information” [20] into IPs (i.e. intonation phrases). The second system, tonicity, has a major role to play in the signaling of information as Given - generally deaccented and out of focus in English- or New -generally made prominent and placed within the focus domain. The informational notion of Focus [8] contributes to this distinction, with all-New domains for Broad Focus and partly-New domains for Narrow Focus, particularly when Given or contrastive information is present. O’Grady [18] redefines the notion of Given by establishing three forms of Givenness: a) recoverable/predictable, b) non-salient (interest) b) shared knowledge. In other words, the Given is not necessarily shared or retrievable information, as it could also constitute information conferred a low degree of relevance by the speaker. In general, this type of information is either deaccented- if presented towards the end of the IP-, or accented, but marked as Given by the choices of tone or key.

In Riverplate Spanish, however, speakers were found to re-accent Given information in a large number of occasions, according to a recent study conducted [6]. The tendency of Spanish towards a non-plastic accentuation [8] is reported by Labastía [14] to be manifested by the general placement of the sentence accent towards the last lexical item regardless of the informational focus domain, except in cases of correction or contrast. Nevertheless, in the recent study mentioned [6], only cases of polar contrast presented deaccentuation of Given information, while other cases of contrast displayed (re-)accentuation of repeated information.

The system of tone in English contributes to the marking of information as New or Shared, from a transactional perspective, with falling tones presenting the message as completing an “increment” and acting as information which is “so far from an unnegotiated set” [3], and rising tones acting as “loops” of information which can be retrieved from the area of common ground. The choice of rising tone helps to mark the information as “background” [13], that is, presenting lower relevance in dependent structures, whereas falling tones generally indicate that the information is “foregrounded” [12] or treated as “major” [19].

Two pieces of research on a small corpus of lectures in General British [6,7] established that most marked themes and also a considerable amount of non-pronoun unmarked topical themes are generally chunked apart from their rhemes. A vast majority of the thematic elements scrutinized in [6] were found to be produced with a fall-rising tone, and were thus projected as “loops”, though presented as “highlighted themes” [19], versus rises, which could be said to merely claim “dominance through continuity”[3]. A limited number of thematic elements carried falling “citation contours as topic markers” [21], and were thus treated as “increments” in themselves. The use of a level tone on several thematic elements marked a shift towards oblique orientation, a text-focussing practice related to creating rhetorical effects, quoting, or declamation.

For Riverplate Spanish, Labastía et al [15] establish that background information is marked through rising tones, and they remark that rise-fall-rises and fall-rises “postpone the evaluation of relevance”. Our
study of a small corpus of lectures in Riverplate Spanish [6] reveals, however, that apart from rising and fall-rising tones, it is in fact the rise-fall tone (including its “truncated” manifestation, Granato, 2005 and Gurlekian, 2010 as cited in [6]) that is more frequently applied onto Given elements in thematic position, in keeping with the findings of Le Gac [16]. The rising versus falling distinction between “loops” and “increments” described for English does not always appear to match the patterns found for Spanish, as the difference between both varieties appears to be both phonetic –allotonic- and phonological –use of rise-falls as “referring” in Spanish. On the other hand, the use of the oblique level tone seems to fulfil the same function as in English, though the typical manifestation of this contour is slightly different.

The differences outlined between General British and Riverplate Spanish predict possible patterns of L1 interference in the accentuation and tone choices of thematic elements by Spanish-speaking users of English as an L2: final Given elements may be found to be accented, and the presence of a rise-fall tone or Spanish allotonic varieties of the rise or the fall-rise may also be applied where General British fall-rises or rises may be expected. The development of metaphonological awareness of these differences for Spanish speakers of English as L2 could facilitate the appropriation of the L2 phonological and phonetic features characterizing the delivery of lecture genres in English.

3. THEMATIC STRUCTURE AND PROSODY IN LECTURES: THE STUDY

3.1. The Context

The study hereby reported was not initially conceived of as a formal piece of research, but rather as a common practice in the Laboratory III course, aimed at allowing students to measure their progress by confronting their reading of the same text at the beginning and the end of the course. The results of these pre- and post-tests are also internally used to reflect upon the tools and procedures used for instruction, and their degrees of effectiveness, particularly in terms of the use of DI and SFL as metalinguistic frameworks. This section will describe the pilot experience with the intention that later studies be formally carried out.

The Laboratory III group selected was made up of ten third year teacher trainees with some previous knowledge of DI. During the eight-month course, the students were trained on the prosodic configuration of different teaching speech genres, analysing different textual and interpersonal linguistic choices from SFL and DI perspectives. Special attention during the training was placed on the role of thematic structure as a means of organizing discourse, and on the development of students’ self-monitoring [22] skills in making appropriate tonality, tonicity and tone choices on these preparatory thematic elements to mark patterns of periodicity. The sequencing of activities during the course generally started from initial collaborative analyses of sample lectures, reflection and imitation of model lectures, followed by controlled practice through the reading aloud of new lecture transcripts, and culminating in freer practice in the production of mini-lectures, where students’ metaphonological awareness and self-regulatory skills were ultimately tested.

3.2 The Test

The abovementioned group was given a short diagnostic task, consisting in the reading aloud of a short lecture extract retrieved from the British Council Professional Podcasts collection [5]. The passage was recorded after some considerable time for preparation, though students were not allowed to write any marks on the text, to ensure a more natural rendering. The test was repeated, with the same passage and characteristics, at the end of the course.

The selected lecture extract contained 23 thematic elements possibly requiring their own IP because of their length and relevance, and 16 of these were marked and unmarked topical themes. 9 of those thematic elements were to be interpreted as being in Narrow Focus and presenting an early nucleus because of either a) the presence of Given information through verbatim or paraphrasing practices b) their contrastive constitution, organising the lecture into three main axes: Past, Present and Future perspectives to life. The text selected allowed for a limited number of options and thus, it constituted a controlled environment to test the students’ metaphonological awareness and self-regulatory habits regarding the recognition and production of appropriate L2 intonation patterns on thematic and rhematic elements, at least on a written script, towards a later application onto spontaneous, unscripted practice.

The production of each thematic element was assessed by the teacher via both impressionistic and acoustic techniques including the use of PRAAT [2], and the analysis was coded onto a table, based on three criteria, thus:
3.3 Discussion and Results

The following tendencies were found in the students’ prosodic treatment of topical themes in the tests:

<table>
<thead>
<tr>
<th>Table 1: Tendencies identified in the pre- and post-tests. Note: percentages represent tokens. Shaded areas mark the post-test results.</th>
</tr>
</thead>
</table>
| Tonicity choices also show levels of improvement, though there a considerable number of students who failed to recognize Given information and contrastive focus in both tests. Even though it is common in English to re-accent Given information, the characteristics of the lecture selected invited an exploitation of the contrasts to provide the right procedural instructions [14]. In fact, it was through these thematic contrasts and progression that this particular lecture was structured, and many of the choices made by students were not successful in acknowledging this.

Tonicity choices for marked topical themes remained constant, and those on unmarked topical themes, which show reduced values on the post-test, are consistent with previous studies [6, 7] of English lectures, in that many lower relevance thematic elements are appended to their rhemes, prosodically speaking. This latter factor would appear to reveal some sort of underlying assessment on the part of the trainees of the communicative hierarchy of the thematic elements to the progression of the lecture read aloud.

4. CONCLUSIONS

In spite of the fact that this study cannot be considered a formal piece of experimental research, the results reveal that metaphonological training from DI and SFL frameworks on the intonation of thematic structure has helped this group of students assess the communicative value of different topical themes and make appropriate L2 intonation choices, at least for this stage of controlled, reading-aloud practice of lecture transcripts. First, students appeared to favour the use of rising tones as a means of presenting the information as either “background” or as a “loop”. They also chose level tones for items of lesser relevance, and turned to oblique orientation accordingly, or even to theme+rheme conflated IPs. In most cases, students also turned to English qualities, thus dropping their Spanish rise-falls in thematic position.

The recognition of patterns of deaccented Given information in English, however, requires further attention tasks so that the choices of tone so accurately made by the students actually match those patterns of periodicity of the Given and the New that tonicity choices can so clearly signal to the audience.

It is believed that a formal study also including assessment tasks of unscripted lecture delivery, could shed light on the degree to which metaphonological awareness and self-regulation have
been achieved by these trainees in less controlled environments.

5. REFERENCES

Self-learning Catalan phonetics. From description to application

Josefina Carrera-Sabaté¹ Imma Creus Beller² Claudia Pons Moll¹
¹Universitat de Barcelona and ²Universitat de Lleida
jcarrera@ub.edu; icreus@xtec.cat; claudia.pons@ub.edu

ABSTRACT

Internet resources for self-learning in the field of phonetics are quite common for majority languages such as English, French and Spanish. In the last decades, though, Internet utilities with the same purpose have also been developed for minority languages such as Catalan. It is the case of some websites that offer material to learn about the different levels of phonetic description (articulation, transmission, and perception) and to practice the pronunciation of Catalan, both of the discrete sounds and of the intonation. The goal of this paper is to describe and compare two available websites: “Els sons del català” (SC) [The sounds of Catalan] [8] and “Guies de pronunciació del català” (GPC) [Catalan pronunciation guides] [9]. In this paper, we explain how these websites work, and how they complement each other. While the former has been designed to provide a complete description of Catalan sounds, the latter has been designed to be applied in the process of learning and teaching Catalan pronunciation, especially by and to non-native Catalan speakers.

Keywords: Catalan phonetics, Catalan pronunciation, self-learning, L2 pronunciation learning

1. INTRODUCTION

The freedom of the individuals to learn and self-manage, to find what they need and to adapt it to their environment is becoming a common practice. A practice that has its origin in the 70’s, when CALL (Computer Assisted Language Learning) are developed and used for the first time. The advantages of this new system were soon made obvious: a greater motivation of the students, an encouragement of the process of individual learning and a dynamic and interactive access to the information (see [12]).

With respect to Catalan, in 2006 a new website, “Els sons del català” (henceforth SC), was created with the aim of providing resources to learn Catalan phonetics autonomously. The SC (available at the address http://www.ub.edu/sonscatala and with free-access) was developed by a team of professors from the University of Barcelona and with the help of advisors from different Catalan universities.

The website tried to fill the existent gap with respect to the materials devoted to self-learning phonetics in other languages (see [13]). Since then, the website has been enlarged, improved and in constant renovation. It has been adapted, for instance, to the new technological needs, such as its access through tablets and cellular phones, or the possibility of self-managing the contents. The SC is essentially aimed at a) assisting in teaching and learning the mechanisms of production, transmission and perception of Catalan sounds; b) helping Catalan graduate and undergraduate students of those degrees that have phonetics and phonology as object of study or as a working means (i.e. Philology, Speech Therapy, Linguistics, Translation, Communication, Journalism, among others) in learning both disciplines autonomously, in accordance with the European Higher Education Area guidelines; c) providing Internet support materials and resources to those professionals working in fields such as language advising, correction, and speech therapy (for more information, see [3]).

Ultimately, the SC offers visual information, with different levels of complexity, about different aspects of Catalan phonetics. In this respect, it has been demonstrated that this type of information works more efficiently with adult learners than with non-adult learners, and it is especially useful for acquiring self-consciousness about the position of the articulatory organs [11]. The disadvantage of using just this kind of system is that it does not contain information about the coarticulation of sounds or the communicative context in which the sounds are used.

In this respect, it is well known that speakers who have to learn a non-mother language are faced with
a series of difficulties, most of them concerning the differences between the nature of their mother language and the language they have to learn. In fact, two factors become essential when determining the problems that L2 learners experience: 1) the differences between L1 and L2 ([10]) and 2) the relevance of the intonation ([6]).

In the case of Catalan, there are some phonetic and phonological contrasting studies, which mainly compare Catalan with Spanish, English and German. The non-Catalan speaking university students who want to work on their Catalan pronunciation in an autonomous and self-regulated way have very few opportunities to do so and the resources at their disposal are certainly limited (there are few proposals in [1], [2], [7], [14], [16]).

The experience in the elaboration of the website SC and the new needs in the university and in the society in general lead the group of teaching innovation FONCAT (Departament de Filologia Catalana, UB) to initiate in 2010 the development of another website, currently addressed to those peninsular Spanish speakers who need or want to learn the pronunciation of Catalan.

The website, called Guies de pronunciació per a no catalanoparlants (henceforth GPC), is also free access and it is available at the address http://www.ub.edu/guiesdepronunciacio/el-web. It has been developed in collaboration with the Universitat Pompeu Fabra (UPF), the Institut del Teatre de Barcelona (IdT) and stuff from the Universitat de Lleida (UdL).

The aim of the website (see [9]) is to work with itineraries biased by the L1 language of the speakers/learners and to promote the acquisition of Catalan intonation patterns to best achieve the Catalan pronunciation.

These two websites, therefore, give an answer to the different needs of society so that Catalan phonetics can be worked on autonomously, either from a theoretical or from an applied point of view. Behind these websites, apparently simple, there is a great effort to assemble and reproduce the phonetic contents in a clear and transparent way.

The big challenge of these websites is to transmit the complexity of phonetics avoiding an excessive simplification of the contents, which would distance the users from the reality. This is why the websites contain real images along with graphic representations. Next we describe and compare the features and utilities of both websites.

2. DESCRIPTION OF THE SC AND GPC

2.1. SC: structure, appearance and contents

The website is structured around a tool bar menu, placed at the top and containing the sections Home, Phonemic charts, Team, Glossary, Links, News, and Funding, and it is available in two languages, Catalan and English.

2.1.1. User-friendly interactive Phonemic Chart

The sounds are presented through a user-friendly interactive Phonemic Chart (containing pop-up windows with explanations about the origin of the symbols, the articulation of the sounds, the meaning and uses of the main diacritics and suprasegmentals, a description of the manner and place of articulation, etc).

The Phonemic Chart has two levels of surfing: the one displayed when the cursor is placed on the symbol and the one that it is displayed when the user clicks on the symbol. When the cursor is placed on a phonetic symbol, a pop-window appears, that contains a description of that sound’s articulatory features as well as the origin of the corresponding phonetic symbol. When the user clicks on a phonetic symbol, a plain window appears that contains three main frames, which correspond to the moments of production and transmission of sounds: a) a video frame, b) a diagram frame, and c) an oscillogram and spectrogram frame (Fig. 1). These main frames, at the same time, are surrounded by further links and additional material.

Figure 1: Plain window of a phonetic symbol.

2.1.2. Additional information about the sounds: Palatograms and DMRI and Parts of the vocal tract

Just below the articulatory diagram in the middle there is a link to the section Palatograms and Dynamic Magnetic Resonance Images (DMRI). The former represent the degree of contact (represented in purple) between the tongue and the palate in the articulation of those sounds that involve these two organs, and the latter are a faithful reproduction of the movements of the articulators in the production of sounds.
Figure 2: Palatograms and Dynamic Magnetic Resonance Images (DMRI).

In the top right part of each of these windows, there is the section Parts of the vocal tract, which show the labels that those receive in an interactive way: placing the cursor on the labels in the left the different parts become highlighted and can be easily identified and recognized (Fig. 3). Active articulators are pink in the labels, whereas passive ones are green. The relevant cavities (oral, pharyngeal, and nasal) in the articulation of sounds are also highlighted when the cursor is placed on the respective label.

Figure 3: Parts of the vocal tract.

2.1.3. Additional resources: Comparison of sounds

Finally clicking on the Comparative sound chart, placed at the top right part of the main Phonemic chart (corresponding to each dialectal variety), the user accesses to a webpage that includes three Flash animated articulatory diagrams by which the user can choose and simultaneously view the articulatory diagrams of three different consonant or vowel sounds in order to compare them.

2.2. GPC: structure, appearance and contents

The website is structured around a tool bar menu, placed at the left and containing the sections Home, Team, Before starting, Languages, Intonation, Links, Faq and News.

2.2.1. Languages section

The Languages section of the website is organized through tables containing the sounds of Catalan, which can be compared to those of other languages, for example those of Spanish. From this comparison, different groups of sounds are classified: 1) those that are common in both languages; 2) those which are exclusive of one of the languages, and 3) those that are shared by both languages but which do not have the same distribution (see Fig. 4). These three possibilities are depicted by an analogy to traffic lights, in the same order: green-colored sounds exist in both languages. Red-colored sounds are those that do not exist in the itinerary language. Amber-colored sounds are those that could be problematic for a non-Catalan speaker (since these sounds appear in different contexts, etc.). The blue-colored sounds are the ones that do not exist in Catalan. Above each sounds, there is a short list of words that contain the sounds, so that learners can associate the phonetic and the orthographic symbols.

Figure 4: Language section: Spanish-Catalan.

For each of the sounds in groups 2 and 3, the user can find a variety of audio and video activities (with regular and slow speed, and organized according to the degree of complexity) that consist in sensitization (listening exercises), explanation, imitation (speaking exercises), practice activities and communicative activities (see Fig. 5). Of special interest is the tool which allows the user to record himself in order to contrast his pronunciation with the prototypical one. In this respect, we are working now in the possibility of automating this tool.

Figure 5: Listening self-correctable exercises.

A collection of Catalan folklore material is included in the practice activities of the site in order to contextualize each problematic sound in a wider context and assimilate it in a more easy and natural way. The Catalan expressions and sayings, and the traditional songs and stories in Catalan try to help the speaker/learner know about different aspects of
Catalan culture. In addition, following the lines of European projects such as the “European Music Portfolio: A Creative Way into Languages”, in which they show how language learning can be benefited from music, the traditional Catalan songs included in this website have been harmonized and arranged following current musical tendencies so that the learning of sounds is more attractive to the speaker (see Fig. 6).

Figure 6: Catalan song to work [z].

2.2.2. Intonation section

Besides this material and given the proved importance of intonation in the process of learning a foreign language, practical exercises about intonation have also been added to the website. These exercises, which are being conducted by a team from the UPF, are organized according to the different sentence modalities found in Catalan, and include the communicative contexts by means of audiovisual material, intonation contours, repetition exercises, and also exercises of association between intonation modalities and the communicative context.

Figure 7: Intonation section.

Figure 8: Intonation: questions.

2.2.3. Before starting section

A few exercises on auditive sensitization were introduced in one of the beginning stages “before starting the project”. These exercises, designed by a professor of the IdT, have as a goal to prepare the users of the website using techniques that relax the body, so that the learning process is more successful (see [17]).

Figure 9: Auditive sensitization exercises

2.3. SC and GPC: application, targets and monitoring

The website SC has been applied with excellent results in the undergraduate courses devoted to Catalan phonetics and phonology in various Catalan universities. The specific results for its application to undergraduate Catalan courses in the academic years 2010/2011 and 2013/2014 can be looked up in [4] and [15], respectively. A Google Analytics account was created since the creation of the website, and the reports show an impressive influx of users: this year 2015 (Jan-June), for instance, the website has been visited by 11,866 users with a total of 17,245 sessions (15,049 from Spain, 775 from US, 114 from Germany, 105 from France, and the rest from other countries such as UK, China, Japan, Russia, Switzerland, Brazil, and Italy). For now, the user’s needs and queries are dealt non-automatically, although we are working on a mechanical tool to monitor its usage and performance.

The website GPC, on the other hand, has been exploded more recently, also with promising results, among students of the Catalan oral courses (Communication and Teaching degrees). Thanks to a continuous work with the website, the students have shown a clear improvement in the pronunciation of specific sounds (see [5]). The next academic year, an experimental action is going to be implemented: the website will be taught to be used as a tool to train, not only learning skills, but also teaching skills following a peer learning approach.

3. CONCLUSIONS

Learning phonetics and the pronunciation of a language is a process just as diverse as the learners. This is why a variety of input materials (visual, sonorous, audiovisual) and a variety of approaches to the subject are both necessary. Together these tools may help in presenting a possibly dry subject in an attractive, easily accessible manner. Herein, we have presented two websites, which are aimed at modernizing and updating the existing materials devoted to learning Catalan phonetics and at the
same time encouraging the process of self-learning in a dynamic, sequenced and multidimensional way. As seen, whereas the website SC is theoretically-sided, in that it is mainly intended to provide the Catalan native students with the necessary information to learn the phonetics of Catalan sounds, and therefore to help them to improve their knowledge about this subject, the website GPC has a more applied perspective in that it is mainly intended to assist non-Catalan native students in their performance of Catalan pronunciation, which, last but not least, is an increasing need among the Catalan society.

4. REFERENCES


5. ACKNOWLEDGEMENTS

This research is supported by several grants from Universitat de Barcelona, Generalitat de Catalunya, Ministerio de Educación, Recercaixa and Universitat Pompeu Fabra.
MAKING YOUR PHONETICS AND PHONOLOGY LESSONS INTERESTING

Alice Yin Wa Chan
City University of Hong Kong
enalice@cityu.edu.hk

ABSTRACT

This paper discusses ways to make the learning of phonetics and phonology interesting to ESL students. Using analogies drawn from students’ daily experiences and cracking jokes related to students’ daily encounters are good ways of arousing students’ interests and sustaining their attention. Evoking some laughter in class can enhance teaching effectiveness without stifling the academic value of a phonetics and phonology lesson.

Keywords: learning of phonetics, analogies, authentic materials, laughter

1. INTRODUCTION

A lot of students think that the learning of phonetics and phonology is boring because of the many theories and technical terms that they need to remember and/or understand. The introduction of the vocal tract chart and the various articulatory organs may even turn a phonetics lesson into a biology-like lesson, which many linguistics students are sick of. How to make a phonetics and phonology lesson interesting is, thus, a difficult task for many linguistics teachers. This paper discusses ways which teachers can use to make the learning of phonetics and phonology appealing to ESL learners.

2. USING ANALOGIES

One way of arousing learners’ interest in learning phonetics and phonology is the use of analogies drawn from their daily or previous learning experiences. One advantage of using analogies is that students can understand abstract concepts more easily and remember them more deeply.

2.1. Distinguishing one speech sound from another

One of the first few topics of phonetics teaching are often devoted to the notion that the shapes and sizes of the vocal tract are different for different speech sounds because of the differences in their sound qualities. This topic is often taught by teachers showing different vocal tract diagrams for different sounds, such as /i:/ having a high tongue and /æ:/ having a low tongue. However, students may wonder why it is necessary to know about the lip shape, tongue height etc. in order to learn the pronunciation of a particular English sound, or why it is necessary to observe the differences in these dimensions when distinguishing between different English speech sounds. Before the technical concepts of tongue height, lip shape, etc. are taught, it is important that students be made aware of the effects of the differences in the sizes or shapes of the vocal tract on the quality of the speech sounds thus produced.

In order to arouse students’ interest and consolidate their understanding, ESL teachers can use analogies from students’ daily life experiences. An analogy can be made by the banging of drums. The sound that we will get by banging a large, round drum is different from the sound that we will get by banging a small, round drum, because the size and shape of the vessel containing the air that is moving makes a difference in the sound produced. However, do teachers need to bring a drum, or different drums, to class as teaching aids? That is of course possible, but not economical, as it is not convenient for teachers to carry a number of drums to class. In my opinion, teachers can help students appreciate the effects of vocal tract size and shape on the quality of speech sounds without having to put excessive efforts in the preparation or carrying of specific teaching materials.

Many students are used to bringing a water bottle to class. These water bottles can be of different sizes, different shapes or even different textures (hard plastics, soft plastics, metal etc.). Some may have long necks, some short necks, and some round bottoms, etc. They are analogous to the vocal tract, with the former being vessels of liquid and the latter a vessel of air. Before introducing the production of speech sounds, teachers can engage students in a discovery process by asking some of them to bang their water bottles. They should be strategic in their selection of participating students. A few students having the same kind of bottles can be selected first and requested to bang their bottles in front of the whole class, one after another. When the selected students were banging their bottles, teachers should focus the class’ attention on the qualities of the
sounds produced and ask students whether the sounds they heard were different or the same. It is more than expected that students will find the sounds very similar, if not exactly the same. A few other students having different kinds of water bottles will then be selected and requested to bang their bottles, also one after another. The same question can be asked after this second round of bottles banging, and this time students are expected to be able to say that the sounds they heard were all different. The ability to notice the differences between the sounds is not enough. Students should also be encouraged to describe the differences between the sounds heard (e.g. some sounds were higher in pitch, some were lower in pitch, some were heavier, etc.) and to suggest the reasons for the differences. After students’ successful grasp of the concepts, teachers can then show the analogy between the water bottles and the vocal tract and relate the production of banging sounds to the production of speech sounds. Vocal tract charts showing the different sizes and shapes of the vocal organs can then be used for demonstration.

One advantage of this analogy is that students will be engaged in some sort of physical activities without the need to know any technical concepts such as those in physics. Teachers only need to alert students to the effects of the different shapes and sizes of the water bottles on the qualities of the sounds produced. The experiments can be done very easily without consuming a lot of time. The materials needed for the experiments are also very handy, so teachers do not need to spend too much time preparing for suitable teaching aids.

2.2. Introducing the concept of allophones

The concept of allophones, the different realizations of a phoneme, (Crystal [2]), is very abstract. Learners are often unaware of the subtle differences between the different allophones of a phoneme, such as the raising of the front of the tongue for the clear [l] (as in laugh) and the raising of the back of the tongue for the dark [l] (as in dull). Even harder to notice are the effects of the phonetic environment on the realizations of a phoneme, such as the effects of the lip shape of the preceding vowel on the realization of the /l/ sound in feel and fool (i.e. The [f] sound in the former has slight lip-spreading in anticipation of the vowel /æ/ whereas the [f] sound in the latter has slight lip-rounding in anticipation of the vowel /uː/). Helping students understand that the different realizations of a phoneme are its allophones and that the different allophones belong to the same phoneme is not an easy task, so using analogies from learners’ daily experiences is again an effective strategy.

A good analogy can be drawn from the different ways of writing English letters (Chan [1]). All students know that there are different ways of writing a letter, so, for example, no matter whether we write A, a, or α, the letter written is still the first letter of the English alphabet (Knowles [3], Roach [4]). Teachers can ask different students to write the letter ‘a’ on the board. It is more than expected that the written versions are all different, albeit to different degrees, from each other. Teachers can then draw an analogy between the different ways of writing an English letter and the different ways of pronouncing an English phoneme, such as the different ways to pronounce the lateral /l/.

Similar analogy can be drawn from the writing of English letters in the discussion of the effects of phonetic environments. When writing a cursive ‘a’ after, say, ‘b’, ‘c’ and ‘q’, certain modifications will be made to the ‘a’ to accommodate the previous letters, in such a way that the ‘a’ after ‘b’ often begins at a higher level than after ‘c’ and ‘q’, and that the connecting line after ‘q’ often has a steeper slope than the connecting line after ‘c’ although in both cases ‘a’ often begins at a lower level than when following ‘b’ (Chan [1]). The following diagram illustrates this:

\[ \text{ba} \quad \text{ca} \quad \text{qa} \]

These kinds of modifications to the writing of a letter as a result of the letters around it are analogous to the influence of phonetic environments on the pronunciation of a sound and the resultant allophones (Chan [1]).

Using such analogies in explaining the concept of allophones is useful not just because the concepts can be introduced by reference to something that students are very familiar with or something that they often encounter in their daily lives, but also because the analogies themselves provide students with a visual representation of an abstract concept, thus facilitating internalization of knowledge.

3. EVOKING LAUGHTER

Not all phonetics and/or phonology concepts can be explained using analogies. Another way to arouse students’ interest and sustain students’ attention is to relate the topic under discussion to students’ daily encounters and evoke some laughter in class.
3.1. Classifications of vowels

As is well-known, vowels can be classified using three different dimensions, namely, tongue height, the shape of the mouth, and frontness and backness of the vowel. Some laughter can be evoked in the discussions of the shape of the mouth and the tongue height. With the advent of advanced technology and the use of smartphones, Instagram and Facebook, many students like to take pictures of themselves. Girls like to adopt different poses, such as protruding their lips to show cuteness, rounding their lips to show unexpectedness, grinning to show delight, opening their mouths widely to show surprise, or clenching their fists and putting their fists in front of their spread lips when trembling. Teachers can ask some students to demonstrate to the whole class what they (or girls) will do to take pictures of themselves when they want to show the different emotions mentioned above. To make the class more fun, teachers can even ask students to take pictures of those poses instantly using their smartphones and show to the whole class by means of a visualizer or other available equipment in the lecture hall. These activities will definitely create a lot of laughter in class. Teachers can then grasp the chance to invite students to describe the characteristics of the mouths in adopting those poses and relate the characteristics to the production of different vowels and the classifications of vowels (e.g. protruding the lips to show cuteness has a characteristic similar to the production of /ɑː/ (see figure 1); rounding the lips to show unexpectedness has a characteristic similar to the production of /ɔː/ (see figure 2); spreading the lips to show delight has a characteristic similar to the production of /ɪ/; opening the mouth to show surprise has a characteristic similar to the production of /ɑː/, etc.)

![Figure 1: protruding the lips](image1)

![Figure 2: rounding the lips](image2)

The photo-taking activity suggested above is not merely a fun-evoking activity. It demonstrates to students clearly that the shape of our articulatory organs differs in the production of different sounds. After engaging in the activity, students should be able to understand the concepts more easily and remember the concepts more deeply.

3.2. Minimal pairs and minimal sets

Another example of phonetics and phonology concept that can be taught by reference to students’ life experiences is the concept of minimal pairs. In Hong Kong, youngsters like to share their (romantic) relationship statuses with their peers by the use of a letter-number notation, such as A0, O1, C2, etc. The letter on the left is the current relationship status, with A being available (i.e. not in a romantic relationship), O being occupied (i.e. in a romantic relationship), and C being complicated (i.e. in a complicated romantic relationship). The letter on the right is the number of relationships they have had so far. Another notation, namely OBA (meaning occupied but available), is used to describe playboys and playgirls who are engaged in a (some) current romantic relationship(s) but are still available for other potential relationships. These notations can be used in minimal pairs identification exercises which require students to identify from a word list words which are in a minimal pair relationship with others. When a word list, say bit, bat, bid, cat, key, kit, is given to students, they are expected to be able to identify that bit is in a minimal pair relationship with both bat and bid, although the latter are not a minimal pair themselves; and while cat and kit are a minimal pair, key is not in a minimal pair relationship with any other words on the list. When going through the exercise and checking the answers with students, teachers can make use of the notations mentioned above to describe the situation. For example, the relationship status of bit is OBA (i.e. although it is already in a minimal pair relationship with bat (being occupied), it treats itself as if it is available and thus engage in another minimal pair relationship with bid); the relationship status of cat and that of kit are both O1 (i.e. They are in a minimal pair relationship with each other, and they have only one relationship, so they are both occupied once); the relationship status of key is A0 (i.e. It is not in a relationship with any word, so it is available for future potential relationships), etc. In a word, minimal pair relationships are made analogous to romantic relationships, and a word that is in a minimal pair relationship with another word (or other words) is made analogous to person engaging in a romantic relationship with another person (or other persons).
The above description will again evoke some laughter among students, as the notations used are exactly those they use in their everyday chats. They will also be thrilled in learning that the everyday language that they use can be employed by a serious academic to describe serious and academic issues such as a phonology concept.

4. CONCLUSIONS

The author of this paper has tried the above methods in her phonetics and phonology class for a number of years. In mid-semester and end-of-semester teaching evaluations, her students often responded warmly with positive feedback on her teaching, commenting that her lessons were enjoyable and that the teacher’s teaching skills were commendable. The teacher herself also enjoyed the lessons and sometimes burst into laughter herself while making those jokes and analogies. Of course it is impossible to evoke laughter for each and every topic of phonetics and phonology, but it is possible to bring some form of amusement to most lessons. To many university lecturers, teaching is seen as a serious and academic task. University lecturers are academics, so they want/need to maintain an academic image. In the authors’ opinion, however, phonetics and phonology teachers (or ESL teachers) should learn not to explain linguistic complexities just using technical terms. It is more meaningful and cost-effective to make their lessons interesting and full of laughter without stifling its academic value.

It should be noted, however, that the author has no intention of suggesting teaching techniques which avoid the use of linguistic terms altogether or which trivialize phonetics and phonology teaching and learning. The strategies suggested in this paper are all meant to be a starting point for the respective topic. More serious discussions of the technicalities involved in each topic should follow the fun-evoking activities. As suggested in 2.1, after the introduction of how one sound is distinguished from another, detailed discussions of the sizes and shapes of the vocal tract and the positions of the various articulators in producing different sounds are indispensable. Teachers should also take the background of their students into consideration in their adoption of teaching strategies, as some students are more ready to grasp purely technical and rather abstract explanations than others. To more advanced learners who have had knowledge of acoustic phonetics, the use of, for example, spectrograms and/or waveforms may provide clearer and even more precise visual representations of the differences between the allophones of the same phoneme than the analogies suggested here. How many fun-evoking activities and analogies are to be used in the classroom, therefore, depends very much on the individual differences of the learner group that a teacher is facing. With the basic principles in mind, teachers could also experiment on other phonetics and phonology topics to discover the applicability of the suggested techniques to other topics for other groups of learners/learners from other cultures.

5. REFERENCES


---

1 The following are some examples of comments students gave in end-of-semester teaching evaluations: “The things are completely new to me and some of my friends, and those are boring. But Alice really makes everything interesting. It's fun to have her class” “The teacher is very interesting and well prepared. She tried her very best to make the 3-hour lecture engaging and she succeeded” “Professor Chan conducts the lesson in a humorous way which makes boring linguistic lessons become enjoyable”.

---
THE LEARNING OF ENGLISH INTONATION BY SPANISH SPEAKERS IN A DISTANCE EDUCATION ENVIRONMENT

Eva Estebas-Vilaplana
Universidad Nacional de Educación a Distancia (Spain) eestebas@flog.uned.es

ABSTRACT
This paper explores the advantages and disadvantages of using TL_ToBI to teach English intonation to Spanish native speakers in a distance learning environment. A study was carried out to compare the production of English intonation patterns by two groups of Spanish students of English phonetics, each of them trained with a different methodology: 1) the British School of intonation and 2) TL_ToBI. Each speaker recorded the following sentence types: declarative, yes-no question, wh-question, and imperative, produced both with neutral and marked readings. The results showed that those speakers trained with TL_ToBI produced more native-like intonation patterns than those instructed with the other model. These findings suggest that a system based on tonal targets and their association to the metrical structure has more benefits for the acquisition of the L2 prosody in a distance learning setting.

Keywords: L2 prosody, English intonation, TL_ToBI, British School, distance learning.

1. INTRODUCTION
Studies on second language acquisition have shown that exposure alone does not suffice to acquire a second language (L2) and that in order to achieve a successful command of the L2 it is crucial to be aware of its main features ([7], [8]). This is known as the Noticing Hypothesis which states that learners cannot learn the features of a language unless they notice them and that that noticing is the essential starting point for acquisition. Thus, L2 students learn about the things that they pay attention to and are aware of but do not process much about the things they do not notice. Even though some studies ([9], [10]) have shown that the Noticing Hypothesis does not always apply and that L2 students, especially children, may learn a second language without any explicit instruction, other studies ([6]) have demonstrated that, in adulthood, learning is enhanced by attention and that there are areas, such as the learning of the L2 pronunciation, that are very difficult to acquire unless students are made aware of the differences and similarities between their native language and the L2. This is especially true for languages whose phonological systems are far apart, such as English and Spanish.

In this study, we examine whether awareness and metalinguistic knowledge can help adult Spanish students of English to acquire some basic intonation patterns of the L2 in a distance learning setting. Two intonation teaching models are compared: the British School and TL_ToBI.

1.1. The British School
The British School (BS) of intonation has been the model mostly used for the teaching of English prosody over the past seventy years ([3], [11]). It is characterized by: 1) a configurational analysis of pitch (i.e. each intonation phrase is divided into a nuclear configuration and a pre-nuclear configuration), 2) the lack of boundary tones, since the last accent accounts for the pitch movement till the end of the contour, 3) a tonal inventory based on pitch trajectories (e.g. rise, fall, fall-rise, etc.), and 4) a notation system with diacritics that outline the pitch movement. An example is presented in (1) for a declarative sentence.

(1) It’s rather exˈpensive.

1.2. TL_ToBI
TL_ToBI (ToBI for Teaching and Learning) ([4]) was designed to teach English intonation in a distance learning setting. It is an adaptation of the ToBI system ([11]) and it combines some of the main tenets of the original ToBI framework with some of the ideas of the BS. Its main characteristics are: 1) the incorporation of visual aids (graphs) along with the audio material to outline the F0 movements, 2) three levels of tonal description (L, H, M), 3) one level of phrasing signalled by a boundary tone, 4) (only) monotonal pitch accents (L, H, M), 5) monotonal (L, H, M) and bitonal boundary tones (HL, LH) and 6) the signalling of the metrical information on the segmental string (by means of *) rather than on the tone itself. An example is presented in (2).
The equivalences between the notation conventions of TL_ToBI and those of the BS of intonation are included in Table 1.

Table 1: TL_ToBI tonal combinations and their correspondences with the conventions of the British School.

<table>
<thead>
<tr>
<th>BS</th>
<th>TL_ToBI</th>
<th>BS</th>
<th>TL_ToBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear tones and diacritics</td>
<td>Pitch accents</td>
<td>Boundary tones</td>
<td>Pre-nuclear tones and diacritics</td>
</tr>
<tr>
<td>*High-fall</td>
<td>H</td>
<td>L</td>
<td>*High</td>
</tr>
<tr>
<td>Low-fall</td>
<td>M</td>
<td>L</td>
<td>Low</td>
</tr>
<tr>
<td>*High-rise</td>
<td>M</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>Low-rise</td>
<td>L</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>*Mid-level</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>*Fall-rise</td>
<td>H</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>*Rise-fall</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

### 2. THE EXPERIMENT

In this paper we compare the productions of English intonation by two groups of Spanish speakers who have been trained with different teaching methodologies (the BS and TL_ToBI), so as to examine which model is more beneficial to acquire some basic English intonation patterns in a distance learning environment.

#### 2.1. Informants

The informants were 746 Spanish students of English phonetics registered in the *English Pronunciation* course offered at the Distance Learning University in Spain (UNED), as part of the Degree in English Studies, in the academic years 2012-2013 and 2013-2014. This course involves 150 hours of distance learning tuition distributed in four months. The aim of the course is to give students guidelines, tools and methodological resources for the autonomous learning of English pronunciation. The syllabus includes: 1) English Vowels, 2) English Consonants, 3) Combination of Sounds and Connected Speech Processes, 4) Stress and Rhythm, and 5) Focus and Intonation. The students have at least a B1+ level of English according to the CEFR. 546 students were registered in 2012-2013 and 657 in 2013-2014. Only the students that sat for the final oral exam took part in this experiment, i.e. 334 students in 2012-2013 and 412 in 2013-2014.

#### 2.2. Teaching methodology and corpus

The students in year 2012-2013 learned English intonation according to the tenets of the BS, whereas those in year 2013-2014 followed the TL_ToBI methodology. The two models could not be tested simultaneously due to the limitations of the distance learning platforms since each course only allows for one virtual space where all students share the same teaching materials and participate in the same forums. The corpus of sentences and the teaching procedures presented in the two academic years were exactly the same, the only difference being the teaching methodology. For each course, students were presented with some theoretical explanations on English intonation following the premises of either the BS or the TL_ToBI model. The same examples and ear-training exercises, annotated with the conventions of one of the two frameworks, were included in the two courses.

Four sentence types were presented to the students: declaratives (D), yes-no questions (YN), wh-questions (WH), and imperatives (I), produced both with neutral (N) and marked intonation (M). An example of each sentence type is presented in Table 2 with the TL_ToBI annotation.

Table 2: English sentence types and examples.

<table>
<thead>
<tr>
<th>Neutral declarative</th>
<th>She can *come on *Monday.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked declarative (with reservations)</td>
<td>She can *come on *Monday.</td>
</tr>
<tr>
<td>Neutral yes-no question</td>
<td>*Are you *coming?</td>
</tr>
<tr>
<td>Marked yes-no question (confirmation seeking)</td>
<td>*Are you *coming?</td>
</tr>
<tr>
<td>Neutral wh-question</td>
<td>*What’s your *name?</td>
</tr>
<tr>
<td>Marked wh-question (polite)</td>
<td>*What’s your *name?</td>
</tr>
<tr>
<td>Neutral imperative</td>
<td>*Don’t *touch it.</td>
</tr>
<tr>
<td>Marked imperative (warning)</td>
<td>*Don’t *touch it.</td>
</tr>
</tbody>
</table>

All sentences were made of one pre-nuclear pitch accent followed by a nuclear configuration which contained one pitch accent and a boundary tone. Each sentence type had the same segmental string for the neutral and the marked readings and they only differed in the intonation patterns. For each sentence type students listened to four examples presented in pairs (neutral and marked). The ear-training exercises consisted of three listening activities of ten entries each in which the students had to identify, discriminate and reproduce the intonation patterns.
2.3. Experimental procedure

The gathering of the data was carried out as part of the course final oral exam. This exam is recorded by means of the E-oral application, which is a software designed at UNED to do oral exams on-line ([5]). Students need to have a computer with a microphone and a webcam. Students can see the questions on the screen and they are asked to read and record a series of phonetic contrasts. Students can repeat the recording a maximum of three times. The oral test is divided in three parts: 1) the production of English vowels, consonants and clusters, 2) the production of English stress, rhythm, focus and intonation patterns, and 3) the reading of a phonetic transcription. For this study, only the data on intonation was analysed.

There were seven models of oral tests which appeared randomly in the application. The types of questions were exactly the same in all models, the only difference being the data included in each question so that students could not know the precise words or sentences they were expected to produce. With respect to the intonation patterns, sentences were presented as in the teaching materials, that is, they were accompanied by the prosodic annotations used in each intonation model. In addition, for each sentence, the expected nuance (e.g. polite, warning, etc.) was also indicated. The recording of the final oral exam was done at the end of the course. Each student had to produce eight sentences as the ones presented in Table 2. A total of 5968 sentences were assessed, 2672 in year 2012-2013 and 3296 in year 2013-2014.

3. DATA ANALYSIS

The students’ productions in the oral test were analysed perceptually by three members of the teaching team, including the main lecturer and two backup tutors, one of them English. For each sentence type they had to decide whether the productions of the Spanish students followed the expected English pitch patterns according to three positions: 1) the nuclear configuration, 2) the pre-nuclear accent, and 3) the whole tone. Before assessing the productions, tutors were instructed about the expected differences and similarities between English and Spanish intonation for the various sentence types, so as they could better perceive cases of possible interference. The pitch configurations presented to the tutors for the two languages are exhibited in Table 3, which shows that pre-nuclear accent is different for all sentence types. The nuclear configuration, on the other hand, is similar in neutral and marked yes-no questions, in marked wh-questions and in marked imperatives.

For each production, the assessments of the three teachers were contrasted. In case of disagreement, sentences were analysed again by the three judges.

Table 3: Differences (≠) and similarities (=) in the pitch patterns of the pre-nuclear and nuclear configurations of various E(nglish) and S(panish) sentences. > indicates a peak in the post-tonic syllable.

<table>
<thead>
<tr>
<th>Sentence types</th>
<th>Language</th>
<th>Pre-nuclear</th>
<th>Nuclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-D</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>H&gt;</td>
<td>≠</td>
</tr>
<tr>
<td>M-D</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>H&gt;</td>
<td>≠</td>
</tr>
<tr>
<td>N-YN</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>L&gt;</td>
<td>=</td>
</tr>
<tr>
<td>M-YN</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>L&gt;</td>
<td>=</td>
</tr>
<tr>
<td>N-WH</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>H&gt;</td>
<td>≠</td>
</tr>
<tr>
<td>M-WH</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>H&gt;</td>
<td>=</td>
</tr>
<tr>
<td>N-I</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>H&gt;</td>
<td>=</td>
</tr>
<tr>
<td>M-I</td>
<td>E</td>
<td>H</td>
<td>≠</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>H&gt;</td>
<td>=</td>
</tr>
</tbody>
</table>

4. RESULTS AND DISCUSSION

The percentages of expected English intonation patterns produced by Spanish students after being trained with one of the two models (BS or TL_ToBI) are presented in Table 4 for the whole tune as well as for the nuclear and the pre-nuclear configurations. The results show that in all cases students trained with TL_ToBI produced more native-like intonation patterns than those instructed with the BS model. Two-tailed t-tests comparing the data showed significant differences in all cases (p<0.01).

Table 4: Percentages (%) of expected English pitch patterns produced by Spanish students trained with BS and TL-ToBI.

<table>
<thead>
<tr>
<th></th>
<th>Whole tune</th>
<th>Nuclear C.</th>
<th>Pre-nuclear C.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS TL-ToBI</td>
<td>52.2</td>
<td>65.1</td>
<td>73.1</td>
</tr>
<tr>
<td>BS TL-ToBI</td>
<td>53.3</td>
<td>67.1</td>
<td></td>
</tr>
<tr>
<td>BS TL-ToBI</td>
<td>53.3</td>
<td>67.1</td>
<td></td>
</tr>
</tbody>
</table>

The percentages of expected English intonation patterns according to sentence type are presented in Figure 1 for the overall tune, in Figure 2 for the nuclear configuration, and in Figure 3 for the pre-nuclear accent. For all variables and for all sentence types, the results show a higher percentage of native-like pitch patterns for those students trained with TL_ToBI. Two-tailed t-tests presented significant differences in all cases (p<0.01). These findings suggest that a system based on visual aids and the association of tonal targets with the metrical structure...
has more benefits in the acquisition of L2 prosody in a distance learning environment.

Furthermore, the percentage of native-like pitch patterns increases for the nuclear configuration, which shows better results than those of the pre-nuclear accent and the whole tune for all sentence types. The comparison between the results of the pre-nuclear accent and those of the whole tune presents very similar patterns. Two-paired t-tests showed non-significant differences in these cases (p>0.01), suggesting that the cause of a foreign intonation is often derived from the unexpected production of the pre-nuclear configuration. In all Figures, the highest percentage of native-like pitch patterns is obtained in questions, with the highest score (over 80%) in the N-YN questions nuclear configuration. This might be explained by the general association of questions with rising patterns.

Figure 1. Percentage of English intonation patterns produced by Spanish speakers trained with the BS and TL_ToBI.

Despite the better results in the L2 intonation patterns of those students trained with TL_ToBI, more work needs to be done to further improve their productions as well as to present them with a wider range of tunes and nuances. Two measures will be taken into account in the following years: 1) to make students aware of the intonation patterns of the L1 along with those of the L2, following the TL_ToBI methodology (i.e. with perceptual and visual aids), and 2) to teach students to assess their own productions with the help of a program which displays the F0 patterns, such as Praat ([2]). Given the importance of self-assessment in a distance education environment, the comparison between the L1 and L2 intonation patterns and the visualization of the student’s own productions can be a further step for the autonomous learning of the L2 prosody.

5. CONCLUSIONS

This paper has shown that awareness of the L2 intonation patterns enhances its learning and acquisition. Two intonation teaching models (the British School and TL_ToBI) were compared to examine which of them is more beneficial to acquire English prosody in a distance learning environment. The results show that students instructed with TL_ToBI produced more native-like intonation patterns, suggesting that a system based on tonal targets and their association to the metrical structure is more helpful for self-tuition.

6. REFERENCES


TECHNIQUES TO IMPROVE THE PERCEPTION OF THE ENGLISH FALL–RISE TONE

Smiljana Komar

University of Ljubljana, Faculty of Arts, English Department
smiljana.komar@ff.uni-lj.si

ABSTRACT
This paper presents the techniques for developing and improving the perception of the English fall–rise tone in L1 speakers of Slovene for whom this is a new pitch movement to acquire in spontaneous speech.

I propose that the teaching begins with a pre–training recognition test, is followed by an exchange of different perception exercises and ends with a post–training recognition test, the purpose of which is to evaluate the students’ perception of intonation, the fall–rise tone in particular.

The comparison of results of pre–training and post–training recognition tests of the fall, rise and fall–rise tones, which I carried out with a group of second–year university students of English, proves that complementing auditory stimuli with visual feed–back is an important step in developing intonation awareness in adult L2 learners.

Keywords: English intonation, fall–rise tone, perception.

1. INTRODUCTION

Acquiring new intonation patterns is often more difficult than learning new sounds. Speakers use intonation more or less automatically, or at least without much conscious planning. Even when asked to think about the intonation, they are often unable to describe the pitch patterns they use. In other words, speakers are not aware of intonation until it goes against their expectations. And this often happens in a foreign language. L1 speakers have significant problems perceiving and repeating L2 pitch movements, even those that are similar to their L1 patterns.

There are very few studies which deal with the acquisition of either L1 or L2 intonation. Cruttenden [2: 300–301] claims that although English babies mimic intonation already in the pre–linguistic babbling, they acquire the more subtle usage of intonation only at the age of ten or eleven. Little is also known about the acquisition of L2 intonation by adults except that it is the most difficult and the very last feature to master.

The relationship between the production and the perception of speech has been more thoroughly investigated. Chun [1: 90], however, suggests that more research is required in the perception of intonation in L2, and that the issues of whether correct perception precedes correct production has to be addressed. Cruttenden [2: 348] proposes that the auditory approach, which gives priority to perception, should be used in intonation teaching.

I believe that developing intonation awareness only by exposure to and imitation of different pitch patterns is futile unless it is supported by a step–by–step explanation of how to listen to intonation and when to use a particular pitch pattern. Only then it can be expected that intonation perception, production and confidence in using it in spontaneous speech will improve.

This paper presents the techniques I use in teaching English intonation to university students, whose L1 is Slovene and are enrolled in the 2nd year of BA studies at the English Department at the Faculty of Arts, University of Ljubljana. The intonation course consists of 30 teaching hours and spreads throughout the whole year.

The points I am going to cover in this paper are the following:
- Pre–training recognition test of three tones (fall, rise, fall–rise)
- Sensitization training
- Post–training recognition test

2. PRE–TRAINING RECOGNITION TEST

One of the first tasks that every teacher of English intonation is faced with is to find out how the students perceive different pitch movements. Often the teacher will find out that the students perceive the pitch movements opposite to the way in which they are actually uttered. For example, the falling pitch movement is often heard as rising, and vice versa. One of the reasons for this is that they confuse the pitch height with the pitch movement. Recognition of the English fall–rise tone is particularly difficult for L1 Slovene speakers since the tone is alien to them.

I regularly carry out recognition tests at the beginning of the second year when the students have no previous knowledge of English intonation. In October 2013, 47 second–year students took part in the recognition test. They were supposed to choose
among three tones: the fall, the rise and the fall–rise. The nuclear syllables were underlined and instructions were given how to mark the tones.

The test consisted of two tasks:

- Recognition of tones in 15 monosyllabic words: \texttt{now, /good, /first, /six, /Spain, grilled, /stay, /true, /soon, /why, /bad, /part, /come, /John's, /France}

- Recognition of tones in 15 intonation phrases (IP): 1) You're \texttt{/right}! || 2) I've re \texttt{paired the windows.} || 3) It's \texttt{Jim, I think.} || 4) \texttt{Was she hurt?} || 5) I \texttt{think so.} || 6) \texttt{Guns don't kill people.} || 7) You were \texttt{there.} || 8) Most of it. || 9) \texttt{Did he finish it?} || 10) \texttt{Are you sure then?} || 11) Pretty \texttt{reasonable.} || 12) \texttt{Would you like to join us?} || 13) So \texttt{what was the trouble.} || 14) Is \texttt{fruit expensive in this country?} || 15) Well \texttt{apples are.}

The distribution of tones was random, but with an equal number of occurrences per task. The selected words and IPs were from Wells [6] and recorded by a trained speaker. It was a conscious decision to read all falling tones as high falls, and all rising tones as low rises. Students heard each word or IP three times.

2.1. Results

As expected, the recognition of tones in monosyllabic words was least problematic which the students confirmed in a short discussion following the testing. All three tones were perceived correctly by more than half of the students.

More problematic was the recognition of tones in IPs where the participants generally confused the pitch height with the pitch movement. This was especially the case when the falling tone was preceded by a low pre–nuclear segment and the rising tone by a high pre–nuclear segment. This stepping up or down in pitch was perceived as the rising or falling pitch movement, respectively.

As for the recognition of the fall–rise tone in IPs, I found out that more than 75% of students recognized it as a simple rise and not as a two–way pitch movement.

3. SENSITIZATION TRAINING

As discussed in [4], recognition of intonation improves to a significant extent if the students are first explained the basic rules of intonation, such as tonality, tonicity and the structure of an IP. Equipped with theoretical knowledge, they can proceed to ear–training exercises the purpose of which is to sensitize the students to different tones and tunes.

The sensitization training consists of:

- Exposure to auditory stimuli
- Upgrading auditory stimuli with visual feedback
- Identification of all three tones in texts.

3.1. Exposure to auditory stimuli

Since the fall–rise tone is a compound tone which consists of a two–way pitch movement, first falling and then rising, it is necessary to begin the sensitization process with the exposure to auditory stimuli of the fall and the rise separately. As these two tones also occur in the Slovene language, parallel examples from L1 may be useful. The sensitization begins with monosyllabic words, is followed by IPs where the tone occurs on the first syllable and is followed by a longer tail, and ends with IPs with broad focus tonicity where the nuclear syllable occurs at the end of the IP. Students are encouraged to listen carefully to the end of the utterance before they decide on the pitch movement. Perception of each tone is first practiced individually, but in the end the two tones are practiced interchangeably in two–turn dialogues.

Once the students have successfully mastered the distinction between the fall and the rise, the fall–rise tone is introduced. This tone is new to Slovene speakers, so a detailed explanation of its realization is necessary, especially in cases when the tone is followed by a longer tail. First, the students are exposed to the fall–rise realized in different words and phrases. In order to illustrate the difference between the fall–rise, the rise and the fall, the students are then exposed to the three tones interchangeably.

3.2. Upgrading auditory stimuli with visual feedback

Study by de Bot [3] has shown that audio–visual feedback is more effective than auditory. Thus it is highly recommendable to upgrade the auditory stimulus with the visual one. This can be done in two steps:

- First, by using the tadpole transcription which represents the speech melody by means of larger or smaller black dots between two lines showing the stressed and accented syllables and the pitch movement of the whole IP;
- Second, by using pitch tracks generated by SFS/WASP [5] which provide audio and visual stimuli simultaneously.
In order to test the recognition of all three tones in a longer text, I decided to add an additional task in which the students listened to a text divided into 15 IPs with marked nuclear syllables:

1) \textit{Welcome} | 2) to \textit{Bellamy’s Restaurant, Ladies and Gentlemen}! | 3) \textit{I’m your \textit{waiter} this evening} | 4) and I’d \textit{like} to go through the \textit{menu} with you. | 5) The \textit{first} course (6) offers a wide \textit{choice} of starters. | 7) I’d \textit{particularly} recommend (8) the \textit{angels on horseback}, (9) the \textit{pumpkin soup} (10) or the \textit{celery soup}. | 11) For the \textit{main} course (12) we have \textit{steak, lamb or fish}, (13) or also a \textit{vegetarian alternative}. | 14) I be\textit{lieve} the \textit{rumpsteak} (15) is \textit{particularly} good tonight. ||

The text and the recording were from Wells [6]. The students listened to the text three times.

4.1. Results

The results presented in graphs 1–3 show a significant improvement in the recognition of all three tones in monosyllabic words.

Graph 1: Recognition of falls in monosyllabic words in pre-training (left column) and post-training (right column) tests.

The recognition of the fall seems to have improved most in words \textit{now, Spain, why} and \textit{John’s}, whereas in the word \textit{soon} the improvement is slight – some students still confuse the fall with the rise.

Graph 2: Recognition of rises in monosyllabic words in pre-training (left column) and post-training (right column) tests.

The recognition of the rise shows also a good deal of improvement, mainly in the sense that the students no longer confuse it with the fall. Instead, they tend to
The recognition of the fall–rise tone has also improved, and quite a good number of students (around 40%) seem to distinguish between a simple rise and the fall–rise.

Graph 7 shows the results of the recognition of all three tones in a text.

Graph 7: Recognition of the three tones in a text.

The recognition of the fall–rise in the text, where it appeared in 7 IPs out of 15, has also improved. But it remains problematic in those IPs when the nuclear syllable is followed by a longer tail: 4 syllables in IP 3 and 6 syllables in IP 7. In these two IPs the fall–rise has been recognized correctly by 35 % of the students, but misperceived as a fall by 25 % and as a rise by 40 %. It can be concluded that the longer the span between the falling and the rising movement, the higher the possibility of its misperception as a simple rise.

The rising tone, which appeared in the text only twice (IPs 8 and 12), has been correctly recognized by 75% of students.

The recognition of the falling tone is quite good, but the success of the recognition depends on the structure of the IP: if there is a high pre–nuclear segment (i.e. head) preceding the falling tone, the recognition of the fall is worse than in those IPs where there is no high head. Thus the recognition of the fall in IPs 2, 4 and 6, which all contain a high head, is between 52% and 75 %, whereas the percentage of correct recognitions in IPs 10, 13 and 15, which contain only one pitch prominent syllable (the nucleus), is between 79% and 82 %.

5. CONCLUSION

In the paper I discuss the techniques to develop and improve the perception and recognition of intonation, in particular the fall–rise tone, in L2 adult learners whose L1 is Slovene. I propose that auditory stimuli should be upgraded by visual ones in the form of the
traditional tadpole transcription, as well as pitch tracks. This audio–visual approach needs to be accompanied by a detailed description of the pitch movements and supported by a large amount of practical examples, as well as the basic theoretical principles underlying English intonation.

6. REFERENCES

LEARNING ENGLISH PHONETICS: PREFERENCES AND ATTITUDES

Pekka Lintunen and Aleksi Mäkilähde

University of Turku, Finland
pekka.lintunen@utu.fi; aleksi.makilahde@utu.fi

ABSTRACT
Phonetics is a commonly taught subject at university level. It may be an explicit part of the curriculum or an implicit element in, for example, foreign language pronunciation teaching. However, there are very few studies on the teaching and learning of phonetics. Moreover, the learner perspective has rarely been investigated although learner opinions and attitudes affect the learning outcomes.

The purpose of this study was to investigate learner attitudes towards phonetics and preferences in the broad themes covered in phonetics. The data were collected with a questionnaire from university students (n=133) of English at the end of a lecture series on English phonetics. According to the results, accent differences was the most preferred theme. Suprasegmental issues and practical phonetic transcription were preferred to segmental issues. Students’ attitudes were positive: phonetics was considered challenging, interesting and useful. Students also thought that phonetics teaching had improved their pronunciation of English.

Keywords: phonetics, EFL, attitudes.

1. INTRODUCTION

Phonetic knowledge is part of the curriculum in many fields. In foreign language programmes, the role of phonetics is to deepen the students’ knowledge of the target language and to support their overall pronunciation skills. This may be realized, for example, by using phonetic notation in text books or by explaining the articulation of challenging target language sounds. Often the practical side, successful L2 (second language) pronunciation skills, is seen as the goal and measure of successful phonetics teaching. However, at university level, curricular aims also include understanding the basic concepts and principles of linguistics, including phonetics. Therefore, in university level foreign language programmes, the role of phonetics courses can be twofold: on the one hand, strengthening the practical language skills of L2 learners and, on the other, providing theoretical information on the target language studied. Especially in the former case, phonetics courses often include or are complemented by practical pronunciation exercises.

It has been noticed that the practical side alone is not enough for future language teachers, for example, but information about the system and processes behind L2 pronunciation are also needed [5, 7, 8, see also 2]. In this area, phonetics is essential. Ashby [3] says in summary that foreign language teachers need “a (good) grasp of articulatory phonetics; a well-trained ear; knowledge of the phonology […] of both the mother tongue(s) […] and the target language”. In other words, they need both theoretical knowledge and practical skills.

Despite the fact that phonetics is taught in many educational programmes, there are surprisingly few studies on the teaching and learning of phonetics. The contents of curricula [e.g. 9] and the methods, resources and technologies used in phonetics teaching have received some attention [e.g. 1, 10, 13, 14], mostly in conference proceedings. What seems to be missing currently is the learner perspective on phonetics. As shown in, for example, L2 acquisition studies, learners’ attitudes, opinions and learning motivation affect the learning outcomes [11]. Therefore, the learner perspective towards phonetics learning should be examined.

Recently, the role of phonetics and phonetics pedagogy has been discussed by Ashby & Ashby [4], who draw a distinction between phonetics as a discipline, often chosen as the field of study by learners, and a tool used when teaching other subjects, such as the pronunciation of foreign languages. In the latter case, some phonetic knowledge is implicitly or explicitly required by the programme. They also emphasise that there is very little research done on phonetics from a pedagogic perspective, whereas the focus has relatively often been on the teaching of pronunciation, and this is what the term ‘phonetics pedagogy’ often refers to. Ashby & Ashby further divide the role of phonetics in university programmes into pure elective, pure obligatory, applied implicit and applied explicit. In their terms, our study focused on learners who were exposed to phonetic information during a pure obligatory phonetics course or as an explicit applied element during pronunciation training. Phonetics will also be important as an applied, mostly implicit aspect in EFL (English as a foreign language) classrooms during the students’ later professional careers.
The purpose of the present study was to focus on a course on English phonetics offered to first-year university students of English as a foreign language in Finland. We were interested in students’ preferences on the topics covered and students’ attitudes towards phonetics.

This study had two research questions:
1. Which themes were the most and least favoured during a course on English Phonetics?
2. What are student attitudes to phonetics? More specifically, how did the students perceive phonetics as a topic to study, its importance for their prospective professions, and did learning phonetics, in their opinion, have a positive effect on their pronunciation skills?

2. METHODOLOGY

The data were collected with a questionnaire at the end of a course on English phonetics at a Finnish university. The students (n=133) answered on a five-step Likert-scale (strongly agree – strongly disagree) whether they agreed with the 19 given statements or not. Furthermore, the contents of the lecture series were divided into 11 categories, and the students were asked to choose their three favourite themes discussed.

The lecture series was obligatory for first-year university students of English and consisted of eleven 90-minute lectures and a final written exam. Students answered the questions before the final exam. The students were also taking a lecture course taught in Finnish on basic phonetics, which strengthened their knowledge of phonetic terminology and introduced the discipline as a whole. According to the curriculum, practical pronunciation exercises were not part of the English phonetics lecture course, but the students were expected to learn how to transcribe individual words in IPA during the course. Therefore, the pronunciation of many example words was discussed during the course. Many students taking the course will become EFL teachers, while some will become translators or work in other language-related fields.

Some students left some questions unanswered or, for instance, did not choose all three favourite themes. There were 3 students (2.3%) who did not choose any themes as their favourite.

3. RESULTS

In this section the results of the study will be presented. First we focus on the preferred themes. Then we look at student attitudes towards phonetics and phonetic knowledge.

3.1. Preferred themes

To answer our first research question, we calculated the answer percentage for each theme introduced during the course. The distribution of preferences is shown in Figure 1. As mentioned, the students were allowed to choose three.

![Figure 1: Percentage of students (n=133) preferring a given theme.](image)

As shown in Figure 1, the students clearly preferred accent differences the most. More than 2/3 of the students, 92 in total, chose this as one of their favourite themes discussed during the course. On the whole, it is clear that there was a lot of variation between the answers, and only one theme was chosen by the majority of students. The second most favoured theme was intonation analysis, which may have been slightly influenced by the fact that it was the last theme to be discussed and was, therefore, fresh in the students’ minds while answering the questionnaire. The third most preferred theme was phonetic transcription, which was a discussed topic but also a skill the students practised during the course. More than 1/3 of the students chose phonetic transcription as one of their preferred themes. In addition, speech rhythm was chosen by more than 1/3 of the students. The remaining themes were chosen by fewer than 1/4 of the students.

When focusing on the least favoured themes in Figure 1, we can see that the consonant and vowel systems were not often chosen by the subjects. Only 5.3% of the students (that is, 7 students) mentioned the consonant system as one of their favourite themes. Voicing, which was discussed in connection with the consonant system, was also very seldom chosen as a preferred topic. Only 9 students (6.8%) chose voicing as a preferred theme. Similarly, the vowel system was not a very popular theme among the students. These three themes were discussed during the early parts of
the course, which may have affected the results. In addition, exploring the phonological systems was one of the main parts of the course, which meant that they were discussed in detail. This may have affected student opinions. Interestingly, places and manners of articulation was a clearly more preferred theme than the vowel and consonant inventories although, naturally, these are very much connected and were discussed during the same lectures.

Excluding accent differences, which cover many phonetic levels, and phonetic transcription, which is mainly skill learning and quite concrete compared to the other themes, it seems that suprasegmental issues were preferred to segmental ones. Especially the phonological inventories were not often chosen as a favourite topic. This may reflect the tendency that suprasegmental issues are considered more crucial for the overall oral proficiency and intelligibility of EFL learners [e.g. 6, 12].

3.2. Attitudes towards phonetics

As for the second research question, we wanted to investigate how students perceived the role of phonetics in their education and future career, students’ attitudes towards phonetics, and if students thought that theoretical phonetics teaching had affected their practical pronunciation skills. We calculated the median values for student answers on the five-step scale. Means were calculated as well to reveal more subtle differences between the statements. Table 1 shows in an order of preference from the most to the least agreed the 19 statements used to reveal student attitudes. The lower the average, the more strongly the students agreed with the statement.

If we focus on the distribution of the answers, the students most agreed with the statement that phonetics has raised their awareness of the English sound system. 88.7% of the students agreed with this statement and only 1.5% disagreed (the distribution of the answers on the five-step scale strongly agree – strongly disagree was 46.6%-42.1%-9.8%-1.5%-0%). Most students (81.2%) also wanted to hear about many varieties of English during the course (distribution 41.4%-39.8%-11.3%-7.5%-0%). These two statements provoked the most positive reactions.

In addition, students also often thought that phonetics had improved their pronunciation (30.8%-39.8%-19.5%-6.8%-3.0%), phonetics is a challenging (28.6%-38.3%-23.3%-9.8%-0%) but interesting (30.8%-30.1%-25.6%-9.0%-4.5%) subject to study, and phonetic knowledge was considered important for their future careers (27.1%-34.6%-23.3%-12.0%-3.0%).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning phonetic transcription has helped my pronunciation</td>
<td>2.56</td>
<td>2</td>
</tr>
<tr>
<td>Phonetics has too many details to remember</td>
<td>2.60</td>
<td>3</td>
</tr>
<tr>
<td>I can list the consonant sounds of English</td>
<td>2.74</td>
<td>3</td>
</tr>
<tr>
<td>I can list the vowel sounds of English</td>
<td>2.83</td>
<td>3</td>
</tr>
<tr>
<td>Learning phonetics has made me more confident about my pronunciation</td>
<td>3.02</td>
<td>3</td>
</tr>
<tr>
<td>Learning phonetics has made me more nervous about my pronunciation</td>
<td>3.14</td>
<td>3</td>
</tr>
<tr>
<td>Finnish learners find phonetics important because they think that languages are pronounced as they are written</td>
<td>3.18</td>
<td>3</td>
</tr>
<tr>
<td>Phonetics is too theoretical</td>
<td>3.29</td>
<td>3</td>
</tr>
<tr>
<td>I think it would be better to focus on larger segments of speech than phonemes when teaching phonetics</td>
<td>3.34</td>
<td>3</td>
</tr>
<tr>
<td>I think we should focus on one variety during phonetics lectures</td>
<td>3.90</td>
<td>4</td>
</tr>
<tr>
<td>I think I will not need phonetics later in my life</td>
<td>3.95</td>
<td>4</td>
</tr>
</tbody>
</table>

According to the answers, the challenge that phonetics creates can be partly explained by the new terms that students are faced with (24.1%-35.3%-22.6%-16.5%-1.5%). On the other hand, most students believe that the improvement in pronunciation skills may have been caused by the information on the manners and places of articulation (21.8%-34.6%-22.6%-15.8%-4.5%) or learning phonetic transcription (19.5%-34.6%-23.3%-15.0%-7.5%).
The other statements caused more mixed or negative reactions. The students agreed slightly more than disagreed with the statement that phonetics has too many details to remember (21.8%-25.6%-29.3%-17.3%-6.0%). In a similar manner, the students more often than not thought that they could list the consonants (15.8%-29.3%-24.1%-26.3%-4.5%) and vowels (13.5%-25.6%-31.6%-23.3%-6.0%) of English. Consonants received slightly more positive reactions than vowels, but this was not statistically significant according to the Wilcoxon Signed Ranks test (p=.173).

The remaining statements were more often disagreed than agreed with. Thus, according to the students’ answers, learning phonetics has not made the students more confident (9.0%-24.8%-30.8%-26.3%-9.0%) or more nervous (11.3%-22.6%-19.5%-33.8%-12.9%) about their pronunciation, although there were also many opposite opinions. Similarly, the Finnish context, the fact that Finnish has a close correspondence between letters and phonemes, was not seen as a reason for the importance of phonetics (4.5%-20.3%-36.8%-21.1%-12.8%), and most students did not consider phonetics too theoretical (6.0%-14.3%-38.3%-27.1%-14.3%). Almost half of the students reacted neutrally towards focusing on larger segments than phonemes (3.0%-9.8%-45.9%-31.6%-9.0%)

The students most clearly disagreed with the statements that only one variety of English should be focused on during the course (3.0%-9.8%-14.3%-39.8%-33.1%) and that phonetics is not needed later in their lives (3.8%-6.8%-16.5%-36.1%-36.8%). Very few students agreed with these statements.

4. DISCUSSION AND CONCLUSIONS

The purpose of this study was to examine the learner perspective towards themes introduced during a course on English phonetics and learner attitudes towards phonetics in general. The results suggested that students were most interested in accent differences and the varieties of English. This was reflected in answers to both research questions.

One limitation of this study was that data collection was limited to the questionnaire. Therefore, the motivations for these results can only be the subject of speculation. The students seemed to prefer many pronunciation models to a single one. During the course, the main differences between the standard British and American English accents were discussed, while other native accents of English were less systematically referred to. Based on the answers, it is not certain if the students were interested in systematic comparisons of standard varieties or individual features which differed between some (non-standard) varieties.

As mentioned above, the teaching order may have affected the most preferred themes. Earlier discussed topics may have been less important in the minds of the students during the data collection process. The themes chosen followed fairly clearly the outline of the course although some themes (such as the consonant system) needed more time than others (such as the weak forms). In addition, phonetic transcription was a very different theme from the others as it was discussed to some extent throughout the course. It was also a concrete learning goal for the students as they knew that they had to transcribe some words in the final exam. Suprasegmental issues may have been preferred to segmental ones as they may have been less familiar to the students, or the students may value fluency over accuracy in their own spoken skills.

Although the students took the course as a required element in their English studies, their attitudes towards phonetics were, on average, quite positive. The learners had learnt about the sound system of English and believed that this had improved their pronunciation. The students considered that phonetics would be important for them later in their lives. Our methodology did not allow us to further analyse the reasons for these answers. The high frequency of neutral responses to some statements may suggest that these statements were not clear to the students or focused on issues the students did not consider important.

As mentioned, phonetics is an often taught discipline, but studies on emphasizing the learner perspective are scarce. To facilitate the learning process and to make teaching more efficient, further studies on learners are needed. These studies could use mixed methods in data collection to deepen the perspective. For instance, learner interviews or learning journals written during taught courses could offer new insights into the subject matter. As the results of this study have suggested, learner attitudes and preferences vary considerably, which could be caused by, for example, learners’ individual learning histories, study motivation or career goals. A qualitative approach could be used to examine these variables.

This study has shed light on the challenges and possibilities that learners face while studying phonetics in a foreign language programme at university level. Other programmes including a phonetic component are still to be investigated. The contribution of the present study is to strengthen our knowledge of phonetics as a field of study and enhance the quality of teaching to meet learner
expectations to support learning and learning motivation. The results can be used in curricula and course structure planning. Learning may also be facilitated by motivating the less preferred topics more thoroughly during contact teaching.

5. REFERENCES


LABORATORY ACTIVITIES FOR LARGE AND ONLINE PHONETICS CLASSES

Timothy Mills¹, Karen Pollock², and Benjamin V. Tucker³

¹Department of Linguistics, ²Department of Communication Sciences and Disorders, University of Alberta

timills@ualberta.ca; karen.pollock@ualberta.ca; btucker@ualberta.ca

ABSTRACT

Phonetics is an inherently lab-oriented topic, involving the investigation and analysis of speech data; but it is often taught in large sections with limited instructional time and other resources, so instructors are unable to engage in the deep interactive explorations of topics that would be ideal. We present three phonetics laboratory activities designed to address this problem. In addition, these activities can engage students in material in ways that are not possible in a primarily lecture-based course; they can provide a source of empirical data for use in longitudinal research; and they can give students a taste of the experimental investigations from which much of the material in their course derives. We discuss pedagogical goals, activity design and grading. We also discuss challenges and successes in the implementation of these activities.

Keywords: laboratory activity, large class, online

1. INTRODUCTION

At our university, as in many programs, an introductory course in phonetics is a core part of a degree in linguistics. Introductory phonetics is a second-year course for undergraduate students and is often the second or third linguistics course taken as part of their program. Phonetics is a prerequisite for many more advanced courses, such as phonology, speech perception, acoustic phonetics, and language acquisition. It also fulfills requirements in other degrees such as English as a Second Language taught in Education. In Canada generally, degrees in speech-language pathology are offered only at the Masters level and an introductory course in phonetics is often a prerequisite for admission. As a common prerequisite, our introductory phonetics class comes under heavy demand, with total annual enrolment often exceeding 150. Per-course enrolment in a normal academic semester often ranges from 50 to 80. These per-course numbers are likely to increase due to lack of funding to offer additional sections.

We have also seen increasing interest, from students at universities where a phonetics course is not offered, in taking this course online. This is related a wider trend internationally, noted in [6], toward more interactive online and digital content in phonetics courses. Although we are aware of several other online phonetics courses that have been offered internationally (for example, see [3, 4]), it appears that no such distance course is currently offered in Canada. We have developed this course in an online format to meet the demands at both the local and national levels.

We work with the belief that a phonetics course should be a laboratory-based course. However a laboratory-based course with either large enrolment numbers or students taking the course at a distance (online) presents a challenge to this belief (see also [2]) and raises a fundamental question: Can we teach this material and continue to provide students with a hands-on learning experience? We do not claim that these laboratory activities are novel in any way and we are aware that many of these types of activities are implemented in many similar classes around the world. In fact much of the content of the current labs is based on examples from others.

The present paper reports our experience in modifying and developing three instructional laboratory activities for online and large-course delivery. These activities have been piloted with students in both on-campus and online courses.

2. INSTRUCTIONAL LABORATORY ACTIVITIES

In this section we provide details about the overall course delivery; then we describe three laboratory activities and their major components. We provide a brief description of the purpose of each assignment, and describe the content of the assignment in detail. We then provide a brief evaluation of the success of each activity and plans for continued refinement.

The course material is delivered in two ways. The first is as a conventional on-campus lecture style course, typically delivered as 3 hours of lectures each week for a 12-week course, or 6 hours per week for a 6-week course. The second is as an online course, delivered using video lectures and extensive electronic dialogue with students (via email, video conference and a chat room). In both formats, assignments and quizzes are given to students through...
our online learning environment, Moodle [7], to complete on their own. Assignments have fixed due dates, and quizzes are available for students to write them over approximately a 48-hour window. All courses include a final examination. The course is designed to offer online students and on-campus students the same content and, as far as practical, a similar experience.

The assignments are designed to address two key learning objectives: they give hands-on experience in discovering and learning about phonetics, and they give students opportunity to apply what they have learned from the material presented in lectures. The three selected laboratory activities address different core topics in our phonetics curriculum: Articulation, Acoustics, and Speech Perception.

2.1. Articulation: “Frankentract”

2.1.1. Target outcome

Seventeen anatomical structures are discussed in our course. Some, such as the lungs and lips, are already familiar to students when they enter the course. Others, such as the cartilages of the larynx or parts of the palate, are not. Students learn to identify these structures, and they learn how these structures are used in the production of speech sounds. The Frankentract activity gets students to think creatively about how anatomy and speech production work together.

Figure 1: Sample “Frankentracts” created by students in introductory phonetics. The one on the left is made of clay; the one on the right is constructed in Minecraft.

2.1.2. Activity structure

This activity asks students to construct an artificial vocal tract using convenient materials in their environment. They are encouraged to build a physical model. They may instead make a digital or hand drawn schematic, create an image of it, and submit the image. See Figure 1 for two examples (reproduced with permission). Along with the image, they submit a write-up that details how the artificial analogues of each anatomical structure covered in class is constructed and how the materials and construction would affect its function in the speech production system. Artistic ability and acoustic function of the construct are not graded.

2.1.3. Discussion

Although drawings would be easier for most students, a significant portion of students (38%) choose to physically construct a model. We take this to indicate that students are enjoying the material, and they are motivated to engage with it beyond what is required in the graded evaluation.

The main difficulty at first for this activity was making the scope clear to students. Having a specific list of which specific structures they are expected to address has helped. So has reminding students that the purpose of the activity is to demonstrate that they understand how the anatomy functions to generate speech and that they won’t be graded for number of hours spent developing the model.

A second challenge, common to all of the activities discussed here, is calibrating the time spent assessing the assignments. Even a cursory scan of each submission to ensure the student has mentioned all 17 structures and made relevant comments becomes time consuming in a class of 50 or more students.

Overall, students show creativity in their constructions and exhibit understanding in their write-ups. This demonstrates to us that this is a successful exercise, from the instructor’s perspective. Informal feedback from students suggests that they find it enjoyable as well.

2.2. Acoustics: “English Vowel Spaces”

2.2.1. Target outcome

The English Vowel Spaces activity connects students directly with the acoustic characteristics of speech sound by having them record their own vowel productions, measure the formants, and reason about how these measurements relate to the patterns discussed in class. A secondary pedagogical goal is the aggregation of data across all students in a class to produce formant plots, to compare against published formant spaces in the literature. This forms an interesting follow-up analysis that can be undertaken in class.
2.2.2. Activity Structure

Students are given written instructions for the assignment on the course website. The instructions direct students through the process of using Praat [4] phonetic analysis software to record twelve standard [hVd] words, measure the first three formant values of the vowels, and record these values. They then discuss the patterns observed with respect to the patterns discussed in class. Does F1 pattern with vowel height as expected? Does F2-F1 distance pattern with vowel frontness as expected? Students are encouraged to discuss apparent anomalies and to speculate on reasons.

They submit a report that includes their formant measurements for each target vowel, as well as a one-page discussion of the patterns they observe in those measurements. They may also choose to input their measurements (along with limited demographic information - see section 3 below) into a form that adds their data to the longitudinal dataset.

2.2.3. Discussion

The main goal of this lab activity is to reinforce the basics of acoustic phonetics. It also provides students with experience interacting with their own data. For an instructor to manually check reported formant values against the sounds is implausible, either in the large on-campus course or in the intensive online course. Feedback on measurement accuracy is limited to noting implausible values, without penalizing student grades. Measurement errors, such as mistaking the voice bar for F1 (and the consequent shift - identifying F1 and F2, etc.) can be discussed in class. However, there is no straightforward way to check whether students have assimilated this instruction.

A plot aggregated across the whole group can then be used as an illustration in class lectures after the data have been collected. A series of R scripts [3] is used to generate a plot in class as in Figure 2 so that students can also see the process.

Figure 2: Sample aggregate vowel plot (logged axes) from student data, with ellipses one standard deviation from the mean.

The use of a Google form for students to add their data to the accumulating pool of measurements has presented a challenge to students and we have found that some time in class or in an online video is needed to demonstrate how they are to carry out this exercise.

This lab activity offers students a chance to apply their vowel acoustics knowledge to their own speech. It further engages students with their own speech and allows them to compare it to others.

2.3. Speech Perception: “Incomplete Neutralization”

2.3.1. Target outcome

This activity is an extension and illustration of patterns of perception that are covered in class. In particular, it gives students an opportunity to see what sort of evidence linguists use to measure and analyse phonetic and phonological phenomena in languages.

2.3.1. Activity Structure

Students are given an experimental setup to explore the “neutralization” of /t/ and /d/ in the environment where both are produced as a flap [ɾ] in Canadian English (after a stressed syllable nucleus and before an unstressed syllable nucleus). The theoretical question of interest is presented, and students are given five word pairs, such as “ladder”/“latter” and “wider”/“whiter”. They are to produce these words naturally (avoiding hyperarticulation, which can interfere with flapping) and have a listener note, for each production, which item in the pair they think they heard. They perform this for five listeners, generating a total of 50 data points (1 production each for 10 words with 5 listeners).

Students are asked to submit a report of their results, including brief introduction, methods, results, and discussion sections. Some of this is simply a restatement of the assignment instructions in their
own words. They also look for patterns and comment on them using their understanding of speech production and perception.

2.3.5. Discussion

Many students find evidence for incomplete neutralization. In addition, Canadian raising has a huge impact on perception in the diphthongs (like “whiter”?/“wider”). Students often find variation stemming from differences in language background (non-native vs. native speakers) and age (older vs. younger). Often half or more of the student population is non-native, which impacts the results.

Because this activity is undertaken under widely varying conditions by different students (different language backgrounds for the speaker and listener, different acoustic environments, etc.), not all students find good evidence of incomplete neutralization. While it is a good illustration for those students who do find incomplete neutralization, its illustrative value for others is unclear. Different stimulus sets or phenomena might be better suited to the pedagogical goals of the activity, but it is unlikely that in a linguistically heterogeneous student population, a single phonological pattern can be found that is broadly reliable. Using flapping, a pattern common in the dominant local dialect (Canadian English), probably gets the best results we can expect.

3. GENERAL DISCUSSION

Phonetics can have a “high ‘wow’ factor – the scope for fun and impressive demonstrations is vast.” [1] However, such demonstrations often require direct interaction that is not always convenient (e.g., in large classes) or even possible (e.g., in online classes).

We have attempted to create engaging activities that give students the opportunity to demonstrate their understanding of the material, to engage directly by producing their own models (Frankentract), data (English Vowel Acoustics), and experimental stimuli (Incomplete Neutralization).

These activities continue to require refinement. One challenge related to these activities is the time required to mark them. Possible solutions include: (1) hiring teaching assistants to carry some of the marking load; (2) switching activities from graded-with-feedback to participation marks (automatic submission credit), while offering model answers; and (3) enacting a peer-marking scheme. Each of these options has advantages and disadvantages in terms of resources, quality and depth of feedback, and logistical complexity.

We are carrying out a review using targeted student surveys as well as analysis of aggregated grade outcomes to determine how successful each activity is in engaging student interest and reinforcing the course content. We are also comparing online versus on-campus delivery of the course, to establish how well the activities translate between different teaching modalities. We plan to continue developing further activities as graded material or optional enrichment exercises.

One of the interesting possibilities of the English Vowel Spaces activity is that it can also be used longitudinally for both research and pedagogical purposes. Assuming the proper ethical procedures have been established and informed consent has been obtained, the data can be gathered and kept for research purposes. These data can be used in two ways: (1) for continued pedagogical purposes – the data from previous classes can be compared to the current class; and (2) as a research object, as documentation of the vowel system of the local dialect and more generally of the students taking the course. Obtaining empirical research data in this way is beset with difficult-to-overcome confounds. Students are a heterogeneous and non-random sample (widely varied language backgrounds and ages, among other things). One major challenge in using these data for research is verifying the accuracy of submitted measurements (for example, there are several errors in the data presented in Figure 2). We are currently trying to devise a means by which students can submit their recordings alongside the measurements, allowing for spot-checks.

4. CONCLUSION

While we are still in the process of formally evaluating the success of these activities and differences in the delivery types, we feel that these activities have been successful in allowing our students to get more hands-on experience with phonetics. This experience is an essential part of a phonetics course, even with the challenges presented by a large-class and online presentation modes. We are committed to creating an engaging, interactive class and not a course where students are only expected memorize material and take a test. We hope in sharing our experience, after the example of [5], that more activities of this sort will be reported and the potential for sharing will increase.

5. REFERENCES


TEACHING A PHONETICS COURSE WITH A HIGH STUDENT ENROLMENT USING A BLENDED-LEARNING FORMAT

Shawn L. Nissen, Ph.D.
Brigham Young University
Provo, Utah USA
shawn_nissen@byu.edu

ABSTRACT

This study examined the efficacy of using a blended-learning (BL) format of instruction to facilitate student learning of phonetics and IPA transcription in a course with a relatively large enrolment (e.g., 60 - 100 students). The effectiveness of the BL instruction was measured through students’ academic performance, course ratings, and survey data collected before and after the new method of instruction was initiated. These data indicate that in many aspects the BL format of phonetic instruction was successful in maintaining student learning and engagement when compared to a more traditional model of instruction.

Keywords: phonetic instruction, blended learning, computer-mediated instruction.

1. INTRODUCTION

Ideally a course in phonetics would be taught with a low student-to-teacher ratio to allow for a high degree of student engagement and individualized mentoring from the instructor. However, in the environment of shrinking higher education budgets and in many instances increased student enrolment, some phonetic instructors are finding themselves teaching courses to a relatively large number of students. For university courses that rely primarily on the memorization and direct recall of information, classrooms with large numbers of students may not hinder learning in a significant manner. However considering that phonetics is a course that involves both the understanding of theoretical concepts and the development of the perceptual ability to distinguish the many phonetic features of a language, classrooms with large numbers of students are not the ideal learning environment. As it is well known to phonetic instructors, to become proficient in clinical transcription, students often require individualized practice, immediate feedback, and a distraction-free listening environment.

Blended learning is an instructional method that combines two modalities for student learning, traditional face-to-face instruction and distributed learning systems through computer-mediated instruction [1]. Thus the aim of BL instruction is not to replace classroom instructors with online content, but rather to emphasize the strengths and minimize the limitations of both methods of instruction [2]. Some phonetic principles might be best taught and understood through the responsiveness and human connection found with face-to-face student and instructor interaction, such as the ability to perceive and phonetically transcribe allophonic variation. However the quantity and quality of such interactions are often limited by the time restraints of the instructor and a lack of basic knowledge on the part of the student.

A well-designed BL curriculum may improve phonetic teaching and learning by increased cost-effectiveness. Studies by universities and corporations, such as IBM and Microsoft, have indicated that BL can provide a significant return on investment in terms of both the time of the instructor and in physical infrastructure costs, e.g. [1, 4, 5, 6]. BL can also provide students with increased access to course information and flexibility in their learning environment. Specific to a phonetics course, it may be highly beneficial for a student to be able to practice IPA transcription at home under headphones, compared to a classroom with visual and auditory distractions. BL techniques might also lead to improved pedagogy [3]. Although many concepts are best taught through face-to-face instruction, some phonetic principles may be more easily communicated to students through computer models, online audio, or peer-to-peer learning strategies [1, 3].

Despite the many advantages to BL, this type of instruction may not be appropriate or effective for some courses. If a blended course is not highly organized and supported, the complexity of the instruction can become a stumbling block for efficient learning. In addition, the impersonal nature of computer-mediated instruction can lead to a disinterest in learning and a tendency toward procrastination [1]. It is important to weigh the advantages and disadvantages of BL for each specific course and instructional situation. Thus the aim of
this study was to provide some additional understanding into the efficacy of using a BL to facilitate student learning of phonetics and IPA transcription in a course with a relatively large enrolment.

2. METHOD

2.1. Description of the Course

The class evaluated in this study was a clinical phonetics course offered to undergraduate students in a department of communication disorders. The course was taught multiple times over a span of four years, one year prior to the initiation of BL and three years following the change in instruction. Each year the course was taught three times per week (Monday, Wednesday, and Friday) for 50 minutes each session over a period of 16 weeks. The number of students enrolled in the course varied between 52 and 84 students per section. The course was designed to present phonetic knowledge relevant to the understanding of how speech sounds are produced and perceived in human languages, as well as assist in the development of the ability to conduct IPA transcription of American English.

2.2. Model of Instruction

2.2.1. Theory and Principles of Phonetics

The theory and principles of phonetics were presented during a large group face-to-face lecture on Monday and Wednesday of each week. During several weeks throughout the semester, these large group lectures were replaced by smaller group instruction to facilitate more student engagement on select topics. At the beginning of the semester students were divided into two groups (1 and 2) based on the first letter of their surname at the time of enrolment in the course. During select weeks, students in each group would only attend lecture on either Monday or Wednesday, but not both. In place of the time students did not attend, they were then assigned to complete an online tutorial and pass a quiz at 80% competency during that week.

2.2.2. IPA Transcription

On Friday of each week students were required to complete a series of alternating online and in-person labs to allow for more individualized instruction in IPA transcription. The in-person labs were held every other Friday to provide students immediate practice and feedback based on their application of material memorized online.

The online lab assignments featured self-paced tutorials, audio examples, and practice transcription exercises. The online tutorials contained short passages of real speech embedded in a Microsoft Powerpoint presentation that could be downloaded from a university website. The audio samples differed in level of difficulty to allow learners the opportunity to practice IPA transcription near their skill level and to more closely simulate real-world applications. The samples were recorded by speakers of various ages and dialects of English. Considering that the course was taught to students majoring in speech-language pathology, samples of disordered speech were also made available for transcription. The online audio samples also varied in terms of speech rate and level of extraneous noise, with samples being recorded in a noise attenuating booth, a classroom, and also in the presence of some background noise. All of the practice exercises were accompanied by a transcription key for students to check their accuracy. When the students felt they had mastered the online material, they were instructed to complete an online assessment with a competency score of at least 80%. Students were allowed to download the tutorials, practice exercises, and take the quiz an unlimited number of times.

2.2.3 Method of Evaluation

The effectiveness of the BL instruction was measured through students’ academic performance, course ratings, and survey data collected before and after the new method of instruction was initiated. Baseline data were collected for one semester using a traditional method of instruction, after which data were collected for three consecutive years using a BL format of instruction. The instructor and the text for the course remained the same throughout the duration of the study.

3. RESULTS

The students’ academic performance before and after the initiation of the BL instruction is listed in Table 1. These data are based on five IPA examinations involving auditory recording of typical speech recordings and two multiple-choice examinations on the theoretical content of the course. These data do not include any extra credit adjustments.

| Table 1: Student academic performance at the end of the course in mean percentages. |
A summary of the course ratings completed by the students enrolled in the course is shown in Table 2. These ratings were submitted by students toward the end of the course, yet prior to the posting of the academic grades. The response rate for each of the courses listed was ≥ 94%.

**Table 2**: Summary of course ratings completed by the students (out of 8)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IPA Assessment %</td>
<td>91.6</td>
<td>86.1</td>
<td>91.4</td>
<td>90.0</td>
</tr>
<tr>
<td>Final Course %</td>
<td>83.9</td>
<td>84.7</td>
<td>85.5</td>
<td>86.3</td>
</tr>
</tbody>
</table>

Figures 1 – 3 include student responses to three questions specifically about the effectiveness of the BL instruction used in the course. These data were averaged across the three semesters following the initiation of BL (i.e., 2012 through 2014).

**Figure 1**: Student responses to the first question regarding the effectiveness of the BL instruction.

**Figure 2**: Student responses to the second question regarding the effectiveness of the BL instruction.

**Figure 3**: The smaller lecture groups created more in-class discussion.

5. REFERENCES


4. DISCUSSION

Data from students’ academic performance, course ratings, and survey responses indicate that the BL format of phonetic instruction was successful in maintaining a similar degree of student learning when compared to a more traditional model of instruction. Generally the academic performance in the course for each group of students’ was similar across pre- and post-BL instruction. However, academic performance did decrease during the first year of implementation of the new method of instruction. This decrease may have been due to the instructor’s inability to smoothly organize and administer the differing elements of the BL format.

Students’ responses suggested that they were generally satisfied with the BL aspects of the course. In particular, students indicated that they liked the online tutorials and practice IPA exercises, as well as the flipped transcription labs that provided a smaller learning environment. However the data also suggested that splitting the theoretical lectures into smaller groups did not necessarily create more in-class discussion and engagement. This change was made to increase student participation, however the amount of participation during podium lectures may have decreased. This may have been because many of their questions were already being addressed in the smaller lab groups.

Considering that the students in the course differed from year to year, direct comparisons of the data collected in this study were somewhat tenuous in nature. However despite this limitation, these data provide some support for the effectiveness of using a BL format of instruction when teaching a phonetics course with a relatively high student enrolment.
ASSESSING THE PHONETIC LEVEL IN L2 VOWELS PRODUCTION WITH THE VISUVO SOFTWARE: CASE OF CZECH LEARNERS ACQUIRING CONTRASTS BETWEEN FRENCH MID VOWELS e/ɛ, ø/œ AND o/ɔ

Nikola Maurová Paillereau
Laboratoire de Phonétique et Phonologie (UMR 7018) CNRS / Paris 3, 75005 Paris, France
nikola.paillereau@mac.com

ABSTRACT
The VisuVo software is designed to visually compare formants of vowels as uttered by different speakers, in isolation and in different contexts or prosodic positions. Both researchers and teachers can use the software. VisuVo aims to evaluate the initial phonetic level of a learner and to trace the progress in L2 vowels realisation compared to a set of pre-recorded references. VisuVo has an interactive and simple interface to handle. Its features and output graphics are illustrated by a study of contrasts between the French mid vowels e/ɛ, ø/œ and o/ɔ produced in isolation and in symmetrical contexts /p, t, k, R/ by 10 native non-southern French speakers representing the reference, and 10 Czech learners of French to be tested.

Keywords: pronunciation assessment, vowels, formants, French as a Foreign Language (FFL), Czech

1. INTRODUCTION

1.1. Teaching and Assessment in Phonetics

Since the appearance of the communicative approach in the 1970s, perfect pronunciation of L2 sounds is no longer an objective. Still, learners may wish to go unnoticed for their foreign accent and to speak, ideally, like natives [3]. Teachers should therefore compare the particularity of phonetic-phonological systems of L1 and L2, their prosody and segments, their phonotactic rules and coarticulation patterns leading to allophonic variations [2]. Coarticulation habits are language dependent [12]. Once the vowels have been inserted in the spoken chain, the formant frequency values in the middle part of the vowels, on one side, and the direction of the formants during the transition between the vowel and the surrounding consonants, on the other side, depend on many factors, such as the language itself, surrounding consonants, prosodic position (initial, final, stressed, etc.), style (hyper- or hyp-o-articulated) and speech rate [11]. The mastery of allophonic variations should therefore be included in the curriculum.

An objective evaluation of an exact phonetic level acquired by a learner is hardly possible on the sole basis of perceptual judgments. Human assessors can be helpful, of course, as they can normalise the incoming speech and categorise sounds without any big difficulty; they cannot, however, perform fine acoustic analyses enabled by modern technological means [1]. The difficulty in objectively assessing different phonetic levels is reflected by the Common European Framework of Reference for Languages (CEFRL)[5], which does not provide clear criteria for distinguishing between six levels of phonological skill (A1, A2, B1, B2, C1 and C2), and confuses levels C1 and C2. Progress in the acquisition of new phonological contrasts is gradual, and teachers should therefore appreciate different stages of phonetic progress [7]. Machines can help and, in this perspective, VisuVo (VISualisation of VOWels) can be considered a tool for teachers to objectively assess different phonetic levels.

1.2. Acoustic Target

To study deviations in L2 pronunciation, researchers and teachers need to refer to a phonetic norm defined by native speakers’ productions [9]. The same vocalic timbre can be achieved with different articulatory strategies [13], so that vowels are more easily defined in acoustic terms, that is in terms of their formant frequencies. The current French segmental system includes 3 nasal vowels and ten oral vowels /i, e, æ, u, o, ɔ, y, ø, ɑ/. Back vowels can be adequately defined by the first two formants (the upper formants have no perceptual weight [6]), front vowels by the first three formants, and the vowel /ɨ/ by the first four formants. A closer distance between F3 and F4 defines the French vowel /ɨ/, so that F4 is important in the acoustic definition of vowels with /ɨ/-like timbre [14].

2. AIM OF THE STUDY

The objective of this paper is to present the VisuVo diagnostic tool. VisuVo is a piece of specialised software which can quickly and interactively generate three types of graphs representing vocalic formants
from a large database. The graphs are vowel triangles, allophonic variations tracers and contrasts diagrams. They aim to highlight acoustic differences in the pronunciation of L2 vowels compared to pre-recorded L1 formant frequency values as a reference, and display learners’ progress.

In this study, VisuVo illustrates three first formants (F1 to F3) of French mid vowels e/ɛ, ɔ/œ and two first formants of ɔ/ø, as produced by 10 Czech learners and 10 native non-southern French speakers (reference). The Czech language has five phonologically short vowels /a, e, i, o, u/ (/æ/ and /o/ being sometimes transcribed in recent Czech literature with /ɛ/ and /ɔ/ and five long ones /aː, eː, iː, oː, uː/. The Czech language (unlike French) does not make contrast between rounded/unrounded front vowels, nor between mid-open/mid-close vowels.

The question is to what extent Czech learners can acquire contrasts between e/ɛ, ɔ/œ and ɔ/ø? Which are the consonantal contexts that facilitate learning of the three contrasts?

3. METHODOLOGY

3.1. Corpus and Speakers

The corpus, derived from a larger body [10], consists of the mid vowels /e, e/æ, o, œ/ and /o, œ/ pronounced:
1. in isolation (i.e. between two short pauses as to avoid coarticulatory phenomena), inserted in carrier sentences, such as « Tôt il a dit <ɔ> comme dans iôt. » (“Soon, he said <u> like in soon.”)
2. in trisyllabic logatoms CVCVCVC (where C = /p, t, k, R/) inserted in carrier sentences, such as « Le mot kaaukaukeu peut bien coller. » (“The word kaaukaukeu matches well.”)

The corpus includes recordings of female speakers without any known hearing problems, namely:
1. Ten native non-southern French, aged between 21 and 48 (M = 28,5), 4 repetitions.
2. Ten Czech advanced learners (Learners 1 to 10), aged between 25 and 28 (M = 25,7), 10 repetitions.

Productions of French natives were used to define the French reference in the present study.

3.2. Recording and Data Processing

All instructions were given in writing. The recordings were made in a quiet room with a headband microphone AKGC 520 L and a sound card Edirol UA 25 connected to a Mac computer. The set of sentences was recorded with a sampling frequency of 44100 Hz and a sampling rate at 16 bits. Recordings were preceded by a training phase.

Ten target vowels of two corporuses were labelled and analysed using the Praat software [4]. The formant values were automatically calculated with the formula To Formant (burg) of the Praat system (analysis window: 25 ms). All values taken semi-automatically were checked and corrected manually, when necessary.

4. RESULTS

4.1. Average Values

The differences in F1, F2 and F3 (in Bark, expressed as a percentage) of the mid vowels realised by the French speakers (FR) and Czech learners (CZ) are summarized in Table 1.

The French native speakers realised the contrast between mid vowels with a systemic difference in F1. According to the consonantal context, the difference in F1 varies between 25 % and 34 % for the pair e/ɛ, between 23 % and 30 % for ɔ/œ, and between 26 % and 31 % for ɔ/ø. The difference in F2 is equal to or less than 6 % for pairs e/ɛ and ɔ/œ, and between 13 % and 22 % for the pair ɔ/ø. The difference in F3 doesn’t exceed 3 % for all pairs.

The Czech learners do not exhibit a difference in F1 and F2 between mid vowels e/ɛ and ɔ/œ, except for F1 of e/ɛ produced in isolation and F2 of ɔ/œ realised in a labial context where the difference represents 5 %. According to Flanagan [1], a 5 % formant difference corresponds to a just noticeable difference for vowel formant frequencies F1 and F2, but it is probably not large enough to distinguish between two different phonemes. In contrast, vowels ɔ/œ are produced with a difference in F1 between 5 % (in certain contexts) and 15 % (in isolation). The difference in F3 never exceeds 1 %.

The average vocal formants of e/ɛ, ɔ/œ and ɔ/ø produced by natives and learners, in different consonantal contexts and prosodic positions (the first, mid and the last vowels in logatoms) is illustrated in Figure 1 generated by VisuVo. Finally, the strengthening of contrasts between e/ɛ and, in particular, ɔ/œ produced by the Czech learners in isolation is illustrated in Figure 2. We note a difference in mean values of F1.

4.2. Individual Variability

The study of 10 replicates of vowels individually performed by 10 learners shows that acoustic cues constituting contrasts vary from learner to learner. A quasi-systematic acoustic difference between pair vowels indicates that learners have created two different categories for the vowels in contrast.
Vowel triangles F1/F2 (in Bark) of Figures 3 and 4 generated with VisuVo make it possible to compare performances of different learners according to phonetic contexts.

In isolation (Figure 3), Learner 2 does not distinguish between e/ɛ, but the majority of occurrences of ø/œ and o/ɔ have different F1 and F2 (yet the difference is lower than in natives). Learner 9 achieves e/ɛ and o/ɔ with a difference in two formants (again, the difference is lower than in natives). The contrast between ø/œ in Learner 9 is based on F2 contrary to F1 in natives.

In a dental context (Figure 4), Learner 2 produces the contrast e/ɛ with a difference in F2 primarily, and does not distinguish between ø/œ and o/ɔ. Learner 3 pronounces e/ɛ with a difference in F1 and F2, but, in contrast to natives, the [ɛ] is produced, on average, with a lower F1 than [e]. The contrasts between ø/œ and o/ɔ are essentially based on F2.

5. CONCLUSION

Native French speakers systematically produce the contrast between mid vowels based on the main acoustic cue F1, traditionally correlated to the degree of aperture. These contrasts are an obvious phonetic difficulty for Czech learners who generally do not distinguish e/ɛ and o/ɔ, and realise the contrast ø/œ with less acoustic difference than natives. This result shows that the difficulty of acquiring four contrastive degrees of aperture instead of three is not the same for all vowels. Indeed, the contrast between phonemically new vowels ø/œ (two new phonemes for Czech learners) is more approximated than the contrasts between phonemically similar vowels e/ɛ and o/ɔ.

A detailed analysis of formant frequencies eased by the use of VisuVo shows that 1) performances in the production of contrasts vary from learner to learner, 2) some contexts appear to facilitate contrasts implementation, and 3) a number of potentially specific problems for a given learner can be easily identified. In order to objectively evaluate different stages of phonetic progress, teachers should be advised to complete their perceptive appreciation by visualisation of formant frequencies, as done easily by VisuVo.

Figure 1: Average formants F1, F2 (F3) (in Bark) of vowels e/ɛ, ø/œ and o/ɔ produced in isolation (horizontal strokes) and in logatoms CVCVCVC (where C = p, t, k, R) in the first, second and the last vowel by 10 native French speakers (cross, 4 repetitions per speaker) and 10 Czech learners (rhombus, 10 repetitions per speaker). Values taken at 1/3, ½ and 2/3 of vowel lengths. The standard deviation is one.

Figure 2: Average formants F1, F2 and F3 (in Bark) of vowels e/ɛ, ø/œ produced in isolation by 10 Czech learners (10 repetitions per speaker). Values taken at 1/3, ½ and 2/3 of vowel lengths. The standard deviation is one.

Figure 3: Triangle F1/F2 (in Bark) of 10 French oral vowels repeated in isolation 4 times by 10 natives (circle) and 10 times by Learner 2 (cross)
and Learner 9 (triangle). Values taken at 1/3, ½ and 2/3 of vowel lengths.

Figure 4: Triangle F1/F2 (in Bark) of 10 French oral vowels repeated in logatoms tVtVtV 4 times by 10 natives (rhombus) and 10 times by Learner 2 (square) and Learner 3 (triangle). Values taken at half of the vowel lengths.

Table 1: Difference of average formants (in percentages) F1, F2 (F3) (in Bark) of vowels e/e, ø/e and ø/ø produced in isolation (0) and in logatoms CVCVVC where C = p, t, k, R by 10 native French speakers (FR, 4 repetitions per speaker) and 10 Czech learners (CZ, 10 repetitions per speaker). Values taken at half of vowel lengths.

<table>
<thead>
<tr>
<th>Ctx</th>
<th>Vw</th>
<th>Diff. F1 %</th>
<th>Diff. F2 %</th>
<th>Diff. F3 %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FR</td>
<td>CZ</td>
<td>FR</td>
<td>CZ</td>
</tr>
<tr>
<td>0</td>
<td>e/e</td>
<td>34 5 4 0 3 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/e</td>
<td>30 15 1 3 1 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/ø</td>
<td>26 4 20 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>e/e</td>
<td>30 4 5 2 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/e</td>
<td>29 7 1 2 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/ø</td>
<td>30 4 21 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>e/e</td>
<td>27 2 3 1 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/e</td>
<td>26 5 2 1 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/ø</td>
<td>28 2 13 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>e/e</td>
<td>25 2 3 1 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/e</td>
<td>23 5 3 1 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/ø</td>
<td>26 2 19 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>e/e</td>
<td>27 2 6 2 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/e</td>
<td>27 6 5 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ø/ø</td>
<td>31 4 22 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. REFERENCES


DIFFERENT DESIGNS OF UNIVERSITY COURSES BASED ON PHONETICS IN THAILAND

Rungpat Roengpitya  
Department of English, Faculty of Liberal Arts, Mahidol University, Thailand  
rungpat@gmail.com, rungpat.roe@mahidol.ac.th

ABSTRACT

Phonetics teaching and learning has been important not only for English as a foreign language (EFL) learners, but also for learners of other fields such as medical disciplines. This paper presents the designs of three different university courses based on phonetics: “Sound System of English (SSE),” “Scientific Experimental Phonetics (SEP),” and “Speech Production and Perception for Language Teaching (SPP)”. The details of these three courses include background, course structures and contents, learners, evaluation, assessment, community services, and career preparation. It is hoped that this study will shed light on to the designs of future courses based on the phonetics teaching and learning.

Keywords: Design, English, Phonetics, Thailand, University.

1. INTRODUCTION

In Thailand, English is the first foreign language (FL) of Thai citizens, who speak Thai as their native language. In terms of education, it was found that Thai learners have different levels of English competence when they reach university level. This is due to the fact that, prior to tertiary level, the curricula of urban and suburban schools vary. Some schools have used English as a medium of instruction, starting from kindergarten or primary level. Some schools include English courses from Grade 1 or as late as Grades 5-6. Thus, the shorter or longer duration of English exposure affects the English competence of Thai learners. Besides, the English content and instructions of the curricula in schools may not equally cover all four skills: listening, speaking, reading, and writing. Basically, English in schools concentrates on the English grammar and some reading and writing with less listening and speaking, though the latter two skills are important for learners (Flowerdew and Miller [6]).

At university level, English courses are offered for undergraduate learners as level (fundamental) courses, a major, and a minor and for postgraduate learners in applied linguistics. English level courses have a focus on the four skills. English courses in the B.A. major and minor programs cover English linguistics, literature, and English for specific purposes (ESP). Besides, the M.A. program in Applied Linguistics has both theoretical and practical English for EFL teaching.

Previously, the phonetics teaching and learning are found in the B.A. and M.A. curricula in the fields of linguistics and applied linguistics, with the hope to help advance learners’ listening and speaking skills. However, other learners who are not majoring in English are introduced to some phonetic practices for the correct English pronunciation in the English level courses. However, this year, a new course in phonetics, added in the general education curricula, is running for the first time. This course aims to provide the basic knowledge in theoretical and practical phonetics to students in medical fields, so that they can use this knowledge in their future medical careers.

In this paper, the aim is to present three university courses based on theoretical and practical phonetics: (1) “Sound System of English” (SSE, thereafter) for B.A. learners in English; (2) “Scientific Experimental Phonetics” (SEP, thereafter) for learners in science and medical fields in international programs (M.D., M.Sc., DDS, etc.); and (3) “Speech Production and Perception for Language Teaching” (SPP, thereafter) for M.A. learners in applied linguistics.

2. METHODOLOGY

In this paper, the three courses: SSE, SEP, and SPP are compared and contrasted, in terms of background, aims and objectives, designs, learners, settings (time, place, and class size), textbooks and self-learning resources, evaluation, assessment, community services, and career preparation.

3. COURSES’ DESCRIPTION

The description of the three courses: SSE, SEP, and SPP is presented below.

3.1. Background

The three courses: SSE, SEP, and SPP have different backgrounds.

The first course (SSE) is a compulsory elective course in the curriculum of the B.A. program in English. Students can choose to take this course in
junior or senior years, after taking other linguistic courses such as Introduction to Linguistics and the English Language, English Morphology and Syntax, and English Phonetics.

The second course (SEP) has been newly added to the university general education curriculum for all science and arts students of the university in the international programs. The design of this course was inspired by the fact that phonetics can be applied to other fields such as medical fields, as in research and paper presentations delivered at the ICPHeS 2011 in Hong Kong.

The third course (SPP) is an elective course in the M.A. international program in Applied Linguistics. Previously, it was titled as “English Phonetics and the Teaching of Listening,” which was renamed to have its scope extended to cover both speech production and perception for language teaching.

3.2. Aims and objectives

The three courses: SSE, SEP, and SPP share the same aims and objectives in that all three courses provide the background in theoretical and practical phonetics, based on English. However, each course has a different focus, as described courses’ description below.

The course description of SSE is “English sound system; phonological theories and analysis; and systems of Englishes spoken around the globe.” The objective of this course is that learners know the English phonetics, understand phonological theories, and systems of Englishes spoken around the World.

The course description of SEP is “Different types of sounds in languages around the world, fundamental concepts of acoustic phonetics, as well as scientific experiments in terms of speech acoustics and speech perception.” The objective of this course is that learners understand the basic concept of human speech, the science of speech sounds, sounds of the World’s languages, and the application of the scientific phonetic knowledge to enhance their future careers.

The course description of SPP is “English phonetics and phonology and the teaching methods of English listening and speaking for EFL learners for an effective communication.” The objective of this course is that learners understand English phonetics and phonology and the teaching methods of English listening and speaking.

3.3. Courses’ designs

Each of the three courses has its own design, based on lecture and credit hours and structures and contents.

3.3.1. Lecture and credit hours

The three courses: SSE, SEP, and SPP have different lecture and credit hours, as in Table 1.

**Table 1: Lecture and credit hours**

<table>
<thead>
<tr>
<th>Items</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Course credits</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. No. of hours/week</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3. No. of lectures</td>
<td>15</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>4. Total hours</td>
<td>45</td>
<td>30</td>
<td>45</td>
</tr>
</tbody>
</table>

It can be seen from Table 1 above that both SSE and SPP courses share the same number of course credits (3 credit hours), the same number of lectures (15 lectures), and total hours (45 hours) per semester. On the other hand, the SEP course for learners in medical fields has only 2 credit hours, 10 lectures (3 hours per week), and 30 hours of total hours per semester.

3.3.2. Structures and contents

Structures and contents of the three courses are listed in Table 2 below. In the table, the letter ‘L’ stands for the word ‘Lecture.’

**Table 2: Structures and contents**

<table>
<thead>
<tr>
<th>Contents</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Fields in linguistics</td>
<td>L1</td>
<td>L1</td>
<td>L1</td>
</tr>
<tr>
<td>2. Phonetics vs. phonology</td>
<td>L2</td>
<td>L1</td>
<td>L1</td>
</tr>
<tr>
<td>(allophones vs. phonemes)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. IPA</td>
<td>L2</td>
<td>L2</td>
<td>L1</td>
</tr>
<tr>
<td>4. Speech organs</td>
<td>L2</td>
<td>L2</td>
<td>L1</td>
</tr>
<tr>
<td>5. The English sound system</td>
<td>L3</td>
<td>L4</td>
<td>L1</td>
</tr>
<tr>
<td>6. Consonants</td>
<td>L3</td>
<td>L4,</td>
<td>L3, 4</td>
</tr>
<tr>
<td>7. Vowels</td>
<td>L3</td>
<td>L6</td>
<td>L5, 6</td>
</tr>
<tr>
<td>8. Suprasegmentals</td>
<td>L4</td>
<td>L7</td>
<td>L7</td>
</tr>
<tr>
<td>9. Acoustics &amp; the Praat program</td>
<td>L12</td>
<td>L3</td>
<td>L11,12</td>
</tr>
<tr>
<td>10. Speech production</td>
<td>L14</td>
<td>L8</td>
<td>L2</td>
</tr>
<tr>
<td>11. Speech perception</td>
<td>L14</td>
<td>L9</td>
<td>L2</td>
</tr>
</tbody>
</table>
In Table 2, all three courses cover the basic contents such as background of linguistics, phonetics, phonology, IPA, speech organs, consonants, vowels, suprasegmentals, and the basic acoustics, including how to use the Praat program. What varies is the focus of each course during the second half of the semester.

In detail, the focus of the SSE course is on phonological theories; the one of the SEP course is on speech production and perception to be applied in learners' future medical careers; and the teaching design and practicum are for the SPP course.

### 3.4. Learners

Learners of the three courses come from different background, as in Table 3.

<table>
<thead>
<tr>
<th>Learners</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>20-22</td>
<td>18</td>
<td>&gt;22</td>
</tr>
<tr>
<td>2. Year of study</td>
<td>3, 4</td>
<td>1</td>
<td>Grad.</td>
</tr>
<tr>
<td>3. Nationality</td>
<td>Thai</td>
<td>Thai &amp; others</td>
<td>Thai &amp; others</td>
</tr>
<tr>
<td>4. L1</td>
<td>Thai</td>
<td>Thai &amp; others</td>
<td>Thai &amp; others</td>
</tr>
<tr>
<td>5. Majors</td>
<td>English</td>
<td>DDS, MD</td>
<td>Applied Ling.</td>
</tr>
<tr>
<td>6. Background in linguistics</td>
<td>&gt; 3</td>
<td>None</td>
<td>Vary courses</td>
</tr>
</tbody>
</table>

From Table 3, learners of the three courses come from various background (age, year of study, nationality, native languages (L1), majors, and background in linguistics).

### 3.5. Settings (time, place, and class size)

The settings of the three courses are time, place, and class size.

All three courses share similarities in terms of the duration of class per week (3 hours per week) and the small class size with up to 30 students per section.

As for the classrooms, the lectures of the SSE and SPP are in small lecture rooms designed for 30 students, but the SEP classes are located in a computer room. In the latter case, each student has one's own computer for doing hands-on activities during each lecture.

### 3.6. Textbooks and self-learning resources

The textbooks and self-learning resources are presented in Table 4.

<table>
<thead>
<tr>
<th>Resources</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ladefoged and Johnson [11]</td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>3. Odden [13]</td>
<td>√</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4. Goh [7]</td>
<td>-</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>5. Poedjosoedarmo</td>
<td>-</td>
<td>-</td>
<td>√</td>
</tr>
<tr>
<td>6. Abramson [1]</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>7. Albery et al. [2]</td>
<td>-</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>8. Baumann-Waengler [3]</td>
<td>-</td>
<td>-</td>
<td>√</td>
</tr>
<tr>
<td>10. Jotikasthira</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>[8, 9]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Ladefoged and Maddieson [10]</td>
<td>-</td>
<td>-</td>
<td>√</td>
</tr>
<tr>
<td>12. Naksakul [12]</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>13. Ohala [14]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>14. Roengpitya [16]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>15. Roengpitya [17]</td>
<td>-</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>16. Thongseiratch et al. [18]</td>
<td>-</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>17. UCLA [19]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>18. Yavaş [20]</td>
<td>√</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
From Table 4, the required textbook is Ladefoged and Johnson [12], and the books No. 2-5 are strongly recommended for students of all three courses. The other textbooks and resources (No. 6-18) are the materials that the instructor uses for the lectures and all are recommended readings for learners to advance their phonetic knowledge via self-study.

3.7. Evaluation

The evaluation criteria is shown in Table 5.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Attendance</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>2. Assignments</td>
<td>15%</td>
<td>15%</td>
<td>35%</td>
</tr>
<tr>
<td>3. Term Paper</td>
<td>20%</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>4. Midterm</td>
<td>30%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>5. Final</td>
<td>30%</td>
<td>30%</td>
<td>-</td>
</tr>
</tbody>
</table>

In Table 5, the evaluation includes attendance, examinations, a term paper, and other assignments.

3.8. Assessment

The assessment of the three courses includes the IPA transcription, theoretical and practical phonetics, and phonological analysis. The details are presented in Tables 6-7 below.

<table>
<thead>
<tr>
<th>Midterm</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. IPA</td>
<td>30%</td>
<td>30%</td>
<td>50%</td>
</tr>
<tr>
<td>2. Theory</td>
<td>50%</td>
<td>70%</td>
<td>50%</td>
</tr>
<tr>
<td>3. Phonological</td>
<td>20%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (points)</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final</th>
<th>SSE</th>
<th>SEP</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Theory</td>
<td>-</td>
<td>20%</td>
<td>-</td>
</tr>
<tr>
<td>2. Phonological</td>
<td>50%</td>
<td>50%</td>
<td>-</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Spectrogram</td>
<td>-</td>
<td>10%</td>
<td>-</td>
</tr>
<tr>
<td>Readings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Practical</td>
<td>-</td>
<td>20%</td>
<td>-</td>
</tr>
<tr>
<td>Phonetics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Practicum</td>
<td>-</td>
<td>-</td>
<td>100%</td>
</tr>
<tr>
<td>Total (points)</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

3.9. Community services

During the academic years 2010-2011, under the annual project of the community services run and supported by Faculty of Liberal Arts, M.A. learners of the SPP course participated in this project and created their own mini-phonetic classes to teach English pronunciation to students in Grades 5-6 in local schools near the university campus in Salaya. M.A. learners have used this experience. They have written this as their final term paper project and can express their own reflections towards the real phonetic teachings in the real setting.

3.10. Career preparation

In terms of career preparation, amongst the three courses, the SSE course will strengthen the phonetic knowledge for learners to use in their higher education and future career e.g., an English teacher. In addition, the phonetic knowledge of the SEP course will help learners understand normal speech and speech disorders of their future patients and to be inspired to invent novel techniques and instruments for a better quality of life of patients. Furthermore, the phonetic teaching and practicum of the SPP course assists learners’ future careers in EFL teaching.

4. DISCUSSION AND CONCLUSION

In Thailand, previously, the field in phonetics was limited to only courses in English or foreign language (FL) studies. However, this year (2015), phonetics teaching and learning at university level has been expanded to be included in the general education curriculum for students in medical fields. This can benefit learners to have a broadened worldview, along with their expertise in breadth, width, and depth.

To conclude, this paper presents how phonetics teaching and learning are embedded in three different university courses, designed for learners both in humanities and science. The three courses share the same contents for the first half of the semester, but have a different focus during the second half of the semester, based on learners’ background and career goals.

It is hoped that this paper will benefit interested individuals and will shed light on to future phonetic teaching and learning in both arts and science.
5. REFERENCES


[19] UCLA Phonetics Lab Data http://phonetics.ucla.edu/


ACKNOWLEDGMENTS

All comments and suggestions from the reviewers and from the PTLC participants are deeply appreciated. The author is grateful for a travel grant from Faculty of Liberal Arts, Mahidol University, for support from Assistant Professor Dr. Kanita Roengpitya, and the training from Professor John J. Ohala at the University of California at Berkeley. All errors are the author’s own.
USING FLIPPED LEARNING TO SUPPORT THE DEVELOPMENT OF TRANSCRIPTION SKILLS AMONG L1 AND L2 ENGLISH-SPEAKING STUDENTS

Jane Setter

University of Reading, UK
j.e.setter@reading.ac.uk

ABSTRACT
“Flipped learning” was used to spend less classroom time on theory and more time on the practice and development of broad phonetic transcription skills in undergraduate and postgraduate English phonetics and phonology lectures. The undergraduate classes comprise home students for whom English is their first language, whereas the postgraduate classes comprise largely overseas students for whom English is a second or foreign language. The data are for two separate year groups in each case and compare performance in two different transcription passages each used in flipped and non-flipped years.

The results show general improvement in the final dictation assessment for both groups in flipped years compared with non-flipped years, although there is more improvement for the undergraduates than for the postgraduates for one of the passages and, in the other, the postgraduate marks worsened.

This finding is discussed in terms of individual and group learning styles.

Keywords: Transcription; flipped learning; English phonetics and phonology.

1. INTRODUCTION
Technology-enhanced “flipped” learning [2] is a technique in which students watch a video or a series of videos based on lecture material at home in order to devote more classroom time to developing practical skills. It is particularly popular in science-based subjects such as Chemistry and Pharmacotherapy, which require a lot of laboratory-based work (see, e.g., [6]). Bishop and Verleger [3] provide a survey of research in this area and indicate that students are, in general, in favour of the approach, but that there is little work which objectively evaluates student learning outcomes.

Studies such as Ashby [1] have shown that flipped learning can be applied to the teaching of phonetics and phonology with favourable results, however. Ashby flipped two out of six topics which appeared as “free choice” questions on the final examination paper for her finalist students’ Phonology module, the remaining four topics being taught using traditional classroom methods. She found that, not only did more students select the flipped topics on the exam, they scored better in answering questions on those topics; this difference in scores was shown to be highly statistically significantly different. In addition, 62% of students reported that they felt they had learned more about the flipped topics than those taught in the traditional manner, and 57% of students thought the subject appropriate to be studied using a flipped approach.

Ashby’s experience with this methodology strongly influenced my decision to try it with my English Phonology classes.

2. CONTEXT
Undergraduate (UG) students entering their second year of study on the BA in English Language at the University of Reading have experienced tutorials in the first year. Second year modules are taught using “interactive lectures” lasting two lecture periods (i.e., 2 x 50 minutes) and not all topics commonly have tutorial support.

The 10-credit (5 ECTs) module English Phonology is a skills-based module on the programme which culminates in a broad phonetic transcription of English (Modern RP / Standard British Southern accent) as part of the assessment. However, constraints on the amount of time spent on practical work in these classes, owing to the necessity to cover theoretical aspects, has resulted in dissatisfaction for both tutor and students in terms of how much practical work students can expect to do under supervision.

Postgraduate (PG) students cover the same material as the UG students in a separate class. Unlike the UG students, who tend to be from the UK (largely from the south of England) and for whom English is their first language (L1), the majority of PG students on our MA programmes in Applied Linguistics and English Language Teaching are from overseas and are second language (L2) speakers of English. Often, no single L1 speaker group dominates; students are typically from Japan, China, Greece, Malaysia, Thailand, Arabic-speaking countries and, more recently, Kazakhstan.
Over the years, I have developed a series of weekly take-home exercises which aim to help students build practical skills in broad phonetic transcription. The solution for these exercises appears one week after the class on our virtual learning environment, Blackboard. Students are encouraged to post any questions or comments to the Discussion Board facility of Blackboard or raise them in class. However, my experience has been that there is very little engagement with these exercises – and certainly with the discussion board – unless they are used in class as a worksheet. There may be a number of reasons for this lack of engagement, including students’ concerns about being able to use metalanguage correctly on the discussion board in this rather technical subject; it has been noted in other studies that use of asynchronous discussion boards can vary in success (see, e.g., [5]).

In addition, students have requested extra tutorial sessions but, for various reasons, it has not been possible to provide these.

Having heard Ashby’s presentation on flipping phonetics in 2011, I saw this as an opportunity to devote more class time to the practical exercises and thereby build the students’ transcription skills. It was also hoped that the approach might support the learning of the PG students who are L2 speakers of English by giving them additional time to digest the lecture content. In this paper, I present data from two years of flipped learning (2013 and 2014) in comparison with data from two non-flipped years (2006 and 2011).

3. THE EXPERIMENT

Dictation to broad phonetic transcription begins in earnest in Week 5 of the 11-week module. It was decided to flip the lectures from Week 5 onwards.

Using Camtasia Studio’s [9] “record the screen” facility, a webcam (Hewlett Packard HD 270P) and a headset microphone (Creative HS-390), I produced vodcasts of the lecture slides, which are in PowerPoint format, along with audio narration and a video of my talking head, which appeared in the top-right of the screen.

I also took the opportunity to include other software and material from the internet which I thought might be of interest to students and/or enhance their learning. I did this for two reasons: firstly, to make links using YouTube videos with other areas students study on the degree (e.g., child language development; accent features) to show the point of studying the subject; and, secondly, to use free software such as UCL’s RT Pitch [8] to give a visual representation of the speech patterns students were learning about. RT Pitch actually works particularly well; the course text for English Phonology is Roach [7] and the intonation diacritics map very well to the visual representation of F0 which appears on the screen.

The vodcasts were edited so they lasted between 10-20 minutes each and saved in MP4 format. There were usually two to three vodcasts for each week. They were then posted to Blackboard one week prior to the class for students to view in their own time. Blackboard tracking statistics indicated that the majority of the students were watching the vodcasts prior to the lecture.

The following topics were covered in the vodcasts using this pedagogic approach: problems in phonemic analysis; aspects of connected speech (assimilation, coalescence, elision, linking and intrusive /r/); and intonation (three sessions).

A sample of the vodcasts can be viewed here: https://www.youtube.com/watch?v=kIAmqXF6bBE&list=PLBJxx2dHkl17FG8kD-LzWJfdFse69ImA5. If you do follow the link, you will see I attempted not to be too “tidy”, following Ashby’s comment that students could find the videos rather sterile.

In class, I went over the slides briefly in the first ten minutes, stopping to check for understanding of key points, before then using the material I had previously provided for take-home exercises as practical classwork, and also going over any of the issues arising from the exercises provided in the first four weeks which students wished to address.

The dictation test takes place in the final week of teaching (Week 11), and has the following format:

- The passage is ten tone units varying in length;
- Students hear the whole passage once;
- They then hear each tone unit ten times, after which they have five minutes to check their work;
- They then hear the passage a further three times;
- After one minute of final checking, students write their name on the script they are handing in and pass it to the tutor.

Students are expected to do the following:

- Produce a broad phonetic transcription according to the conventions in the text book;
- Indicate the onset to the head pattern and whether it is a high or low head using suitable diacritics;
- Underline the tonic syllable in each tone unit and allocate a suitable tone diacritic to it.

Marks are calculated out of 100. The marking scheme is as follows:
• One mark is deducted for each error up to a total of 10 errors per line;
• Stress errors are ignored, unless the student has failed to identify the onset to the head of the tone unit or put a stress mark on a syllable containing a weak vowel (happy or thank you) or single schwa vowel (in the latter case, two marks are deducted);
• If students have the right tone in the wrong place or vice versa they only lose one mark unless it conflicts with the previous point.

The success of this approach was evaluated using the dictation scores from years when the classes had been flipped (2013 and 2014) and years when they had not (2006 and 2011). I commonly recycle dictation passages over a period of three to four years and so comparative data was available; the passages from 2006 and 2013 are the “fire” passage (Dictation 1), whereas the passages from 2011 and 2014 are the “winter” passage (Dictation 2).

A total of 97 UG students (2013 = 61; 2014 = 36) and 54 PG students (2013 = 20; 2014 = 34) were taught using the flipped learning approach. The data is compared with 86 UG students (2006 = 28; 2011 = 58) and 54 PG students (2006 = 33; 2011 = 21) in non-flipped years.

4. RESULTS

Table 1 presents average scores along with standard deviation (SD) and number of participants (n); Figure 1 gives the scores in visual format with an indication of statistical significance.

Table 1: Average scores for UG and PG students on dictation tests in flipped and non-flipped years.

<table>
<thead>
<tr>
<th></th>
<th>UG</th>
<th>PG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-flipped</td>
<td>Flipped</td>
</tr>
<tr>
<td>Dict 1</td>
<td>60.6 (SD=16 n=28)</td>
<td>65.5 (SD=23.6 n=61)</td>
</tr>
<tr>
<td>Dict 2</td>
<td>66.5 (SD=20 n=58)</td>
<td>65.5 (SD=23 n=61)</td>
</tr>
<tr>
<td>Dict 2</td>
<td>65.5 (SD=23 n=61)</td>
<td>69.12 (SD=19 n=36)</td>
</tr>
<tr>
<td></td>
<td>60.6 (SD=16 n=28)</td>
<td>65.5 (SD=23.6 n=61)</td>
</tr>
<tr>
<td></td>
<td>66.5 (SD=20 n=58)</td>
<td>65.5 (SD=23 n=61)</td>
</tr>
<tr>
<td></td>
<td>65.5 (SD=23 n=61)</td>
<td>69.12 (SD=19 n=36)</td>
</tr>
<tr>
<td></td>
<td>60.6 (SD=16 n=28)</td>
<td>65.5 (SD=23.6 n=61)</td>
</tr>
<tr>
<td></td>
<td>66.5 (SD=20 n=58)</td>
<td>65.5 (SD=23 n=61)</td>
</tr>
<tr>
<td></td>
<td>65.5 (SD=23 n=61)</td>
<td>69.12 (SD=19 n=36)</td>
</tr>
<tr>
<td></td>
<td>60.6 (SD=16 n=28)</td>
<td>65.5 (SD=23.6 n=61)</td>
</tr>
<tr>
<td></td>
<td>66.5 (SD=20 n=58)</td>
<td>65.5 (SD=23 n=61)</td>
</tr>
<tr>
<td></td>
<td>65.5 (SD=23 n=61)</td>
<td>69.12 (SD=19 n=36)</td>
</tr>
<tr>
<td></td>
<td>60.6 (SD=16 n=28)</td>
<td>65.5 (SD=23.6 n=61)</td>
</tr>
<tr>
<td></td>
<td>66.5 (SD=20 n=58)</td>
<td>65.5 (SD=23 n=61)</td>
</tr>
<tr>
<td></td>
<td>65.5 (SD=23 n=61)</td>
<td>69.12 (SD=19 n=36)</td>
</tr>
</tbody>
</table>

Of immediate note is that, while the UG students’ scores improved across the board when flipped learning was adopted, the PG students’ scores improved in 2013 in comparison with 2006 but fell in 2014 in comparison with 2011.

Statistical analysis using t-tests in SPSS indicated that the UG scores taken as a whole were highly significantly different from the PG scores taken as a whole (p<0.01, t=3.068, df=287,1). When scores for each dictation passage were examined, the only statistically significant difference was between the UG and PG scores for flipped Dictation 2 (p<0.05, t=2.428, df=68,1). In addition, scores for flipped or non-flipped years were not significantly different within the UG or PG cohorts.

![Figure 1: Average scores for UG and PG students on dictation tests in flipped and non-flipped years.](image)

It is also worth noting that Dictation 1 (2006 and 2013) appears to be more difficult than Dictation 2 (2011 and 2014), certainly where the UG students are concerned and also looking at the non-flipped data for the PG students. There is no statistically significant difference here, however.

5. QUALITATIVE DATA

Although, unlike Ashby’s [1] study, there is no data directly comparing students’ impressions of flipped learning with traditional methods, I do have comments from students’ module feedback questionnaires. The response was extremely positive from the UG students and tentatively positive from the PG students. There is no statistically significant difference here, however.

Here is a sample of the positive comments:

- I think this format is perfect for this subject.
- Flipped learning allows students to watch your lectures in their own time, at their own pace.
- I found it useful to be able to stop the vodcast/lecture at trickier points and either watch the slide again, take more notes or do
some reading around the point for further information.

- [...] we were able to spend time discussing any questions people had [in class] and then move onto transcription practice for the upcoming exam.
- I would much rather do the transcriptions in class with the support of classmates and the lecturer than alone, at home.
- Flipped learning is brilliant!

6. DISCUSSION

While it is good to see the UG students and one group of PG students’ scores improving on the dictation assessment following the adoption of flipped learning for part of the module, it is disappointing to note that the 2014 PG group did worse than the 2011 PG group in the same dictation passage, and also that the PG scores are lower across the board. I had hoped that the flipped classroom might support L2 PG students’ learning; the current University of Reading teaching policy includes strong support of the incorporation of technology-enhanced learning (TEL), and flipped learning is very much seen as part of this.

There are, in my opinion, a number of possible reasons why the PG groups did not perform as well as the UG groups overall, which include: learning styles / cultures and diversity; exposure to technology; age; studying in a non-L1 environment; and general group ability.

While the UG group is largely homogeneous, the PG group is diverse in a number of different ways, one of which is learning style/culture. For Chinese students or others from an East or Southeast Asian background, philosophies such as Confucianism consider the teacher to be an unquestionable authority and the expectation is that learning will take place in a traditional classroom setting. It is possible that the flipped learning approach is simply not traditional enough for many students in the class.

In addition, the UG students have all grown up with technology and are therefore “native users” of it. The PG students, some of whom are considerably older than the UGs, may not have had this experience and may, therefore, be less able to adapt to it; TEL in the form of flipped learning may not necessarily be the optimum form of tuition for these students.

The challenge of learning through the medium of a language which is not one’s L1 is also an issue. While closer examination of the dictation scripts would be required to make any definite claims, many L2 PG students have much more difficulty with English vowels than the L1 UG students. There is, therefore, the whole issue of L1 transfer to consider when undertaking an exercise such as this. In the IPA Certificate, L2 students’ transcriptions are marked slightly differently to take this into account; here, they are not.

Finally, I look at group ability. To do this, and as a very rough comparator, I have the scores from a different but similarly technical module for comparison: English Grammar. The 2014 PG group’s average score for English Grammar was 63.53%, whereas the 2013 group’s score was 62.5%. As far as that comparison goes, it would seem that ability in Grammar does not predict ability in Phonology – not a surprise, perhaps. It will be interesting to look at the overall degree scores of the 2014 group in comparison with the 2013 group once they have finished the programme, as it is certainly the impression of staff that the 2013 group was more studious. What it is possible to say is that the standard deviation scores indicate that there was more variability between students in the 2014 group (SD = 22.4) in comparison with the 2013 group (SD 20.1), although not by much.

7. CONCLUSION

Focus on the UG results and the 2013 PG group give support to Ashby’s [1] results that flipped learning can be beneficial in the teaching and learning of phonetics and phonology.

Ashby’s study looked at performance in essay questions, whereas this paper looks at dictation scores. A study comparing the two types of assessment in a single phonetics and phonology class could give insights into whether improvements are seen in both assessment types or whether flipping is more suited to supporting learning in practical assessments.

8. ACKNOWLEDGEMENTS

I would like to thank all the students who took part. This research was supported by a University of Reading Partnerships in Learning and Teaching (PLanT) award.

9. REFERENCES

TEACHABILITY AND LEARNABILITY OF ENGLISH TONICITY FOR JAPANESE JUNIOR HIGH SCHOOL STUDENTS

Yusuke Shibata*, Masaki Taniguchi**, Tamikazu Date***

Kochi University*, Kochi University**, Poole Gakuin University***
yusukeshibata@kochi-u.ac.jp*, tamasaki@kochi-u.ac.jp**, t.date@kcc.zaq.ne.jp***

ABSTRACT

Pronunciation teaching is largely neglected in English classrooms in Japan, nonetheless intonation teaching. It may be due to the teachers’ lack of confidence and the complexity of English intonation. The aim of this study is to examine the teachability and learnability of English intonation, specifically tonicity, in junior high school English classrooms in Japan. According to an experiment conducted, it is therefore safe to say that tonicity is teachable and learnable in English classrooms in Japan. The next step in this line of research is to examine how we can train teachers to teach intonation.

Keywords:
English tonicity, teachability, learnability, Japanese learners, nuclear accents

1. INTRODUCTION

Pronunciation teaching in Japan faces a number of problems. A serious one is that the NS intelligibility of English words uttered by Japanese university students is about 50% [1]. This is caused by the lack of adequate pronunciation training provided in junior and senior high schools. According to a survey conducted with Japanese university students [2], 93.1% of them said they had either no or little English pronunciation training in junior high school English classrooms and 93.0% in senior high school.

One of the reasons why such training is not provided is the teachers’ lack of confidence [3]. There are no guidelines as to what to teach and how to teach it when it comes to pronunciation as opposed to vocabulary teaching. Even the Course of Study for Junior High Schools: Foreign Languages (English) published by MEXT [4] does not give accurate guidelines on pronunciation teaching. It is not clear what is actually meant by “Pronouncing English sounds correctly, paying attention to such basic phonetic features as stress, intonation, and pauses.” Intonation has a very complex system. This paper focuses on intonation amongst all the phonetic features, and specifically on tonicity, and aims to examine whether tonicity is actually teachable and learnable in Japan’s junior high school English classrooms.

2. IMPORTANCE OF TONICITY IN ENGLISH

Tonicity (nuclear tone placement) is an important system of intonation that can determine the meaning of an utterance. Knowledge and practical ability with regards to tonicity can affect one’s intelligibility significantly [5]. [5] says, “As concerns intonation, speakers of English repeatedly face three types of decision as they speak. They are: how to break the material up into chunks, what is to be accented, and what tones are to be used. These linguistic intonation systems are known respectively as tonality, tonicity and tone. We refer to them as the three Ts.” [5] also says that tonicity is probably the most important of all of the systems of intonation.

3. METHODOLOGY

3.1. Subjects

The subjects were junior high students from two different schools in Nagasaki Prefecture, Japan. A total of 76 students (44 male and 32 female) in their first to third years (i.e., 7th to 9th grades, ages 12 to 15) participated in the experiment involving a lecture and a test. All speak Japanese as their first language with no record of hearing impairment.

3.2. Material

3.2.1. Examples used for demonstrations and practice

The following nine sets of dialogues were used to practise how a nuclear tone, which is a high fall, changes locations, depending on the context.

Dialogue 1
A: Do you play /sports/?
B: No, but I watch /baseball/.

Dialogue 2
A: Do you play /baseball/?
B: No, but I /watch/ baseball.
Dialogue 3
A: Do you like /animals?
B: I like small /dogs.

Dialogue 4
A: Do you like /dogs?
B: I like small /dogs.

Dialogue 5
A: Does your father speak /English?
B: No, but my father speaks Ko/\orean.

Dialogue 6
A: Does your mother speak Ko/\orean?
B: No, but my /father speaks Korean.

Dialogue 7
A: Do you have a /car?
B: No, but I /want a car.

Dialogue 8
A: Did you have a hot /summer?
B: We had a /cool summer.

Dialogue 9
A: Did you take the /picture?
B: My /sister took the picture.

3.2.2. Test materials

The following nine sets of dialogues were used to test how well the subjects could recognize different locations of nuclear placement marked by a high fall tone. The students were asked to listen to each set of pre-recorded question and four responses to it and choose the one with correct nuclear placement. Prior to the experiment, the dialogues were recorded with the Audio-Technica AT2035 connected to the Zoom H6 at a sample rate of 44.1khz and a 16 bit depth. The following were the nine questions followed by four responses to each:

Q1: Did you see the /man?
A: My /sister saw the man.
B: My /sister saw the man.
C: My /sister saw the man. (correct)
D: My /sister saw the /man.

Q2: Is that your new /bike?
A: It’s my /old bike. (correct)
B: It’s my /old bike.
C: It’s my old /bike.
D: It’s my old /bike.

Q3: Does your mother teach /science?
A: My /mother studies science.

Q4: Do you have a red /car?
A: I have a blue /car.
B: I have a blue car.
C: I have a blue car.
D: I have a blue car. (correct)

Q5: Can you write Ko/\orean?
A: I can read Ko\orean. (correct)
B: I can read Ko\orean.
C: I can read Ko\orean.
D: I can read Ko\orean.

Q6: Does your brother play the pi/\ano?
A: My sister plays the pi\ano.
B: My sister plays the piano. (correct)
C: My sister plays the piano.
D: My sister plays the piano.

Q7: Does your mother have long /hair?
A: She has short hair.
B: She has short hair. (correct)
C: She has short hair.
D: She has short hair.

Q8: Does your mother cook /dinner?
A: My father cooks dinner.
B: My father cooks dinner. (correct)
C: My father cooks dinner.
D: My father cooks dinner.

Q9: Do you like your new /school?
A: I love my new school.
B: I love my new school.
C: I love my new school.
D: I love my new school. (correct)

The vocabulary and grammar in all of the dialogues above were carefully examined so that even the first year junior high school informants would have no difficulty understanding the meanings. In cases where they had not been taught certain vocabulary and grammar, extra explanation was provided, such as that the word “saw” is the past tense of “see” and that “Korean” refers to the language spoken in Korea.

3.3. Lesson procedures

Prior to the listening experiment, the authors provided an explanation of intonation by comparing it with music, by referring to it as the melody of
speech. The authors then gave a lecture on the basics of English tonicity in which the centre of focus was the nuclear accent placements in different contexts. After demonstrating the first six sets of dialogues (3.2.1), the students were asked to perform them by themselves. Three further sets of examples (3.2.1) were used to reinforce the learning. WASP [6] (a free program for the recording, display and analysis of speech) was used to show a visual display of the changes in the fundamental frequency. The authors also used broad hand gestures and jumping that accompanied the speech, showing the location of the nucleus as well as the movement of the pitch.

3.4. Listening test

Towards the end of each fifty-minute long lesson, a listening test was conducted to assess the students’ understanding and recognition of the nuclear placement. The short dialogues (3.2.2) were played from the Fostex PM 0.1 loud speaker connected to Zoom H6. Subjects listened to each question and four possible responses (as shown above) as they marked their answers on a sheet of paper provided. They were asked to choose the best response out of the four choices which only varied in their nuclear placement. After their answer sheets were collected, the correct answers and explanations were provided to the students.

4. RESULTS

Graph 1: Results of the Tonicity Listening Test: Average Percentages of Correct and Incorrect Answers given by Students of Different Genders

Graph 2: Results of the Tonicity Listening Test: Average Percentages of Correct and Incorrect Answers given by Students of Different Grades

Graphs 1 to 3 above show the results of the test, in which Q1 to Q9 refer to the nine sets of questions in 3.2.2; Year 1 refers to students in the first year of junior high school, Year 2 second year, and Year 3 third year; and Male refers to the male students, and Female the female students.

5. DISCUSSION

According to the results shown in all the graphs, the overall percentage of correct answers was 71.8%, which may not be too high but was certainly higher
than the authors’ expectations, which were less than 50%, because we had been told by the teachers of the students that they had had no specific knowledge or training in tonicity. They expressed surprise when they heard the results of the current experiment. There was no significant difference in the ability between male and female students, either, as shown in Graph 1.

When the data was divided into three groups of students according to their ages, as shown in Graph 2, the average percentage of correct answers given by the first-year students was 70.0%; that of the second-year students was 77.8%; and that of the third-year students was 76.9%. There was no significant difference among them, which proves that these students had similar levels of knowledge and recognition competence at the time of the experiment of the current research.

It proves that tonicity is certainly teachable and learnable for junior high school students in Japan. The students seemed to understand the explanation of tonicity quite well in the first half of the session when the authors displayed and demonstrated physically the nucleus placement in the dialogues provided. The authors used both WASP [6] and broad hand gestures to show the location of the nucleus as well as the movement of the pitch. One of the teachers of the students commented that even though some of the students did not do too well in general English exams, they seemed to catch on very well in this experiment. It is quite surprising that tonicity is not generally taught in English classrooms in Japan, in spite of the fact that, according to [5], it is one of the most essential elements of intonation that determine the meaning of the utterances and that it is one of the weakest points for Japanese and some other foreign learners of English.

6. CONCLUSION

The aim of this study was to examine the teachability and learnability of English intonation, specifically tonicity, in English classrooms in Japanese junior high schools. According to our experiment conducted, the results were in the affirmative, although not perfectly so. The percentage of the students’ understanding and recognition of nucleus placement was over 70%, which was much higher than the expectations of the authors and their teachers. It is therefore safe to say that tonicity is teachable and learnable in junior high school English classrooms in Japan if properly introduced. The next step in this line of research is to examine how we can train teachers of Japanese junior high schools to teach intonation, specifically tonicity. The teachers of the junior high school students who participated in the experiment observed how the authors attempted explaining tonicity to them, using hand gestures as well as PowerPoint slides that showed the location and movement of the nuclear tones. It is hoped that junior high school teachers, by observing this kind of experiment and English intonation class that the current authors conducted for this paper, have also learned how to teach tonicity and realized the teachability and learnability of tonicity. More preferably, such courses as the UCL Summer Course in English Phonetics should be attended by all teachers and prospective teachers of English, and UCL phoneticians should be invited to give lectures in Japan more often.

7. ACKNOWLEDGEMENTS

This study was made possible with the aid of JSPS Kakenhi Grant Number 25580132, a research project led by Prof Akiyo (Funatsu) Joto, to whom the authors’ sincerest gratitude is due.

8. REFERENCES

[4] MEXT (2008) COURSE OF STUDY for Junior High Schools Foreign Languages (English)
[6] WASP (a free program for the recording, display and analysis of speech) http://www.phon.ucl.ac.uk/resource/sfs/wasp.htm
PERCEPTION OF FOUR GENERAL BRITISH VOWELS BY SLOVENIAN UNIVERSITY STUDENTS OF ENGLISH AS A FOREIGN LANGUAGE

Andrej Stopar

University of Ljubljana, Slovenia
andrej.stopar@ff.uni-lj.si

ABSTRACT
The paper presents a study on the perception of four General British (GB) vowels – the DRESS-vowel /e/, the TRAP-vowel /æ/, the STRUT-vowel /ʌ/ and the NURSE-vowel /ɜ/ – by 101 Slovenian university students of English as a Foreign Language. The analysis includes the students’ performance, the differences between the strong and the weak performers, and the respondents’ most frequent misperceptions. The results show that production difficulties mirror perception difficulties; that the misperceptions of L2 sounds are predictable; and that there is a correlation between the least successfully perceived GB vowels and the ones most typically heard as interchangeable. While the findings are mostly of pedagogical value for teachers of GB pronunciation, they are also used to address the potential impact of a recent change in the GB vowel notation in Gimson’s Pronunciation of English.

Keywords: perception of vowels, GB vowels, Slovenian vowels, vowel notation

1. INTRODUCTION AND LITERATURE OVERVIEW

1.1. Perception of Vowels

Successful foreign language (FL) learners recognize (and produce) the vowels of their target language reliably and consistently. Frequently, they are expected to acquire the ability to discriminate between contrasts that do not exist in their native language (L1). To achieve this, they still predominantly rely on their existing linguistic background – that is, on the phonetic system of their native language.

Some authors, for instance Flege [4], suggest that the phonetic categories of L1 are fixed, which prevents FL learners from creating new categories that are based on phonemic features. Hence, in Flege’s framework, a phoneme that does not exist in L1 can form a new category only when based on some non-phonemic feature of the FL. A similar dependence on L1 is described by Best [2], who suggests that the new phonemes can be assimilated in different ways, while relying on one (or more) phonemic categories provided by L1.

Regardless of the approach, research has shown that it is possible for FL learners to learn to perceive new contrasts (see, for instance, Escudero’s [5] study on the contrast between the Scottish /t/ and /l/). Also noteworthy are the research findings that establish positive correlations between (vowel) perception and (vowel) production – see Bion, Escudero, Rauber, and Baptista [3] and Baker and Trofimovich [1].

1.2. Slovenian and the General British Vowel Systems Compared

Following various studies conducted by Komar, Petek, Srebot-Rejec, and Šuštaršič, see [7], [8], [9], [10], and [11], it can be established that one of the fundamental differences between the GB and Slovenian vowel systems is in the number of vowel phonemes. The two vowel systems are presented in Figure 1.

In [9: 9], Šuštaršič observes that for each of the GB vowel pairs /i: - i, u: - o, a: - æ/ Slovenian ‘only has a (relatively) close front and a (relatively) close back vowel /i, u/ and an open central vowel /a/, which are all relatively long in accented and relatively short in unaccented position’.

As may be expected, the lack of a number of contrasts in the Slovenian vowel system represents a
challenge for Slovenian learners of English. A typical example is the opposition of Slovenian phonemes /el/ and /ɛl/, on the one hand, and the GB phonemes /æ/ and /æ/ (now transcribed as /æ/ in Cruttenden [4]) on the other. Despite being close to the GB close-mid open-mid /æ/ and the open-mid to open /æ/, the Slovenian close-mid and open-mid front vowels /el/ and /ɛl/ are only rarely used as counterparts. Suštaršič suggests that this is due to Slovenian /el/ being ‘perceived by Slovene speakers as too close to be applied for the English /æ/’ [9:10]. This proximity may result in perception problems that have been observed in student transcriptions (see [9:83]), or in both GB vowels being produced as the Slovenian /ɛl/, which is an undesired intrusion of Slovenian features into English since ‘it interferes with the maintenance of contrasts carrying a high functional load’ [4:335-336].

The present study explores whether the same phenomenon is mirrored in the perception of the two GB vowels. The experiment is extended to the GB vowels /æ/ and /æː/: the central open /æ/ is studied because of its relative closeness to the increasingly more open /æ/ (see [4:122]), and the central mid /æː/ because of its closeness to the centralized and often slightly raised /æ/.

2. THE STUDY

2.1. Participants

The study included a total of 101 participants. All participants were first-year BA students of English at the Faculty of Arts, University of Ljubljana, with a similar background in foreign language instruction: in order to enrol in their BA programme they had to pass the national secondary school-leaving examination in English, which is aligned with the Common European Framework of Reference for Languages as a B2-level examination. Their ages ranged from 18 to 23, the average being 19.1 (σ=0.8, μ1/2=19). 78.2% of the participants were female, which is representative of the gender structure at the department. The study only included the responses of the participants who declared their first language as Slovenian.

In December 2014, when the study took place, the participants had already been familiarized with the IPA symbols for GB via eight 45-minute lectures and the same number of (practical) seminar classes in English pronunciation.

2.2. Instruments and Procedure

The participants were asked to complete a questionnaire which consisted of two sections. In the first part they were required to indicate their age, gender and mother tongue. The second section was a task that asked them to listen to the recordings of seventeen monosyllabic words and to identify the vowels embedded in these words.

The participants were informed that some of the words were real words, while others were non-sense words. Also, their choice of vowels was limited to four GB phonemes, the DRESS-vowel /æ/, the TRAP-vowel /æ/, the STRUT-vowel /æː/, and the NURSE-vowel /æː/. The relevant vowels were clearly presented with their IPA symbols and corresponding examples (monosyllabic, high-frequency words containing the mentioned vowels; both the spelling and the transcription were given).

The words used in the task were listed using IPA transcription, with a gap in the place of the vowel that the participants were asked to identify (for instance, ‘Word 1: /k ____ t/’). The list included two real words and two non-sense words for each of the vowels mentioned above. Among the four words representing each of the studied vowels, two ended in a voiced, and two in a voiceless consonant. To avoid any speculation by the participants on the potentially balanced distribution of the vowels in the task, an additional, seventeenth, non-sense word was added to the list, which contained the TRAP-vowel.

The words in the DRESS set included /gret/, /med/, /streɪ/, and /hæk/. The TRAP-vowel words were /bæʊ/, /dæp/, /kænl/, /læg/, and /mepl/. The STRUT set included /læd/, /kæv/, /kæp/, and /næm/. And, finally, the words with the NURSE-vowel were /bæd/, /læːk/, /plæːp/, and /æːgl/.

The recordings were made by a native speaker of English with training in phonetics. The words were played three times before the questionnaires were collected.

The responses of the participants were collected and analysed using Microsoft Excel.

2.3. Research Questions

RQ1: How do Slovenian university students of English as a foreign language perform in their perception of GB vowels /el/, /ɛl/, /æ/, and /æː/?

RQ2: Are there differences between the strong performers and the weak performers with regard to the perception of the observed vowels?

RQ3: What are the most frequent substitutes for the wrongly identified vowels?

RQ4: Are there any differences in the participants’ perception of real and non-sense words?
2.4. Results

The average performance of the participants was at 71.1%. Table 1 shows their performance with regard to the individual GB vowels (RQ1).

Table 1: Perception of /æ/, /ɛ/, /ʌ/ by Slovenian students of English (N=101)

<table>
<thead>
<tr>
<th>GB Vowel</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ɜ/</td>
<td>86.1</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>74.0</td>
</tr>
<tr>
<td>/æ/</td>
<td>65.8</td>
</tr>
<tr>
<td>/æ/</td>
<td>61.0</td>
</tr>
<tr>
<td>Average</td>
<td>71.1</td>
</tr>
</tbody>
</table>

Focussing on the strong performers, specifically on the top third of the participants (N=33), the average for the correctly perceived vowels is 87.6%. The weak group of students (N=33), in contrast, achieved an average result of 54.7%. Despite the significant difference in the groups’ average performance (see RQ2), the facility values for their perception of the four vowels exhibit the same rankings, with /æ/ being the most challenging and /þ:/ the easiest to perceive. The comparison between the two subgroups of participants also shows that the results of the strong performers are noticeably more equally distributed (see Table 2).

Table 2: Perception of /æ/, /æ/, /ʌ/, and /þ:/ by the strong (N=33) and the weak performers (N=33)

<table>
<thead>
<tr>
<th>GB Vowel</th>
<th>Strong Performers (% Correct)</th>
<th>Weak Performers (% Correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ɜ/</td>
<td>93.2</td>
<td>72.7</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>89.4</td>
<td>63.6</td>
</tr>
<tr>
<td>/æ/</td>
<td>84.1</td>
<td>48.5</td>
</tr>
<tr>
<td>/æ/</td>
<td>83.6</td>
<td>34.0</td>
</tr>
<tr>
<td>Average</td>
<td>87.6</td>
<td>54.7</td>
</tr>
<tr>
<td>σ</td>
<td>4.6</td>
<td>17.1</td>
</tr>
</tbody>
</table>

Table 3 presents the most frequent substitutes that the participants chose for the wrongly perceived vowels (RQ3). Particularly noticeable is that the least successfully perceived vowels (/æ/ and /æ/) are also the ones that are most typically heard as interchangeable.

Table 3: The most frequent substitutes for the wrongly perceived vowels /æ/, /æ/, /ʌ/, and /þ:/ (N=101)

<table>
<thead>
<tr>
<th>GB Vowel</th>
<th>Substitute (% of All Respondents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ɜ/</td>
<td>/lo/ (7.2)</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>/æ/ (17.8)</td>
</tr>
<tr>
<td>/æ/</td>
<td>/æ/ (24.5)</td>
</tr>
<tr>
<td>/æ/</td>
<td>/lo/ (20.8)</td>
</tr>
</tbody>
</table>

With regard to the differences in the perception of the real and non-sense words (RQ4), it can be concluded that the perception of vowels in non-sense words was consistently more challenging for the participants – the average facility values were 67.0% for non-sense words, and 75.7% for real words. This is additionally confirmed by examining the data for individual vowels: in all but one case, the most poorly perceived words were always non-sense words (the exception being the word <cut> /kʌt/, which had the lowest facility value among the four STRUT-words).

3. DISCUSSION

The results confirm that the frequently observed production difficulties Slovenian students have with the vowels /æ/ and /æ/ are also mirrored in the perception of these phonemes. While the performance of the weakest students is somewhat surprising for university students of English (merely 34% correct for /æ/), the relatively poor result is predictable considering the differences between the Slovenian and GB vowel systems. This contrast is additionally quantified and illustrated by the inclusion of /ʌ/ and /þ:/ in the study – the latter two vowels are more readily assimilated into the existing Slovenian vowel categories, so the students performed noticeably better in perceiving them.

The examination of the wrongly perceived vowels further supports the prediction that the GB vowels /æ/ and /æ/ are frequently perceived/produced as interchangeable by Slovenian speakers. Also expected is the substitution of the GB central vowel /ɜ/ with the relatively centralized /ɜ/. The more interesting result, though, is the one that points to /ɜ/ being frequently mistaken for /æ/. An explanation can be found in Cruttenden [4], who mentions that the tendency of /æ/ and /ɜ/ to become more open leads to the occasional neutralization of the two vowels [4: 120].

From a pedagogical perspective, however, this last result is worrying – the confusion of /æ/ and /ɜ/ combined with the change in notation (from /æ/ to /lo/) that was introduced in the latest edition of Gimson’s Pronunciation of English (see [4]) could lead to a new set of challenges for Slovenian teachers and learners of GB pronunciation and transcription. Namely, the new notation may be interpreted by the Slovenian university students as one requiring from them to
consistently produce the more open variant of /æ/.
And this, in turn, may cause them to encounter
difficulties in the perception of /æ/ as uttered by those
GB speakers who still produce it as an open-mid to
open vowel. In other words, the often discussed
Slovenian problem of distinguishing between <pet>
/pet/ and <pat> /pat/ may now become the problem of
distinguishing between <pet> /pet/, <pat> /pet/,
and <putt> /pi:t/.

4. CONCLUSION

The study presents how Slovenian students of English
perceive four GB vowels embedded in monosyllabic,
non-sense and real words. The findings confirm the
significant role that the L1 vowel system plays in
learner acquisition of FL vowels: the existing L1
vowel categories can both aid as well as distract the
learners on their way towards proficient FL
pronunciation.

The average result for the successful perceptions of
GB vowels /æ/, /æ/, /ai/, and /i:/ is at 71.1%. The
TRAP-vowel is the most difficult for L2 learners to
perceive (61.0%), followed by the STRUT-vowel
(74.0%), the DRESS-vowel (65.8%), and the NURSE-
vowel as the easiest to perceive (86.1%). Comparing
the differences between the strong performers and the
weak performers, it can be established that the order
of difficulty for the observed vowels is the same for
both groups. Furthermore, the analysis also
demonstrates that there is a correlation between the
least successfully perceived GB vowels (/æ/ and /ɛ/)
and the ones most typically heard as interchangeable.
And finally, the study demonstrates the impact of
word frequency (or the lack of it) on the perception of
vowels: non-sense words were consistently perceived
less successfully.

The results of the research can also be considered
a practical tool for Slovenian teachers of English as
an FL, since they indicate which of the observed GB
vowels should be given additional classroom
attention. Moreover, the very experiment presented in
the article can be used for teaching purposes: the
identification of FL vowels without reliance on
orthography is a frequently recommended teaching
method (see, for instance, [4: 347]). Complemented
by other teaching techniques, as well as formal
instruction in phonetics and phonology, perception
exercises can help FL learners achieve the perception
performance levels similar to the ones observed in the
strongest members of our sample.

With the above in mind, the presented research
should be extended to other GB vowels and their
production.

5. REFERENCES

[1] Baker, W., Trofimovich, P. 2006. Perceptual paths to
accurate production of L2 vowels: the role of individual
differences. International Review of Applied Linguistics in
Language Teaching, 44(3), 231-250.
speech perception. In Strange, W. (ed.). Speech
Perception and Linguistic Experience: Theoretical and
Methodological Issues. Baltimore: York Press, 171-
203.
Category formation and the role of spectral quality in
the perception and production of English front vowels.
Pittsburgh, United States.
Contrasts: Acoustic Cue Reliance in the Development of
New Contrasts. New Sounds 2000: Proceedings of
the Fourth International Symposium on the Acquisition of
Second Language Speech (ed. by A. James and J.
theory, findings and problems. In Strange, W. (ed.).
Speech Perception and Linguistic Experience:
Theoretical and Methodological Issues. Baltimore:
York Press.
Analysis of Contemporary Vowels of the Standard
Slovenian Language. Fourth International Conference
on Spoken Language Processing, October 3-6,
duration in Standard Slovene. An acoustic and
linguistic investigation. Munich: Verlag Otto Sagner.
phonetic and phonemic analysis and its application in
Teaching English phonetics and phonology. Ljubljana:
Znanstveni inštitut Filozofske Fakultete.
qualities of English and Slovene. Studies in Contrastive
Linguistics. Proceedings of the 3rd International
Contrastive Linguistics Conference, Santiago de
Compostela, September 2003, 601-608.
of the IPA: Slovene. Journal of the International
Phonetic Association 25/2. Dublin: IPA.
AN ANALYSIS OF JAPANESE JUNIOR HIGH SCHOOL TEXTBOOKS AS PRONUNCIATION TEACHING MATERIALS

Junko Sugimoto; Yoko Uchida

University of the Sacred Heart, Tokyo; Tokyo University of Marine Science and Technology
sugimoto@u-sacred-heart.ac.jp; uchidayo@kaiyodai.ac.jp

ABSTRACT
As a preliminary step towards conducting a teacher survey on pronunciation teaching in order to design a phonetics course for current and prospective English teachers in Japan, this paper analyses six government-approved English textbooks used in junior high schools. The purposes are to examine the types of pronunciation activities in the textbooks and to discuss the knowledge of phonetics required for Japanese teachers of English. The analysis reveals trends such as an emphasis on suprasegmental features and the incorporation of phonics. Although details of the textbook activities vary greatly, most of the textbooks cover core features of the English sound system and provide a satisfactory number of pronunciation exercises. To make effective use of textbooks as pronunciation teaching materials, however, teachers need to have a solid grounding in phonetics.

Keywords: pronunciation teaching, phonetic symbols, phonics, suprasegmentals, teacher training

1. INTRODUCTION

Pronunciation teaching is one of the biggest challenges that Japanese teachers of English face. Morley explains that a teacher needs to help students as a coach, and just like coaches in other areas such as debate, drama, voice, music, and sports, a pronunciation/speaking coach ‘supplies information, gives models from time to time, offers cues, suggestions and constructive feedback about performance, sets high standards, provides a wide variety of practice opportunities, and overall supports and encourages the learner’ [5 p. 507].

To fulfil all these roles, teachers need to be equipped with a variety of skills and knowledge. They need to know practical techniques and strategies as well as understand the sound systems of the target language and the learners’ first language.

This paper analyses six government-approved English textbooks currently used in Japanese junior high schools. Compared to the past editions, the current textbooks seem to have more exercises. The aims of the analysis are to categorise and assess the pronunciation-related activities that are included in the textbooks and to identify the knowledge of phonetics that teachers need to possess in order to teach pronunciation more effectively and with confidence.

Our ultimate goal is to describe the status quo of English pronunciation teaching in junior high school classes in Japan, investigate teachers’ concerns and needs, and design a phonetics course specifically tailored for current and prospective English teachers.

2. METHOD

The following is the exhaustive list of junior high school English textbook series approved by the Ministry of Education, Culture, Sports, Science and Technology of Japan [4] in 2011, and to be used for the school years 2012–2015:

1. New Crown (Sanseido); hereafter NC-1, 2, 3
   Takahashi, S., et al. 2011
2. New Horizon (Tokyo Shoseki); NH-1, 2, 3
   Kasashima, J., Seki, N. et al. 2011
3. One World (Kyoiku Shuppan); OW-1, 2, 3
   Matsumoto, S., et al. 2011
4. Sunshine (Kairyudo); SS-1, 2, 3
   Matsuhata, K., Sano, M. et al. 2011
5. Total English (Gakko Tosho); TE-1, 2, 3
   Yada, H., Yoshida, K. et al. 2011
6. Columbus 21 (Mitsumura Tosho); C21-1, 2, 3
   Togo, K. et al. 2011

Each series consists of three volumes, corresponding to the three grades in Japanese junior high school.

As for the choice of an English variety as a model, the only government guideline [4] is to adopt a ‘contemporary standard pronunciation’, and no specific variety is recommended. Yet, all the textbooks seem to be in favour of American English, and accordingly, General American is adopted.

The authors thoroughly collected and compiled the descriptions relevant to pronunciation in the textbooks. Each item was then categorised and given a label according to its nature: For example, to a description involving the distinction between the vowels in hot and hat, a sound pair many Japanese learners of English have difficulty with [8], the label LOT-TRAP (based on Wells’ lexical sets [10]) was assigned. Some of the labels used are as follows:

Vowels: FLEECE-KIT, START-NURSE
Consonants: /l/-/b/, /t/-/d/, /s/-/f/, clear/dark /l/
Phonics: long/short vowels, silent letters, the two sounds of c and g
Connected speech: linking, elision, assimilation
Stress and rhythm: content/function words, word stress, compounds, rhythm
Intonation: tone unit, tone type, nucleus

All the textbooks have explanations and/or demonstrations of letter-sound correspondences based on phonics rules, which play a certain role in pronunciation teaching in Japan. For this reason, descriptions pertaining to phonics are included in the analysis even though phonics is not a phonetic phenomenon. Based on the results, the specific elements of phonetic knowledge that teachers are expected to know were extracted.

When referring to explanations given in the textbooks (usually in Japanese) in the following discussion, the authors provide literal translations in italics without any modification of the symbols to indicate pronunciation. For example, in all six textbooks, brackets ([ ]) instead of slashes (/ /) are used along with phonemic symbols, most of which are in accordance with the IPA. However, small capitals are not used (e.g. cake [k戒k] for /kék/), presumably so as not to confuse the students with unfamiliar symbols.

3. RESULTS

3.1. Segments

Having a five-vowel system in their mother tongue, Japanese learners of English often experience difficulties in distinguishing/identifying English vowels. Yet, the approach to the problem differs among textbooks. While two textbooks (OW, C21) deal with all the vowel categories at least once in their activities, one textbook (TE) focuses on a limited number of vowels (e.g. FLEECE, FACE, MOUTH). Two textbooks (NH, SS) present about 10 difficult minimal pairs (e.g. STRUT-LOT, GOOSE-FOOT, SQUARE-NEAR) in their activities. Overall, very little is mentioned about weak vowels or the difference between strong and weak vowels (only once in NC, NH, and OW, respectively).

As for consonants, certain ones (e.g. /l/, /v/, /ð/, /s/) are repeatedly dealt with whereas others (e.g. /p/, /ʃ/, /ə/, /m/) are rarely mentioned. In addition to the problematic consonant pairs /ʃ/-/s/ and /θ/-/v/, voiceless-voiced pairs such as /ʃ/-/v/, /θ/-/ð/, /p/-/b/, and /f/-/v/ are dealt with in the majority of the textbooks even though they are not difficult for Japanese speakers to distinguish. As for allophones, three textbooks explain the difference between clear and dark /l/ (NC, SS, C21), but no mention is made of /l/-voicing or allophones of voiceless stops.

Three textbooks (NC, NH, C21) attempt a phonetic explanation of difficult consonants by making use of the sagittal section of the vocal tract and/or by providing articulatory explanations, taking a sound in Japanese as a reference point (e.g. to pronounce /r/, say /b without the tip of the tongue touching the roof of the mouth). Although all the textbooks explain basic phonics rules [3] such as the pronunciation of one-letter consonants, short and long vowels, many of the two-letter consonants (e.g. ch, ng, sh, th) and vowels (e.g. ea, ey, ow, oo), and the two sounds of c and g, the thoroughness of these descriptions differs. The most extreme case is TE, in which all the explanations of segments are devoted to showing which spelling corresponds to which sound along with some sample words (e.g. ea [i:] season, sneaker, eaten; ow [au] how, flower, allow; aw [a:] draw, saw, law; a-e [ei] cake, name, face), while no information is provided on the phonetic features or articulation of the sounds.

Even in other textbooks where the description of phonics and phonetics is more balanced, the level of detail varies. For example, in explaining the two sounds of c and g, rules are explicitly indicated in NC: Look at the sample words carefully. 'C' and 'g' are pronounced as [s] and [dʒ], respectively, when followed by either 'e', 'i', or 'y' [NC-2 p. 116]. On the other hand, in all the other textbooks, only sample words are listed and no rules are mentioned.

3.2. Connected speech

All six textbooks emphasise the importance of linking, especially in the following two positions: a consonant followed by a vowel (e.g. group_of, look_at) and a consonant followed by another consonant that shares the same place of articulation (e.g. with_them, went_to). An alveolar nasal followed by a vowel (e.g. one_of, can_I) is one example of a problematic case for Japanese learners since they tend to replace a word-final /n/ with a moraic nasal.

Among the sound changes, coalescent assimilation seems to be the most favoured, and it is treated in four of the textbooks (NC, NH, SS, C21). One textbook (NH) uses katakana transliteration (e.g. an egg アネッガ, meet you ミーチュー). Three textbooks (NH, OW, C21) introduce elision, but its explanation needs improvement. For example, a laterally released word-final plosive is included as an example of elision (e.g. good luck).
The explicit presentation of the rules of sound changes is limited. Usually, the textbook instructions tell learners to simply focus on how the words are joined together to keep the rhythm, and they only urge learners to practise.

3.3. Stress and rhythm

To signal word stress, all six textbooks use the standard method of Japanese dictionaries and teaching materials, that is, putting a stress mark directly on the vowel symbol or letter (e.g. guitár [ɡɪtɑ́ːr]). Popular exercises include pattern-matching (e.g. music ɡɪtɑ́ː, arrive ɡɪtɪv) (NC, OW) and comparison of the stress patterns of English words with those of corresponding loanwords written in katakana words (NH, SS). The treatment of other features, such as compound stress and stress shift, is fairly limited.

Although the wording and amount of explanation vary, four textbooks (NC, SS, OW, C21) mention that English words can be divided into two groups: words that tend to receive stress and words that do not. Rhythm is one of the most strongly emphasised features. All of the textbooks try to visually represent the contrast of strong and weak syllables in English rhythm (e.g. Nice to meet you ɡɪtɪv ɡɪtɪv ɡɪtɪv). However, the allocation of strong syllables is not always consistent, and there are questionable examples such as placing unnecessary stress on personal pronouns.

None of the textbooks use technical terms such as ‘stress-timing’ or ‘content/function words’. Also, although each circle in word stress or rhythmic pattern-matching exercises corresponds to a syllable, no textbooks mention the notion of the syllable.

3.4. Intonation

In regard to intonation, all the textbooks explain the relationship between sentence type and the choice of tone: for example, the use of a falling tone in wh-questions and rising tone in yes-no questions. Three textbooks (NC, NH, OW) explicitly present the rules, whereas the others only give examples and expect students to deduce the rules.

The practice of dividing long sentences into tone units is introduced in five textbooks (NC, NH OW, SS, C21), but the number of exercises is insufficient. Only one textbook (NH) mentions that tone units correspond to information and grammatical units, whereas the others simply present sentences with boundary symbols and ask students to pronounce them accordingly.

The placement of the nucleus is another important aspect of intonation, but its treatment is also restricted. Four textbooks (NC, NH, SS, C21) include a few exercises on this topic, and three (NH, SS, C21) explain that the nucleus should be placed on words that represent ‘important information’ or ‘the word that answers the question’.

4. DISCUSSION

4.1. Phonetic symbols

Although a phonetic transcription accompanies every newly introduced word, it is merely supplementary information about how the word sounds, and there are no activities on the use of the phonetic symbols in any of the textbooks. This may be partially because the writing system of Japanese is not alphabetic and many Japanese learners find it difficult to interpret what each symbol represents.

An alternative way to show which letter(s) represent which sound(s) is the phonics system, which has recently gained more popularity in Japanese classrooms, possibly because there is no intervention of another set of symbols and the relationship between spelling and sound is more direct and explicit.

Care must be taken in the use of phonics, however, because there are a number of words that do not follow the rules, especially at the introductory level (e.g. the pronunciation of g in girl and get). Teachers have to be able to tell rule-conforming words from exceptions.

4.2. The use of katakana

The use of katakana is peculiar to teaching materials targeting Japanese learners. In an analysis of the previous editions of the textbooks, NC was reported to use katakana transcription alongside the phonetic transcription in both the body of the textbook and the appendix [9].

In the current editions of the textbooks, however, katakana is only sporadically used for supplementary purposes. For example, four textbooks (NC, NH, SS, C21) enable learners to compare the pronunciation of English words with corresponding loanwords written in katakana, and one textbook (NH) uses katakana characters when giving an articulatory description of English sounds that do not exist in the Japanese sound system, and when explaining sound changes (see Sections 3.1 and 3.2). The overall use of katakana has decreased.

4.3. Focus on suprasegmentals

The textbooks provide a substantial number of exercises for suprasegmental features. This reflects the government’s Curriculum Guidelines, which clearly state that students need to understand and
become familiar with the basic characteristics of English such as stress, intonation, and phrasing to enhance their English speaking and listening skills [4].

The guidelines also specify the following points to be incorporated in the teaching of English sounds: (i) contemporary standard pronunciation, (ii) sound changes that result from linking, (iii) basics of stress patterns in words, phrases, and sentences, (iv) basics of intonation in sentences, and (v) basics of phrasing in sentences [4]. Four of these points are concerned with pronunciation at the sentence and suprasegmental levels. The importance of teaching suprasegmental features has been emphasised in the past few decades [1] [5], and the shift of focus from segments to suprasegmentals has also influenced the government guidelines in Japan.

Among the content on suprasegmental features, more exercises on nucleus placement should be included, given that it is considered one of the most important features in communication [2]. Teachers need to fully understand the rules of English rhythm and intonation to be able to guide students through these exercises.

4.4. Activity types

Among the six textbooks, variations are observed in the ways the same phonetic phenomenon is approached. In explaining /l/-/tʃ/, for example, one textbook (OW) simply lists words containing the two consonants, while another textbook (NH) attempts step-by-step instructions for learning the difference, pointing out that the distinction does not exist in Japanese and introducing a tongue twister as an exercise. Yet another textbook (NC) goes one step further by inviting students to work in pairs [NC-1 p. 73]:

Listen carefully to the ‘l’ and ‘r’ in the following words, and practise. Work in pairs and find out what is going on with the tip of the tongue when pronouncing the two sounds:

like, leg, live, leave, right, red, river, really

NC is characteristic in that its activities often contain a process of guiding students to figure out phonetic differences that might otherwise be too minute for native speakers of Japanese to notice, and to discover the rules themselves.

5. CONCLUSION

This study investigated six government-approved textbooks for junior high schools in Japan. It revealed that the majority of the textbooks cover the elements in phonetics necessary to enable students to understand the basic characteristics of English sounds, with an emphasis on suprasegmentals, although the presentation of materials varies.

Needless to say, teachers are the key to students’ successful understanding. Teachers need to have sufficient knowledge in order to effectively use the textbooks as pronunciation teaching materials.

Unfortunately, however, many junior and senior high school teachers in Japan have little background in phonetics since it is not a required subject in teacher training curricula at Japanese universities [7]. Each university can decide whether to include it in the curriculum or not. As a consequence, a phonetics course may be optional or obligatory, and its length may be one semester or a whole year. In this situation, it can be assumed that quite a few teachers teach in the classroom without the necessary knowledge and skills.

The next step of our project is to conduct a questionnaire survey on a considerable scale. Teachers currently teaching in public junior high schools will be asked to participate in the survey, which will ask questions about the time they spend teaching pronunciation, their understanding of the phonetics elements identified in the present study, and concerns they may have in pronunciation instruction.

Once the results of the survey are obtained and the teachers’ needs are clarified, we will design a phonetics course for prospective English teachers and teachers who have not previously learned about phonetics. The course will first cover the basic concepts of phonetics and basic characteristics of English in comparison with Japanese. Without this knowledge, teachers will not be able to facilitate students’ practice in an effective and beneficial way, nor theoretically understand the problem areas for Japanese learners. For example, to teach the exercises in accent and rhythm, teachers need to understand the notion of the syllable and the differences between English and Japanese syllable structures. The course will also provide training in practical skills such as recognising students’ problems, giving constructive advice, and evaluating student progress.

Phonetics is indispensable for English teachers since it helps them guide student learning and improve their own pronunciation skills so as to provide the students with a good model. The authors strongly hope that the benefits of learning phonetics will be properly acknowledged by people involved in English education in Japan.

6. REFERENCES


CROSS-LINGUISTIC PERCEPTION OF ENGLISH /p/-/b/ IN NOISE BY TWO LEVELS OF JAPANESE LEARNERS OF ENGLISH

Shinichi Tokuma¹, Won Tokuma²

¹Chuo University, Tokyo, 192-0393, Japan
²Seijo University, Tokyo, 157-8511, Japan
tokuma@tamacc.chuo-u.ac.jp

ABSTRACT
This study reinvestigates the results of two previous studies on the effect of different types of clear speech on cross-linguistic perception with varying levels of noise, and examines how the level of English proficiency affects the perception. An English minimal pair beach-peach, extracted from two types of clear speech in the LUCID corpus, is presented to the two levels of Japanese learners of English for identification at various SNR levels. The results endorse the findings of the previous studies and show that a higher proficiency in English has a positive influence on perception.

Keywords: L2 perception in noise, L2 proficiency

1. INTRODUCTION
When an L1 speaker talks to an L2 listener who has a low command of the target language, the speaker employs a special mode commonly called ‘clear speech’ to facilitate the communication, by slowing down the speech rate, increasing the pitch range or trying to pronounce each sound clearly. The anecdotal experience tells us that this strategy works, and this is indeed the case, supported by numerous studies, such as [2, 8, 9]. In particular, [8] investigated the perception of English minimal pairs /p/-/b/ and /s/-/ʃ/ by L1 and L2 speakers of English, and showed that when presented in noise, words produced to help L2 interlocutors were perceived more easily by them than those for L1 interlocutors to counteract babble noise. Furthermore, their subsequent study [9] demonstrated that there were differences in the robustness to the noise among speech sounds: English /p/ was more robust to babble noise than English /b/ when they were perceived by native speakers of Japanese, which is attributed to the strong influence of the acoustic characteristics of Japanese /b/.

However, do these results hold for all native speakers of Japanese? In these studies, Japanese university students with lower-intermediate level English participated in the perceptual experiment and none of them had studied English intensively or lived in an English-speaking country. Provided that students with a higher proficiency of English participate in the same or a similar experiment, do they produce different results because of their higher level proficiency, or do their results stay the same due to the strong intervention of their L1, Japanese? [7] investigated the perception of English F0 peak-delay by Japanese intermediate and advanced learners of English, and it was shown that the difference in results was noticeable but small, suggesting that there was a strong influence of Japanese. It is yet to be examined, however, whether these findings on prosody can be applicable to the perception of phonemes.

Furthermore, there is another issue to be addressed regarding [8, 9]. The original babble noise employed there was obtained from [10], but although its SNR was varied, the segment taken was from the identical position in time. As the spectral characteristics of babble noise is not flat in the frequency axis, the perceptual results can be influenced by the particular spectral peak and/or valleys of the noise segment. Hence, in order to consolidate the findings, it is now necessary to re-examine whether the results of [8, 9] can be reproduced with babble noise that is extracted from another temporal position and therefore has a different spectral characteristics.

The present study investigated the issues mentioned above by administering a perceptual experiment. The speech materials used were identical as those used in [8, 9]: tokens that were intended to combat the L2 condition and babble noise (of the L1 condition) taken from LUCID corpus [1]. However, in this experiment, three segments of babble noise were taken from a different time position. High-intermediate Japanese learners of English participated in the experiment.
The detail of the experiment is described in the following section.

2. EXPERIMENT

2.1. Materials

In the LUCID corpus, a diapix task was employed, where two speakers in a separate room talked to each other to spot the differences in the pictures via headsets with a microphone. The speakers A and B in a pair completed the task in three adverse speech conditions: when the speaker A’s voice was passed through a three-channel noise-excited vocoder, or it was mixed with multi-talker babble noise (Babble), and the speaker B was a low-proficiency non-native English speaker (L2). The speakers were all native speakers of Southern British English, except in the L2 condition, where the speaker B’s were native speakers of Chinese (Mandarin/Taiwanese) or Korean with a low English command.

In this study, materials in the L2 condition were selected, as well as those in the Babble condition, whose L1 perception data is presented in [5]. The materials were taken from the speech of the speaker A. The tasks of the LUCID were designed to elicit a set of minimal or near-minimal pairs with sounds /b/-/p/ and /s/-/ʃ/. In this study, one minimal pair, beach-peach was chosen as stimuli to compare the results with the previous studies [8, 9]. Also, tokens from two speakers (one male, one female) were selected for stimuli as in [8, 9]. Since the speakers in the LUCID corpus completed either the tasks of the L2 condition or the Babble condition [1], the words of the minimal pairs came from the identical speaker for each L2 / Babble condition, but not across the two conditions. Each token for the individual minimal-pair word was extracted from the corpus, and in a few cases where there was more than one token that were suitable for stimuli, one of them was selected randomly. All tokens were normalised to a fixed intensity level, using Speech Filing System (ver. 4.10), and then were presented as the stimuli with babble noise.

The original noise, 100-talker babble noise, was taken from [10] at three different temporal positions. Wide-band and narrow-band sound spectrograms of each extracted noise were obtained and examined, and they showed that these noise tokens were not identical in frequency characteristics. These babble noise segments were then manipulated by Speech Filing System, and added to beach-peach tokens at five levels of SNR: -15dB, -10dB, -5dB, 0dB and +5dB. This process produced 60 stimulus tokens for each stimulus condition (L2 / Babble): 2 token types (beach-peach) x 2 LUCID speakers x 3 babble noise types x 5 SNR levels (-15dB / -10dB / -5dB / 0dB / +5dB).

2.2. Participants

Two groups of first-year Japanese undergraduate students at the Faculty of Commerce Chuo University in Tokyo, aged 18-20, participated in the experiment, with 7 and 8 members respectively. The participants were tested altogether in the same Language Laboratory room (see details below). None of them had reported on hearing or language impairment. They were all monolinguals. They had studied English as a foreign language for at least six years at school. They had all taken an on-line English exam called GTEC, which was used to streamline English classes at the Faculty. Their GTEC scores were between 280 and 325, which can be converted to TOEFL iBT scores of 72-84 or B2-C1 level in CEFR scale. Therefore, their English abilities were judged to be at the upper-intermediate / lower-advanced level. They were not paid for their participation.

2.3. Procedure

Because the task of this study was L2 perception as in [7, 8], the participants were instructed to identify the whole word to reduce their burden.

The participants were tested in a quiet Language Laboratory room at Chuo University, at the same time, with stimuli played through covered-ear headphones at a level adjusted by them. None of them reported that their attention had been diverted by noise, or by the presence of other participants. Each group was assigned with either of two blocks of stimuli. Specifically, one group (7 members) listened to the stimuli in the L2 condition, while the other group (8) listened to those in the Babble condition.

The participants were seated individually in front of laptop PCs, and their task was to listen to the stimulus token, at a different level of SNR, through headphones and, using a mouse, click the correct word in a minimal pair that appeared on the screen, as quickly and accurately as possible. The locations of the words on the screen, i.e. whether a word appeared on the right or left, were also randomised, and each stimulus token was played in a random order to the participants twice. In total,
there were 120 presentations per participant: 2 token types (beach-peach) x 2 LUCID speakers x 3 babble noise types x 5 SNR levels (-15dB, -10dB, -5dB, 0dB and +5dB) x 2 button positions. The whole experiment process was controlled by Praat (ver. 5.4), utilising Experiment MFC objects.

The experiment was preceded by the task demonstration by one of the authors designed to make the participants familiar with the experimental setting and the nature of the stimuli. Care was taken to ensure that the participants knew that they must not click a word on a screen before the whole word was played, through instructions before and after the demonstration.

2.4. Results

As the first stage of the analysis, the mean percentages of correct responses were calculated for each token type (beach-peach), stimulus condition (L2/Babble) and SNR level (-15dB / -10dB / -5dB / 0dB / +5dB).

Figures 1 to 4 show the mean percentages of the correct responses for each token type. Figures 1 and 2 are for the L2 condition (i.e. the interlocutor was an L2 speaker in the LUCID corpus) and Figures 3 and 4 for the Babble condition (i.e. the speaker’s voice in LUCID was originally mixed with babble noise). In the Figures, the results of the current study are shown in red as ‘High E’, while the results of the lower-intermediate students from the previous study [9] are also displayed in blue as ‘Low E’.

Overall, Figures 1 to 4 demonstrate a high rate of correct responses when the token was presented at a lower level of babble noise. However, as the SNR increases, the percentages of correct responses decrease for all tokens although the decrease rate is much smaller for the peach token. The Figures also show that for Japanese L2 listeners, perception of English /p/ is more robust to noise than the perceptual scores of the initial /b/ regardless of the stimulus condition, and that on the whole, the words produced to help L2 interlocutors are perceived more easily by them than the words produced for L1 interlocutors to counteract speech babble noise. These results are in concordance with the previous studies, which is a strong proof to endorse the finding of [8, 9].

Furthermore, Figures 1-4 show the difference in the perceptual score according to the English proficiency level of the participants: in Figures 1 and 3, the mean percentages of correct identification scores for /b/ are higher, which suggests that for Japanese upper-intermediate / lower-advanced learners of English, perception of /b/ is more robust to noise than for their lower-intermediate counterparts. This observation is also supported by the higher mean correct identification score of /p/ in +5dB tokens of Babble condition. In Figure 4, the identification rate of the current participants is 88 % while the rate drops to 77 % for the lower-intermediate participants.
3. DISCUSSION AND IMPLICATIONS ON TEACHING

The results of the experiment demonstrate that although the higher level of English proficiency has little influence on prosodic perception, as shown in [7], the difference in the level of English proficiency affects the perception of English /p/ and /b/: it has a positive impact on the segmental perception, especially for /b/. This implies that the reverse effect of Japanese, their L1, on their segmental perception can be reduced. The improvement is impressive: as mentioned before, in [5], the perception of /p/-/b/ in noise by native English speakers was already tested in the almost identical condition (tokens were presented in the +0dB Babble noise condition), and the average rate of correct responses was 89% for /p/-/b/, and the perceptual rate is 90% in our experiment.

The results also endorse the results of the previous studies, even if a different masking noise was utilised as a masker.

Furthermore, if you look at the backgrounds of the participants in detail, it offers an interesting insight: none of them had lived or studied abroad, and their English command had been built in Japanese L2 classrooms. This means that, without intensive or specific ear training, if the students are motivated and receive usual amount of general English input, they can make good progress in listening to unfamiliar L2 sounds (i.e. English /b/ that sounds like /p/ in Japanese).

This is far from obvious because, as all the experienced EFL teachers are aware, teaching segmental listening is not straightforward in many circumstances. Also, in realistic L2 classrooms, teachers do not have sufficient time for intensive sessions of ear-training, and even if they do, it is often not very effective: as [3] state, ‘adult listeners find it very difficult indeed to improve their performance … although huge amounts of training can bring about some improvement.’ (p.72)

The fact that this English /p/-/b/ contrast is relatively easy to overcome with general input from ordinary classroom teaching shows that teachers with limited resources and time can easily motivate beginner or pre-intermediate students and improve their perception of English sounds. More difficult and time-consuming phonetic contrasts such as /d/-/l/ are to be dealt with at the advanced level, since the pre-intermediate learners would find the contrasts impossible to make and this could discourage them to acquire the correct phonetic knowledge of English sounds.

4. REFERENCES

PHONETICS INSTRUCTION IN L2 FRENCH: CONTRIBUTIONS OF SEGMENTS, PROSODY AND FLUENCY TO SPEECH RATINGS

Pavel Trofimovich¹, Sara Kennedy¹, Josée Blanchet²

¹Concordia University, Montreal, Canada; ²Université du Québec à Montréal, Canada
pavel.trofimovich@concordia.ca; sara.kennedy@concordia.ca; blanchet.josee@uqam.ca

ABSTRACT

This study focused on the speech of 30 adult learners of French as a second language (L2) in a 15-week pronunciation course, investigating the relationship between instruction and listener-based ratings of accent, comprehensibility, and fluency before and after instruction in read-aloud and extemporaneous (picture description) speaking tasks. Results showed that the learners improved in all speech ratings, especially in extemporaneous speaking. Results also revealed that accent ratings were linked to prosody (intonation accuracy, pitch range), while fluency and comprehensibility ratings were additionally linked to fluency phenomena (length of fluent run, hesitation rate). We discuss implications of these findings for L2 pronunciation learning and links between instruction, listener-rated dimensions of speech, and performance in different tasks.

Keywords: French, speech rating, prosody, accent, fluency, comprehensibility, pronunciation teaching.

1. INTRODUCTION

Although second language (L2) pronunciation can no longer be described as a neglected aspect of teaching and research, many questions still remain for learners of languages other than English about the links between instruction and development. For example, many textbooks for L2 French cover pronunciation (e.g. [1]), yet little is known about the development of French pronunciation in instructed learners. Teachers and learners thus rely on intuition, course materials, and past experience to guide their teaching and learning. This study aimed to fill this gap by investigating pronunciation development of adult L2 French learners over time, with the overall goal of contributing to the knowledge base about how L2 French learners’ pronunciation development is linked to pronunciation instruction.

1.1. Pronunciation instruction in L2 French

Prior research on L2 French pronunciation has often used one-time measures of learner speech, targeting learner proficiency [2] or different learning contexts [13]. To date, the few longitudinal studies have been set in university contexts [5, 11]. For example, [8] tracked the use of liaison (as in mes amis ‘my friends’ spoken as [me-za-mi]) over one year in a weekly three-hour French language and literature class at a Korean university. The learners showed a significant increase in production of obligatory and optional liaison in word pairs. Yet tying learning gains to the nature of instruction is problematic in this case because the instruction was not described. To sum up, instructional research on French pronunciation presently provides limited evidence of instruction–pronunciation links, because there are few studies and most provide little methodological detail, precluding a clear interpretation of findings.

1.2. Motivation for the current study

To address these shortcomings, we recently explored the effectiveness of phonetics teaching in a 15-week L2 French course targeting segments, prosody, fluency, and connected speech processes (e.g. enchaînement) [10]. We found improvements in learners’ segmental and intonation accuracy, use of enchaînement, pitch range, and number of hesitations. However, it was unclear whether the reported gains in segmental accuracy, prosody, and fluency are linked to L2 speech characteristics which are perceptible to listeners. Therefore, in this study, we revisited the data from our original study to examine the impact of phonetics instruction on listener-based ratings of accent (native-likeness), comprehensibility (ease of understanding), and fluency (smoothness of speech delivery) in L2 French speech by the learners before and after instruction. The research questions were:

1. Do L2 French learners improve in listener-based ratings of accent, comprehensibility, and fluency following phonetics instruction?
2. Which segment, prosody, and fluency aspects of learner speech are associated with these listener-based ratings?

2. METHOD

2.1. Participants

The participants were 30 adult learners of L2 French (23 women) in an intermediate-level listening and speaking course at a French-medium university in Quebec, Canada. The learners, with a mean age of
35.8 years (27-52) and a mean length of residence in Quebec of 3.2 years (0.3-10), came from diverse language backgrounds: Mandarin (11), Russian (7), Farsi (3), Cantonese, Portuguese, Spanish (2 each), Korean, Malay, and Romanian (1 each).

### 2.2. Instruction

The 15-week listening and speaking course met once per week for three hours, with about one hour devoted to practice in a multimedia lab. The instructor was a native speaker of Quebec French with a graduate degree in applied linguistics and 12 years of teaching experience. The instruction, tar-geting speaking and listening, focused on segmental and suprasegmental aspects of French. The main focus was on connected speech processes, which included enchaînement and liaison (defined below), and on developing fluency and prosody through work on phrasal stress (rhythmic groups) and into-nation. For enchaînement and liaison, the emphasis was on comprehension but learners were encouraged to produce them through practice. For phrasal stress and intonation, the emphasis was on fluid delivery of speech, with practice involving both controlled output recorded in the lab and guided tasks (e.g. practice a scene from a play). In a typical pedagogic sequence, each topic was covered in one class meeting and reviewed during the following class. Each meeting started with a discovery activity, followed by the teacher’s explanation of the targeted aspect, then by controlled practice. The learners then practiced the targeted aspect through communicative and fluency tasks (e.g. role plays, shadowing). Lab-based dictation or production tasks involved short sentences illustrating the targeted aspects.

### 2.3. Tasks and procedure

Learner production was analyzed in two tasks, used at pre-test and post-test. The first task was a read-aloud story (163 words), which involved an ex-change between a woman standing in a ticket line and a man who wanted to cut into the line. The dialogue between the two characters (five turns, nine sentences) was preceded and followed by a three-sentence narrative. The sentences were about 10-15 words long (M = 11 words), and 90% of all vocabulary were among the first 1000 most frequent words in French [6]. The second task was an oral picture description based on an eight-panel image sequence. The task, used widely in the elicitation of spontaneous L2 production [7], featured two people who bumped into each other on a street corner, accidentally exchanged the identical suitcases, but realized their mistake only later.

The tasks were administered twice, in Week 3 as a pre-test and in Week 15 as a post-test, using the same equipment, instructions, and procedure. The learners recorded their speech in a multimedia lab using interactive software [4]. For the read-aloud task, they received a copy of the text and had a 2 min period to review the text. They had 150 s to re-cord the text. For picture description, the learners received a copy of the picture story entitled *Erreur sur la valise* (Suitcase mixup), to contextualize the story’s central element. They then had 2 min to re-view the images and 5 min to record their narrative.

### 2.4. Speech ratings

Pre- and post-test excerpts (20 s) were rated by 20 native French listeners (28.2 years old; 13 females) for accent (1 = accent marqué ‘heavy accent’, 9 = pas d’accent ‘not accented’), comprehensibility (1 = difficile à comprendre ‘hard to understand’, 9 = facile à comprendre ‘easy to understand’), and fluency (1 = pas du tout courant ‘dysfluent’, 9 = couramment ‘fluent’). Listeners (students in linguistics, education, or psychology, with no formal training in L2 phonetics) were first given definitions of each construct, then rated five practice files. They worked at their own pace, playing each consecutive file and recording their ratings in the booklet, with replays permitted. Listeners showed high consistency (Cronbach’s alpha) for accent (a = .89–.93), comprehensibility (a = .93–.96), and fluency (a = .94–.97), so mean scores were computed per speaker by averaging across listener ratings for each rated construct.

### 2.5. Speech measures

The audio recordings of both tasks, considered along with the transcripts, were then analyzed for seven measures reflecting the course aims and content:

1. **Segmental errors**: single segment additions, deletions, or substitutions (e.g. *tu ‘you’* spoken as *tou ‘all’*), and spelling-based mis-takes (e.g. *coup ‘hit* spoken as *coupe ‘cut’*). The measure was a ratio of all segment errors over all words produced.

2. **Intonation errors**: inappropriate pitch moves, with rising or flat contours signalling clo-sure; and falling or flat contours used for an expected signal of continuity (e.g. *Chacun son tour, monsieur! ‘Wait for your turn, sir!’ with *tour* spoken with a falling pitch). This measure was a ratio of the total number of inappropriate intonation contours produced over the total number of expected contours.

3. **Enchaînement use**: a successful consonant-to-vowel or vowel-to-vowel link (e.g. *il a ‘he*
has’ becomes [i-la] and tu as ‘you have’ becomes [ty-a]). Enchaînement required an expected rise-fall pattern signalling a word boundary (i.e. il ʁɑ̃tɑ̃-iʎɛʁ ‘he will go’) and had to involve a continued, unbroken phonation (absence of perceptible pausing). This measure was a ratio of the total number of successfully produced enchaînement over the total number of contexts where such links could potentially occur.

4. Liaison use: obligatory liaison between two words (e.g. between personal pronouns and verbs, determiners and nouns). Liaison had to involve a proper grammatical context (e.g. nous avons ‘we have’ [nu-za-vɔ̃]), accurate phonetic realization of the linked consonant (e.g. [d] produced as [t] as in grand homme ‘tall man’ [ɡʁɑ̃-tuɑ̃m]) and an expected rise-fall pattern signalling a word boundary, with no perceptible pausing between words. This measure was a ratio of the total number of successfully realized obligatory liaison out of the total number of contexts for obligatory liaison in each learner’s production.

5. Pitch range: difference between highest and lowest fundamental frequency (F0) values, extracted from a pitch tracker display [4]. This measure was to capture the degree of pitch range for each learner, in absolute terms, on the assumption that narrower pitch ranges characterize flat, monotonous delivery and wider ranges describe lively, animated speech (see [14]).

6. Mean length of run (MLR): mean number of syllables produced between two adjacent filled or unfilled pauses of 400 ms or longer, following [12].

7. Speech hesitations: all dysfluencies such as filled and unfilled pauses (e.g. ils vont… euh… au travail ‘they are going… hmm… to work’, where ‘…’ = unfilled pause), and pauses inside a rhythmic group (e.g. il y a un… panneau ‘there is a… sign’). This measure was a ratio of the total number of hesitations over all syllables in each sample.

All measures were coded by native-speaking trained coders. Agreement reached 98-100%.

3. RESULTS

3.1. Speech ratings across task and time

We first compared listener ratings through ANOVAs, with task (read-aloud, picture) and time (pretest, post-test) as repeated measures. For accent, there was only a significant effect of time \(F(1, 29) = 5.84, p = .022\), with learners improving in accent, albeit modestly, in both tasks (Fig. 1).

**Figure 1:** Accent (error bars = ±1 SE).

For comprehensibility, there was a significant task \(\times\) time interaction \(F(1, 29) = 5.33, p = .028\), such that the learners improved in comprehensibility only in the picture task (Fig. 2).

**Figure 2:** Comprehensibility (error bars = ±1 SE).

For fluency, there were significant effects of task \(F(1, 29) = 20.33, p < .001\) and time \(F(1, 29) = 5.97, p = .021\), with greater fluency in the read-aloud than picture task and an increase in fluency, mostly in the picture task (Fig. 3).

**Figure 3:** Fluency (error bars = ±1 SE).
3.2. Linguistic contributions to ratings

We then explored contributions of the seven speech measures to listener ratings via partial correlations, carried out between each rating set (accent, fluency, comprehensibility) and each speech measure at post-test, with the relevant pre-test measure partialled out. By controlling initial performance, we examined the extent to which each speech measure was related to listener ratings at the end of the course.

**Table 1:** Partial correlations between accent ratings and individual speech measures from the post-test, with the relevant speech measure from the pre-test partialled out (*p < .05, **p < .01).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Read-aloud</th>
<th>Picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmental errors</td>
<td>−.26</td>
<td>−.19</td>
</tr>
<tr>
<td>Intonation errors</td>
<td>−.42*</td>
<td>−.43*</td>
</tr>
<tr>
<td>Enchaînement</td>
<td>−.06</td>
<td>−.04</td>
</tr>
<tr>
<td>Liaison</td>
<td>.13</td>
<td>−.07</td>
</tr>
<tr>
<td>Pitch range</td>
<td>.26</td>
<td>−.36*</td>
</tr>
<tr>
<td>MLR</td>
<td>.18</td>
<td>.22</td>
</tr>
<tr>
<td>Hesitations</td>
<td>−.11</td>
<td>−.14</td>
</tr>
</tbody>
</table>

For accent (Table 1), less accented L2 speech was linked to fewer intonation errors (both tasks) and a narrower pitch range (picture task).

**Table 2:** Partial correlations between comprehensibility ratings and individual speech measures from the post-test, with the relevant speech measure from the pre-test partialled out (*p < .05, **p < .01).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Read-aloud</th>
<th>Picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmental errors</td>
<td>−.23</td>
<td>−.03</td>
</tr>
<tr>
<td>Intonation errors</td>
<td>−.37*</td>
<td>−.36*</td>
</tr>
<tr>
<td>Enchaînement</td>
<td>−.29</td>
<td>−.01</td>
</tr>
<tr>
<td>Liaison</td>
<td>.13</td>
<td>.10</td>
</tr>
<tr>
<td>Pitch range</td>
<td>.10</td>
<td>−.31*</td>
</tr>
<tr>
<td>MLR</td>
<td>.31*</td>
<td>.08</td>
</tr>
<tr>
<td>Hesitations</td>
<td>−.04</td>
<td>−.45**</td>
</tr>
</tbody>
</table>

For comprehensibility and fluency (Tables 2, 3), more comprehensible and fluent speech was linked to fewer intonation errors (both tasks), longer fluent speech runs (read-aloud), narrower pitch range and fewer hesitations (picture task).

**Table 3:** Partial correlations between fluency ratings and individual speech measures from the post-test, with the relevant speech measure from the pre-test partialled out (*p < .05, **p < .01).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Read-aloud</th>
<th>Picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmental errors</td>
<td>−.14</td>
<td>.05</td>
</tr>
<tr>
<td>Intonation errors</td>
<td>−.38*</td>
<td>−.37*</td>
</tr>
</tbody>
</table>

4. DISCUSSION

The current findings showed that the learners’ gains in intonation accuracy, pitch range, and hesitation rate reported in our earlier study [10] are also associated with a measurable pre- to post-test improvement in ratings of accent, comprehensibility, and fluency. This finding is noteworthy as it implies that focused phonetics instruction has an impact beyond specific aspects of L2 speech, contributing to listener judgments of accent, comprehensibility, and fluency.

Results also showed that listener-based accent, comprehensibility, and fluency ratings after focused phonetics instruction in an L2 speaking and listening course are linked to several aspects of speech targeted through instruction. Less accented speech was linked to fewer intonation errors and a narrower pitch range, while more fluent and comprehensible L2 output was additionally related to longer fluent speech runs and fewer hesitations. Contrary to our prediction, a narrower pitch range was associated with higher speech ratings, suggesting that an exaggerated pitch range, although typical of lively, animated speech, might lead listeners to downgrade their evaluations.

These results also contribute to ongoing research efforts to isolate linguistic aspects of L2 speech associated with listener ratings of accent, fluency, and comprehensibility, especially across tasks [e.g. 9, 15]. Our results imply a distinction between ratings of accent and ratings of fluency and comprehensibility, in that a wider range of speech measures was associated with the latter ratings, especially in spontaneous production. This result is supported by correlations among the three sets of ratings in the picture task, with fluency and comprehensibility sharing 81% of variance (r = .90); in contrast, these two ratings shared only 42% of variance with accent (r = .65 in each case). Our findings are generally encouraging for both researchers and teachers as they suggest that L2 pronunciation, despite the inherent difficulty it poses for adult learners, is a skill which can be learned in a classroom context.

5. REFERENCES


THE EFFECTS OF HIGH-VARIABILITY PHONETIC TRAINING ON CANTONESE ESL LEARNERS’ PRODUCTION OF ENGLISH /ɪ/-/iː/ CONTRAST - AN ACOUSTIC ANALYSIS

Janice Wing Sze Wong
Hong Kong Baptist University
janicewong@hkbu.edu.hk

ABSTRACT

This study investigated the effects of High Variability Phonetic Training (HVPT), a perception-only training, on Cantonese ESL learners’ production of English /ɪ/-/iː/ contrast in terms of the first and second formant frequencies as well as the vowel duration. A total of 17 subjects participated in this study and they were all trained over 20 sessions of the HVPT. The results showed that after training, the male learners could improve their production of the two vowels in terms of F1, F2 and vowel duration; the female subjects only improved in terms of F1 and vowel duration. This suggested that perceptual learning can be transferred to the production domain and the subjects could generally improve their productions in terms of both vowel height and vowel length. More individual variations were observed in the change of vowel frontness/backness. This study showed that after training, the awareness of the differences, particularly for the length, between the target vowel contrasts was raised.

Keywords: High Variability Phonetic Training, speech acoustics, speech perception and production

1. INTRODUCTION

A number of studies (e.g. [3], [10]-[11], [13], [18]-[19], [21]-[23]) have shown that the perception and production of non-native contrasts could be improved after phonetic training, one of which was the High Variability Phonetic Training (HVPT). HVPT is a perception-only phonetic training approach which emphasizes the use of multiple speakers and various phonetic contexts to increase the stimulus variability of the natural training minimal pairs. Subjects under training can be exposed to the natural tokens produced by different speakers. This approach was firstly adopted in a series of research studies ([3], [14]-[16]) in investigating the effectiveness of stimulus variability on the perceptual training of /ɪ/ and /iː/ contrast of Japanese speakers of English. In the last two decades, this approach has been extensively researched and many displayed encouraging results in improving the subjects’ perceptual performance of other confusing non-native contrasts and most of them showed that even though only training in the perception domain was given, the production of the contrasts could be improved. However, many of these studies only reported the production performance in terms of native speakers’ judgement, without delving into the acoustic dimension to see what aspects of the vowels produced by the subjects have changed.

The present study has chosen to investigate the production learning of English vowels /ɪ/ and /iː/ by Hong Kong Cantonese speakers of English. This pair of vowels has been reported as posing difficulties in the L2 learners’ production [5], [9], [17]. Even though most of the learners can distinguish the two vowels in terms of the vowel length, they are not aware that the two vowels differ also in terms of vowel quality. A couple of HVPT studies (e.g. [21], [23]) reported positive results in the modification of /ɪ/-/iː/ production by Hong Kong Cantonese speakers after training, but the acoustic properties of the production data were seldom analysed. This study aimed to complement the understanding of the effectiveness of the HVPT paradigm as well as the product of learning through a perception-only training paradigm.

2. METHODOLOGY

2.1. Participants

A total of 17 native Hong Kong Cantonese-speaking secondary school students (9 females and 8 males) aged around 16 to 17 were recruited to participate in the current experiment. They completed production pretest and posttest as well as 20 perceptual training sessions. They had not resided in any English-speaking countries before and the average age of learning English as an L2 was 3.71 (SD = .36), an average of 12.70 years (SD = .34). They reported no history of hearing or speaking impairment.

For the perceptual training and tests, another 6 native Received Pronunciation English speakers (3 females and 3 males) were recruited to produce the perceptual stimuli. Their ages ranged from 20 to 45.
2.2. Production Test Design

The aim of this study was to test how the HVPT paradigm can affect the production of the front vowel non-native contrast /ɪ/-/iː/ in terms of F1 and F2 values as well the vowel duration. The same production test was given to all the subjects before and after the 20 sessions of HVPT training.

All the subjects were given a word list with 20 words with the target vowels (10 /ɪ/ and 10 /iː/) plus 10 distractors. These target words were only a subset of the perceptual training set. They recorded the words in isolation, one at a time by reading from a screen and into a headset-mounted microphone with Adobe Audition 1.5 software for digitization (sampling rate at 44.1 kHz). The instructions for this production pretest were offered to the subjects in the form of five practice trials and they had to produce them with natural loudness and speaking rate. They were not provided with any audio prompts or instructions before or during the recording. They could also pause and resume during the recording based on their own pace. The test took around 15 minutes to complete.

2.3. Stimulus Materials

Six native Received Pronunciation English speakers produced all the stimuli used in the HVPT training. All of them produced all the minimal word pairs used in the training. All words were CVC monosyllabic words with different onsets and codas.

Each stimulus word was produced three times to avoid intra-speaker variability. All the apparatus and setting used in the preparation of perceptual materials were the same as those used in the production tests.

2.4. Perceptual Training Design

A total of 40 stimuli (20 /ɪ/ and 20 /iː/) were presented to the subjects. All the stimuli were produced by six different native speakers to enhance the stimulus variability. The words were randomized in terms of speakers and word order in each session. The subjects were trained on a two-alternative forced choice paradigm. The stimuli were one of the counterparts in a minimal word pair contrasting the two vowels, e.g. among “bid” and “bead,” only one of them was chosen for the test item. During training, immediate feedback was given; at the end of each session, their total scores were also shown. Each session required around 10 minutes to complete. All the subjects participated in two sessions each day for 10 consecutive days.

2.5. Method of Analysis

All the production data in the pretest and posttest was analysed acoustically by using the Praat speech analysis software [1]. The duration of the vowels was measured whereas the first two formant frequencies were gauged at the midpoint to evaluate how similar or different the vowel productions were before and after training.

In total, there were 680 tokens (17 subjects × 20 words × 2 tests) produced by all 17 subjects. Yet, 23 tokens were not analysed since the subjects produced them as other words. Tokens produced by male and female subjects were analysed separately.

3. RESULTS

3.1. Vowel height: F1 values

Figure 1 juxtaposes the F1 values in English /ɪ/ and /iː/ productions by both male and female subjects at pretest and posttest.

3.2. Vowel duration

Figure 2 juxtaposes the vowel duration (in ms) in English /ɪ/ and /iː/ productions by both male and female subjects at pretest and posttest.

For male participants, a two-way repeated measures ANOVA with Test (pretest and posttest) and Vowel (/ɪ/ and /iː/) as factors revealed significant main effects of Test [F(1,153) = 7.31, p = .01] and Vowel [F(1,153) = 52.7, p < .001] due to the change of F1 values before and after the training. The interaction Test × Vowel was also significant [F(1,153) = 70.24, p < .001]. Planned comparisons revealed that the F1 of /ɪ/ increased by 77.77 Hz after training (p < .001) whereas that of /iː/ decreased by 41.30 Hz after training (p < .001). The F1 values of the two vowels were not significantly different in the pretest (p = .519). This result suggests that the male subjects have started to distinguish the two vowels and have learnt the vowel height differences between the two target vowels.

While for the female participants, only the main effect of Vowel was significant [F(1,166) = 157.39, p < .001]; the effect of Test was not (p = .524). Yet, the interaction Test × Vowel was significant [F(1,166) =
16.87, \( p < .001 \). Planned comparisons showed that the F1 increased by 20.54 Hz (\( p = .044 \)) for vowel English /i/ after training whereas the F1 decreased by 27.82 Hz for /iː/ (\( p < .001 \)). Yet, the female subjects could distinguish the two vowels at the pretest (F1 of /i/ minus /iː/ = 61.14 Hz; \( p < .001 \)); it was only the difference in F1 of the two vowels that was increased in the posttest (F1 of /i/ minus /iː/ = 109.56 Hz; \( p < .001 \)). This result shows that after training, the female subjects had more awareness of the difference of the two vowels in terms of vowel height and they tended to increase the vowel height differences between the two vowels.

### 3.2. Vowel frontness/backness: F2 values

Figure 2 shows the target vowel production performance in terms of F2 values by both male and female subjects at pretest and posttest.

**Figure 2:** Boxplots of F2 in English /i/ and /iː/ productions by both male (left) and female (right) subjects at pretest and posttest.

For male participants, a two-way repeated measures ANOVA with Test (pretest and posttest) and Vowel (/i/ and /iː/) as factors revealed that there was only a significant main effect of Vowel \( [F(1,153) = 85.09, p < .001] \) but not for Test (\( p = .822 \)). The interaction Test × Vowel was significant \( [F(1,153) = 9.58, p = .003] \) and planned comparisons showed that the F2 of /i/ was decreased by 87.58 Hz after training (\( p = .032 \)) whereas that of /iː/ was increased by 76.35 Hz after training (\( p = .024 \)). However, in the pretest, the male subjects could already make a distinction between the two target vowels (F2 of /iː/ minus /i/ = 183.07 Hz; \( p < .001 \)). The difference in F2 of the two vowels /iː/ minus /i/ was increased to 346.98 Hz after training, suggesting that the male subjects were more aware of the differences of the two vowels in terms of frontness and backness of the vowels.

While for female participants, the same ANOVA showed that only a significant main effect of Vowel was observed \( [F(1,166) = 10.03, p = .002] \) because the two vowels were well distinguished in terms of vowel frontness/backness by the subjects. However, that of Test was not robust (\( p = .285 \)) and neither was the interaction Test × Vowel (\( p = .577 \)). This result showed that even after training, the female subjects generally did not alter the F2 much in the production of the two vowels. It is also worth noting that more individual variation was observed among the female subjects than the male ones.

### 3.3. Vowel duration

Figure 3 displays the vowel duration values of English /i/ and /iː/ productions by both male and female subjects at pretest and posttest.

**Figure 3:** Boxplots of vowel duration in English /i/ and /iː/ productions by both male (left) and female (right) subjects at pretest and posttest.

Both male and females subjects performed similarly. A two-way repeated measures ANOVA with Test (pretest and posttest) and Vowel (/i/ and /iː/) as factors revealed significant main effects of Test \( [F(1,153) = 15.69, p < .001] \) and Vowel \( [F(1,153) = 166.14, p < .001] \) for the male subjects. The interaction Test × Vowel was also robust \( [F(1,153) = 15.94, p < .001] \), due to the fact that the subjects produced a longer vowel /iː/ in the posttest than in the pretest (an increase of 55.45 ms, \( p < .001 \)), although the subjects already distinguished the two vowels in terms of vowel length in the pretest.

The same ANOVA was conducted on the data of the female subjects. The main effects of Test \( [F(1,166) = 24.27, p < .001] \) and Vowel \( [F(1,166) = 175.55, p < .001] \) were robust and so did the interaction Test × Vowel \( [F(1,166) = 25.87, p < .001] \). Similarly, the female subjects lengthened the vowel /iː/ in the posttest (an increase of 50.31 ms, \( p < .001 \)). The result showed that all the subjects tended to exaggerate the vowel length of the long vowel /iː/ and after training they were more aware of the differences in terms of vowel length between the two target vowels.
4. DISCUSSION

This study has complemented previous training studies by showing what aspects of the English/i/ and /ɪ/ productions the subjects had changed after HVPT training. The results display that after 20 sessions of HVPT, both male and female subjects could improve their production of both /i/ and /ɪ/. Concrete learning in production can be observed as the acoustic values have changed and more distinction was found between the two vowels. From previous studies, the two English vowels were found to be overlapped in terms of F1 and F2 since this contrast does not exist in Cantonese phonological system. Only a similar pair (/i/ and its surface form [i] with overlapping acoustic properties) exists in Cantonese. Thus, the pretest result is in accord with SLM [8] as L2 sounds which are similar to the L1 system are more difficult to acquire. The promising results of the present study then show that the HVPT is effective in modifying a difficult contrast. Highly variable stimuli are beneficial in leading to a significant level of success in non-native contrast acquisition as the stimuli could promote selective attention of the subjects. Offering a wider range of stimuli to the subjects is also believed to be able to expand the exemplar space along the dimensions where the two vowels differ and shrink along the dimensions that do not show distinctions for the two vowels, thus leading to more accurate productions.

The results also show that the awareness of the subjects towards vowel height and vowel length was raised, whereas that of vowel frontness/backness show inconclusive results because the female subjects did not demonstrate significant changes in the production of the two vowels in terms of F2. Although this study did not aim to investigate what cues the subjects relied on, the present results still suggest that the subjects had more awareness and reliance on vowel height and vowel length as cues to help distinguish the two vowels. The subjects even exaggerated the vowel length of the long vowel in the posttest although they could distinguish the pair in terms of vowel duration in the pretest. This result hints that the subjects tended to rely heavily on duration as the cues to distinguish a contrast, similar to previous reports (e.g. [2], [4], [6], [7], [12]). However, Zhang, Peng and Wang [24] discovered that Cantonese speakers were affected more by vowel quality cues than durational cues when they perceived Cantonese vowels. After training through the HVPT paradigm, the subjects also improved the ability to distinguish between the differences of the two vowels in terms of vowel height. This piece of result might lend some support to the above study that duration alone may not be an efficient cue in categorization by Cantonese speakers in both L1 and L2; rather, vowel quality cues overrode the effect of duration. Further experiments particularly on the cues that this group of speakers rely on are demanded. This will reveal more about the link between perception and production as well.

Useful pedagogical information was obtained from the positive results in the study which indicated that there exists the possibility of successfully training ESL learners to produce a non-native contrast more accurately. While maintaining the viability of application and simplicity of procedures of the training was also of high importance for language teachers, HVPT is practical enough to be adopted easily in schools or learning centers since it is not a complicated design. This paradigm can even be setup online for learners who are willing to receive training at their own pace and time. It will be worth investigating the optimal number of training sessions and training intensity which can successfully promote successful learning so that the training can be adopted in L2 classrooms and bring the most benefits to both learners and teachers.

Future studies can further assess the learning of the subjects by adopting and comparing different training types such as production training, audio-visual training etc. The adoption of assessment tasks such as identification test, category discrimination test, cue weighting test, etc. can also be investigated as they all tap into different aspects of L2 vowel perception and production. More efforts should also be directed at individual variability among the subjects both before and after training. The improvement of the subjects shown in the present study only gave an evaluation of the HVPT paradigm at the aggregate level without considering individual performances. From this study, there are clearly some learners who benefited from training more than others. Some previous studies (e.g. [20]) have also shown that individual’s performance at the pretest is a good predictor of learning outcomes, and thus more efficacious training paradigms can be developed and variability found in different training studies can be eliminated when individual differences of subjects are taken into account.
5. REFERENCES


