Vol. 11.1 2013

Research in Language

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

RL Vol. 11.1 2013

Research in Language

edited by
EWA WANIEK-KLIMCZAK

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

ŁÓDŹ 2013

EDITING AND TYPESETTING
Przemysław Ostalski

COVER DESIGN
Barbara Grzejszczak

Printed directly from camera-ready materials provided to the Łódź University Press

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego
Wydanie I. W.06237.13.0.C

ISSN 1731-7533

Online version is a primary one (vol. 11.1)

Wydawnictwo Uniwersytetu Łódzkiego
90-131 Łódź, ul. Lindleya 8
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl
tel. (42) 665 58 63, faks (42) 665 58 62

Druk i oprawa: Quick Druk

Contents

EDITORIAL TO RIL 11.1

The current issue of Research in Language brings together contributions exploring different aspects of second language accent studies, a steadily developing field of applied linguistics, which continues to inspire researchers and teachers alike. The approaches taken by the contributors to this issue bear witness to two major perspectives developed in the field: second language phonetics/phonology and instructed learning and teaching of the pronunciation of English. While the papers explore various aspects of non-native accents of English, they all report on primary research based on data from Polish, Czech, French, Finnish and Greek speakers of English, with native speakers of English used as a reference point in many cases. A wide range of methods are employed, including quantitative acoustic studies investigating the effect of imitation (**Rojczyk**, **Berger** and **Porzuczek**, **Zając**), an acoustic analysis of spectral characteristics (**Volín**, **Weingartová** and **Skarnitzl**), an auditory analysis of contextually conditioned stress marking (**Horgues**), a quantitative study of the effect of anxiety on success in pronunciation learning (**Baran-Łucarz**), a questionnaire based exploration of pronunciation-related views and attitudes (**Tergujeff**) and finally, the study of perception of gated casual speech (**Shockey** and **Ćavar**).

The first two contributions focus on the same approach and language context, as they investigate the effect of phonetic imitation in Polish learners of English, with **Arkadiusz Rojczyk**, **Andrzej Porzuczek** and **Marcin Bergier** examining the effect of immediate and distracted imitation of unreleased plosives, and **Magdalena Zając** exploring this effect with respect to the durational characteristics of English vowels. Both papers concentrate on phonetic characteristics of English believed to be difficult for Polish learners; by using imitation, the authors investigate the extent to which speech accommodation may affect the pronunciation of non-native speakers. Interestingly, although the results of both studies verify the hypothesis that phonetic (sub)segmental features chosen for the analysis can be imitated by non-native speakers, the degree and direction of imitation is strongly related to experimental conditions. Searching for stable parameters which would make it possible to account for differences in the production of the vowel schwa in English and Czech, **Jan Volín**, **Lenka Weingartová** and **Radek Skarnitzl** move to a suprasegmental level of analysis for which the studied vowel is crucial. Having conducted a detailed acoustic study, they propose the use of the distribution of acoustic energy in the vowel spectrum as the most reliable measure distinguishing between native and non-native speakers. Concentrating on prosodic parameters, **Celine Horgues** continues the suprasegmental theme by investigating the effect of intonational contexts on the relative difficulty in implementing the stress pattern of English by native speakers of French; unlike previous studies, however, she uses perceptive judgements to verify her hypotheses. The auditory approach moves us towards pedagogically-oriented studies, which tend to rely on the assessment of speech as performed by interlocutors (or teachers) in natural or classroom discourse. The study by **Małgorzata Baran-Łucarz** attempts to determine optimal conditions for phonetic

production by means of examining the relationship between phonetic learning anxiety and success in pronunciation of English in the case of advanced Polish learners, with pronunciation assessed from text and word-list reading. The attitude towards pronunciation training and more generally, the views concerning pronunciation itself have been explored by **Elina Tergujeff**, who asked Finnish school students to talk about their experiences regarding English pronunciation and their attitudes towards it. Adopting a qualitative perspective, the study provides insights into the way young teenagers perceive the role of formal education and their out-of-school English input as decisive in the development of intelligibility and fluency in English. Finally, the study by **Linda Shockey** and **Małgorzata Ćavar** adopts a yet another perspective: searching for reasons conditioning varied success in English casual speech perception by learners from different language backgrounds, the researchers propose to explore the native-language characteristics of the learners. Thus, the final contribution to the volume combines a traditional contrastive analysis based approach with an innovative research programme calling for a data-based analysis of the first languages before formulating predictions as to their effect on second language phonetics and phonology.

EWA WANIEK-KLIMCZAK
University of Łódź
ewaklim@uni.lodz.pl

IMMEDIATE AND DISTRACTED IMITATION IN SECOND-LANGUAGE SPEECH: UNRELEASED PLOSIVES IN ENGLISH

ARKADIUSZ ROJCZYK
University of Silesia, Poland
arkadiusz.rojczyk@us.edu.pl

ANDRZEJ PORZUCZEK
University of Silesia, Poland
ampj24@wp.pl

MARCIN BERGIER
University of Silesia, Poland
emberg@o2.pl

Abstract
The paper investigates immediate and distracted imitation in second-language speech using unreleased plosives. Unreleased plosives are fairly frequently found in English sequences of two stops. Polish, on the other hand, is characterised by a significant rate of releases in such sequences. This cross-linguistic difference served as material to look into how and to what extent non-native properties of sounds can be produced in immediate and distracted imitation. Thirteen native speakers of Polish first read and then imitated sequences of words with two stops straddling the word boundary. Stimuli for imitation had no release of the first stop. The results revealed that (1) a non-native feature such as the lack of the release burst can be imitated; (2) distracting imitation impedes imitative performance; (3) the type of a sequence interacts with the magnitude of an imitative effect

Keywords: imitation, plosives, unreleased, distraction.

1. Introduction

Human beings have an inborn capacity to reproduce the actions and intentions of others (Hauser 1996; Honorof et al. 2011; Nagell et al. 1993; Whiten and Custance 1996). This imitative tendency starts immediately after birth (Meltzoff and Moore 1999) - for instance, twelve-week old infants already imitate ambient vocalic sounds (Kuhl and Meltzoff 1996) - and appears to reach its climax between two to five years of age (Horner and Whiten 2005). Those early imitative reactions are logically linked with language acquisition processes which encourage children to acquire language from their caretakers and peer group (Chambers 1992; Babel 2012). The automatic imitative behaviour observed in humans appears to have a neurophysiological basis in the

architecture of mirror neurons which make up an action-observation matching system (Rizzolatti and Craighero 2004; Rizzolatti et al. 2001; Schwartz et al. 2012). It is suggested that the human mirror-neuron system creates parity between the speaker and the listener, which is a prerequisite for successful imitation (Arbib 2005; Gentilucci and Corballis 2006; Rizzolatti and Arbib 1998). It is achieved by activation in brain areas responsible for planning and production of speech during auditory or visual perception of speech (Pekkola et al. 2006; Pulvermuller et al. 2006; Skipper et al. 2007; Wilson and Iacoboni 2006; Wilson et al. 2004).

The speech-imitative behaviour persists into adulthood and serves many sociolinguistic functions. For example, adults relatively easily acquire features of a dialect in the new surrounding (Delvaux and Soquet 2007; Evans and Iverson 2007; Munro et al. 1999; Trudgill 1986). Such imitation seems to be driven by the need to assimilate with a new positively-evaluated social group, however Bourhis and Giles (1997) also observed a dialect divergence conditioned by a negative affective attitude towards a particular dialect group. Talkers also exhibit imitative tendencies in various communicative interactions to express similarity (Shepard et al. 2001), attraction (Byrne 1971), to gain approval (Street and Giles 1982), or to increase one's intelligibility (Triandis and Triandis 1960).

2. Imitation in speech

Phonetic imitation, also referred to as phonetic convergence or phonetic accommodation, is the process in which a talker takes on acoustic characteristics of their interacting partner (Babel 2012). A variety of phonetic features have been reported to undergo imitative convergence, such as accent, speaking rate, intensity, pitch, variation of frequency bands, long-term average spectra, frequency of pauses, and utterance length (Giles et al. 1991; Goldinger 1997; Gregory 1990; Gregory and Webster 1996; Gregory et al. 1993, 1997, 2001; Namy et al. 2002; Natale 1975; Pardo et al. 2012). Other studies have concentrated on VOT as a temporal parameter that undergoes assimilation as a result of exposure to the model talker. Shockley et al. (2004) demonstrated significant VOT imitation of voiceless plosives in words with artificially extended VOT values. More recently, Nielsen (2011) showed that not only are longer VOTs imitated by talkers but also that this re-modelled feature can be generalized to other plosives. Moreover, imitation in this study was found to be selective, in that it did not occur for reduced VOT and depended on the frequency of tested lexical items. Significant imitation of the model talker has also been reported for vowels, as expressed by formant frequencies of individual productions (Babel 2010; Evans and Iverson 2007; Pardo 2010; Pardo et al. 2010, 2012). Here again, the degree of convergence was modulated by selectivity, in that only some vowels in different linguistic items were imitated. Finally, Honorof et al. (2011) reported convergence of articulatory gestures in imitated allophonic instances /l/, expressed as the distance between F2 and F1.

All the observed imitative tendencies in speech are captured by non-abstract theories of linguistic representations. In this view, fine-grained linguistic and non-linguistic phonetic features available in the speech signal are preserved in perception to make a set of exemplars that forms a perceptual category (Hintzman 1986; Nosofsky 1986). This is

the main assumption of exemplar-based models of speech perception (Coleman 2002; Johnson 1997; Pierrehumbert 2006). Although those models do not rule out completely the possibility of modularity in formation of categories, they predict that mental phonological representations of words encode both allophonic variability and speaker-specific information. Following this reasoning, imitation of speech emerges as a natural process in which the listener perceives and reproduces fine-grained phonetic features provided by the model talker. Even if such features differ from the listener's canonical representations, they are not filtered out or discarded, but rather they are successfully delivered in imitation. Such accommodation from perception to production is considered to be both automatic (Gentilucci and Bernardis 2007, but see Nielsen 2011 and Mitterer and Ernestus 2008 for selectivity) and quick (Fowler et al. 2003; Honorof 2011).

Imitation is an undisputed factor in acquisition of second-language speech. Successful production of non-native sound categories should logically arise from effective imitation of patterns absent in one's native language. Previous research in our lab has shown that acoustic features defined by cross-linguistic differences between Polish and English can be imitated to a significant degree by learners when shadowing after the model talker. Rojczyk (2012a) had Polish learners of English imitate the English low-front vowel /æ/ in a rapid shadowing task. This vowel is reported to be difficult to acquire for Poles: it is equally likely to be assimilated by two Polish neighbouring vowels /e/ and /a/. Productions in two tasks were compared: reading of words with the target vowel presented as a list with a view to establishing a baseline condition and imitations of the same words delivered binaurally. F1 and F2 were measured for all productions and the Euclidean distance to the model vowel frequencies was calculated to express the degree of convergence. The results showed that the learners significantly converged their productions of the target vowels with the model. It was taken as evidence that imitation can override the influence of native categories in production of new sound categories. In another study, Rojczyk (2012b) used VOT as another Polish-English typological difference that emerges in Polish pronunciation of English. While English /p, t, k/ are characterised by long-lag VOT values, Polish /p, t, k/ use short-lag VOT values. As mentioned earlier, this difference surfaces in Polish pronunciation of English as observable underaspiration of English voiceless stops. In this study, production in three tasks was compared: (1) reading of English words with /p, t, k/ word-finally as a baseline condition; (2) immediate imitation of those words pronounced with native-like long VOTs; (3) distracted imitation in which the imitators were required to read a digit presented on the screen after hearing a model word and prior to imitation. The results revealed significant increase in VOTs in immediate imitation and intermediate values for distracted imitation. These results were interpreted to indicate that immediate imitation may bypass the influence of native articulatory habits and that distraction in imitation results in incomplete recovery of native phonetic patterns.

3. Release burst in stop sequences in English and Polish

English and Polish differ considerably in the frequency of release bursts both word-finally and when preceding another stop. Many textbooks on English phonetics observe that English stops tend to be unreleased when followed by another stop or affricate (e.g.

Abercrombie 1967; Gimson 2001; Jones 1956; Ladefoged 1975; Roach 2000). Experimental research has supported this observation, however its magnitude is not so great as may be expected. Crystal and House (1988a) found 59% of English stops without the release burst in all sentence positions. Randolph (1989 reported in Byrd 1993) reported that in word-final position English stops are mostly unreleased. Byrd (1993) analysed data from the TIMIT database and found 40.3% of releases in stops. Davidson (2010) investigated spontaneous speech from the National Public Radio and found the frequency of the unreleased stops in pre-stop and pre-pausal position between 50% and 60%. The actual frequency of unreleased variants depends on many factors. Bilabial stops tend to be more often unreleased, followed by alveolars and velars (Byrd 1993; Crystal and House 1988b). Voiceless stops have a stronger tendency to include and acoustically measurable release burst than voiced stops (Crystal and House 1988a, b, but see Byrd 1993). Finally, women have been reported to release stops more often than men (Byrd 1992, 1993).

Polish stops are generally described as invariably released except when they precede another homorganic stop (Dukiewicz and Sawicka 1995; Jassem 1974; Kopczyński 1977; Wierzchowska 1980). Rojczyk (2008) studied experimentally the putative tendency to unrelease Polish stops in same-place clusters. Stops were matched across a word boundary in two-word phrases and sentences. The results revealed that in the case of an intervening word boundary stops were released more than 50% of the time in homorganic clusters. The actual context significantly influenced the frequency of release bursts. Stops inserted in short two-word phrases were more frequently released than stops in sentences. All this leads to the conclusion that Polish differs from English in the tendency to unrelease stops and that this will have consequences on Polish pronunciation of English. Indeed, observations by experienced teachers of English pronunciation indicate that Polish learners have noticeably more frequent and stronger releases in English stop sequences than English native speakers and that this contributes to the perception of their speech as non-native. Textbooks on English pronunciation tailored for Polish learners include exercises in this area (Bałutowa 1974; Mańkowska et al. 2009; Sobkowiak 2001). Although controlling for the lack of release is initially difficult, appropriate phonetic training and instructing can yield positive effects on ultimate performance (Bergier 2010).

4. The current study

The current study investigates the degree of imitation of English unreleased stops by Polish learners. It consists of three tasks: (1) reading of phrases presented as a list to establish a baseline frequency of releases in the studied group; (2) immediate imitation of the unreleased sequences provided by the model talker; (3) distracted imitation of the unreleased sequences in which listeners are required to read a digit after hearing a model stimulus and prior to imitation. Accordingly, the research questions are formulated as follows:

1. Is the lack of the release burst imitated in immediate imitation calculated as a significant decrease in the frequency of releases compared to list-reading?

2. Does distraction in imitation impede the performance compared to immediate imitation or does it block imitative behaviours altogether as compared to list-reading?
3. Does the type of a sequence - same place of articulation vs. different place of articulation - interact with the magnitude of imitation?

Previous research showed that delaying imitation reduces its degree (Goldinger 1998). In this study, we decided to use distraction, which we suggest poses a greater challenge on listeners. While simple delaying extends the time interval between auditory input and articulatory production, it does not interfere with articulatory planning because no other cognitively taxing processes are included. On the other hand, distraction, in which the response is not only delayed but subjects also engaged in reading digits, provides both an increase in cognitive processing and articulatory resetting between hearing and imitating target stimuli. We therefore assumed that a greater challenge to imitators would yield more reliable results on how long-lasting perceptual traces of the lack of release were in the studied group.

Gender was not included as an independent variable largely due to the fact that some of the used statistical procedures used were non-parametric for nominal variables and did not allow inclusion of more than one independent variable. We predict, however, that the investigation of how gender interacts with imitation of unreleased plosives would yield interesting results. In earlier studies women were reported to converge to a larger extent to the model talker compared to men (Namy et al. 2002; Pardo 2006).

4.1 Participants

Thirteen native speakers of Polish (eight females and five males) were included in the study. They ranged in age from 20 to 21. All participants were students at the Institute of English, University of Silesia. None of them had had any prior phonetic training concerning unreleased stops in English. They did not have any reported speech or hearing disorders.

4.2 Materials

The stimuli used in the experiment were nine two-word noun phrases in which stop sequences were matched across the word boundary (Table 1).

ca**p P**at	tha**t p**an	bla**ck p**ack
ta**p t**ap	tha**t t**ap	bla**ck t**ap
la**p c**at	tha**t c**at	bla**ck c**ap

Table 1. Stimulus phrases used in the experiment. Bolding indicates combinations of stops straddling word boundaries

All nouns were preceded by either a verb, adjective or determiner and had a natural focus stress on the second word. The intonation pattern was uniform with a H* H*L-L%

contour. Only voiceless stops were selected for two reasons. First, voiceless plosion is more conspicuous in a spectrographic display because it is not attenuated by concomitant periodicity. Second, voiceless plosives have been previously reported to be more frequently released (Crystal and House 1988a, b). Considering a greater tendency to release voiceless plosives rather than voiced plosives, it was assumed that using voiceless plosives would be a more challenging task on the participants and thus a more sensitive metric of the occurrence of imitative behaviour. The stimuli reflected all possible combinations of places of articulation: bilabial, alveolar and velar (3x3=9) and contained only one, low front vowel /æ/.

All stimuli were recorded by the second author, a qualified phonetician, using the recording specification described below. The stimuli were created by saving individual sound files in a computer. Next, spectrographic analysis in Praat (Boersma 2001) was used to inspect target sequences of two stops. No release bursts were detected in the preceding stops. The duration of each stop sequence was durationally normalized to range from 180 ms to 190 ms. Finally, all stimuli were peak normalized to 70 dB SPL.

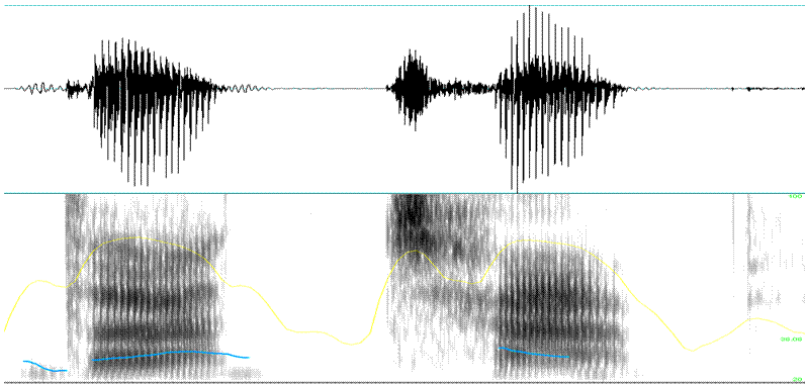

Figure 1. Waveform and spectrogram of the phrase *that tap*. No detectable release burst of the first stop in a sequence.

4.3 Procedure and recording

The experiment took place in the Acoustic-Phonetic Laboratory at the Institute of English, University of Silesia. As described earlier, data were collected from three blocked tasks. The first task was reading the list of phrases presented orthographically to establish a baseline frequency of release bursts in the studied material. The words were flashed sequentially on a monitor screen in 54-point black font in the middle of the white screen. Seven foil phrases were randomly dispersed among target phrases to distract the participants' attention from the object of the study. The second block was immediate shadowing after the model talker in which participants were instructed that upon hearing the model pronunciation they were to immediately repeat it. The orthographic representations of the phrases were also sequentially flashed during imitation. The approximate interval between complete imitation and the onset of the next phrase was 1 sec. The third block was distracted imitation. The participants were instructed that they

would hear the model pronunciation, next they would read a digit flashed in the centre of the screen, and finally that they were to imitate the phrase. The interval between playing the model voice, displaying a digit, and flashing an orthographic representation of the imitated phrase was also approximately 1 sec. Tasks 2 and 3 were counterbalanced by participants to avoid a carry-over effect from one task to another.

The recordings were made in a sound-proof booth with a monitor screen located in front of a participant. The signal was captured with a headset dynamic microphone Sennheiser HMD 26, preamplified with USBPre2 (Sound Devices) into .wav format with the sampling rate 48 kHz, 24-bit quantization. The model voice was delivered binaurally through high-quality headphones built in the headset at a comfortable listening level.

4.4 Measurements

All measurements were made using waveform and spectrogram displays available in Praat (Boersma 2001). Prior to any quantifications of data, it was necessary to define the acoustic criteria for classification of measured tokens as released or unreleased. Introductory analyses revealed a number of cases with visible weak energy spikes in the spectrogram but which, at the same time, gave no auditory impressions. Henderson and Repp (1982) listed five stages of the unreleased-released continuum: (1) unreleased; (2) silently released; (3) inaudibly released; (4) weakly released; (5) strongly released. As a result, we decided to classify our tokens as released when they belonged to stages (4) and (5) from Henderson and Repp (1982). In other words, stops were classified as released when they had auditorily detectable burst and it was manifested as the sudden rise of energy visible as acoustic transients in waveform and spectrogram. Other tokens were classified as unreleased.

It is interesting to note that release bursts measured in the current data were characterised by significant variability. While the force of the burst appeared to be, to a large extent, individual, some task effects were observed, especially in distracted imitation. Some sequences had a very long compression phase of the first stop followed by relatively long release, sometimes even exceeding 100 ms. This may show articulatory attempts to hold the compression and proceed to another stop, which however were not successful and ended in a strong release of pent-up air. It is, however, for future experiments to see how detailed acoustic properties of release may reveal imitative behaviour.

Measurements were divided into two main types for statistical analyses. Nominal measurements classified tokens as either released or unreleased. Ratio measurements identified the time duration of the burst expressed in ms. It was assumed that measuring the duration of bursts might be a more sensitive metric of whether imitation occurred or not. Duration of the burst was defined as the time interval between the onset of the rise of energy following a silent period of compression to its offset indicated by a complete drop of energy signalling compression for the next stop. A total number of measured tokens was 351 (13 participants x 9 sequences x 3 tasks).

4.5. Results and analysis

Two types of statistical tests were run to calculate both nominal and ratio data. For nominal observation, i.e. a stop may be released or unreleased, Cochran Q test was applied. It is an alternative to one-way within-subject ANOVA when the dependent variable is dichotomous. For duration measurements in ms a two-way mixed ANOVA 2x3 was designed with 2 levels of a between-subject variable (place of articulation: homorganic / heterorganic) and 3 levels of within-subject variable (task: list reading / immediate imitation / distracted imitation).

Figure 2 shows the overall proportion of release bursts in all three tasks as well as the proportions broken down into homorganic and heterorganic clusters.

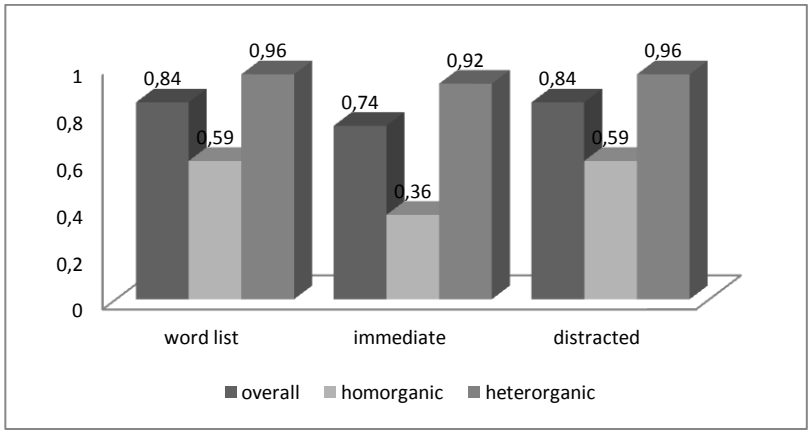

Figure 2. The proportion of release bursts in all three tasks: overall, homorganic clusters and heterorganic clusters.

For an overall number of release bursts the Cochran Q test revealed significant influence of task on releasing a stop [x^2 (2) = 10.67, $p < .01$]. This effect was achieved by reduction of release bursts in immediate imitation (74%) compared to word list (84%) and distracted imitation (84%). Breaking down data into homorganic and heterorganic clusters showed different magnitude of contribution to the main effect of task. For homorganic clusters there was again a significant effect of task on releasing a stop [x^2 (2) = 9, $p < .05$]. The number of release bursts decreased in immediate imitation (26%) relative to word list (59%) and distracted imitation (59%). No statistically significant effect of task was found for heterorganic clusters [x^2 (2) = 2, *ns*]. Summing up the analysis of the frequency of release busts, the following observations may be formulated. Immediate imitation reduced the number of release bursts relative to the baseline articulatory habits. This reduction was mostly contributed to by homorganic sequences. When imitation was distracted it did not lead to the reduction of release bursts.

The analysis of duration of bursts in ms was predicted to be a more sensitive measure of imitative behaviour, in that it would be able to register slight reductions of plosion force which were disregarded in a dichotomous released / unreleased measure. Figure 3

shows the overall mean durations of bursts in ms in all three tasks as well as durations broken down into homorganic and heterorganic clusters.

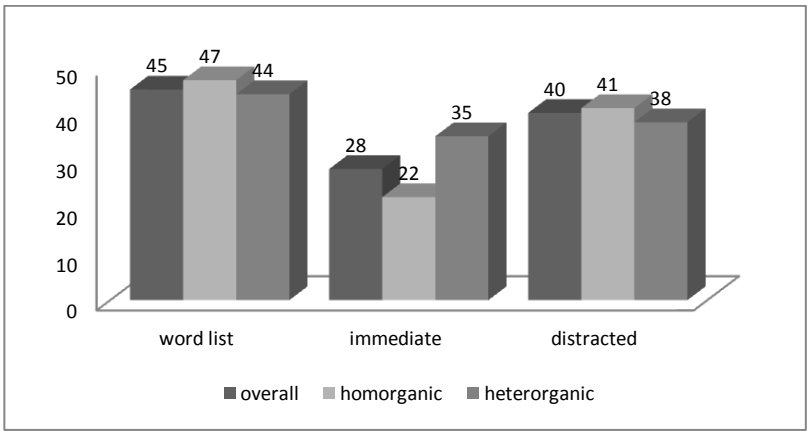

Figure 3. Mean durations of release bursts in all three tasks: overall, homorganic clusters and heterorganic clusters.

The main effect of task on duration of the release bursts was highly significant [$F (2, 230) = 15.86$, $p < .001$]. Post hoc Bonferroni tests revealed that this effect was mainly achieved by significant reduction of durations in immediate imitation (28 ms) compared to baseline list-reading (45 ms) ($p < .001$) and distracted imitation (40 ms) ($p < .05$). When imitation was distracted, durations of release burst tended to be lower than in a baseline list-reading task, however not significantly (both $p > .05$).

The task x cluster type interaction was significant [$F (2, 230) = 4.38$, $p < .05$], indicating that the effect of task on durations varied in magnitude depending on whether the cluster was homorganic or heterorganic. Post hoc Bonferroni tests revealed that significant reduction of burst durations observed in immediate imitation was mostly contributed to by homorganic clusters in which durations of bursts dropped from 47 ms to 22 ms ($p < .001$). Although heterorganic clusters also demonstrated reduction in immediate imitation (35 ms) compared to list-reading (44 ms), it was not statistically significant ($p > .05$). All comparisons between immediate imitation and distracted imitation for either homorganic or heterorganic clusters were non-significant. The same lack of significance was also found for comparisons between baseline list-reading and distracted imitation, which indicates that distraction in imitation resets articulatory patterns to their default status.

5. General discussion

The current study investigated how two types of imitation - immediate and distracted - modify the pronunciation of FL learners. Unreleased plosives in two-stop sequences in English were chosen, because English and Polish differ in the frequency of releases in such sequences. Polish learners of English were exposed to the auditory model that

produced unreleased stops in two imitation tasks: immediate imitation in which shadowing commenced immediately after the auditory input and distracted imitation in which participants were instructed to read a digit after the auditory input and prior to shadowing. Two types of measurements were used, nominal released / unreleased and durational in which the durations of bursts were expressed in ms. The results were expected to show if, and to what extent, unreleased stops can be imitated compared to baseline list-reading.

The analysed data allow us to answer questions formulated before the experiment.

1. Is the lack of the release burst imitated in immediate imitation calculated as a significant decrease in the frequency of releases compared to list-reading?
The answer is positive. Both nominal and durational measures revealed that the participants reacted to the auditory input and modified their productions to converge with the model talker. The frequency of bursts and their durations significantly decreased in immediate imitation compared to baseline list-reading. It demonstrates that the lack of release can be imitated even by talkers whose native language releases stops in stop sequences. Future studies should look more closely into individual variability because, as shown in the current data, participants differed in their initial tendency to release and in how they reacted to the auditory model in imitation. Some participants had a high initial release frequency and did not observably reduce it in imitation. Other participants showed a similar initial high release frequency but reduced it as a result of auditory exposure. Yet others had a relatively low release frequency initially and either reduced it or not in imitation. It is for future research to investigate more thoroughly the extent of individual variability both in a default release pattern and in the release pattern in imitation and to seek explanation for those idiosyncrasies.

2. Does distraction in imitation impede the performance compared to immediate imitation or does it block imitative behaviours altogether as compared to list-reading?
The answer is positive. Current results showed that distracting participants by asking them to read digits after the auditory exposure and prior to imitation reduced significantly imitation effect. Although this is evident in calculations of statistical significance, it is worth noting that the imitation effect was not absent altogether. Both the frequency of bursts and their durations were lower for distracted imitation relative to baseline list-reading. It points to some weak remnants of auditory traces despite distraction, as evidenced by participants' productions. As mentioned earlier, some productions in distracted imitation were characterised by long compression phases and sudden strong and long releases. It may be interpreted to mean that the participants attempted to imitate the lack of release by extending durationally the hold phase, but were finally unsuccessful, which resulted in strong releases of compressed air. In order to verify this tendency, future studies may make use of more sensitive acoustic metrics.

Another point that merits discussion is the very nature of distracted imitation compared to delayed imitation. While delayed imitation extends the time interval between auditory exposure and production, distraction provides two additional parameters: cognitive taxing and articulatory resetting. Cognitive taxing is caused by the need to perceive and recognise the digit to plan articulatory commands for its production. Articulatory resetting is a product of a new plan for articulation. This is not

the case for delayed imitation in which articulatory resetting does not occur because participants are inactive between the exposure and production. This difference does not seem to be satisfactorily explained in the imitation literature. While it seems safe to assume that distraction will cause more disturbance to imitation than delaying, the actual magnitude of this inhibition is not clear. Future studies should incorporate the distinction between delaying and distraction as a variable to disentangle their effect on imitation.

3. Does the type of a sequence - same place of articulation vs. different place of articulation - interact with the magnitude of imitation?
The answer is positive. Homorganic clusters, unlike heterorganic clusters, had a significant contribution to the observed effect of imitation. We ascribe it to the fact that, as discussed earlier, stops in homorganic clusters in Polish can be optionally unreleased. If homorganic unreleased stops are classified as allophones of released stops in Polish, it is clear evidence that what is most readily imitated is the allophonic variant which occurs in participants' native language. Previous research has demonstrated that allophonic experience from learners' native language improves both perception of non-native contrasts as well as their learning. (Best and Strange 1992; Halle et al. 1999; Jamieson and Moroson 1986; Jenkins and Yeni-Komshian 1995; Kondaurova and Francis 2008; McAllister et al. 2002; McClaskey et al. 1983; Pisoni et al. 1982; Pruitt et al. 2006). The current results suggest that such experience also contributes to imitation. It is not surprising considering the fact that successful imitation is an important factor in developing perception of non-native sounds and in successful acquisition of their production. A special status of native language allophonic variants reported here appears to confirm their special status in acquisition of second-language speech in general.

References

Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Arbib, M. A. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioural and Brain Sciences* 28: 105-167.

Babel, M. 2010. Dialect convergence and divergence in New Zealand English. *Language in Society* 39: 437-456.

Babel, M. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40: 177-189.

Bałutowa, B. 1974. *Wymowa Angielska dla Wszystkich*. Warszawa: Wiedza Powszechna.

Bergier, M. 2010. *The occurrence of unreleased oral stops in English voice agreeing plosive clusters straddling word boundaries. Production experiments with Polish advanced learners of English.* M. A. Thesis. University of Silesia.

Best, C. T. and W. Strange. 1992. Effects of language-specific phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305-330.

Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 10: 341-345.

Bourhis, R. Y. and H. Giles. 1977. The language of intergroup distinctiveness. In H. Giles (ed.) *Language, Ethnicity, and Intergroup Relations*. London: Academic Press: 119-136.

Byrd, D. 1992. Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America* 92: 593-596.

Byrd, D. 1993. 54,000 American stops. *UCLA Working Papers in Phonetics* 83: 1-19.

Byrne, D. 1971. *The Attraction Paradigm*. New York: Academic Press.

Chambers, J. 1992. Dialect acquisition. *Language* 68: 673-705.

Coleman, J. 2002. Phonetic representations in the mental lexicon. In J. Durand and B. Laks (eds) *Phonetics, Phonology and Cognition*. Oxford: Oxford University Press: 96-130.

Crystal, T. H. and A. S. House. 1988a. The duration of American-English stop consonants. *Journal of Phonetics* 16: 285-294.

Crystal, T. H. and A. S. House. 1988b Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America* 83: 1553-1573.

Davidson, L. 2010. Variation in stop releases in American English spontaneous speech. *Journal of the Acoustical Society of America* 128: 2458.

Delvaux, V. and A. Soquet. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64: 145-173.

Dukiewicz, L. and I. Sawicka. 1995. Fonetyka i fonologia. In H. Wróbel (ed.) *Gramatyka i Fonologia Języka Polskiego*. Kraków: Instytut Języka Polskiego PAN.

Evans, B. G. and P. Iverson. 2007. Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America* 121: 3814-3826.

Fowler, C. A., J. M. Brown, L. Sabadini and J. Weihing. 2003. Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language* 49: 396-413.

Gentilucci, B. R. Bernardis. 2007. Imitation during phoneme production. *Neuropsychologia* 45: 608-615.

Gentilucci, B. and M. C. Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience and Behavioral Reviews* 30: 949-960.

Giles, H., J. Coupland and N. Coupland. 1991. *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: Cambridge University Press.

Gimson, A. C. 2001. *Gimson's Pronunciation of English*. 6th edition. Rev. by A. Cruttenden. London: Arnold.

Goldinger, S. D. 1997. Perception and production in an episodic lexicon. In K. Johnson and J. W. Mullennix (eds) *Talker Variability in Speech Processing*. San Diego: Academic Press: 33-66.

Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105: 251-279.

Gregory, S. W. 1990. Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior* 14: 237-251.

Gregory, S. W., K. Dagan and S. Webster. 1997. Evaluating the relation between vocal accommodation in conversational partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior* 21: 23-43.

Gregory, S. W., B. E. Green, R. M. Carrothers and K. A. Dagan. 2001. Verifying the primacy of voice fundamental frequency in social status accommodation. *Language Communication* 21: 37-60.

Gregory, S. W. and S. Webster. 1996. A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status predictions. *Journal of Personality and Social Psychology* 70: 1231-1240.

Gregory, S. W., S. Webster and G. Huang. 1993. Voice pitch and amplitude convergence as a metric of quality in dyadic interviews. *Language Communication* 13: 195-217.

Halle, P. A., C. T. Best and A. Levitt. 1999. Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics* 27: 281-306.

Hauser, M. D. 1996. *The Evolution of Communication*. Cambridge, MA: MIT Press.

Henderson, J. B. and B. H. Repp. 1982. Is a stop consonant released when followed by another stop consonant? *Phonetica* 39: 71-82.

Hintzman, D. L. 1986. "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93: 411-428.

Honorof, D. N., J. Weihing and C. A. Fowler. 2011. Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39: 18-38.

Horner, V. and A. Whiten. 2005. Casual knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition* 8: 164-181.

Jamieson, D. G. and D. E. Moroson. 1986. Training non-native speech contrasts in adults: Acquisition of the English /ð/ - /Θ/ contrast by francophones. *Perception and Psychophysics* 40: 205-215.

Jassem, W. 1974. *Fonetyka Języka Angielskiego*. 4th edition. Warszawa: PWN.

Jenkins, J. J. and G. H. Yeni-Komshian. 1995. Cross-language speech perception: Perspective and promise. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Speech Research*. York: York Press: 463-479.

Johnson, K. 1997. Speech perception without speaker normalization: An exemplar model. In K. Johnson and J. W. Mullennix (eds) *Talker Variability in Speech Processing*. San Diego: Academic Press: 145-166.

Jones, D. 1956. *An Outline of English Phonetics*. Cambridge: Heffer and Sons.

Kondaurova, M. V. and A. L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124: 3959-3971.

Kopczyński, A. 1977. *Polish and American English Consonant Phonemes*. Warszawa: PWN.

Kuhl, P. and A. N. Meltzoff. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America* 100: 2425-2438.

Ladefoged, P. 1975. *A Course in Phonetics*. Orlando: Harcourt Brace.

Mańkowska, A., M. Nowacka and M. Kłoczowska. 2009. *"How Much Wood Would a Woodchuck Chuck?": English Pronunciation Practice Book*. Kraków: Wydawnictwo WSE.

McAllister, R., F. E. Flege and T. Piske. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30: 229-258.

McClaskey, C. L., D. B. Pisoni and T. D. Carrell. 1983. Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics* 34: 323-330.

Meltzoff, A. and M. Moore. 1999. Persons and representation: Why infant imitation is important for theories of human development. In J. Nadel and G. Butterworth (eds) *Imitation in Infancy*. Cambridge: Cambridge University Press: 9-35.

Mitterer, H. and M. Ernestus. 2008. The link between speech perception and production is phonological and abstract: Evidence form the shadowing task. *Cognition* 109: 168-173.

Munro, M. J., T. M. Derwing, J. E. Flege. 1999. Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics* 27: 385-403.

Nagell, K., K. Olguin and M. Tomasello. 1993. Processes of social learning in tool use in chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*. 107: 174-186.

Namy, L. L., L. C. Nygaard and D. Sauerteig. 2002. Gender differences in vocal accommodation: The role of perception. *Journal and Language and Social Psychology* 21: 422-432.

Natale, M. 1975. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology* 32: 790-804.

Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39: 132-14.

Nosofsky, R. M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115: 39-57.

Pardo, J. S. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119: 2382-2392.

Pardo, J. S. 2010. Expressing oneself in conversational interaction. In E. Morsella (ed) *Expressing Oneself/Expressing One's Self. Communication, Cognition, Language, and Identity*. New York: Psychology Press: 183-196.

Pardo, J. S., I. Cajori Jay and R. M. Krauss. 2010. Conversational role influences speech imitation. *Attention, Perception and Psychophysics* 72: 2254-2264.

Pardo, J. S., R. Gibbons, A. Suppes and R. M. Krauss. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40: 190-197.

Pekkola, J., M. Laasonen, V. Ojanen, T. Autti, L. P. Jaaskelainen, T. Kujala et al. 2006. Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: An fMRI study at 3T. *NeuroImage* 29: 797-807.

Pierrehumbert, J. B. 2006. The next toolkit. *Journal of Phonetics* 34: 516-530.

Pisoni, D. B., R. N. Aslin, A. J. Perey and B. L. Hennessy. 1982. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8: 297-314.

Pruitt, P. and J. van Santen. 2006. Training the perception of Hindi dental and retroflex stops by native speakers of English and Japanese. *Journal of the Acoustical Society of America* 119: 1684-1696.

Pulvermuller, F., M. Huss, F. Kherif, F. Moscoso del Prado Martin, O. Hauk and Y. Shtyrov. 2006. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America* 103: 7865-7870.

Randolph, M. 1989. *Syllable-based Constraints on Properties of English Sounds*. Doctoral Dissertation, Massachusetts Institute of Technology.

Rizzolatti, G. and M. A. Arbib. 1998. Language within our grasp. *Trends in Neurosciences* 21: 188-194.

Rizzolatti, G. and L. Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 21: 169-192.

Rizzolatti, G., L. Fogassi and V. Gallese. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2: 661-670.

Roach, P. *English Phonetics and Phonology*. 3rd edition. Cambridge: Cambridge University Press.

Rojczyk, A. 2008. Release burst in Polish homorganic stop geminates. *Linguistica Silesiana* 29: 75-86.

Rojczyk, A. 2012a. Spontaneous phonetic imitation of L2 vowels in a rapid shadowing task. Paper presented at Pronunciation in Second Language Learning and Teaching 2012. Vancouver Canada, August 24-25 2012.

Rojczyk, A. 2012b. Phonetic and phonological mode in second-language speech: VOT imitation. Paper presented at EUROSLA 2012, Poznań Poland, September 5-8 2012.

Schwartz, J-L., A. Basirat, L. Ménard and M. Sato. 2012. The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics* 25: 336-354.

Shepard, C. A., H. Giles and B. A. Le Poire. 2001. Communication accommodation theory. In W. P. Robinson and H. Giles (eds) *The New Handbook of Language and Social Psychology*. Chichester: John Wiley & Sons.

Shockley, K., L. Sabadini and C. A. Fowler. 2004. Imitation in shadowing words. *Perception and Psychophysics* 66: 422-429.

Skipper, J. I., V. Van Wassenhove, H. C. Nusbaum and S. L. Small. 2007. Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex* 17: 2387-2399.

Sobkowiak, W. 2001. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Street, R. L. and H. Giles. 1982. Speech accommodation theory: A social cognitive approach to language and speech behavior. In M. Roloff and C. Berger (eds) *Social Cognition and Communication*. Beverly Hills: Sage.

Triandis, H. C. and L. M. Triandis. 1960. Race, social class, religion, and nationality as determinants of social distance. *Journal of Abnormal and Social Psychology* 61: 110-118.

Trudgill, P. 1986. *Dialects in Contact*. New York: Blackwell Publishing.

Whiten, A. and D. M. Custance. 1996. Studies of imitation in chimpanzees and children. In C. M. Heyes and B. G. Galef (eds) *Social Learning in Animals: The Roots of Culture*. San Diego: Academic Press: 291-318.

Wierzchowska, B. 1980. *Fonetyka i Fonologia Języka Polskiego*. Wrocław: PAN.

Wilson, S. M. and M. Iacoboni. 2006. Neural responses to non-native phonemes varying in producibility: Evidence for the sensimotor nature of speech perception. *NeuroImage* 33: 316-325.

Wilson, S. M., A. P. Saygin, M. I. Sereno and M. Iacoboni. 2004. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience* 7: 701-702.

PHONETIC IMITATION OF VOWEL DURATION IN L2 SPEECH

MAGDALENA ZAJĄC
University of Łódź
zajac1234@gmail.com

Abstract
This paper reports the results of a pilot study concerned with phonetic imitation in the speech of Polish learners of English. The purpose of the study was to investigate whether native speakers of Polish imitate the length of English vowels and to determine whether the extent of phonetic imitation may be influenced by the model talker being a native or a non-native speaker of English. The participants were asked to perform an auditory naming task in which they indentified objects and actions presented on a set of photos twice, with and without the imitation task. The imitation task was further sub-divided depending on the model talker being a native or non-native speaker of English (a native Southern British English speaker and a native Polish speaker fluent in English). As the aim was to investigate the variability in durational characteristics of English vowels, the series of front vowels /æ e ɪ iː/ were analysed in the shortening and lengthening b_t vs. b_d contexts. The results of the study show that the participants imitated the length of the investigated vowels as a result of exposure to the two model talkers. The data suggest that the degree of imitation was mediated both by linguistic and social factors and that the direction of convergence might have been affected by the participants' attitude toward L2 pronunciation.

Keywords: phonetic imitation, phonetic convergence, phonetic accommodation, L2 pronunciation, vowel lenght in L2 pronunciation, pre-fortis clipping in L2 pronunciation, Polish learners of English, social factors affecting phonetic imitation, linguistic factors affecting phonetic imitation.

1. Phonetic imitation

Phonetic imitation, sometimes termed phonetic convergence or phonetic accommodation (e.g. Babel, 2009; Pardo, 2010; Kim et al., 2011), can be defined as the process in which a talker takes on acoustic characteristics of another individual as a result of exposure to his or her speech (Babel, 2011). Many instances of this phenomenon have been reported to take place in "cooperative, socially rich, dyadic interactions" (Babel, 2011: 178). For example, Gregory and Webster (1996, in Babel, 2009) examined F0 convergence in the interviews from the Larry King Live television programme and observed that the show's host accommodated more toward guests with higher social status. Bilous and Krauss (1988, in Pardo, 2010) examined convergence in spontaneous conversations between same-gender and mixed-gender dyads. One of the observations they made was that both

male and female participants converged in average utterance length and frequency of pauses. More recently, Pardo et al. (2010) asked pairs of participants to give each other instructions and cooperate in order to complete a map task. The degree of phonetic imitation was calculated by using perceptual similarity judgments, measures of articulation rate and measures of vowel spectra. It was found that some of the subjects imitated the speech of their conversational partners and that a talker's gender and his or her role in the interaction affected the degree of phonetic convergence.

Phonetic imitation has also been found to take place in "socially minimal situations where talkers are simply producing single words" (Babel, 2011: 178). In a study by Babel (2009), the subjects read and then repeated a series of words containing different English vowels after two model talkers. The results showed that phonetic imitation did take place and that factors such as implicit racial biases and attractiveness ratings influenced the degree of convergence. In a similar study, Babel (2010) investigated whether New Zealand English speakers imitated the speech of an Australian talker. The participants performed and auditory naming task, the stimuli were single-word productions from the Australian model talker. It was found that the participants imitated the model talker and that "[s]ocial biases about how a participant feels about a speaker predicted the extent of accommodation" (Babel, 2010: 437). In Nielsen's (2011) study, the subjects listened to a model talker producing a series of words with extended VOT values and were then asked to read the words. The results indicated that after exposure to the model talker's speech, the participants produced significantly longer VOTs.

As referred to above, phonetic imitation may be affected by various social factors, such as gender, model talker's perceived attractiveness or the subject's implicit attitude towards race. As stated by Babel (2009: 23) "a talker's social knowledge and desires mediate the strength and nature of convergence in language". Nonetheless, it has been observed that phonetic imitation can also be conditioned by linguistic factors. For instance, Babel (2009) observed that participants in her study imitated /æ/ and /ɑ/ to a greater extent than other investigated vowels. Analogously, the results of the study on NZE speakers (Babel, 2010) revealed that not all analysed vowels were imitated to the same extent. Nielsen (2011), on the other hand, discovered that lexical frequency had an effect on the degree of VOT imitation and that productions with reduced VOT were not imitated.

2. Phonetic imitation in non-native speech

An interesting issue related to phonetic imitation is whether or not it occurs in non-native speech. Kim et al. (2011) investigated phonetic convergence in conversations between subjects who had either the same or different regional dialects, and between native and non-native speakers of English. The degree of imitation was measured by asking an independent group of listeners to judge the similarity of utterance samples taken from one participant to the utterance samples taken from his or her conversational partner. The results of the study demonstrated that "sharing the same language and dialect was the only condition amongst the three language distance conditions where phonetic convergence was likely to occur" (Kim et al., 2011: 139). The authors attributed the apparent lack of phonetic convergence on the part of the non-native talkers who interacted with the native speakers to the fact that "the extra demands of second

language production may have interfered with any alignment process" (Kim et al, 2011: 143).

Conversely, in an earlier study, Beebe (1981, in Zuengler, 1991) explored the pronunciation of Chinese-Thai children and adults, who were interviewed in Thai by an ethnically Chinese and an ethnically Thai interlocutor. Beebe (ibid.) analysed the subjects' L2 pronunciation of six Thai vowels and found that they realised five of these vowels significantly more Thai-like when talking to the Thai interviewer. Most recently, Rojczyk (2012a) demonstrated that native Polish speakers of English imitated the realisation of /æ/ when required to immediately repeat a series of words after a native English model talker. In addition, Rojczyk (2012b) found Polish learners of English imitated English VOT values under similar conditions.

3. The Current Study

Given the relative scarcity of studies on phonetic imitation in non-native speech as well as their varied results, the issue seems an interesting and important subject for research. The main purpose of the study reported here was to investigate whether native speakers of Polish imitate the length of selected English vowels. The second goal, related to the issue of phonetic convergence being mediated by various social factors, was to explore whether the extent of imitation is influenced by the model talker being a native or a non-native speaker of English. Finally, the study aimed to determine if any potential imitation was selective from a linguistic perspective, i.e. to check whether the imitation of vowel length by Polish learners of English is affected by linguistic factors.

3.1 Variables

The dependent variables under investigation were the durational characteristics of English /æ e iː ɪ/, which were analysed in the shortening and lengthening b_t and b_d contexts. Such variables were chosen as pre-fortis clipping, a feature characteristic of English pronunciation, may cause difficulties for Polish learners. As explained by Waniek-Klimczak (1998: 397):

> Vowel duration is used in English at the phonological level as an inherent feature of individual vowels, enhancing the articulatory differences for individual vowel phonemes, and at the phonetic implementation level as a cue for voicing of the following obstruent; consequently, English can be claimed to be vowel-length sensitive, as compared with languages like Polish, which are vowel-length insensitive due to the lack of phonological use of inherent vowel duration or the use of vowel lengthening / shortening cue for consonant voicing.

In other words, vowel duration in shortening and lengthening contexts was selected as a variable in order to explore whether articulatory habits typical of the participants' L1 would prevent them from imitating this L2 feature.

3.2 Participants

Twenty native speakers of Polish, twelve females and 8 males, took part in the study. The subjects were all first-year students of English Studies, recruited from the University of Lodz in Poland. Two model talkers were included in the study, one of them a native speaker of Southern British English, the other a native speaker of Polish, fluent in English but speaking with a relatively heavy foreign-accent. Both model talkers were male and in their mid-twenties.

3.3 Stimulus

The following eight monosyllabic words were selected as stimuli: *bad, bat, bed, bet, bead, beat, bid, bit*. The model talkers were recorded while reading the analysed words in carrier sentences (*I'm saying ____ again*). This was done to prevent the model talkers from using a special intonation pattern associated with reading word lists (Ladefoged, 2003). The stimuli were extracted from the recordings and presented as isolated words in the imitation (shadowing) task.

Vowel durations in the model talkers' productions were calculated, making up a total of 16 measurements. The obtained data is presented in Table 1. The abbreviations NM and NNM denote the native model talker and the non-native model talker respectively; b_d and b_t represent the voiced and voiceless contexts. As expected, the native model talker used noticeably longer vowels in the voiced context in each of the analysed pairs of words. The vowels in the non-native model's productions, on the other hand, were shorter in the voiced environments in two instances, /iː/ and /ɪ/.

	NM		NNM	
vowel	b_d	b_t	b_d	b_t
æ	140	98	145	128
e	127	77	138	94
iː	167	145	114	118
ɪ	103	81	81	105

Table 1. Vowel durations in the model talkers productions.

3.4 Procedure

The experiment consisted of three tasks: a written matching exercise, an auditory naming task, and a shadowing task. First, each participant was given a sheet of paper with the eight analysed words and a set of black-and-white photos (see Appendix A), and then asked to match the words with the photos that represented them. This exercise was designed to ascertain that the subjects knew all the words and their meaning before being asked to produce them.

Next, the eight photos from the matching exercise were presented sequentially on a computer monitor in five-second intervals. The subjects were instructed to name what they see in each photo using only the words from the written matching exercise. This was done to establish the participants' baseline productions of the analysed words. The photos were presented in a random order, which was the same for all participants.

In the shadowing task, sixteen photos were presented sequentially on a computer monitor in five-second intervals. Each photo was accompanied by either the native or the non-native model's voice pronouncing the word that was represented in the photo, i.e. each of the eight photos from the two previous task was shown twice, once with the native model's voice and once with the non-native model's voice. The photos were shown in a random order, which was the same for all subjects. As in the previous task, the participants were instructed to identify what they see, the difference being that this time they were required to listen to another person pronouncing the words before producing them themselves. It is important to note that the subjects were never explicitly instructed to imitate what they heard.

The words spoken by the model talkers in the shadowing task were presented together with the photos to make certain that the participants knew to which word they were listening at a particular moment. This was essential for the study because the non-native speaker's realization of /æ/ and /e/ closely resembled Polish /ɛ/, while his /ɪ/ and /iː/ appear to have been realized as Polish /i/. Without being able to see the photos, this would render it exceedingly difficult for the participants to determine whether the Polish model talker was saying *bad* or *bed*, *bead* or *bid*, *bat* or *bet*, etc, which, in turn, would make it impossible to correctly interpret the results of the study.

3.5 Data analysis

Vowel durations in baseline and shadowed productions were measured for each participant, giving a total of 24 vowel length measurements for each subject (8 words x 3 productions = 24 measurements). The obtained data was analysed in two stages. The first stage consisted of comparing vowel durations in the three productions (baseline, shadowing after the native model, shadowing after the non-native model). Its purpose was to examine whether the participants imitated vowel length in individual words. The second stage involved examining vowel durations in the shortening and lengthening contexts and comparing them across the three productions. This was done to establish whether exposure to the model talkers' pronunciation made the subjects modify the degree of pre-fortis clipping. The data obtained for the model talkers (see section 3.3.) was also included in the analysis as it was integral to the process of interpreting the results.

3.6 Results

The following two tables show mean vowel durations in the subjects' baseline productions contrasted with vowel durations in the model talkers' productions (NM and NNM stand for native model talker and non-native model talker respectively). The mean

durations in the baseline productions are, on the whole, noticeably longer than the durations in the model talkers' productions. The only exception is the word *beat*, in which the mean vowel duration in the baseline is very close to that of the native model. It is also worth mentioning that, as opposed to the native model talker and the participants, the non-native speaker's vowel durations are in two cases shorter in the *voiced* context (*bead* vs. *beat, bid* vs. *bit*).

word	vowel	baseline	NM	NNM
bad	æ	202	140	145
bed	e	194	127	138
bead	iː	205	167	114
bid	ɪ	140	103	81

Table 2. Mean vowel durations in baseline b_d productions contrasted with vowel durations in model talkers' b_d productions (in milliseconds).

word	vowel	baseline	NM	NNM
bat	æ	162	98	128
bet	e	143	77	94
beat	iː	148	145	118
bit	ɪ	138	81	105

Table 3. Mean vowel durations in baseline b_t productions contrasted with vowel durations in model talkers' b_t productions (in milliseconds).

Tables 4 and 5 show the subjects' mean vowel durations before and after hearing the native model's speech (NM stands for native model talker). The significance of the differences between the mean values was calculated by carrying out one-tailed paired-samples t-tests. The probability levels for non-chance difference between the mean values are tabulated in the last column of each table. The data in the two tables shows that the general tendency among the participants was to decrease vowel length after exposure to the native talker's pronunciation. In addition, the differences between the means are statistically significant in nearly all of the investigated words. These findings imply that the subjects systematically imitated the durational characteristics of most of the native model's vowels, except for the ones in *bid* and *beat*.

word	vowel	baseline N=20	shadowing NM N=20	p
bad	æ	202 (46)	160 (31)	**0.000****
bed	e	194 (44)	160 (40)	**0.001****
bead	iː	205 (45)	184 (33)	**0.008****

word	vowel	baseline N=20	shadowing NM N=20	p
bid	ɪ	140 (32)	131 (29)	0.108

Table 4. Mean vowel durations in the b_d context (in milliseconds; SD given in brackets).

word	vowel	baseline N=20	shadowing NM N=20	p
bat	æ	162 (38)	143 (25)	**0.011***
bet	e	143 (25)	111 (26)	**0.000****
beat	iː	148 (36)	141 (28)	0.207
bit	ɪ	138 (42)	106 (21)	**0.002****

Table 5. Mean vowel durations in the b_t context (in milliseconds; SD given in brackets).

Mean vowel durations before and after exposure to the non-native model's speech are presented in Tables 6 and 7 (NNM stands for non-native model talker). The results of one-tailed paired-samples t-tests are tabulated in the last column of each table and show the probability levels for non-chance difference between the mean values. As can be seen, mean vowel durations in the shadowing task are generally shorter than those in baseline productions and almost all of the differences in means are statistically significant. These results indicate that the participants converged toward the non-native model talker by decreasing vowel length. The only irregularity in the data is the lack of systematic imitation in the case of *beat*.

word	vowel	baseline N=20	shadowing NNM N=20	p
bad	æ	202 (46)	170 (33)	**0.001****
bed	e	194 (44)	164 (29)	**0.000****
bead	iː	205 (45)	162 (34)	**0.000****
bid	ɪ	140 (32)	125 (22)	**0.039***

Table 6. Mean vowel durations in the b_d context (in milliseconds; SD given in brackets).

word	vowel	baseline N=20	shadowing NNM N=20	p
bat	æ	162 (38)	136 (23)	**0.001****
bet	e	143 (25)	119 (19)	**0.000****
beat	iː	148 (36)	132 (30)	0.066
bit	ɪ	138 (42)	111 (26)	**0.003****

Table 7. Mean vowel durations in the b_t context (in milliseconds; SD given in brackets).

Table 8 shows mean vowel durations in the shortening and lengthening contexts under three different conditions: in the baseline productions, after hearing the native model's pronunciation (shadowing NM) and after hearing the non-native model's pronunciation (shadowing NNM). One-tailed paired-samples t-tests were conducted to calculate the significance of the differences between the mean values. The probability levels for non-chance difference between the values are presented in the last column of each table. The results reveal that the subjects generally tended to shorten the investigated vowels in the voiceless context. Interestingly, although the mean length of /ɪ/ in the baseline productions is almost the same in both the shortening and the lengthening environment, in the shadowed productions the same vowel is significantly shorter in the b_t context. This may signify that the participants converged toward the native speaker by increasing the amount of pre-fortis clipping. However, the same is not true in the case of the non-native model talker, as the results show that he produced a longer /ɪ/ in the voiceless context. Also, after listening to the non-native speaker's pronunciation, the subjects used a significantly shorter /iː/ in the b_t context despite the fact that the model talker shortened /iː/ in the b_d environment. These observations suggest that, in terms of pre-fortis clipping, the subjects *converged* toward the native speaker and *diverged* from the non-native speaker.

	baseline			shadowing NM			shadowing NNM		
vowel	b_d N=20	b_t N=20	p	b_d N=20	b_t N=20	p	b_d N=20	b_t N=20	p
æ	202 (46)	162 (38)	**0.000****	160 (31)	143 (25)	**0.001****	170 (33)	136 (23)	**0.000****
e	194 (44)	143 (25)	**0.000****	160 (40)	111 (26)	**0.000****	164 (29)	119 (19)	**0.000****
iː	205 (45)	148 (36)	**0.000****	184 (33)	141 (28)	**0.000****	162 (34)	132 (30)	**0.000****
ɪ	140 (32)	138 (42)	0.423	131 (29)	106 (21)	**0.000****	125 (22)	111 (26)	**0.031***

Table 8. Participants' mean vowel durations under three conditions (in milliseconds; SD given in brackets).

4. Discussion

One of the findings of the study was that in baseline productions the participants used considerably longer vowels than both model talkers. It seems that the large discrepancy between vowel durations stems from the fact that the model talkers were recorded while reading the analysed words in frame sentences (see section 3.3.), which might have caused them to produce the stimuli with a faster speech tempo. The participants produced the words in isolation and, as a consequence, might have used a slower tempo. This way another independent variable was unintentionally introduced in the study. One way of resolving this problem in follow-up studies would be to make the elicitation procedure the same for the model talkers and the participants, i.e. let the model talkers

familiarise themselves with the words selected as stimuli and then ask them to use the words to identify what they see in a set of photos.

As far as individual words are concerned, the results indicate that the participants imitated both model talkers by significantly decreasing vowel length in the shadowing task. Nonetheless, some inconsistencies were also found in the data. Firstly, the subjects did not significantly shorten the vowels in *bid* and *beat* after listening to the native speaker. The apparent lack of systematic imitation of the native model's *beat* could be attributed to the fact that mean vowel length in the baseline production of this word was already very close to that of the native model (see Table 2), thus providing the subjects with no context for imitation. The lack of regularity in the case of *bid*, on the other hand, is particularly intriguing. Unfortunately, no convincing explanation for this finding was found. Another interesting observation was that the subjects did not significantly decrease vowel duration in *beat* after hearing the non-native speaker. Similarly as with *bid*, it proved difficult to explain the lack of systematic convergence toward the non-native model. Nonetheless, the fact that the participants imitated vowel length in some of the investigated words to a greater extent than in others suggests that imitation might have been linguistically-selective. If so, the obtained data supports the observations made by Babel (2009, 2010) and Nielsen (2011) (see section 1).

As regards examining vowel durations across shortening and lengthening contexts, the results imply that exposure to the model talkers' pronunciation caused the subjects to modify the degree of pre-fortis clipping in their pronunciation. The participants appear to have converged toward the native speaker by increasing the vowel length difference between *bit* and *bid* in the shadowing task; they also seem to have diverged from the non-native speaker by maintaining a longer /iː/ in the b_d environment and increasing the vowel length difference between *bid* and *bit*. This indicates that the subjects might have realised that one of the model talkers spoke with a foreign accent and that the other was a native speaker of English. If so, it is probable that they diverged from the NN model in order to distance themselves from other foreign-accented speakers, whereas their convergence toward the N model was the result of a desire to sound more native-like. Hence, it seems that the direction of phonetic imitation might have been influenced by the participants' attitude toward L2 pronunciation. These findings seem to endorse the claim that phonetic imitation may be affected by social factors (Babel, 2009). In this case, it appears that imitation was to some extent mediated by the model talker's status as a native/non-native speaker of English and the subjects' desire to sound native-like. Notice also that social aspects seem to have a bearing on the degree of imitation even if the experiment takes place in "socially minimal situations where talkers are simply producing single words" (Babel, 2011: 178).

5. Conclusions

The results of the study indicate that Polish learners of English are able to imitate durational characteristics of English vowels as a result of exposure to the speech of different model talkers. The obtained data suggest that, just as in the case of native speakers, phonetic imitation in L2 speech may be selective from both a linguistic and a social perspective. It was also found that the direction of convergence may be influenced by the participants' attitude toward L2 pronunciation. Finally, the results of the study

show that phonetic imitation in non-native speech can take place in socially minimal situations.

References

Babel, M. 2009. *Phonetic and Social Selectivity in Speech Accommodation*. Unpublished PhD dissertation.

Babel, M. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39, 437–456. DOI: 10.1017/S0047404510000400

Babel, M. 2011. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40, 177–189. DOI: 10.1016/j.wocn.2011.09.001

Beebe, L. 1981. Social and Situational Factors Affecting the Communicative Strategy of Dialect Code-Switching. *International Journal of the Sociology of Language* 32, 139-149.

Goldinger, S. 1998. Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological Review* 105, 251- 279.

Honorof, D., Weihing, J. and Fowler, C.A. 2011. Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39, 18–38. DOI: 10.1016/j.wocn.2010.10.007

Kim, M., Horton, W. S. and Bradlow, A. R. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology* 2, 125-156. DOI: 10.1515/LABPHON.2011.004

Ladefoged, P. 2003. *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell Publishing.

Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39, 132-142. DOI: 10.1016/j.wocn.2010.12.007

Pardo J. 2010. Expressing oneself in conversational interaction. In Morsella, E. (ed) *Expressing oneself/expressing one's self: Communication, cognition, language, and identity*. London: Psychology Press, 183–196. DOI: 10.3758/APP.72.8.2254

Pardo, J., Cajori Jay, I. and Krauss, R. M. 2010. Conversational role influences speech imitation. *Attention, Perception, & Psychophysics* 72 (8), 2254-2264.

Rojczyk, A. 2012a. *Spontaneous phonetic imitation of L2 vowels in a rapid shadowing task*. Poster presented at PSLLT 2012 - Pronunciation in Second Language Learning and Teaching Conference, Vancouver, Canada, 24-25 August.

Rojczyk, A. 2012b. *Phonetic and phonological mode in second-language speech: VOT imitation*. Paper presented at EuroSLA22 – 22nd Annual Conference of the European Second Language Association, Poznań, Poland, 5-8 September.

Waniek-Klimczak, E. 1998. Sources of variability in L2 phonology: vowel duration in 'Polish English'. In Catherine Gruber, C., Higgins, D., Olson, K. and Wysocki, T. (eds) *CLS 34/2: The Panels*. Chicago Linguistic Society, 387-401.

Zuengler, J. 1991. Accommodation in Native-Non-native Interactions: Going Beyond the "What" to the "Why" in Second-Language Research. In Coupland, J., Coupland N. and Giles H. (eds) *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: Cambridge University Press, 223-244.

Appendix A

The written matching exercise

Dopasuj poniższe angielskie słowa do przedstawiających je obrazków:

bead bad bat beat bet bid bit bed

a).....

b).....

c).....

d).....

e).....

f).....

g).....

h).....

SPECTRAL CHARACTERISTICS OF SCHWA IN CZECH ACCENTED ENGLISH

JAN VOLÍN
Metropolitan University Prague
Jan.Volin@ff.cuni.cz

LENKA WEINGARTOVÁ
Institute of Phonetics in Prague
Lenka.Weingartova@ff.cuni.cz

RADEK SKARNITZL
Institute of Phonetics in Prague
Radek.Skarnitzl@ff.cuni.cz

Abstract
The English central mid lax vowel (i.e., schwa) often contributes considerably to the sound differences between native and non-native speech. Many foreign speakers of English fail to reduce certain underlying vowels to schwa, which, on the suprasegmental level of description, affects the perceived rhythm of their speech. However, the problem of capturing quantitatively the differences between native and non-native schwa poses difficulties that, to this day, have been tackled only partially. We offer a technique of measurement in the acoustic domain that has not been probed properly as yet: the distribution of acoustic energy in the vowel spectrum. Our results show that spectral slope features measured in weak vowels discriminate between Czech and British speakers of English quite reliably. Moreover, the measurements of formant bandwidths turned out to be useful for the same task, albeit less direct.

Keywords: foreign accent, prominence, schwa, spectral slope.

1. Introduction

Foreign accents manifest themselves in various domains of speech architecture. In the segmental area, individual vowels and consonants may display interesting deviations from what is perceived as standard or canonical for native speech, while the specific suprasegmental or prosodic features cause no less remarkable differences in stress, intonation or rhythm. The consequences of sounding foreign can be manifold: from moderate amusement on the part of the native listener to outright prejudice on the attitudinal level, or, in the domain of information flow, from increased strain in speech processing to miscomprehension leading to the breakdown of communication.

Derwing and Munro (2009) summarize some of their observations on the social consequences of foreign accented speech in their general overview and point out that it is

the lack of deeper understanding which sustains negative attitudes to foreign accents. Lev-Ari and Keysar (2010) highlight another source of complications for speakers with a foreign accent: their experiment showed that the credibility of a speaker's statements is affected by the degree of foreign accentedness. The assertions in sentences pronounced by foreigners were perceived as less true than the identical propositions spoken by native speakers of English. However, the design of their study enabled the authors to question the exclusive influence of prejudice. Having analyzed their results, they suggested that the decrease of credibility was connected with increased processing load. This effect can be observed even in the visual domain: if some printed statements are more difficult to read than others due to the colour and size of the font, the perceivers tend to assume they are less true (Lev-Ari and Keysar, 2010). Be that as it may, it is clear that not only areas like second language education, but also forensic practice and security or business would profit from a deeper insight into the mechanisms which underlie communication in non-native languages.

Our study focuses on the English lax central mid vowel, also known as *schwa* or reduced vowel. Due to its unique properties, both phonetic and phonological, this vowel is responsible for specific perceptual effects in the sphere of speech rhythm (along with stress and intonation, even if its impact there is less direct). In this sense, the segmental and suprasegmental domains are brought together quite firmly in our study, through schwa as the object of investigation.

We will not enter the dispute over the phonemic status of schwa. It is clear that, on the one hand, this vowel may participate in a phonological opposition with another vowel: *affect – effect*: /əˈfekt/ – /ɪˈfekt/; *boxers – boxes*: /ˈbɒksəz/ – /ˈbɒksɪz/, or with a phonological zero: *data - date*: /deɪtə/ – /deɪt/, on the other hand, there are countless examples of positionally conditioned alternation of schwa with other vowels: *academy – academic*: /əˈkædəmɪ/ – /ækəˈdemɪk/; *land – England*: /ˈlænd/ – /ˈɪŋgləɳd/, etc. Depending on the stress-pattern of the word, speakers may produce either schwa or a full vowel in the given position: the vowel between /d/ and /m/ in the word *academy* can be described as reduced, while the same position in the word *academic* is occupied by the peripheral front mid vowel.

This type of positional conditioning, however, manifests itself not only in derivational processes on the word level, but also on the phrase level. The metrical structure of syntactic units requires a certain arrangement of strong and weak positions. There are about forty monosyllabic grammatical words in English which regularly occur in unstressed positions and, consequently, surface in the so-called weak forms, which in most cases means with schwa as their syllable nucleus. Such words served as the material in our study (see section Method) since despite their small number in the lexicon they are very frequent in texts. They include prepositions (*at, for, from, of, to*), conjunctions (*and, but, than, as*), pronouns (*you, your, her, them,* etc.), auxiliary and modal verbs (*are, were, was, have, has, do, does, would, can,* etc.), and determiners (*a, the, some*). It has to be emphasised that these words sometimes do occur in their strong forms (e.g., when syntactically stranded or under focus) and when they do, full or peripheral vowels are employed in their pronunciation.

Due to the objective of our present study, we can leave open the question of whether schwa is an independent phoneme or an allophone of other vowels. As Ashby *et al.* (1995) pointed out, to most practically minded people the important fact is that schwa is an essential component of the sound patterning of English. According to Fry, almost one

quarter of all vowels in continuous speech correspond to schwa. Just for comparison, if a learner of English does not pronounce the front open vowel /æ/ (also known as *ash*) correctly, only one in twenty-seven vowels is affected in running texts (Fry, 1947).

To a laboratory acoustician, the male formant frequencies of schwa are F1 = 500 Hz, F2 = 1,500 Hz, and F3 = 2,500 Hz. The ideal female values are about fifteen percent higher. However, the evidence collected over the years of phonetic research confirms that the formant values of schwa retrieved from continuous speech are very variable and sensitive to context (e.g., Lindblom, 1963; Browman and Goldstein, 1992; Flemming and Johnson, 2007). Some of these studies showed that the formant positions of schwa are influenced not only by the neighbouring consonants, but also by vowels in the neighbouring syllables. Thus, the listener's feeling of the weak syllable nuclei being properly reduced or not does not seem to be correlated with the ideal position of formants in the "acoustically pure" schwa. Moreover, Barry (1998) put forward some evidence that computational methods used to determine the properties of schwa from formant frequencies were not entirely correct. He also confirmed that the contextual influence on formants is related to articulation rate.

Despite the fact that formant frequencies are often the only spectral descriptors of vowels (or sonorants in general) offered by textbooks on speech acoustics, we have to ask whether there is indeed no other information in the spectrum that could be correlated with the "proper" sound of schwa. A cue that such information should exist is provided by the metrical role of the English schwa. It is a vowel with the inherent feature [+ *unstressed*] and in the four-level analysis of syllabic prominences it is associated with the lowest level. It has been pointed out previously that the salience of the vowel is reflected by its spectral slope (also spectral tilt or spectral balance). The brighter sounding vowels display more moderate decay of spectral energies in higher-frequency regions or, the other way round, weaker sounding vowels have steeper decrease of amplitude towards the higher frequencies (Sluijter and van Heuven, 1996; Gobl and Ní Chasaide, 2003).

One of the major problems is that there is no single established method of quantifying spectral slope. One of the early attempts to provide an index of spectral balance was that of Britta Hammarberg and her colleagues who used the difference between the energy peaks (maxima) in the 0–2 kHz and 2–5 kHz frequency bands (Hammarberg *et al.*, 1980: 448). Various modifications were later suggested to this approach. The so-called α measure is based on the ratio between the sound energy above and below 1,000 Hz (e.g., Sundberg and Nordenberg, 2006). In their overview, Hanson *et al.* (2001) discuss several further measures, each with some advantages and disadvantages in the context of various objectives (usually the detection of voice disorders or quantifying long-term average spectra). After experimenting with individual vowels produced by healthy voices, we decided to modify the existing measures (Volín and Zimmermann, 2011), as described below in the Method.

2. Method

Our dataset consisted of recordings of three female native Southern Standard British English speakers and three female native Czech speakers. None of them had a speech

impediment or reported any hearing disorder. Both the native and Czech speakers of English were not professional speakers, and their age ranged from 20 to 23 years. They were asked to read out a single news bulletin taken from a broadcast of the BBC World Service. The three Czech subjects were selected on the basis of two criteria: they had to be capable of reading the relatively difficult English text fluently but, at the same time, they had to exhibit consistently a relatively overt Czech accent.

The British speakers were recorded with a portable professional device Edirol HR-09, with a sampling frequency of 48 kHz and 16-bit resolution. Later, these recordings were resampled to 32 kHz. The recordings of Czech speakers were made in a sound-treated studio of the Institute of Phonetics in Prague with an electret microphone IMG ECM 2000, soundcard SB Audigy 2 ZS, 32-kHz sampling frequency and 16-bit resolution.

This material yielded approximately 4–4.5 minutes of speech for each of the subjects. The utterances were cut into breath-groups and manually labelled in Praat (Boersma and Weenink, 2012) by experienced phoneticians.

The spoken texts contained a total of 1,452 occurrences of schwa, of which 692 were found in monosyllabic grammatical words only (see above). These were selected for further analyses, yielding about 115 schwas per speaker.

The extracted parameters were as follows:

- average F1 and F2: measured in the middle third of each vowel
- average bandwidths of F1 and F2: measured in the middle third of each vowel
- spectral slope: measured as band energy and band density difference in the frequency bands of 350–1,100 Hz and 2,300–5,500 Hz in the middle of each vowel

The extraction of all parameters was done with Praat scripts.

Our method of determining the values of spectral slopes deserves some explanation. Similarly to Hammarberg *et al.* (1980), we calculate the difference in energy between specific frequency bands. In the study of Volín and Zimmermann (2011), a similar method was successfully used to distinguish stressed from unstressed vowels of three Czech speakers. The authors hypothesized that excluding the bands which correspond to F0 and F2 will improve the discriminatory power of the method, since the enormous spectral energy of F0 could cloud the relevant spectral measurements, and the highly variable energies in the F2 region are expected to code mainly vowel identity.

In the current study, the Praat predefined functions of band energy difference and band density difference were used, which calculate the sum (in case of the former) or the average (in case of the latter) energy in the given high-frequency band and subtract it from the low-frequency band. Figure 1 illustrates the measurement.

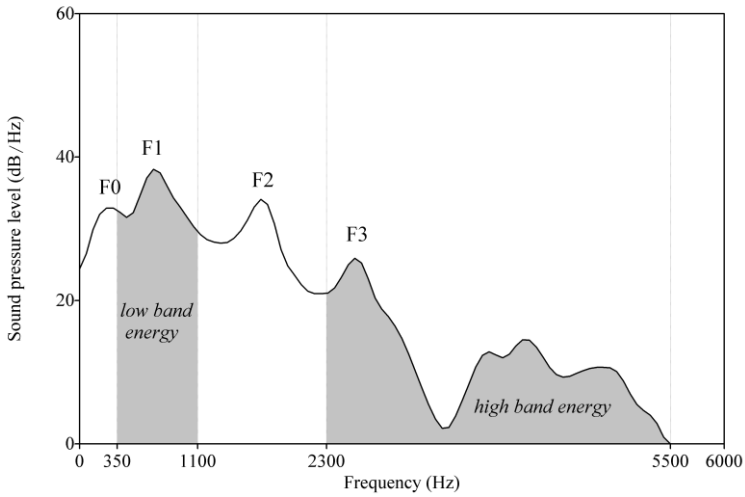

Figure 1. Spectrum of a vowel with highlighted low (350–1,100 Hz) and high (2,300–5,500 Hz) frequency bands. The ranges of F0 and F2 are excluded from the measurement (see text).

To assess the statistical validity of the results, one-way analysis of variance (ANOVA) was employed with the extracted parameters as dependent variables and L1 of the speaker (English/Czech) as a factor.

3. Results

The values of the first two formants did not turn out to be a reliable correlate of the difference between British and Czech speakers. In both cases, the results were non-significant at the level of $p = 0.05$ ($F(1, 690) = 0.0011$; $p = 0.97$ for F1 and $F(1, 690) = 2.62$; $p = 0.11$ for F2).

The bandwidths performed substantially better: both F1 and F2 bandwidths were able to distinguish the British from Czech speakers with high statistical significance (see Figure 2a-b). The ANOVA result for F1 bandwidth was $F(1, 690) = 6.14$; $p = 0.013$ and for F2 bandwidth $F(1, 690) = 12.96$; $p < 0.001$.

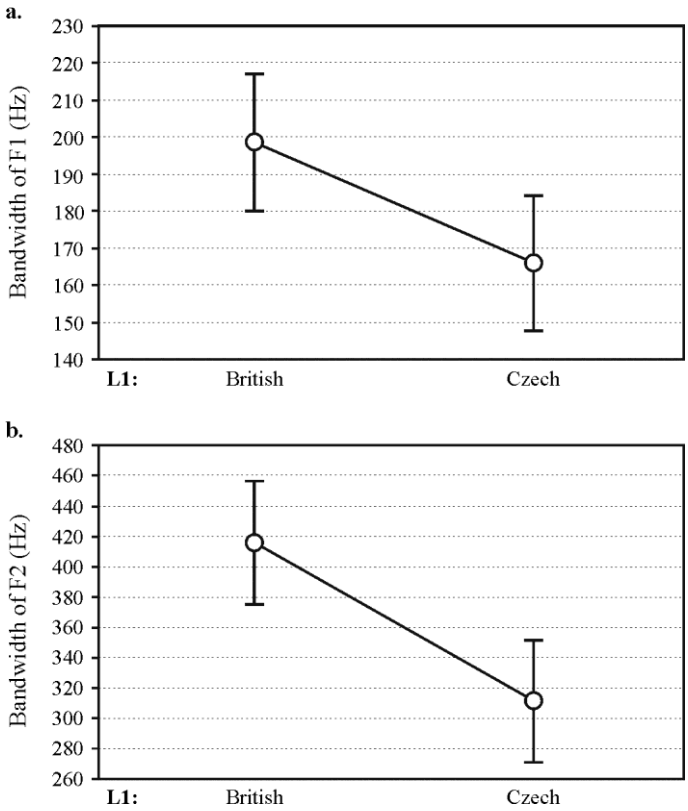

Figure 2. Average values of F1 bandwidth (a.) and F2 bandwidth (b.) of British *vs.* Czech schwas. Whiskers denote the 0.95 confidence interval.

In both cases, our Czech speakers produced narrower formant bandwidths, which should generally be interpreted as stronger or better defined formants. Nevertheless, these results have to be rationalized with some caution, since the difference in recording conditions of the two groups of speakers (i.e., the frequency response of the recording room) may have slightly influenced the bandwidth values.

The two measures of spectral slope, band energy and band density difference, yielded very similar results and almost identical values of the test criterion F in the inferential analysis of variance. Since density seems to be a better option for computational reasons (it is not sensitive to changes in the frequency band boundaries), we will only present band density difference results for the sake of simplicity. As can be seen in Figure 3, the difference in spectral slope between British and Czech speakers also turned out to be significant ($F(1, 690) = 5.13$; $p = 0.024$).

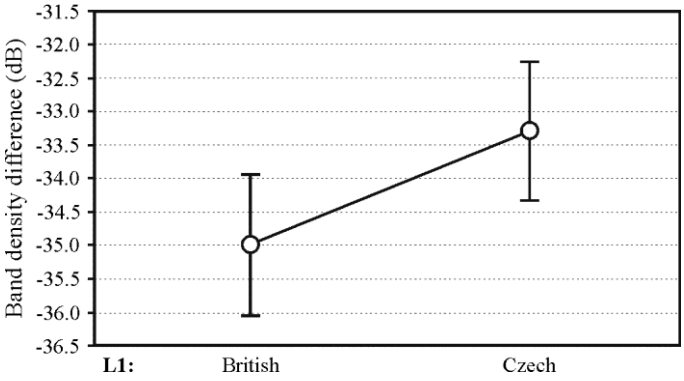

Figure 3. Average values of band density difference (spectral slope measure) of British *vs.* Czech schwas. Whiskers denote the 0.95 confidence interval.

Figure 3 shows the difference in average values of spectral balance between the British and Czech speakers, where the Czech speakers produced values closer to zero (the scale being negative) of the band density difference, meaning a flatter spectral slope. The spectral slope in schwas of the British speakers was steeper, with less energy in higher frequencies. This is found in less prominent vowel sounds.

4. Discussion

The English central mid lax vowel is relatively frequent in spoken texts. Due to its metrical role (the lowest level of prominence), it is an important element of the rhythmic structure of English. Czech speakers use typologically different rhythm in their mother tongue, and vowel reduction does not take place in the standard pronunciation of Czech. It can be expected that their foreign grasp of English would reflect this situation.

Our study showed that the formant values, despite being the primary descriptors of vowel quality, do not allow for discrimination between native (British) and Czech speakers of English. However, the measurements of spectral slope led to a statistical difference between the average schwa production of British and Czech speakers who were reading out identical texts under identical conditions from the point of view of the communicational context. The Czech schwas displayed a less steep decrease of energy towards the higher frequencies, which means more prominent vowel sounds.

The current results also revealed a significant difference in formant bandwidths. It was especially the second formant that was produced wider by the native speakers than by the non-natives. Since greater bandwidths generally indicate weaker formants, it could be argued that this result is a confirmation of the same phenomenon as the spectral slope measurements. We believe it is. However, it has to be remembered that determining formant bandwidths is computationally much more vulnerable than detecting energy in the spectrum and, also, the formant bandwidths might be more sensitive to recording conditions.

In our future research, we would like to take a closer look at individual analyzed items in the recorded texts. For instance, it could be the case that prepositions are more

useful in discrimination between British and Czech speakers than conjunctions due to the fact that they are more consistent in occurring at weak metrical positions. Some speakers tend to hesitate on conjunctions or fortify them as a signal of a new syntactic unit. Similar, but less obvious relationships might be found for auxiliary verbs and pronouns. In addition, our material also provided over 700 schwas in polysyllabic autosemantic words. These vowels will be analyzed using the same methodology as was used for the current sample after a system of finer classification is developed for the lexical items.

Acknowledgements

This study was supported by the Programme of Scientific Areas Development at Charles University in Prague (PRVOUK), subsection 10 – Linguistics: Social Group Variation. The second author was supported by Internal Grant of the Faculty of Arts, Charles University in Prague (VG192).

References

Ashby, M., Ashby, P., Baldwin, J., Holmes, F., House, J. and Maidment, J. 1995. Broad transcription in phonetic training. *Proceedings of the XIIIth ICPhS*: 170–173. Stockholm: IPA.

Barry, W. J. 1998. Time as a factor in the acoustic variation of schwa. *Proceedings of 5th Int. Conf. on Spoken Language Processing*, Sydney: 3071–3074.

Boersma, P. and Weenink, D. 2012. *Praat: doing phonetics by computer* (version 5.3.14). Retrieved from http://www.praat.org/.

Browman, C. P. and Goldstein, L. 1992. "Targetless" schwa: an articulatory analysis. In G. J. Docherty and D. R. Ladd (eds.) *Papers in Laboratory Phonology II*. Cambridge: CUP: 26–56.

Derwing, T. M. and Munro, M. J. 2009. Putting accent in its place: rethinking obstacles to communication. *Language Teaching* 42 (4): 476–490. DOI: 10.1017/S026144480800551X

Flemming, E. and S. Johnson. 2007. Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association* 37: 83–96. DOI: 10.1017/S0025100306002817

Fry, D. B. 1947. The frequency of occurrence of speech sounds in Southern English. *Archives Néerlandaises de Phonétique Expérimentale* 20: 103–106.

Gobl, Ch. and Ní Chasaide A. 2003. The role of voice quality in communicating emotion, mood and attitude. *Speech Communication* 40: 189–212. DOI: 10.1016/S0167-6393(02)00082-1

Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J. and Wedin, L. 1980. Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngologica* 90: 441–451.

Hanson, H. M., Stevens, K. N., Kuo, H-K. J, Chen, M.Y. and Slifka, J. 2001. Towards models of phonation. *Journal of Phonetics* 29: 451–480. DOI: 10.1006/jpho.2001.0146

Lev-Ari, S. and Keysar, B. 2010. Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology* 46: 1093–1096. DOI: 10.1016/j.jesp.2010.05.025

Lindblom, B. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35: 143–162.

Nakatani, L. H., O'Connor, K. D. and Aston, C. H. 1981. Prosodic aspects of American English speech rhythm. *Phonetica* 38: 84–105.

Sluijter, A. M. C. and Van Heuven, V. J. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100: 2471–2485.

Sundberg, J. and Nordenberg, M. 2006. Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech. *Journal of the Acoustical Society of America* 120/1: 453–457. DOI: 10.1121/1.2208451

Volín, J. and Zimmermann, J. 2011. Spectral slope parameters and detection of word stress. *Technical Computing Prague*: 125 – 5 pages. Praha: Humusoft.

• **Research in Language, 2013, vol. 11.1** • DOI 10.2478/v10015-012-0006-8

FRENCH LEARNERS OF L2 ENGLISH: INTONATION BOUNDARIES AND THE MARKING OF LEXICAL STRESS

CELINE HORGUES
Université Sorbonne Nouvelle- Paris 3
celine.horgues@univ-paris3.fr

Abstract
In English, prosodic parameters play a major role at two main levels. First, they indicate the intonation at the level of the utterance by marking the distinction between sentence types (statements *vs* questions) and they are related – although more or less directly- to the informational and grammatical structures of the utterance. Secondly, prosodic cues also contribute to marking the stress pattern at the level of the word (word stress or lexical stress).

Even if it is useful to dissociate these two levels theoretically, when looking at their phonetic implementation in an utterance, it soon appears that the exact same prosodic cues are used (namely fundamental frequency, duration, and intensity). Contrary to what happens in tone languages, there is no pre-set prosodic configuration attached to each word in English. Yet, words in discourse retain a relative accentual independence even though the exact prosodic implementation of word stress depends on the specific intonational context expressed in a given utterance (Pierrehumbert, 1980).

In French, stress pertains to the level of the group of words rather than to the individual word, which has no real accentual autonomy. Therefore, it is not surprising that French learners of English are faced with a major challenge: how to ensure the marking of lexical stress while, at the same time, using the same prosodic cues to indicate the intonational structure of the utterance.

My hypothesis is that some intonational contexts impose a bigger constraint on French learners of English than others. These particularly challenging contexts are the final position at the boundary of non-final clause, or the boundary of a rising interrogative. Other contexts, like the quotation form or the final position of a statement, are less challenging for the intonational marking of lexical stress.

To test my hypothesis, I collected passages of read speech by thirteen upper intermediate/advanced French learners of English along with the same passage read by ten native English speakers. Two trisyllabics carrying primary stress on the second syllable (*com*ˈ*puter*, *pro*ˈ*tection*) were placed in a series of intonational contexts under observation. The test-words were then extracted and submitted to native English listeners. The perceptual results show that the predicted 'challenging' contexts indeed caused substantial instability in the learners' placement of lexical stress as perceived by native English listeners.

Keywords: French-accented English, prosody, word stress, intonational boundaries.

1. Introduction

When looking into the acquisition of prosody, researchers and teachers are faced with an intriguing paradox: prosody is one of the first language features to be acquired in one's mother tongue (Konopczynski, 1991, Dupoux et al., 2003) and its acquisition possibly starts even *before* birth (Kuhl and Iverson, 1995), but when later learning a second language (henceforth L2), prosody remains a challenge even at the later stages of L2 learning (Tahta et al., 1981 and Hewings, 1998[1]).

French learners of L2 English are known for being easily identified (mainly but not exclusively) through the prosodic specificities of their interlanguage, which make them perceptively distinct from native English speakers. A previous perception experiment based on low-pass filtered speech showed that when native English listeners were asked to listen to the same sentence read by native speakers and by French learners of L2 English[2], prosody alone was a sufficient cue for their identification of French accent in English (Horgues, 2010).

One of the questions at stake is the origin of the observed prosodic deviations from L2 norms: Do they result from the prosodic interference from the L1 (in our case L1-French)? Or are they influenced by other external factors which are not specific to the L1?

I use the term 'prosody' in its broadest sense: encompassing the phenomena of word stress and sentence stress, rhythm, intonation, and related phenomena like pauses and segmental cues attached to stressed *vs* unstressed syllables [see Vaissière, 2002].

Among the prosodic characteristics of French-accented English, the perception and production of English word stress has long been an acknowledged difficulty for French learners. Some psycholinguists have even referred to it as a case of 'stress-deafness' (Dupoux and Peperkamp, 2002). But it is now worth looking at the question of the realisation of English word stress by French learners - not as an isolated phenomenon - but in its interaction with other prosodic levels in an utterance: especially the intonational level in the utterance, which I call the supra-lexical level, *i.e.* above the level of the individual word. It is interesting to observe how the phonetic realisations of these two levels (both lexical and supra-lexical) are closely intertwined.

My hypothesis is that some intonational contexts are more challenging than others for the prosodic marking of word stress by French learners. L1 prosody (French prosody) being largely (but not solely) responsible for these difficulties, I will start by giving a comparative overview of the prosodic features of L1-French and L1-English relevant to the present study. I will subsequently expand on what I mean by challenging contexts, before presenting the experimental setup I devised to test my hypothesis. The experimental results I collected will finally point to possible pedagogical implications.

---

[1] Hewings (1998: 318) : "*Intonation features of a foreign accent are the last and perhaps the most difficult to eradicate (the prosodic system of a language is the last to be mastered, it tends to become fossilised in advanced learners and it remains to characterise learners as having a foreign accent after segmental deviations from the target language have been eradicated)*."

[2] Even when based on a relatively short sentence like: "*Henry looked at him rather unpleasantly"*. Some distractors (the same sentence read by a German and an Italian learner) were also presented to the listeners.

2. Comparing some prosodic features of L1-French and L1-English

Prosodic analysis is a complex area due to the fact that the same basic acoustic cues (mainly the fundamental frequency, the duration, and the intensity, and sometimes vowel quality) implement the form of different linguistic levels. Indeed, not only do they ensure the marking of intonation at the level of the utterance (syntactic and informational structures) but they are also used to indicate word stress. However, the relation between sentence intonation and word stress is not the same in French and in English.

2.1 French: word stress and intonational boundaries

Following, Troutbetzkoy (1939) and Garde (1968), the traditional opposition between so-called 'free-stress' or 'mobile stress' languages (like English) and so-called 'fixed-stress' languages (like French) is in fact misleading.

In particular it is difficult to say that word stress is fixed in French since stress does not systematically affect the last syllable of words. Incidentally, perception experiments have shown that French listeners obtain a low rate of agreement when asked to identify the placement of word stress in their mother tongue and word stress in French has been described as more "fleeting" (*fuyant)* than in other romance languages (Fónagy, 1980). Phoneticians thus now prefer to speak about *group accent* rather than *word stress* when describing French (Vaissière, 2002 and Martin, 2009). They generally agree that stress in French belongs to the level of the *group of words* as opposed to the individual word, which has no real accentual autonomy. Accent is associated with the rhythmic group boundaries: the last syllable in a prosodic unit is stressed. It functions as a demarcation accent indicating the final boundary and it is sometimes supplemented by an optional initial accent to mark the start of the unit, therefore creating an accentual bridge between the group-initial and the group final accent. The acoustic correlates of the final demarcation accent in French are a clear lengthening of the final syllable accompanied by an ample F0 rise on that same syllable. French phoneticians like Rossi (1980) and Vaissière (2002) have described this as "*un syncrétisme", a merger* between lexical and supra-lexical prosody in French.

2.2 English: word stress and intonational boundaries

In English, however word stress is part of the identity of individual words and, as such, is fairly stable. Contrary to what happens in tone languages however, there is no pre-set melodic pattern attached to each word. Pitch movements have been described as the main correlates of word stress (the main *transducer of stress* for Pierrehumbert, 1980), but the exact shape of a term's melodic contour is influenced by the status of this term in the intonation unit (nucleus, head, post-nucleus, etc...) and by the associated tone in the case of a nucleus accent (fall, rise, fall rise). Many studies emphasize the impact of the

general intonation contour over the realisation of word stress in discourse, *i.e.* 'accent' (see Pierrehumbert, 1980 and Gray, 2001[3] for example). Pierrehumbert explained:

> The relation of F0 to stress is not as direct as this. A word with a given stress pattern could have any of a number of different F0 contours depending on the intonation pattern that was being used. A given F0 pattern could be compatible with more than one conclusion about the location of stress. (Pierrehumbert, 1980: 103)

However, even though their prosodic realisations are relatively variable and context-dependent, words in discourse retain a relative accentual autonomy whose exact prosodic implementation depends on the specific intonational context expressed in a given utterance.

Considering these differences between French and English, it is not surprising that French learners of L2 English are faced with a major challenge: how can the speaker ensure the marking of lexical stress while, at the same time, using the same prosodic cues to indicate the intonation of the utterance?

The hypothesis I propose is that the prosodic realisation of certain syntactic-informational contexts is responsible for a particular difficulty in the pronunciation word stress patterns by French learners. More precisely, the correct placement of word stress will be made unstable and therefore its perception by native listeners will no longer be guaranteed. I call these contexts 'challenging' contexts.

3. Defining the challenging contexts

I define challenging contexts as linguistic contexts where, in L1-French, prosody is the prime marker of the syntactic and/or informational structures of the utterance. As previously posed, the prosodic cues (especially duration and fundamental frequency) are strongly constrained in French to indicate :

- Firstly, the boundaries of syntactic and/or informational units, and more particularly the right-hand boundary (or final boundary[4]). In French, what has been described as the *major* continuation rise ("*continuation majeure"*) occurs at the boundary of a dependent (non-final) unit in a sentence. These ample continuation rises are perceptively very prevalent when listening to French (see figure 3 below).

In English, continuation contours have received less attention from phoneticians and there seems to be more variation in the direction and shape of the associated pitch movements with a rising, a falling-rising, and even falling contour being found (see Cruttenden, 2001 and Wells, 2006). More importantly, these boundary contours do not

[3] Gray (2001) also underlines the difficulty of talking of a once-and-for-all hierarchy of the acoustic cues of lexical stress or sentence stress without considering the intonation over the whole tone-unit.

[4] Although left-boundary marking through F0 resetting on the first syllable of the unit also exists and it is associated with microprosodic features on initial consonants for example.

generally merge with the prosodic marking of lexical stress contrary to what happens in French (see Chen, 2003 and Martin, 2009[5]).

- Prosodic cues also help to mark sentence-type (*statement vs question)* and maybe more so in French than in English, since in conversational French questions frequently have the syntactic form of a statement and the rising intonation is the only indication that a question is being asked.
- In French, cases where intonation alone is the marker of linguistic focus are very rare (contrary to English). Indeed, informational focus is often expressed through a combination of syntactic, lexical and intonational devices together (Martin, 2009).

4. Experimental set-up and results

4.1 The corpus, the speakers and listeners

To test my hypothesis, I proposed seven linguistic contexts whose prosodic realisation might be more or less challenging for French learners with regards to the marking of lexical stress. These contexts are numbered from one to seven and can be found in the full dialogue of Appendix A (note that the speakers were presented with the normal, unnumbered version of the script). My predictions were that challenging contexts for French learners would be: context C1 (end position in a non-final tone-unit or *major continuation*), context C3 (end position in a rising interrogative), and contexts C2 and C4/5[6] (respectively corresponding to cases of post-focus deaccentuation and intonational focus). On the contrary, I predicted that contexts CIT (citation form in an anchor sentence), C6 (mid-position in an intonation unit) and C7 (end position in a final declarative tone-unit) would represent less of a challenge for the prosodic marking of lexical stress by French learners of English (see Appendix B). In this paper I will focus mainly on the analysis of the contexts corresponding to unit boundaries (*i.e.* C1, C3, C7).

I designed a dialogue which was intended to be read by two speakers: a native English speaker and a French learner. The dialogue was written so that two trisyllabics (*com*ˈ*puter*, *pro*ˈ*tection*) would be found in all seven contexts. These two words were chosen because they are simple, common terms that learners know at this stage, and because they bear primary stress on their second syllable (thus allowing for the possibility of stress misplacement on the first or the last syllable).

---

[5] Martin (2009: 191) explains that due to this merger in French, French learners of Italian might confuse a boundary contour and word stress. (« *Les contours de continuation majeure, positionnés en français sur la dernière syllabe des syntagmes importants, sont réalisés dans les autres langues sur deux syllabes distinctes, la tonique par un contour plat et la finale par un contour fortement montant (à moins que la tonique ne soit en position finale). Les auditeurs francophones ne connaissant pas l'italien identifient alors ce contour comme celui de leur propre langue et attribuent faussement un accent final aux grands groupes syntaxiques. »)*

[6] Initially, I had dissociated cases of intonational focus (C4) and contrastive intonational focus (C5). Since the phonetic implementations of these two types of focus on the test-words by native English speakers were extremely similar, I subsequently decided to gather them under one single category of intonational focus (C4/C5).

The speakers were thirteen French learners of English (second-year students of English at a Parisian university, with an upper-intermediary or advanced level in English). The control group of native speakers was composed of ten speakers representing a variety of British, American, Australian and Irish speakers. Dialectal homogeneity was not an important criterion in this test, the assumption being that, beyond intra-language variation, listeners would perceive some variation as non-native (so-called 'deviations') whereas other variation would be tolerated and interpreted as native variants.

Parts of this read corpus were then used as a basis for a perception test. The test-words (*protection, computer*) read by the French learners and by the control group of native speakers were extracted from their context and presented in isolation to native listeners of English. The words were presented on a computer with a headset. The listeners were asked to identify the syllable they heard as the most prominent in the different test-words (the first, second or third syllable). To avoid listener fatigue, 9 sub-sets of stimuli were created and each was presented to four different listeners.

4.2 The perceptual results

Table 1 below presents the perceptual results drawn from the native-English listeners' perceived placement of lexical prominence in the two types of stimuli: perceptual judgments of the learners' productions of the two test-words appear in the top part of the graph, and in the bottom part of the graph appear the listeners' perceptual judgments of the stimuli produced by the control group of native speakers. As a control condition, listeners were also presented with native speaker productions.

On the horizontal axis, each bar corresponds to the context the word was extracted from, starting with C1 at the left of the horizontal axis. The different types of pattern (black, checked or dotted) convey information about the type of assessment made by the listeners. The perception of the 'correct' or predicted placement of prominence on the second syllable is not coded (e.g. *comPUter).* Cases when no consensus emerged from the listeners' answers, which I called cases of 'perceptive indermination', are coded with the dotted pattern in the answer bars. Answers relating to a perceived placement of prominence on the first syllable (initial stress *COMputer*) appear in a checked pattern and perceived prominence on the final syllable is coded in black (*compuTER*).

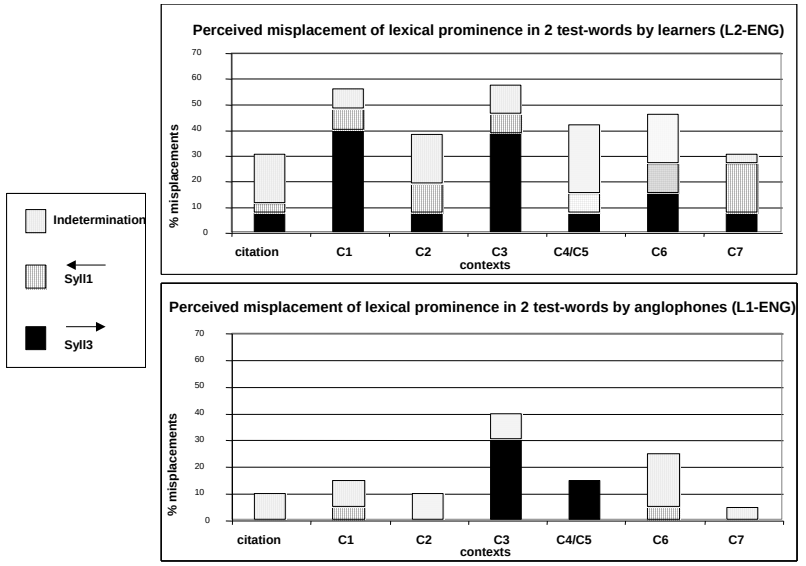

Table 1. Perception answers provided by native-English listeners on two types of stimuli: test-words produced by French-learners of English (top part) and by native-English speakers (bottom part)

The perceptual results show that:

- Perception of learners' stimuli resulted in twice as many indeterminate judgements as when they were based on native stimuli (indeterminate judgements correspond to the dotted areas).
- The answers relating to the perception of prominence on the first and on the last syllable can be grouped under one category since they reflect instances of *perceived misplacement* of lexical prominence. These cases of perceived misplacement are far more frequent for learners' stimuli than native speakers' stimuli (2.5 times more frequent[7]).

It might appear surprising that native productions were not exempt from judgements of misplacement, although these instances remained very exceptional. One token in particular was responsible for these unexpected judgments: *ComputerC3*. One reason for this might be that the question asked is not a neutral, traditional *Yes/No* question aiming at eliciting information. Projecting very subjective and evaluative overtones, many native speakers produced this question as an exclamative, echo-question (meaning: *this can't be true*!). The associated pitch rise was therefore realised as an exceptionally ample and high melodic rise. This observation also shows that the concepts of *word stress* (a theoretical, phonological concept) and *perceived lexical prominence* (a psychoacoustic, perceptive phenomenon) sometimes do not overlap completely.

[7] A Chi test showed that answers reflecting an incorrect placement (syllable1, syllable2 or indetermination) based on the learners' productions were signicantly different from those based on the native speakers' productions. Chi2(1) = 24.2 with $p<0.001$, N1=182, N2=140.

As regards the perception of learners' stimuli, some contexts led to more perceived misplacements than others. This was more particularly the case for contexts corresponding to tone-unit boundaries such as C1 (located at the major continuation at the boundary of a dependent unit in an utterance) and C3 (the end of a rising interrogative contour). C1 and C3 were each responsible for fifteen perceived misplacements out of twenty-six tokens. These contexts therefore favoured substantial instability in lexical prominence as perceived by native English listeners. On the other hand, other contexts favoured more stability, and in that sense, they should be considered less challenging for learners. This is particularly true of the quotation form or the end of a final (non-dependent) unit in a declarative sentence (C7).

The more challenging set of contexts were C1, C3 and C6, and the less challenging set were Citation, C2, C4/5 and C7. The difference between the perception of challenging and non challenging contexts is statistically significant[8].

Parts of these results did not exactly match my initial predictions however. This can be explained by the discrepancy between the projected interpretations prompted by the dialogue and the actual realisations the learners produced. For example, contexts C2 and C4/5 had to be considered not relevant for the analysis of the learners' productions quite simply because the learners proved unable to produce the expected intonational focus and post-focus deaccentuation prompted by the dialogue. Contrary to native speakers - who produced intonation patterns matching the projected meanings - the learners systematically produced broad focus where narrow focus was expected, therefore realising intonation patterns similar to C7 instead (the end of a neutral declarative sentence). Interestingly, this turned out to be the case even when intonational focusing and deaccenting were visually encouraged through the use of italics on the nuclear syllable ("It's a shame because I'm quite sure my computer already *has* a virus protection"). This observation shows that the central role of intonation to mark focus in English is clearly not part of the French learners' active competence at this stage.

What is more, context C6, which theoretically did not correspond to a tone-unit boundary, was often granted a separate tone-unit by learners and therefore produced with a similar pattern as C1 (the end boundary of a non-final tone unit). This tendency for learners to oversegment L2 speech is frequent and not language-specific, being caused by the cognitive overload of speaking in a foreign language.

Ultimately then, what seems to have been be the determining factor to account for the learners' difficulty in marking lexical stress was whether the test-word was situated at the boundary of a tone-unit or not (in the actual realisations of learners). Let's now turn to the acoustic specificities of the most challenging contexts for French learners regarding the marking of lexical prominence: contexts C1 and C3.

[8] A Chi test showed the set of most challenging contexts (C1, C3 and C6) received significantly more incorrect placement answers than the set of less challenging contexts (C2, C4/C5, C7 and Citation). $Chi2(1) = 5.3$ with $p<0.05$, $N1 = 78$ and $N2 = 104$.

4.3 Acoustic analysis of contexts C1 and C3:

a) Context 1:

It corresponds to the continuation contour at the boundary of a dependent unit (end position in a non final tone-unit).

Eg- *//The idea of a good* ***proTECtion*** **(C1)** */ is to guarantee that your computer doesn't get infected by a VIRus//*

Eg- *//Well, the only protection on my* ***comPUter*** **(C1)** */ is just a basic anti-VIRus software, I think//*

First it is interesting to observe a clear difference in the orientation of the contours produced by French learners and native English speakers. Almost all ten native English speakers produced a fall on the last two syllables of the test-word in this context (from the mid-level to the bottom level of their pitch-ranges). The slope of the falls they produced was between 1.7 and 2.9 semi-tones. The fall was sometimes followed by a very small additional rising hook as in the example below (figure 1). The solid line corresponds to the Fo curve and the dotted line corresponds to the intensity curve.

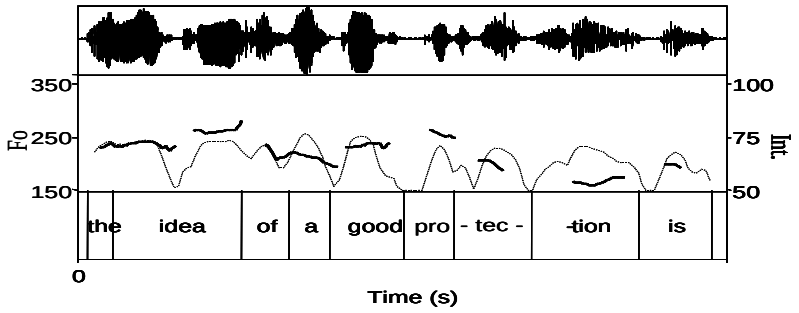

Figure 1. Native speaker, InfoN13protectionC1 (Fo range: 150-300 Hz, intensity range: 50-100 dB)

The temporal ratio of last syllable relative to the duration of the whole word (*computer, protection)* is similar to its ratio when produced by the same speakers in its quotation form. Therefore, there was no lengthening[9] of the last syllable *–tion/-ter* when compared to the quotation form. A decrease in intensity between the second syllable and the last syllable of the test-words was realised by only half the native speakers and it did not seem to be as important a prosodic cue as the frequency variations for the marking lexical prominence on the test-words.

However, in the same linguistic context, most French learners produced a rise on the last two syllables. A falling contour similar to the native speakers' favoured melodic pattern was produced by only a quarter of learners for *computerC1* and one third for *protectionC1.* The rest of the learners realised a melodic rise between the second and the last syllable of the test-words reaching the mid-high level of their pitch ranges.

[9] For the control group of native speakers, the mean lengthening ratio of the last syllable (compared to the quotation form) was –5% for *computerC*1 and –3% for *protectionC1.*

Figure 2 below gives an illustration of the typical pattern realised by the French learners of the corpus. The learners lengthened the last syllable of *computerC1* relative to the temporal pattern observed in the quotation form but there was no substantial lenghthening for *protectionC1*[10].

The decrease in intensity between the second and the last syllable is not clearly realised by the group of learners.

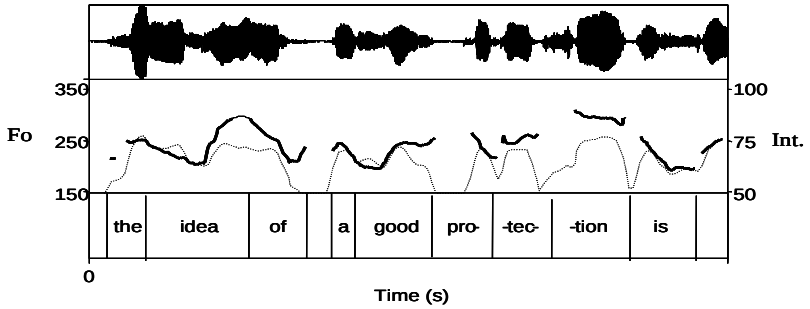

Figure 2. French learner, Info20protectionC1 (Fo range: 150-350 Hz, intensity range: 50-100 dB)

As previously seen, intensity might not be relevant cue to look at since it was not a systematic correlate of lexical prominence in the native speakers' productions.

It was highly revealing to compare the prosodic trends observed in the learners' productions with control data collected (for all the thirteen learners) in similar linguistic contexts in their mother tongue (French). This is particular clear in the case of '*protection'* because the same lexical item was used in the French version of the dialogue.

Eg. (L1-French) // *le principe d'un bon système de* ***protectionC1/*** *c'est de garantir que votre ordinateur ne soit pas infecté par des virus //*

Indeed, the continuation contour was realised as a systematic rise in L1 French in this C1 context. Figure 3 below shows that for the same learner as previously presented, the continuation contour is almost identical to what she produced in L2 English. The shape, alignement and pitch range of the learners' contours are very similar to what happens in her L1, which strongly supports the idea of prosodic interference of L1 French in this case.

[10] Their average lengthening ratio for *computerC1* was 9.8% and 0.5% for *protectionC1.*

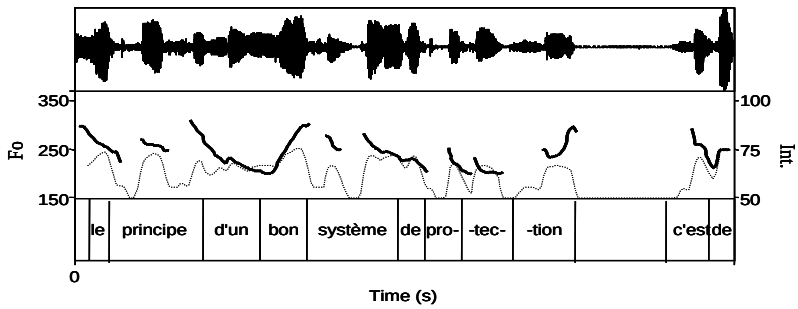

Figure 3. French speaker (L1-French), Info20protectionC1 (Fo range: 150-350 Hz, intensity range: 50-100 dB)

Therefore the acoustic parameters responsible for the listeners' perceived prominence on the last syllable in learners' productions are certainly related to some lenghthening of the final syllable in the test-words coupled with a deep and prominent rise starting either on the very last syllable or else stretching over the whole word. Intensity could not be considered to be a consistent cue distinguishing native speakers' and learners' productions. The analysis of the first two formants of the final target vowel /ə/ did not allow me to show less centralisation in learners' productions than in native speakers' productions, due to the big variability in formant values among speakers in a same group.

The most important prosodic cues accounting for the difference between learners and native speakers' production were: the duration and pitch patterns on the last syllable. Significantly, the other most challenging context (context C3) shares most of these prosodic specificities.

b) Context C3:

It corresponds to the end of a rising interrogative.

Eg- //*Is your PC equipped with any type of computer* ***proTECtion*** **(C3)**?//

Eg- ***/With your*** comPUter (C3)**?**//

I will only deal with the realisation of *protectionC3* here because, as previously mentioned, *computerC3* turned out not to be interpreted as a straightforward question and was therefore discarded from the acoustic analysis. Interesting observations can be made when looking at the oriention of *Yes/No questions* by the two groups of speakers (which also emerged in another production test I carried out with French learners of L2 English, Horgues, 2010).

All thirteen French learners produced a rise as the default contour on the *Yes/No questions* of the corpus whereas the contours realised by native English speakers were more varied. Indeed, only half the native speakers chose a simple rise, while the other half realised a fall-rise or else a simple fall. In the native speakers' productions, there seems to have been room for variation depending on the associated interpretation the speaker projected in a given utterance. When comparing rising contours by learners and English speakers, the alignment of the final rise also differs between the two groups. Indeed, the start of the rise is in strict alignment with the syllable carrying primary lexical stress in native productions, the rise being initiated only on the stressed syllable

which is anchored at the bottom level. In the learners' productions, however, the rise is not strictly aligned with the theoretically stressed syllable: it rather stretches over the whole word (compare figure 4 and figure 5 below).

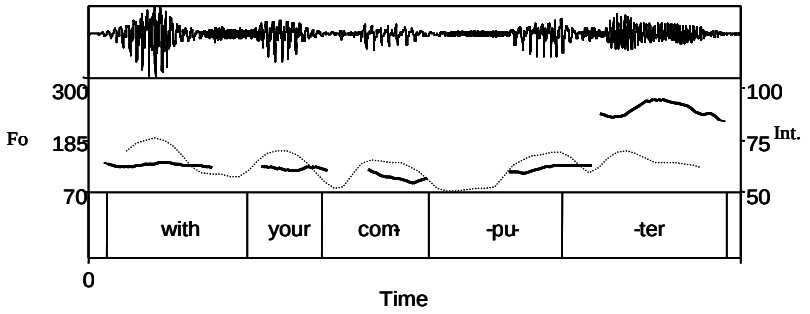

Figure 4. Native speaker, InfoN12computer (Fo range: 70-300 Hz, intensity range: 50-100 dB)

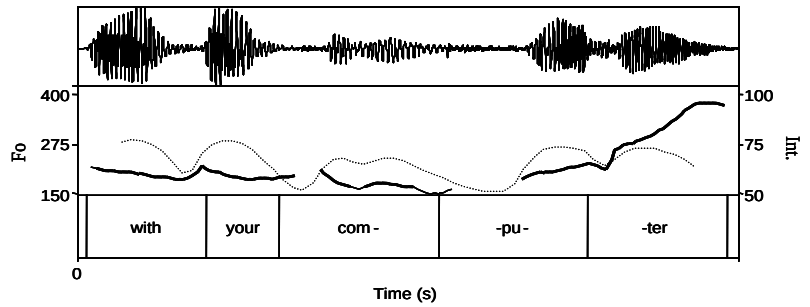

Figure 5. French learner, Info26computerC3 (Fo range: 70-400 Hz, intensity range: 50-100 dB)

What contexts C1 and C3 have in common is the way learners realise them: through a lengthening of the last syllable, associated with a pitch rise and an increase in intensity. Interestingly, another type of tone-unit boundary (C7, the end of a declarative statement) did not represent the same level of difficulty, due to the fact that the associated pitch contour is falling and that there is a natural decrease of intensity at the end of a falling declarative statement.

5. Conclusions and pedagogical perspectives

The results of the perception test investigating perceived lexical prominence confirmed that most prosodic contexts that were predicted to be challenging for French learners of L2 English indeed caused substantial instability in the perception of lexical prominence by native English listeners.

This tends to show that, at this level of competence, even though most learners already have a phonological representation of a word's stress pattern, its phonetic implementation should not be taken for granted. Indeed, some particular prosodic contexts make the implementation quite unstable although the pronunciation of the citation form is correct. We cannot simply blame the so-called 'stress deafness' of French learners of English. I think this would be too simplistic as it would suggest their complete inability to hear and produce word stress in English, which is not the case. What rather appeared in this experiment is that their ability to mark and maintain word stress prosodically was endangered or challenged in some particular contexts more than others. The prosodic constraints imposed at the supra-lexical (intonational) level seem to have an impact on the misplacement of lexical stress. French learners find it particularly difficult to realise the necessary *dissociation* of prosodic cues required to mark both lexical prominence and intonational patterns across the whole utterance.

The incorrect realisation of lexical stress patterns by learners is not inconsequential since it is widely acknowledged that it can cause a loss of intelligibility in L2 speech. Revealingly, most cases of misplacement in this experiment resulted in a stress-shift towards the last syllable of the test words (therefore: *protecTION, compuTER*). As Field (2005) has shown, right-hand stress-shifts have a particular negative impact on the segmentation of continuous speech by native English listeners (possibly more so than left-hand shifts). Intelligibility and lexical access are therefore seriously impaired.

Supra-lexical prosody not only imposes constraints on the realisation of stress patterns by learners but it can also represent a learning challenge for the pronunciation or the perception of certain phonemes or of intonational focus. For instance, Gray (2001) has shown that a rising contour or a falling-rising contour in English represented a challenge for French learners' perception of minimal phonemic pairs (*bead/bid),* lexical stress patterns or sentence stress patterns.

Once again, these results should encourage EFL teachers to depart from a form of teaching based on presenting word stress patterns in isolation. Theoretical knowledge about a word's stress pattern and the correct pronunciation of its quotation form is no guarantee that learners will be able to realise the expected lexical prominence when that same word is used in different discourse contexts. One suggestion would be to take more into account the impact of intonational constraints in the teaching of English phonology and phonetics. When teaching French learners the perception and production of English stress patterns for example, we might envisage a progressive exposition starting with less challenging prosodic contexts before moving to more challenging ones.

References

Bolinger, D. L. 1978. Intonation Across Languages, *Universals of Human Language, Greenberg,* J.H. (ed.) California, Stanford University Press, vol.2, 471-524.

Hirst, D. J. and Di Cristo A. 1998. *Intonation Systems: a Survey of Twenty Languages,* Cambridge: Cambridge University Press.

Chen, A. 2003. Language Dependence in Continuation Intonation, *Proceedings of the 15th International Congress of Phonetic Sciences,* Barcelona: UAB, 1069–1072.

Cruttenden, A. 1997. *Intonation,* Cambridge : Cambridge University Press.

Dupoux, E. and S. Peperkamp, S. 2002. A typology of Stress 'Deafness', in C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology 7*, Berlin: Mouton de Gruyter, 203-240.

Dupoux, E. ; Pallier, C. ; Mehler, J. ; Sebastian, N. (2003b) A Destressing 'Deafness' in French ?, *Journal of Memory and Language*, 36, p. 406-421.

Delattre, P. 1965. *Comparing the Phonetic Features of English, French, German and Spanish: an Interim Report,* Heidelberg : Julius Groos.

Field, J. 2005. Intelligibility and the Listener: the Role of Lexical Stress, *TESOL Quarterly, 39(3*), 399-423.

Fónagy, I. 1980. L'accent en français: accent probabilitaire, in Fónagy, I. and P. Léon. (eds.), L'accent en français contemporain, *Studia Phonetica*, 15, 123-227.

Frost, D. 2011. Stress-Cues to Relative Prominence in English and French: a Perceptual Study, *Journal of the International Phonetic Association (41/1)*

Garde, P. 1968. *L'accent,* Paris: Presses Universitaires de France.

Grover, C., Jamieson D. and M. Dobrovolski. 1987. Intonation in English, French and German: Perception and Production, *Language and Speech*, 30(3), 277-295.

Gray, M. 2001. Place et rôle de l'intonation dans l'apprentissage de l'anglais, unpublished PhD dissertation, Université Paris 12.

Hewings, M. 1998. Intonation Choices in the English Intonation of Non-Native Speakers, an Exploratory Study, in Sánchez-Macarro, A ; Carter, R. (eds.), *Linguistic Choice Across Genres: Variation in Spoken and Written English,* Amsterdam : John Benjamins. p 317-336.

Horgues, C. 2010. *Prosodie de l'accent français et perception par des auditeurs anglophones,* unpublished PhD dissertation, Université Paris Diderot-Paris 7.

Konopczynski, G. 1991. *Le Langage émergent: Aspects Vocaux et Mélodiques*, Hamburg: Ed. Buske Verlag.

Kuhl, P. K. ; Iverson, P. 1995. Linguistic Experience and the 'Perceptual Magnet Effect', in Strange, *Speech Perception and Linguistic Experience, Issues in Cross-language Research,* p. 121-154.

Martin, P. 2009. *Intonation du français,* Paris: Armand Colin.

Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*, PhD Thesis, MIT, Indiana University Linguistics Club

Rossi, M. 1980. Le français, langue sans accent?, in Fónagy, I. and P. Léon (eds.), L'accent en français contemporain, *Studia phonetica*, 15, Ottawa: Didier, 13-52.

Tahta, S. ; Wood, M. ; Loewenthal, K. 1981. "Foreign Accents: Factors Relating to Transfer of Accent from the First Language to a Second Language", *Language and Speech,* 24, p. 265-272.

Troubetzkoy, N. S. 1939. Grundzügue der Phonologie, *Travaux du Cercle Linguistique de Prague*, 7; French translation by Cantineau, J.: Principes de phonologie, Paris: Klincksieck, 1949.

Vaissière, J. 2002. Cross-linguistic Prosodic Transcription: French versus English, *Problemy I metody eksperimental'no-foneticheskih issleddovanij,* St;-Petersburg: St-Petersburg State University, 147-164.

Wells, J. 2006. *English Intonation*: *an Introduction*, London: Longman.

Wenk, B. J. and F. Wioland. 1982. Is French Really Syllable-Timed?, *Journal of Phonetics*, 10(2), 193-216.

Appendix A

Full script of the dialogue read and enacted by a native English speaker and a French learner of L2 English:
[on the phone, a customer (C) and a hotline operator (Op)]

(C)-Is this the **Computer (C6)** Protection Assistance Service?
(Op)-Yes Madam/Sir, what can I do for you?
(C)-Oh Hello! I'm calling to report a serious problem with my **computer (C7).**
You see, a few years ago, my brother Henry gave me his old **computer (C1**) and I've been using it at home. It's a PC.
I also have a brand new one at the office and that one works perfectly well.
(Op) -OK….Well I might be able to help you if you actually tell me what is wrong with your **computer (C2).** By the way, do you mind if our service records our conversation? This is for….
(C)-Um…well, collecting and recording details about your customers is one thing, but going as far as recording a conversation is a little surprising…
In fact, it is not particularly pleasant, but if you say you're doing this to improve your service…I suppose I don't mind…
(Op) -Thank you, Madam/Sir. We'll try to do our best to help. Let me start off with a few questions: Is your PC equipped with any type of computer **protection (C3)?**
(C)- any computer what?
(Op) - computer **protection (C4/C5).** You know, the best way to make sure that your PC is safe from viruses is to install a reliable system of **protection (C7).**
Does your PC have one?
(C)-Well, the only **protection (C6)** on my **computer (C1)** is just a basic anti-virus software, I think. I can't remember the name now.
The thing is, it makes the machine really slow and after all…do I really need this stuff?
(Op) - Well, as I said the idea of a good **protection (C1)** is to guarantee that your computer doesn't get infected by a virus.
You see, a virus attack could seriously damage it.
(C)-Um….I suppose it has its advantages….but…do you mean I need to buy a new one? It's a shame because I'm quite sure my computer already *has* a virus **protection (C2)**.
(Op) - We'll see about that. First, could you give me the details of your **computer's (C6)** brand and serial number please?
(C)-Um…well… you know, it's now become rather difficult to have a conversation.
(Op) -With your **computer (C3)?**
(C)-Of course not!!! You really don't get it, do you? Do you really think I'm stupid enough to try and have a conversation with a machine!
I wasn't talking about my **computer (C4/C5)**… I'm talking about Henry, my brother!!
(Op) - So, you're not on particularly good terms with your brother anymore, from what I understand?
(C)- No….and I remember quite clearly the conversation that we had about this damn computer. I haven't spoken to him since.
(Op) -I'm sorry about that. What I can say is that it's definitely not always easy to have a calm conversation about computers….
(C)- No, you're probably right…
(Op)- Especially because for some reason, people tend to get completely worked up about technical problems!

Appendix B

➢ **More challenging contexts?**

*** C1- End of a dependent (non-final) tone-unit in the utterance**

Eg- *//The idea of a good pro**TEC**tion* **(C1)** */ is to guarantee that your computer doesn't get infected by a VIRus//*

Eg- *//Well, the only protection on my com**PU**ter* **(C1)** */ is just a basic anti-VIRus software, I think//*

*** C2- Deaccentuated post-nuclear position (*tail* of the tone-unit in the British theory)**

Eg- *//It's a shame because i'm quite sure my computer already HAS a virus protection* **(C2)//**

Eg- *//Ok…Well i might be able to help you/ if you actually tell me what is WRONG with your computer* **(C2)//**

*** C3- Final position of a rising interrogative contour**

Eg- //*I your PC equipped with any type of computer pro**TEC**tion* **(C3)** ?//

Eg- //*With your comPU**ter*** (C3)**?//**

* **C4- Intonational focus** or **C5- Intonational contrastive focus**

Eg- //*any computer **WHAT**?*// (*Op) - //Computer pro**TEC**tion* **(C4/5) //**

Eg- //*I wasn't talking about my com**PU**ter* **(C4/C5)** */I'm talking about HENry/ my BROther//*

➢ **Less challenging contexts?**

*** CIT- Quotation/citation form (in anchor sentence)**

Eg- *// That's a pro**TEC**tion, she said //* or *//That's a com**PU**ter, she said//*

*** C6- Mid-position in a tone-unit :**

Eg : *//Well, the only pro**tec**tion* **(C6)** *on my comPUter/ is just a basic anti-VIRus software//*

Eg : *//FIRST,/ could you give me the details of your com**pu**ter's* **(C6)** *BRAND/ and SERial number, please?//*

*** C7- Final position in last tone-unit of a declarative statement**

Eg : //*You KNOW/ the best way to make sure that your pc is safe from VIRuses/ is to install a reliable system of pro**TEC**tion* **(C7)//**

Eg : *//Oh helLO/ I'm calling to report a serious problem with my com**PU**ter* **(C7)//**

Transcription conventions:

Test-words are underlined

Bold type: prominence on the test-words' syllables.

Capital letters: nucleus (tonic syllable) in the tone-unit.

Test-words are underlined

PHONETICS LEARNING ANXIETY – RESULTS OF A PRELIMINARY STUDY

MAŁGORZATA BARAN-ŁUCARZ
School of English, University of Wroclaw
mbaran-lucarz@ifa.uni.wroc.pl

Abstract
The main aim of this paper is to verify the assumption that pronunciation learning during a course of phonetics is hindered by the feeling of anxiety (Phonetics Learning Anxiety) experienced by foreign language (FL) learners studying English as their major at universities or colleges. A study carried out among 32 students of the School of English at Wroclaw University (Poland) revealed a significant negative correlation of moderate strength between the subjects' level of Phonetics Learning Anxiety (PhLA) and their attainments on pronunciation tests (sentence, passage and word reading) conducted after a 45-hour (30x90-minute lessons) course of practical phonetics. The detrimental effect of PhLA on pronunciation learning was further supported by t-tests, in which the pronunciation of high anxiety subjects was found to be at a significantly lower level than that of low anxiety students.

The Phonetics Learning Anxiety Scale, a 44-item questionnaire based on a 6-point Likert scale, designed for the purpose of the research sheds light on the nature of this peculiar type of apprehension experienced by advanced FL learners in a specific educational context (i.e. a traditional classroom, rather than a language or computer laboratory), in which the major focus is on pronunciation practice. The obtained quantitative data imply that such factors as fear of negative evaluation (represented by general oral performance apprehension and concern over pronunciation mistakes, pronunciation self-image, pronunciation self-efficacy and self-assessment) and beliefs about the nature of FL pronunciation learning are significant sources of PhLA. Anxiety about the transcription test (IPA Test Anxiety) - one of the other hypothetical determinants of PhLA - did not prove to be correlated with the general level of Phonetics Learning Anxiety.

Keywords: Phonetics Learning Anxiety (PhLA), pronunciation self-image, self-efficacy and self-assessment, fear of negative evaluation.

1. Introduction

Numerous studies examining language anxiety (LA) have proven its detrimental influence on both FL learning and performance (see Horwitz 2010). Objective data show that the feeling of apprehension experienced by students is related to their pronunciation level (e.g. Horwitz et al. 1986; Price 1991; Young 1992) and, even more strongly, to their perceived pronunciation skills (Baran-Łucarz 2011). As many learners explain (e.g. Price 1991: 105), the anxiety experienced in the FL classroom is caused mainly by their

"great embarrassment" resulting from the belief of having a "terrible accent." Additionally, most observations reveal that the most anxiety-provoking task is oral performance in front of the whole class.

Taking into account the facts presented above, we may presuppose that a FL course of phonetics is particularly anxiety-breeding, especially when run with a group of students (rather than in the form of one-to-one tuition) in a traditional classroom (i.e. not in a computer or language laboratory). It seems that the feeling of uneasiness and worry might explain why despite high motivation to achieve native-like levels of pronunciation, predisposition for language learning, and phonetic competence, the progress of some learners is slow or hardly noticeable.

To ensure whether indeed anxiety is an important debilitative factor not allowing some students to benefit from a practical course of phonetics, and to elucidate the nature of the construct of anxiety experienced by learners during a course of phonetics, an empirical study was designed and conducted in June 2012 among first-year university students of English philology, whose results are presented in this publication.

The paper opens with a brief presentation of concepts that served as a ground for constructing the PhLAS, i.e. the concept of anxiety and language anxiety, and an overview of most important studies examining the influence of LA on FL learning and performance. Next is an attempt to define Phonetics Learning Anxiety. The second part of the paper constitutes a report on the aforementioned study, in which the research methodology, participants, research questions and hypotheses, and instruments are thoroughly explained. The discussion of quantitative data is followed by conclusions, suggestions for further research, and brief practical implications.

2. Theoretical background

2.1 The concept of anxiety

In the early definitions of anxiety, it is found to be synonymous with the phenomenon of fear (Piechurska-Kuciel 2008). For example, Darwin (1965/1872) suggested that anxiety derived directly from the expectation of suffering, while Lewis (1970: 77) considered it to be "an emotional state, with the subjectively experienced quality of fear as a closely related emotion." Rathus (1987) defines anxiety as "a negative emotion characterised by persistent fear and dread" (cf. Piechurska-Kuciel 2008: 27).

In modern psychology, anxiety is perceived as an unpleasant feeling that, unlike fear, may lack a direct source from the outside world. Friedman and Bendas-Jacob (1997: 1035) define anxiety as "a sense of discomfort and worry regarding an undefined threat," which may have not only a physical and physiological nature, but also be related to one's more or less conscious anticipation of his/her self-concept being damaged by "internal, real or imagined dangers" (Lesse 1988: 332). Finally, as contemporary psychologists explain, the difference between fear and anxiety lies in the defensive behaviors evoked by situations or stimuli considered dangerous. While in the case of fear, moving away from the threat ("active avoidance or fleeing") is usually observed, anxiety may lead an individual towards the danger ("approach or fighting") or to

"withholding entering the dangerous situation (passive avoidance or freezing)" (Piechurska-Kuciel 2008: 28). This feature of anxiety differentiating it from fear is emphasized by Riskind et al. (2000: 873), who depict anxiety as "an anticipatory state of active preparation for dealing with threat."

Contemporary psychologists (e.g. Pekrun 1992; Vasa and Pine 2004) usually describe anxiety as a construct consisting of three components, i.e. cognitive, physiological and behavioural. The former refers to how individuals approach and process situations, information and stimuli that they consider threatening (e.g. Ruiz-Caballero and Bermudez 1997). According to Pekrun (1992), anxiety arises when events are appraised as threatening and one's capacity of dealing with them as poor. Furthermore, anxiety is said to lead to several easily observable negative physiological/somatic symptoms. Many of them are caused directly by hormonal changes, which "lead to motor tension that can be observed in shakiness, jitteriness, muscle aches, inability to relax, fidgeting and restlessness" (Piechurska-Kuciel 2008: 29 after Wade and Tavris 1990). Additionally, anxiety may result in the feeling of panic, which reveals itself in different ways depending upon the individual, e.g. in chills, heart pounding, dry mouth, clammy hands or dizziness, just to mention a few (Scovel 1991). Symptoms of bodily tension, such as self-touching, closed body positions or leaning away can also be observed in anxious individuals (Burgoon and Koper 1984). The third component of anxiety – behavioral – reveals itself in irritability, impatience and behaviours typical for avoiding threatening situations, such as withdrawal or task avoidance (Kennerly 1990).

When discussing the construct of anxiety in the context of learning, it seems most vital to explain how it affects cognitive processing. First of all, observations prove that it leads to easy distraction, problems with concentration, limited creativity, increased response rate at the expense of accuracy, and reduced short-term memory capacities (Piechurska-Kuciel 2008). All these difficulties are related to the fact that anxiety causes attention narrowing and difficulties with attention control (e.g. Broadbent and Broadbent 1988). More specifically, cues and stimuli that automatically and intrusively draw the attention of an anxious person, pulling him/her away from the learning task, are those characterized by ambiguity, since they are perceived by him/her as potential sources of danger (Mathews et al. 1997). The constant process of scanning the learning material for stimuli and information considered threatening by the individual makes it difficult for him/her to focus on the proper task. As Eysenck (1997) clarifies, cognitive concern about one's performance (i.e. worry – one of the components of anxiety, next to emotionality) uses up cognitive resources required for storing and processing information, handicapping the attentional capacities and disabling effective handling of tasks, particularly those heavily dependent on storage and processing resources. Moreover, anxiety is said to interfere with information processing at all three levels, i.e. input, central processing and output stages (Piechurska-Kuciel 2008). In other words, an anxious learner will reveal difficulties with taking in, analyzing and retrieving new stimuli and information. Eysenck and Calvo (1992) further explain that anxiety causes ineffective functioning of the phonological loop responsible for temporary storage, and of the central executive coordinating the activity of the working memory. All these cognitive difficulties of an anxious person are said to inhibit his/her learning, making it less efficient, by demanding from him/her to put more effort than a low anxiety learner

to attain the same results (Ashcraft and Kirk 2001; Mathews 1990). The arguments described above constitute the main principles of the interference models of anxiety.

2.2 Language anxiety and its role in FL learning

Although interest in affective factors as important determinants of learning can be observed in mid-20th century, it took a few more decades for the construct of anxiety to be more thoroughly examined in the field of SLA. After first attempts of defining anxiety in reference to FL learning, by transferring other types of anxiety into the FL learning context, a unique and specific concept of language anxiety (LA) and tool to measure it was proposed in 1986 by Horwitz, Horwitz and Cope. They characterize it is "a distinct complex of self-perceptions, beliefs, feelings, and behaviours related to classroom learning arising from the uniqueness of the language learning process" (Horwitz et al. 1986: 128). Furthermore, it is agreed that LA, i.e. "derogatory self-related cognitions…, feelings of apprehension, and physiological responses such as increased heart rate," can be experienced by FL learners both in academic and social contexts, in the situation of both learning and using the target language (TL) that has not been fully mastered (Gardner and MacIntyre 1993: 5).

As the Foreign Language Classroom Anxiety Scale (Horwitz et al. 1986) reveals, LA is said to be related to three other types of performance anxieties, i.e. communication apprehension, fear of negative evaluation and test anxiety. The first component refers to the "discomfort in talking in front of others … caused by the belief in one's inability to express oneself fully or to understand what another person says" (Gregersen and Horwitz 2002: 562). Fear of negative evaluation is defined as "an apprehension about others' evaluations, avoidance of evaluative situations, and the expectation that others would evaluate [us] negatively" (Watson and Friend 1969: 450). Finally, the third type of anxiety considered to be connected to LA, i.e. test anxiety, though related specifically to the academic context of test taking, stems from the more general fear of failure caused by lack of certainty about one's ability or knowledge evaluated via tests.

Many studies have proven a negative correlation of moderate strength between LA and either course grades or outcomes on oral, vocabulary and grammar tests (e.g. Aida 1994; Bailey 1983; MacIntyre and Gardner 1989; Phillips 1992; Saito and Samimy 1996; Spielman and Radnofsky 2001). Some (e.g. MacIntyre 1999) even claim that LA is the strongest predictor of success in FL learning. As data of many studies show, the most anxiety-breeding skill is speaking, particularly when the oral task is to take place in front of other students. Moreover, the aspect that learners worry about most of all, being afraid of appearing ridiculous, is pronunciation (e.g. Phillips 1992; Price 1991; Young 1992).

Several types of language-specific anxieties have been identified, e.g. listening (comprehension) anxiety (Kim 2005; Vogely 1999), writing anxiety/apprehension (Cheng et al. 1999; Hilleson 1996), reading anxiety (Argamon and Abu-Rabia 2002; Saito et al. 1999), speaking anxiety (Woodrow 2006), or grammar anxiety (VanPatten and Glass 1999). However, so far no instrument has been designed to examine specifically the nature of pronunciation anxiety, which could address feelings evoked by

the way one sounds or looks like when talking in a FL or worries experienced when learning/practising FL pronunciation.

Numerous studies in the field of LA have been devoted to identifying its causes. Young (1991: 427) posits the existence of "six potential sources of LA," such as personal and interpersonal anxieties, learner beliefs about FL learning, instructor beliefs about language teaching, instructor-learner interactions, classroom procedures, and language testing. From among them, it is personal and interpersonal anxieties, embracing self-perceptions of the FL learner/user, that have been the most frequently recognized (e.g. Bailey 1983; Gardner and MacIntyre 1993; Onwuegbuzie et al. 1999; Piechurska-Kuciel 2008). The observation is supported by Krashen (1982, cf. Young 1991: 427), who states "the more I think about self-esteem, the more impressed I am with its impact. This is what causes anxiety in a lot of people." The important role of self-perceptions in the development of anxiety is also underlined by Young (1991: 427), who concludes: learners "with a self-perceived low ability level in a foreign or second language are the likeliest candidates for language anxiety."

There are data proving a significant systematic relationship between learners' perceived pronunciation level and their degree of LA and listening anxiety. In research conducted by Baran-Łucarz (2011; 2013) a negative correlation of moderate strength was observed between students' FL pronunciation self-assessment and their LA level ($r = -.49$ at $p < .0005$) and listening anxiety level ($r = -.46$ at $p < .005$). In both cases the quantitative data pointing to the importance of pronunciation self-perceptions were further verified by responses of chosen subjects provided in semi-structured interviews.

3. The model of Phonetics Learning Anxiety

On the basis of the vast body of research conducted on anxiety and LA, a model of Phonetics Learning Anxiety (PhLA) has been proposed, hoping it may contribute to a better understanding of why some FL learners do not benefit as much as they could from a practical course of phonetics. PhLA may be defined as an apprehension experienced during a class of phonetics, evidenced by cognitive, physiological/somatic, and behavioral symptoms. Besides the general level of PhLA represented by the attitude towards the course of phonetics and the three observable types of reactions mentioned above, a few constructs underpinning the phenomenon of anxiety experienced during a practical course of phonetics are suggested. The proposed model of PhLA is depicted by Figure 1 below.

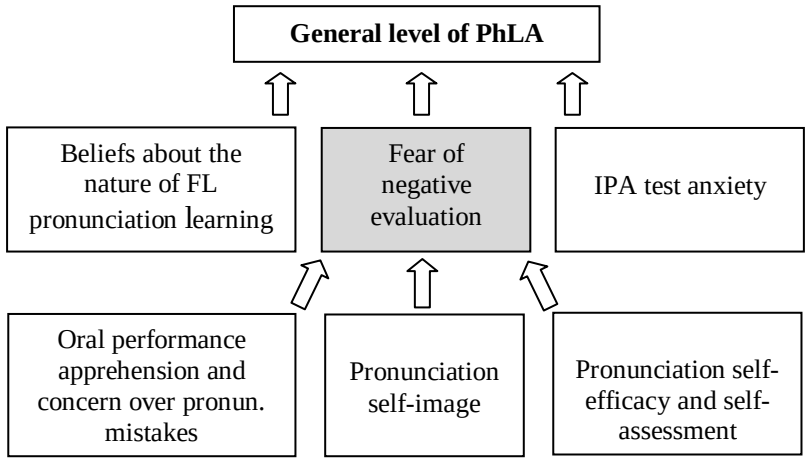

Figure 1. A hypothetical model of Phonetics Learning Anxiety

Relying on the outcomes of earlier studies on language anxiety, it is posited that the most important correlate/subcomponent of PhLA is the fear of negative evaluation. This construct, in turn, is assumed to be shaped by three relatively independent factors, i.e. general apprehension about oral performance and concern over FL pronunciation mistakes, pronunciation self-image related to one's appearance (the way one thinks he/she looks and sounds like) when speaking in a FL and acceptance of the perceived self-image, and finally pronunciation self-efficacy and self-assessment, i.e. beliefs about one's abilities needed to master a FL pronunciation and one's perceived level of pronunciation, both estimated usually in reference to other classmates. Next to the fear of being negatively evaluated, it is also beliefs about the nature of FL pronunciation learning and anxiety caused by transcription tests that have been assumed to determine the learners' feeling of apprehension during phonetics classes.

4. The study

4.1 Research questions and hypotheses

To find out whether PhLA significantly determines how much students benefit from a course of phonetics and to shed light on the nature of the construct, an empirical study has been carried out. The following main research questions and hypotheses were forwarded:

Q1: Does PhLA determine the learners' level of pronunciation measured after a course of phonetics?

H1a: There is a significant negative relationship between the learners' degree of PhLA and their pronunciation level measured after a course of phonetics.

H1b: The pronunciation level measured after the course of phonetics is significantly lower in the case of students revealing high PhLA than in the case of students revealing low PhLA.

Q2: What are the correlates/subcomponents of the PhLA?

H2a: There is a significant positive relationship between the learners' fear of negative evaluation and their level of PhLA.

$H2a_{I}$: There is a significant positive relationship between the learners' degree of oral performance apprehension/concern over pronunciation mistakes and their level of PhLA.

$H2a_{II}$: There is a significant positive relationship between the learners' degree of negative pronunciation self-image and their level of PhLA.

$H2a_{III}$: There is a significant positive relationship between the learners' degree of negative pronunciation self-efficacy/self-assessment and their level of PhLA.

H2b: There is a significant positive relationship between the learners' beliefs about the difficulty of FL pronunciation learning and their level of PhLA.

H2c: There is a significant positive relationship between the learners' IPA test anxiety and their level of PhLA.

4.2 Participants

The research was carried out in two groups of first-year extramural students of the University of Wroclaw, majoring in English (N=32). Among the participants there were 10 males and 22 females. All of them were native speakers of Polish, aged 19-24. Having passed the extended level of the high school leaving examination in English (Matura) with a score of at least 70%, they mostly represented an upper-intermediate level (B2), though some seemed better than the others (C1). None of the participants had spent more than three months in an English-speaking country. Before starting university education most of them were taught by Polish teachers of English who limited pronunciation practice of their students to repetition of new words and correction of mispronunciations.

At the time the data for the study were being collected, the subjects were just finishing their practical course of phonetics (altogether approximately thirty 90-minute lessons). All of them were either highly (19%) or very highly (81%) motivated to achieve a native-like level of pronunciation. When the model of English pronunciation is concerned, 88% declared they were more attracted to Received Pronunciation (RP), while the remaining 12% chose General American (GA) as the norm to approximate.

4.3 Description of the phonetics course

The phonetics classes were predominantly practical, since the aim of the course was to help students improve their pronunciation habits rather than to raise their theoretical knowledge on complex phonetic and phonological issues. While the second year of phonetics in our institute focuses more on suprasegmentals, pronunciation at lexical level and further remedy work, the first year is devoted mainly to practising the articulation of segments, word stress and most important aspects of connected speech (weak forms, assimilations, elisions, linkings).

Before the practical part of each lesson began, the new sound was introduced, i.e. its place and manner of articulation were provided by means of various techniques appealing to different modalities and senses of the students. In the theoretical part of the class the inductive approach was used, that is, the learners tried to observe and come up with their own hypotheses about how particular segments in English are pronounced and how they differ from Polish counterparts. The practical part of the lesson would usually start with a few exercises warming up the articulators, borrowed from speech pathology. What followed was repeating words and sentences in which the sound appeared in different contexts. Finally humorous dialogues were listened to and then read aloud in lockstep (i.e. as a whole group) and pairs. Sometimes after having practised reading with a friend, a student would read the text aloud to the rest of the group. The controlled tasks were supplemented with game-like activities from various sources, songs, and presentations of students. It is important to add that while the learners were practising reading dialogues in pairs, the teacher monitored their work, coming up to each pair and offering further help if needed. Any time a pronunciation deviated significantly from the correct version, the teacher advised how to position the articulators, modeled the proper form and encouraged repetition until the production approximated the proper articulation of the sound. When the learner showed discomfort and reluctance to articulate the segment or word after the teacher, he/she was encouraged to see the instructor after the class, so as to practise the difficult area of pronunciation individually, without the presence of other students. Additionally, about 15 minutes of each lesson were devoted to transcribing difficult vocabulary items, which the learners were required to know for the written tests taken approximately once a month.

As suggested above, the students were allowed to choose either RP or GA as their goal. Consequently, the features of the two norms were presented in a detailed manner in the first semester, and consistency in using one of them was required both in articulation and in written transcription tests.

4.4 Instruments

Three main instruments were designed for the purpose of this research, i.e. the Introductory Questionnaire, Phonetics Learning Anxiety Scale (PhLAS), Pronunciation Attainment Test (PAT), and IPA Tests. The major characteristics of the tools are presented thoroughly below.

4.4.1 Introductory Questionnaire

The measure was a pen-and-pencil questionnaire that the students filled out during the first class of phonetics. It opened with a question about the participants' level of motivation to reach native-like pronunciation. It consisted in the students assessing their level of motivation by writing in a digit from 1 to 5, representing the strength of their desire to speak with a native-like accent, where 1 meant 'I definitely do not want to speak with an English native-like accent', while 5 – 'I want to achieve native-like accent very much'. The remaining seven open questions enquired about the subjects' prior FL learning experience, e.g. about the intensity of learning, ways of having practised pronunciation during their FL courses and individually (autonomously), or about visits abroad. Moreover, the questionnaire required from the subjects to reflect on the difficulties they believed they had with English pronunciation.

The goal of applying the tool, besides getting to know my students better, was to control certain variables and exclude from the study, if necessary, those learners who stood out from the rest, e.g. had spent a longer period of time in an English-speaking country, or had regularly been provided with formal instruction on pronunciation and intensive practise in this FL aspect. However, the responses showed that data collected from all the students could be taken into account in the research.

4.4.2 The Phonetics Learning Anxiety Scale

The aim of PhLAS was to measure the anxiety level experienced during the course of phonetics and to verify the hypotheses H2a – H2c concerning the proposed subcomponents/correlates of PhLA. The instrument had the form a self-report 44-item questionnaire based on a 6-point Likert scale, which required the testees to agree/disagree to various extent with the provided statements. Since it was intended to be distributed among students of Polish origin, the language of the instrument was Polish. The wording of 13 statements in the entire battery required a reversed scoring key to be used.

The first part of PhLAS, consisting of 15 items, was aimed at measuring the general level of PhLA without enquiring about the possible reasons for them feeling anxious. The higher the score the students obtained, the higher their general level of PhLA was considered to be. The items addressed the attitudes toward the course of phonetics and the typical symptoms testifying to the existence of anxiety. Below are examples of statements from each category.

<u>General level of PhLA (15 items)</u>

- Items addressing the **attitude** towards the class of phonetics, e.g.:
  1. I prefer other classes than that of phonetics.
  12. Phonetics is one of my favourite classes. (reversed scoring)

- Items addressing **cognitive symptoms** of anxiety, e.g.:
  7. I am so nervous that I can't hear the new sounds or word stress properly.

8. While reading aloud I am so embarrassed that I can't focus on my pronunciation and control it.

- Items addressing **somatic symptoms** of anxiety, e.g.:
  2. I feel more comfortable and less tense at most of the other courses.
  10. Usually I feel relaxed at the phonetics class. (reversed scoring)
  14. I feel relief when the class of phonetics is coming to an end.

- Items addressing **behavioural symptoms** of anxiety, e.g.:
  4. I often volunteer during the classes of phonetics. (reversed scoring)
  9. I would feel less anxious if the classes of phonetics had the form of individual training with the teacher.
  13. If I didn't have to, I wouldn't attend the classes of phonetics.

The second part of the questionnaire consisted of items addressing concepts that were assumed to be important correlates/subcomponents of the PhLA construct. Consequently, statements referring to oral performance apprehension and concern for pronunciation mistakes, pronunciation self-image and self-efficacy/self-assessment, beliefs about the difficulty of FL pronunciation learning, and IPA test anxiety were formulated. Below are a few examples of the items.

Fear of negative evaluation (20 items)

- Items addressing **oral performance apprehension and concern over pronunciation mistakes** (10 items), e.g.:
  16. I feel my heart pounding when the teacher corrects my pronunciation in lockstep.
  20. I don't like to read aloud in front of the whole class.
  17. I am not bothered about the pronunciation mistakes I make. (reversed scoring)
  29. I feel tense and uneasy knowing that other students are listening to me reading or repeating sth. in English.
  36. I feel more embarrassed committing a pronunciation mistake than any other type of mistake.

- Items addressing **pronunciation self-image** (6 items), e.g.:
  33. I like to talk or sing to myself in English. (reversed scoring)
  37. I think I sound ridiculous pronouncing English sounds and words the way they should be pronounced.
  44. It seems to me that I sound terrible when I pronounce English sounds and words 'in the Polish manner'.
  35. I don't like listening to myself reading aloud or speaking in English.

- Items addressing **pronunciation self-efficacy and self-assessment** (4 items), e.g.:
  18. I find it more difficult to improve my pronunciation than other aspect/skills of English.
  34. I don't have a talent to master FL pronunciation.
  39. Other students have a better pronunciation than I.

41. I am satisfied with my pronunciation level. (reversed scoring)

Beliefs about the nature (difficulty) of FL pronunciation learning (4 items), e.g.:

29. The pronunciation of English is very difficult for Poles.

32. I think pronunciation is the easiest FL aspect to master. (reversed scoring)

IPA test anxiety (5 items), e.g.:

19. Phonetics classes would be enjoyable if there were no transcription tests.

23. I feel more stressed reading aloud than writing an IPA test.

31. Even if I am well-prepared to the test, I am so nervous writing it that I can't recall the transcription of many words.

As the examples of the items show, though they addressed particular correlates/subcomponents of PhLA, they were at the same time usually formulated from the perspective of the typical cognitive, somatic or behavioural symptoms of anxiety. The scoring key of this part of the PhLAS is analogous to the one used in the case of the general level of PhLA, i.e. the stronger the testees agreed with the statements, the more points they gained (with the exception of a few items in which a reversed scoring key was used) for particular correlates.

The main questionnaire (PhLAS) was followed by three additional questions. In the first one the testees were asked to self-assess their level of PhLA, by finishing the statement 'The level of discomfort/ uneasiness/anxiety/fear I experience during the classes of phonetics is usually…' with one of the provided options 'very high', 'high', 'rather high', 'rather low', 'low', 'very low'. The two proceeding open questions asked about (hypothetical) causes of their anxiety and suggestions on how it could be lowered. Due to space limitations, the analysis of these responses is to be offered in one of the forthcoming papers.

4.4.3 The Pronunciation Attainment Test

At the end of the one-year course of phonetics the students' pronunciation was assessed with the use of four oral tests. Although most of the evaluation took place during the actual performance, the testing was recorded by means of a voice recorder (SONY ICD-UX300) and microphone, and then downloaded onto a personal computer, so as to make it possible to return to any fragment of the recording in case any doubts with assessment appeared.

The first two tests – Sentence Reading (T_1) and Text Reading (T_2) – were aimed at evaluating the participants' habits of pronouncing segments practised during the course. The former consisted in the participants reading sets of sentences, in each of which a particular consonant or vowel occurred several times in various positions (initial, internal, final). In the latter test the subjects were handed a text (taken from Celce-Murcia et al. 2008: 398) that they could browse through quickly and then were asked to read aloud, knowing they would have to summarise it. The idea of this procedure was to free the student from focusing on accuracy in pronunciation, by making them draw attention to meaning, and in this way to force them to rely on their pronunciation habits.

In both tests an atomistic approach in assessment was used, with 0 to 3 points distributed for each segment depending on the frequency of producing it properly, i.e. as English native speakers do. The tasks were assumed to differ in the amount of monitoring and control they allowed for, with T_2 being less controllable than T_1. The overall score for Text Reading was additionally determined by word pronunciation and fluency in reading. Serious problems in these areas resulted in the students losing some points for T_2. Additionally, Text Reading allowed to diagnose the subjects' consistency in using one of the accents - RP or GA.

The next two tests – T_3 and T_4 – consisted in the participants reading aloud lists of words. This time it is not pronunciation habits but word pronunciation that was assessed. One list (L_1) – T_3 – consisted of words that were practised during the course and appeared on transcription tests. The exact content of the list was not revealed to the participants. The other list (L_2) – T_4 – was well-known to the subjects. It consisted of words commonly mispronounced by Poles (based on Sobkowiak 1996). In the case of both tests, 1 point was given for each word properly pronounced.

4.4.4 IPA Tests

In this study results of four written tests taken by each subject during the second semester were taken into account and used for further analysis. The major part of each test consisted in transcribing 20 – 30 individual words using the International Phonetic Alphabet (IPA). For each word properly transcribed 3 points could be achieved. Every mistake, including misplacing or lack of the main stress, resulted in losing a point. The minimum result for passing the test was 70%.

4.4.5 Further assessment procedures

Since all the tests constituted the basis for formal assessment of the students in the course of phonetics, the raw scores obtained by each learner were always transformed into grades. For further statistical analysis, the grades were changed into an 8-point interval scale, as Table 1 below displays.

Grades	Points
2	1
3-	2
3	3
3+	4
4	5
4+	6
5-	7
5	8

Table 1. Transforming grades into points

4.5 Presentation and discussion of results

4.5.1 Descriptive statistics

Having gathered all the necessary data with the batteries described above, descriptive statistics for the outcomes on the PhLAS, PAT and IPA Tests were computed (see Table 2 and Table 3).

	T_1 (SR)	T_2 (PR)		Sum	$T_3(L_1)$	$T_4(L_2)$	Sum	Total	IPA tests
		Overall	RP/GA						
Min-Max	1-8	1-8	1-8	3-24	1-8	1-8	2-16	5-40	4-32
L-H	1-8	2-8	3-8	8-23	1-7	2-8	3-15	11-38	4-29
Mean	4.75	4.91	5.72	15.38	3.53	5.34	8.88	24.25	13.29
Median	5	5	6	16	4	5	9	24	13
SD	1.98	1.75	1.57	4.9	1.54	1.66	2.8	7.26	6.6

Table 2. Descriptive statistics for results of PAT

The figures in Table 2 imply that after the 45-hour course in phonetics focusing on practise of segments, word stress, weak forms, consistency in RP/GA and a few aspects of connected speech, there still is a lot of space for improvement in these areas of pronunciation. Unexpectedly, in the case of the less controllable test (T_2), the scores were not lower than in the case of the more controllable test (T_1). This may suggest that the proper articulation of some of the elements of the English phonetic system have become automatic, but this cannot yet be advanced with any certainty. Among the possible explanations for such outcomes may be also the assessment of the samples. Despite the fact that the judge (the phonetics course teacher and, at the same time, the author of this paper) did her best to identify all the pronunciation errors that appeared in the text read by each subject and to use an objective atomistic approach to calculate the final scores, the assessment might have been less accurate (i.e. more lenient) than in the case of sentence reading (T_1).

The data indicate that students did better with the phonetic system than with pronunciation of difficult vocabulary items that appeared during the course. Finally, the results of the IPA tests reveal that many students have still difficulties with transcription. When the average PhLA level of all the participants is concerned, it appears to be relatively low (mean = 43.22pts; min/max = 15-90pts;), with the bell curve shifted slightly more towards the lower scores (low-high = 31-60)

	General level of PhLA	Fear of negative evaluation			Beliefs	IPA test anx
		Oral perf. app/concern o. pron. mist	Pron. self-image	Pron. self-efficacy		
Min-Max	15-90	10-60	6-36	4-24	4-24	5-30
L-H	31-60	26-50	10-24	6-18	7-22	13-29
Mean	43.22	32.91	17.28	12.28	15.34	20.72
Median	42.5	33.0	17.0	12.5	15.0	21.5
SD	8.43	6.71	3.46	3.00	3.56	3.03

Table 3. Descriptive statistics for results of PhLAS

The descriptive statistics in both tables show that all the assumptions (the scales, independence, linearity and normality assumptions) underlying Pearson correlation have been met, which allows us to proceed to further statistical analysis of data.

4.5.2 Correlation analysis and t-test

Table 4 displays the results of Pearson correlation computed between the general level of PhLA and outcomes for the subtests of the Pronunciation Attainment Test and for the IPA Tests. The same calculations have been made between the total scores on the PhLA and the measures of pronunciation

	PAT Total	Pron. habits				Word pron.			IPA tests
		T_1	T_2	RP/GA	Sum	$T_3(L_1)$	$T_4(L_2)$	Sum	
General level of PhLA	-.45	-.38	-.45	-.36	-.43	-.30	-.39	-.44	-.36
PhLA total	-.36	-.36	-.49	-.23	-.44	-.18	-.22	-.23	-.14

df=30
p<.05 rcrit.=.2960
p< .025 rcrit.=.3494
p<.01 rcrit.=.4093
p<.005 rctit.=.4487

Table 4. Pearson product-moment correlation coefficients between the general level of PhLA, PhLA Total, and components of PAT and IPA test results

When the general level of PhLA is concerned, in all cases the relationship proved to be statistically significant of moderate strength, with the correlation between the Total for PAT r=-.45 at p<.005, which allows to accept hypothesis H1a. The outcomes show that the relationship is stronger in the case of pronunciation habits than word pronunciation. The weakest relationship appeared between anxiety and the results on IPA tests. From

among all the coefficients, the highest was achieved in the case of pronunciation measured in the less controllable reading test (T2). Consequently, it may be concluded that the higher the anxiety, the lower the level of automatic articulation. An explanation for this may lie in the cognitive effects of anxiety on learning and typical behavioral reactions of anxious learners, such as avoidance of practice in and outside classroom. When the PhLA total is concerned, a systematic relationship appeared only in the case of the acquired pronunciation habits. This phenomenon may be due to the statements addressing the correlates of phonetics learning anxiety being included in the PhLA total score. Their relation with the general level of PhLA is explained in the following section. To verify hypothesis H1b, the subjects were categorized into two groups. Those who scored above the median for the general level of PhLA were considered the highly anxious subjects (n=15), while those who scored below the median constituted the group of low PhLA learners (n=16). After computing the means and SDs for the high and low PhLA participants, the t values were calculated. The results are displayed in Table 5.

		PAT Total	Pron. habits				Word pron.			IPA tests
			T_1	T_2	RP/GA	**Sum**	T3(L1)	T4(L2)	**Sum**	
Mean	high PhLA	**20.80**	3.87	4.75	5.13	**13.07**	3.00	4.73	**7.30**	11.13
	low PhLA	**27.19**	5.67	5.62	6.31	**17.25**	4.00	5.93	**9.94**	15.69
SD	high PhLA	**6.32**	1.88	1.49	1.64	**4.67**	1.41	1.58	**2.34**	5.94
	low PhLA	**4.80**	1.56	1.71	1.35	**2.57**	1.60	1.61	**2.93**	6.82
t_{obs}		**3.182**	2.882	2.702	2.189	**3.119**	1.844	2.099	**2.302**	1.977

df=29
p<. 10 tcrit.=1.311
p<.05 tcrit.=1.699
p<.025 tcrit.=2.045
p<.01 tcrit.=2.462
p<.005 tcrit.=2.756
(for one-tailed test)

Table 5. Results of the independent t-test computed for components of PAT and IPA Tests

The outcomes showed the existence of a statistically significant difference between the pronunciation of high and low PhLA subjects. In each case, the scores obtained by the low anxiety subjects were higher than those of the high anxiety learners, which justifies the acceptance of hypothesis H1b. As in the case of correlation, the most meaningful results were found in the case of pronunciation habits rather than word pronunciation. The least meaningful difference between high and low PhLA scores, though still statistically significant, appeared in the case of IPA tests.

4.5.3 Correlates/subcomponents of the PhLA

Having verified the presupposition about progress in pronunciation being related to the level of anxiety, the correlates of the construct could be determined. This was done by computing the Pearson product-moment correlation between the general level of PhLAS and each of the groups of items representing particular hypothetical correlates, i.e. oral performance apprehension (OPA)/concern over pronunciation mistakes (COPM), pronunciation self-image and self-efficacy/self-assessment, which altogether are assumed to constitute the fear of negative evaluation, beliefs about the nature of FL pronunciation learning, and IPA test anxiety. The results are presented in Table 6.

	General level of PhLA	**Cronbach alpha**
1. Fear of neg. eval.	.82	.56/.63
OPA/COPM	.65	.77
pron. self-image	.70	.46/.70
pron. self-efifacy	.60	.61
2. Beliefs	.46	.71
3. IPA test anx	-.23	.25
4. General level of PhLA	1	.84
5. PhLA Total	.78	.87

df=30
p<.05 rcrit.=.2960
p<.005 rctit.=.4487

Table 6. Pearson product-moment coefficients between the general level of PhLA and correlates of the PhLAS; internal consistency of the general level of PhLA and its correlates

The outcomes show that PhLA is most strongly related to the fear of negative evaluation (r=.82), with all its components being significantly correlated with anxiety at a high level (from r=.60 in the case of pronunciation self-efficacy/self-assessment to r=.70 for pronunciation self-image). It is also the beliefs that are significantly related to the level of PhLA, though at a moderate level (r=.46). However, no systematic relationship appeared between the general level of PhLA and IPA test anxiety.
Consequently, while hypotheses $H2a_I$ – $H2a_{III}$ and H2b can be accepted, hypothesis H2c must be rejected. To examine the internal consistency of the whole 44-item PhLAS, its first part, measuring the general level of anxiety, and of each of the correlates/subcomponents of the construct, the Cronbach alpha was computed. The outcomes revealed high consistency in the case of the PhLAS Total (.87) and general level of PhLA (.84). However, when the items addressing particular correlates of PhLAS are concerned, not always were the coefficients at a satisfactory level. The low level of

internal consistency and its negative direction in the case of IPA test anxiety might have resulted from the structure of a few statements in which the feeling of discomfort caused by IPA tests was contrasted with the fear connected with oral performance, e.g. *'24. I prefer reading aloud to writing transcription tests'.* Low internal consistency (r=.40) was also found in the case of pronunciation self-image. Such a result was caused by the last question, i.e. *'44. I think I sound terrible pronouncing English sounds and words in the 'Polish manner',* with which most of the students *'rather agreed'* or *'agreed'*. Excluding the statement raises the internal consistency of this subcomponent to .70 and of the fear of negative evaluation to .63.

4.5.3 Results of further observations

Trying to have a closer look at how the responses of high and low anxiety participants differed, the mean answers for each item provided by the two groups were compared. Below is a table presenting some of the statements, in the case of which differences between the average responses of high and low anxiety students were observed. The average answers computed from the marked digits from 1 to 6 corresponded to the degree of agreeing/disagreeing with the statements.

The distribution of answers evidently shows a tendency for the feelings of low and high PhLA students to vary. For example, while the highly anxious participants are bothered by the fact that other students can hear them making pronunciation mistakes, the low PhLA students do not mind being assessed by their classmates (see statements 30 and 32). Moreover, interesting discrepancies may be observed in accepting oneself speaking Polglish (statement 44) and in considering oneself sounding ridiculous when trying to use proper English pronunciation (statements 37 and 38). As the average answers show, the low anxiety subjects seem to be more prone to accept their new 'selves' in the 'English shoes' than the high PhLA learners. Finally, it seems worth drawing attention to statement 39, which reveals that low pronunciation self-assessment is more likely to be found among highly anxious students.

Item	Low PhLA (N=16)	High PhLA (N=15)
3. I get nervous when I have to read aloud in English.	no	rather yes
6. When I read aloud in English at the class of phonetics I get so nervous that my performance is at a much lower level than when I read aloud at home.	no	rather yes
9. I would feel less stressed if the phonetics classes had the form of individual meetings with the teacher.	no	yes
10. Usually I feel relaxed during the classes of phonetics.	yes	rather not
11. I feel relaxed reading aloud during the classes of phonetics.	rather yes	no
15. When reading aloud I read fast so as to be freed from the task as soon as possible.	*no*	*rather yes*
28. I feel ashamed of myself in front of the teacher when making the same pronunciation mistakes.	*rather yes*	*rather not*
30. I feel nervous knowing that other students are listening to me.	*no*	*rather yes*
32. I'd rather my classmates did not hear me making pronunciation mistakes.	*rather not*	*yes*
34. I prefer reading aloud to writing transcription tests.	*rather yes*	*rather not*
36. I feel more embarrassed making a pronunciation mistake than any other type of mistake.	*rather not*	*rather yes*
37. I seems to me that I sound ridiculous/silly pronouncing sounds and words the way they should be pronounced in English.	*no*	*rather yes*
38. I'm more prone to accept myself speaking English with a Polish accent than trying to speak with a proper (nativelike) English accent.	*definitely not*	*rather not*
39. Other students have a much better pronunciation than I.	*rather not*	*rather yes*
44. I think I sound terrible pronouncing English sounds and words in the 'Polish manner'.	*yes*	*rather yes*

Table 6. 'Average answers' of low and high PhLA participants to chosen items of the PhLAS

5. Conclusions and further research directions

The aim of this paper was to throw some light on the construct of phonetics learning anxiety and to examine whether it has an influence on how much students benefit from a course of phonetics. The results of the reported preliminary research show that PhLA is indeed an important factor determining success in FL pronunciation learning. A significant negative correlation of moderate strength ($r = -.45$ at $p<.005$) was found between the level of PhLA and scores on the PAT. The relationship between the two variables appeared to be stronger in the case of pronunciation habits than word pronunciation. A connection was also observed between the PhLA and results of transcription tests, although the strength of relationship was weaker than in the case of PAT scores. The importance of PhLA in pronunciation learning has been also confirmed by t-tests, which proved the pronunciation measured by PAT and the ability to transcribe words using IPA to be at a significantly higher level in the case of low PhLA students

than high PhLA learners. Moreover, the t-tests supported the fact that anxiety affects learning the FL phonetic system more than word pronunciation. This outcome can be further explained by results of examining the correlates of PhLA.

The strongest correlates/subcomponents of anxiety experienced during the course of phonetics were pronunciation self-image, referring to the way the learner believes he/she looks and sounds like when speaking a FL and whether he/she accepts his self-image, pronunciation self-efficacy and self-assessment and oral performance apprehension/concern over pronunciation mistakes. Together these factors, considered to result in the fear of negative evaluation, were strongly correlated with the level of anxiety accompanying pronunciation practice during a phonetics course ($r=.82$). The analysis of responses to particular items provided by high and low PhLA subjects lend support to the premise that the most important sources of anxiety belong to the group of students' self-perceptions, e.g. low pronunciation self-efficacy/self-assessment and negative pronunciation self-image, which presumably lead to the fear of negative evaluation. Such outcomes of the research go hand in hand with results of other studies (Young 1991) in which students' personal and interpersonal anxieties were observed to be most frequent causes of LA. Next to the fear of negative evaluation, it is also beliefs about the nature of FL pronunciation learning that revealed a moderate correlation between the level of PhLA ($r=.46$ at $p<.005$). Finally, the hypothetical model of PhLA assumed IPA test anxiety to be another subcomponent of PhLA. This, however, was not confirmed in the research.

A more thorough analysis of the results leads to the conclusion that problems with pronouncing the FL phonetic system (e.g. segments and aspects of connected speech) cause more anxiety than correctness of pronunciation at lexical level.

It needs to be emphasized that the ideas about the construct of phonetics learning anxiety presented in this paper are yet preliminary and that the reported outcomes of the empirical research ought to be viewed with utmost caution. Not only must the study be replicated among a larger group of students, but also both instruments require several amendments and validation. With respect to the PhLAS, changes need to be introduced, among others, to statements enquiring about IPA test anxiety, many of which demanded that the students estimate the fear connected with transcription tests in reference to oral performance apprehension. Moreover, since the use of IPA is an important part of most classes of phonetics, it seems worth substituting IPA test anxiety with a wider construct of transcription anxiety that would include both transcription practice and test anxiety. Additionally, what might be considered is adding or replacing some items with a few direct statements about experiencing fear of being negatively evaluated by other members of the class or the teacher. The battery diagnosing the participants' pronunciation can also be further worked on. Finally, the pronunciation scores would be undoubtedly more reliable if assessed by a few objective judges, preferably phoneticians sharing the same L1 as the subjects and native speakers.

Designing a valid and reliable tool measuring the level of anxiety experienced during a phonetics course and showing its sources could help phonetics instructors make the course less anxiety-breeding, and thus make it more effective. What could further aid in understanding the emotions accompanying FL phonetics learning is complementing the quantitative data with qualitative research consisting in, e.g. conducting semi-structured interviews with high and low PhLA students or gathering their reflections provided

directly after phonetics classes on what made them anxious during the lesson, why and how the negative feelings could be reduced. Finally, an experiment can be carried out to verify empirically what indeed decreases the students' level of phonetics learning anxiety.

Despite the fact that the data achieved in this study need further verification, it may be suggested that certain aspects of teaching ought to be very carefully planned and controlled by the phonetics teacher. Among the areas worth being considered by the instructor are the following: classroom dynamics, teacher-student rapport, interaction patterns (group, pair, individual performance), manners of providing feedback, correction techniques, pace of work, goal-setting, strategy training and ways of raising self-perceptions of students. It seems that a sense of security during the lesson and a positive pronunciation self-image are necessary conditions for some learners to benefit from the practical course of phonetics.

References

Aida, Y. 1994. Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal* 78: 155-168.

Argamon, O. and S. Abu-Rabia. 2002. The influence of language anxiety on English reading and writing tasks among native Hebrew speakers. *Language, Culture and Curriculum* 15: 143-160.

Ashcraft, M. H. and E. P. Kirk. 2001. The relationship among working memory, math anxiety and performance. *Journal of Experimental Psychology*: General 130: 224-237.

Bailey, K. M. 1983. Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies. In H. W. Seliger and M. H. Long (eds) *Classroom Oriented Research in Second Language Acquisition*. Rowley, MA: Newbury House: 67-103.

Baran-Łucarz, M. 2011. The relationship between language anxiety and the actual and perceived levels of FL pronunciation. *Studies in Second Language Learning and Teaching* 1(4): 491-514.

Baran-Łucarz, M. 2013. Foreign language pronunciation and listening anxiety – a preliminary study. In E. Piechurska-Kuciel and E. Szymańska-Czaplak (eds) *Language in Cognition and Affect, Second Language Learning and Teaching.* Springer-Verlag Berlin Heidelberg: 1-20. DOI: 10.1007/978-3-642-35305-5_15.

Broadbent, d. E.and M. Broadbent. 1998. Anxiety and attentional bias: State and Trait. *Cognition and Emotion* 2: 165-183.

Burgoon, J. I. and R. J. Koper. 1984. Nonverbal; and relational communication associated with reticence. *Human Communication* 10: 601-626.

Celce-Murcia, M., D. Brinton, and J. Goodwin. 2000. *Teaching Pronunciation. A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.

Cheng, Y.-S, Horwitz, E. K. and D. L. Schallert. 1999. Language anxiety: Differentiating writing and speaking components. *Language Learning* 49: 417-446.

Darwin, C. 1965/1872. *The Expression of Emotions in Man and Animals.* Chicago: University of Chicago Press.

Eysenck, M. W. 1997. *Anxiety and Cognition: A Unified Theory*. Hove: Psychology Press.

Eysenck, M. W. and M. G. Calvo. 1992. Anxiety and performance: The processing efficiency theory. *Cognition and Emotion* 6: 409-434.

Friedman, I. A. and O. Bendas-Jacob. 1997. Measuring perceived test anxiety in adolescents: A self-report scale. *Education and Psychological Measurement* 57: 1035-1047.

Gardner, R. C. and P. D. MacIntyre. 1993. A student's contribution to second-language learning. Part II: Affective variables. *Language Teaching* 26: 1-11.

Gregersen, T and E. K. Horwitz. 2002. Language learning and perfectionism: anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal* 86(4): 562-570. DOI: 10.1111/1540-4781.00161.

Hilleson, M. 1996. 'I want to talk with them, but I don't want them to hear': An introspective study of second language anxiety in an English-medium school. In K. M. Bailey and D. Nunan (eds) *Voices from the Language Classroom*. Cambridge: Cambridge University Press: 248-277.

Horwitz, E. K. 1988. The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal* 72: 283-294.

Horwitz, E. K. 2010. Research timeline. Foreign and second language anxiety. *Language Teaching* 43(2): 154-167. DOI:10.1017/S026144480 999036X

Horwitz, E. K., M. Horwitz and J. A. Cope. 1986. Foreign language classroom anxiety. *The Modern Language Journal* 70: 125-132.

Kennerly, H. 1990. *Managing Anxiety. A Training Manual*. Cary, NC, USA: Oxford University Press.

Kim, J. 2005. The reliability and validity of a foreign Language Learning Anxiety Scale. *Korean Journal of English Language and Linguistics* 5: 213-235.

Lesse, S. 1982. The relationship of anxiety to depression. *American Journal of Psychotherapy* 36: 332-349.

Lewis, A. 1970. The ambiguous word 'anxiety'. *International Journal of Psychiatry* 9: 62-79.

MacIntyre, P. D. 1999. Language anxiety: A review of the research for language teachers. In D. J. Young (ed) *Affect in foreign language and second language learning. A practical guide to creating a low-anxiety classroom atmosphere.* Boston: McGraw-Hill: 24-45.

MacIntyre, P. D. and R. C. Gardner. 1989. Anxiety and second language learning: Toward a theoretical clarification. *Language Learning* 39: 251-275.

Mathews, A. 1990. Why worry? The cognitive function of anxiety. *Behavior Research and Therapy* 28: 455-468.

Mathews, A., Mackintosh, B. and E. P. Fulcher. 1997. Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences* 1: 340-346.

Onwuegbuzie, A. J., Bailey, P. and C. E. Daley. 1999. Factors associated with foreign language anxiety. *Applied Linguistics* 20(2): 217-239.

Pekrun, R. 1992. Expectancy-value theory of anxiety: Overview and implications. In D. Forgays and T. Sosnowski (eds) *Anxiety: Recent Developments in Cognitive, Psychological and Health Research.* Washington, DC: Hemisphere: 23-39.

Piechurska-Kuciel, E. 2008. *Language Anxiety in Secondary Grammar School Students*. Opole: Wydawnictwo Uniwersytetu Opolskiego.

Phillips, E. M. 1992. The effects of language anxiety on student's oral test performance and attitudes. *The Modern Language Journal* 76: 14-26.

Price, M. L. 1991. The subjective experience of foreign language anxiety: Interviews with highly anxious students. In E. K. Horwitz and D. J. Young (eds) *Language Anxiety: From Theory and Research to Classroom Implications.* Upper Saddle River, NJ: Prentice Hall: 101-108.

Rathus, S. A. 1987. *Psychology.* New York: CBS College Publishing.

Riskind, J. H., Williams, N. L., Gessner, T. L., Chrosniak, L. D. and J. M. Cortina. 2000. The looming maladaptive style: Anxiety danger and schematic processing. *Journal of Personality and Social Psychology* 79: 837-852.

Ruiz-Caballero, J. A. and J. Bermudez. 1997. Anxiety and attention: Is there an attentional bias for positive emotional stimuli*? Journal of General Psychology* 124: 194-211.

Saito, Y. and K. Samimy. 1996. Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate, and advanced-level college student of Japanese. *Foreign Language Annals* 29: 239-248.

Saito, Y., Garza, T. J. And E. K. Horwitz. 1999. Foreign language reading anxiety. *The Modern Language Journal* 83: 202-218.

Scovel, T. 1991. The effect of affect on foreign language learning: A review of the anxiety research. In E. K. Horwitz and D. J. Young (eds) *Language Anxiety: From Theory and Research to Classroom Implications.* Upper Saddle River, NJ: Prentice Hall: 15-23.

Sobkowiak, W. 1996. *English Phonetics for Poles.* Poznań: Bene Nati.

Spielman, G. and M. L. Radnofsky. 2001. Learning language under tension. New directions from a qualitative study. *The Modern Language Journal* 85: 259-278.

VanPatten, B. and W. R. Glass. 1999. Grammar learning as a source of language anxiety: A discussion. In D. J. Young (ed) *Affect in Foreign Language and Second Language Learning. A practical Guide to Creating a Low-anxiety Classroom Atmosphere*. Boston: McGraw-Hill: 89-105.

Vasa, R. A. and D. S. Pine. 2004. Neurobiology in anxiety disorders in children and adolescents. In T. R. Morris and J. S. March (eds) *Anxiety Disorders in Children and Adolescents.* New York: Guilford Press: 3-26.

Vogely, A. 1999. Addressing listening comprehension anxiety. In D. J. Young (ed) *Affect in Foreign and Second Language Learning. A Practical Guide to Creating a Low-anxiety Classroom Atmosphere*. Boston: McGraw-Hill: 106-123.

Wade, C. and C. Tavris. 1990. *Psychology*. 2nd ed. New York: Harper & Row.

Watson, D. and R. Friend. 1969. Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology* 33: 448-457.

Woodrow, L. 2006. Anxiety and speaking English as a second language. *Sage Publications* 37(3): 308-328. DOI: 10.1177/00336882060 71315

Young, D. J. 1991. Creating a low-anxiety classroom environment: What does the anxiety research suggest? *The Modern Language Journal* 75: 426-439.

Young, D. J. 1992. Language anxiety from the foreign language specialist's perspective: Interviews with Krashen, Omaggio Hadley, Terrell, and Rardin. *Foreign Language Annals* 25: 157-172.

LEARNER PERSPECTIVE ON ENGLISH PRONUNCIATION TEACHING IN AN EFL CONTEXT

ELINA TERGUJEFF
University of Jyväskylä
elina.tergujeff@jyu.fi

Abstract
This paper reports on an interview study with EFL learners that aimed to explore learners' perceptions and views on English pronunciation teaching. The participants of the present study were ten EFL learners studying in the public educational system of Finland. Six of the participants were pupils attending basic education class nine, i.e. 15- to 16-year-old lower secondary level pupils. Two were primary level pupils attending basic education class four (aged 10), and two were upper secondary school pupils (aged 18). The interviews were thematic, and the learners were encouraged to speak freely about the English pronunciation teaching they were receiving and their opinions on this. In addition, they were asked to discuss their goals in English pronunciation, and to consider their pronunciation learning in class and out of class. The interviews were part of a wider study, mapping English pronunciation teaching practices in the context of Finnish schools.

On the basis of the findings, the learners do not seem to have aspirations to native-like pronunciation, but rather aim at achieving intelligible and fluent speech. Only few reported an accent preference (British or American). The primary level learners expressed satisfaction with the amount of pronunciation teaching, whereas most of the lower and upper secondary level learners claimed that pronunciation teaching was insufficient. Despite their criticisms of their pronunciation teaching, the learners reported that they had learnt English pronunciation at school. In addition, many of the learners described learning pronunciation outside school, e.g. through media and personal encounters.

Keywords: English as a foreign language, EFL, pronunciation teaching, interview.

1. Introduction

In Finland, English is taught in schools as a foreign language. It is the most popular language study option and, according to statistics, almost all schoolchildren study English as their first foreign language, beginning their study of English already at the primary level (Kumpulainen 2010, 88–89; for more information about the Finnish educational system, see ibid., 222). Although English has no official status in Finland, globalisation and the media have brought English into the everyday lives of Finns, also outside of the field of education: English is heard and seen in the linguistic landscape, needed in working life, and used in leisure activities, especially by youth. Moreover, Finns generally have a positive attitude to English, and they do not consider it a threat to their native languages or culture. (Leppänen et al. 2011.) The present study uses Finnish

schools as an example of an EFL context of English pronunciation teaching, and is motivated by the claims made about the possible neglect of teaching in this area, both in Finland (Lintunen 2004, 215; Iivonen 2005, 46) and internationally (e.g. Fraser 2000, Gilbert 2010). Also, it adds a learner perspective to the series of studies in English pronunciation teaching in Finland conducted by the author (Tergujeff 2010, 2012a, 2012b).

This article is part of a larger study on English pronunciation teaching in the context of Finnish schools, focussing here on the perspective of learners on the topic. More specifically, the paper reports on an interview study with EFL learners that aimed to explore learners' perceptions and views on English pronunciation teaching in the Finnish school context from primary to upper secondary level. The study sought to answer the following research questions:

1. *What do Finnish learners indicate as their goals in English pronunciation?*
2. *In the learners' view, how is English pronunciation taught in Finnish schools?*
3. *How do Finnish learners evaluate the English pronunciation teaching they are receiving, and their learning of English pronunciation?*

The study addresses several issues related to pronunciation teaching and learning: learners' goals, teaching practices, and learners' evaluations of their teaching. With respect to teaching practices, the teaching of phonemic script has special focus in this study. It has been suggested that phonemic transcription is a beneficial learning tool for Finnish learners of English (see Lintunen 2004). Because the present study addresses such a wide variety of issues, a separate literature review is not given here, but relevant previous research is discussed in section three in connection with the analysis.

2. The present study

The participants of the present study were ten EFL learners, studying English in the public educational system in Finland. Six of the participants were pupils attending basic education class nine, i.e. 15- to 16-year-old lower secondary level pupils. Two were primary level pupils attending basic education class four (aged 10), and two were upper secondary school pupils (aged 18). Participant information is presented in Table 1. The pupils came from three different schools, and one of the pupils (marked with *) studied English with a special education teacher separately from the rest of his class. The names have been changed to ensure participant anonymity. All of the participants volunteered to take part in the study, and signed a written consent allowing the interviews to be used for research purposes. In the case of the under-aged participants, the consent forms were signed by their guardians.

participant	***level***	***school***
Maria	primary	A
Hanna	primary	A
Anna	lower secondary	B
Liisa	lower secondary	B

participant	*level*	*school*
Emma	lower secondary	B
Selma	lower secondary	B
Valtteri	lower secondary	B
Lassi	lower secondary	B*
Suvi	upper secondary	C
Linda	upper secondary	C

Table 1. Participant information.

To answer the research questions set for the present study, thematic interviews were conducted with the participants. In the interviews, the learners were encouraged to speak freely about the English pronunciation teaching they were receiving and their opinions on this. In addition, they were asked to discuss their goals in English pronunciation, and to evaluate their pronunciation learning in class and out of class. As stimuli for the discussion, the interviewees' own EFL textbooks and a list of words in phonemic script were used. The interviews were framed such that the researcher told the interviewees that she did not know how English pronunciation was taught in Finnish schools and considered the pupils as the experts best able to provide her with this information (Fontana & Prokos 2007, 70). The interviews were conducted in the learners' native tongue, i.e. in Finnish. In this article, I refer to the original Finnish-language data, but translations into English are also provided. The interviews took place in various surroundings: at the learners' school, on the premises of the researcher's institution, and also at the home of one of the interviewees (the youngest participants were interviewed in the home of one of them to reduce possible nervousness on the part of the children). In the school context, appropriate permissions were asked from and granted by the head of school and the municipal education authorities.

The interviews were audio-recorded and transcribed by the author for qualitative content analysis (Kvale 2007, 105). The content analysis was applied to the data to identify the central thematic categories. Six categories emerged: (1) pronunciation goals, (2) pronunciation exercises in textbooks, (3) pronunciation teaching practices, (4) pronunciation models, (5) amount of pronunciation teaching, and (6) pronunciation learning. Conclusions were drawn for each of the six categories based on interpretations of the interview excerpts. To avoid researcher bias, researcher triangulation (Denzin 1978) was carried out: two established researchers gave their analysis of the data in addition to the author's. The purpose of triangulation was also to gain a deeper understanding of the topic by discussing the data in a group.

3. Analysis and discussion

The results of the qualitative analysis of this interview study are described below. The analysis is discussed in connection with the results. The quotes illustrate recurrent or otherwise interesting themes spotted in the analysis.

3.1 Learners' pronunciation goals

A review of the previous research on learners' goals in English pronunciation, accent preferences and attitudes towards accents reveal a number of interesting results. Many learners seem to have negative attitudes towards (their own) non-native and *outer circle* (Kachru 1985) varieties (e.g. Pihko 1997, Dalton-Puffer et al. 1997), and they often prefer an accent that is familiar to them: in Europe, this seems to be British Received Pronunciation (e.g. Dalton-Puffer et al. 1997, Genoz & Garcia Lecumberri 1999, Waniek-Klimczak & Klimczak 2005). Learners' aspirations to learn a native-like pronunciation have been recorded in both ESL (Derwing 2003) and EFL (e.g. Janicka et al. 2005) environments. However, in a survey of Polish EFL learners by Waniek-Klimczak (1997), only a minority wished to sound native-like.

In the present study, the majority of the learners reported fluency and intelligibility as their main goals in English pronunciation. In addition, they did not have ambitions to the production of a specific variety, as pointed out by the following learner:

> (1) "[Haluaisin oppia] hyvää englantia, ymmärrettävää englantia. Se menee monesti semmoseksi suomen englanniksi, semmoseksi töksähteleväksi, mutta haluaisin osata semmoista sujuvaa ja ei sillä oo väliä onkse brittiä vai amerikkalaista mutta kunhan se olis oikeen sellasta ymmärrettävää ja sujuvaa."
> *([I would like to learn] good English, intelligible English. I often slip into a kind of Finnish English, awkward-sounding, but I would like to be fluent. It doesn't matter whether it's British or American, as long as it's genuinely intelligible and fluent.)*
> (Anna, lower secondary level)

For the learners interviewed for the present study, native-like pronunciation does not seem to be a goal. Some of the learners pointed out that it does not bother them if people hear that they have a foreign accent, while one learner clearly stated his wish to be identified as a Finn (cf. Jones 2001 on accent as a reflection of identity), when asked whether he would find it desirable to speak without a foreign accent:

> (2) "Ei se hienoa olis. Haluan korostaa sitä että en ole brittiläinen vaan olen suomalainen."
> *("No, it wouldn't be nice. I want to emphasise that I'm not British but a Finn.")*
> (Valtteri, lower secondary level)

If native-like pronunciation was mentioned by the learners, they referred to it as if it were only wishful thinking. A couple of learners considered it "nice" if they could speak like a native speaker, but this was still not their main goal. They emphasised intelligibility and fluency, and also stated that it did not matter if listeners notice their foreign accent. The results perhaps reflect the general change in attitudes towards non-native accents and accented speech: it is widely accepted to speak English with a foreign accent as long as it does not compromise intelligibility (cf. the work of Jenkins, e.g. Jenkins 2000). It has also been suggested that in English pronunciation teaching in Europe, the use of "a type of International English" as pronunciation model is gaining a foothold (Henderson et al. 2012, Tergujeff 2012b).

3.2 Pronunciation exercises in textbooks

Textbooks play an important role in foreign language teaching. In the Finnish context, the dominance of textbooks over other teaching materials has been shown in a survey by Luukka et al. (2008), and, with particular reference to English pronunciation teaching, by Tergujeff (2012b). In the present study, the learners' textbooks were used as stimuli for discussion in the interviews by asking them to introduce typical pronunciation exercises in their textbooks. This was not an easy task because it seemed that pronunciation is not a frequent textbook topic. Many had to struggle to find pronunciation exercises:

(3) ”Emmä tiiä onks täälä semmosia. -- Ehkä enemmän just niissä yläasteen ku lukion. Miten mä en löydä täältä niinku yhtään mitään?”
(“I don't know if there are any. -- Maybe there were more in the lower secondary level books. How come I don't find anything?”)
(Suvi, upper secondary level)

Research-based information on the relative proportion of pronunciation exercises in Finnish EFL textbooks is not available. A classification of pronunciation teaching materials in Finnish EFL textbooks, however, is available in a recent textbook analysis (Tergujeff 2010). This textbook analysis revealed a range of pronunciation teaching materials: phonetic training, reading aloud, imitation, rhymes, rules and instructions, awareness-raising activities, spelling, dictation, and ear training. In the learners' view, the range of pronunciation teaching materials in their EFL textbooks seems to be narrower than was indicated by the textbook analysis. The learners mentioned word stress exercises, in which the learners listen to words and mark the correct stress placement, as a frequent exercise type in the lower secondary school textbooks that they used. Another frequently mentioned exercise type was a list of words and expressions (presented in a text box) preceding a text. In these lists, words and expressions from the text are highlighted before the text is studied. The lists can be listened to on the CD accompanying the teacher's book; according to the learners a typical classroom procedure is repeating words and expressions aloud together as pronunciation practice.

(4) “No tuossa on siitä mihin se painotus tulee. Ja tuossahan on noita merkkejä että miten se äännetään. -- Ja sit se on tässä tekstikirjassa ku on ennen tekstejä näitä laatikoita niin nämä me käydään aina läpi”.
(“Here's one on where the stress falls. And here are symbols showing how it's pronounced. -- And in the textbook there are boxes like this before each text, and we always study them.”)
(Liisa, lower secondary level)

The primary level pupils mentioned that their textbook includes a CD. The pupil's CD is a concise version of the teacher's CD, and usually features the audio version of the key texts of the textbook. Maria stated that the CD was specifically for pronunciation practice at home:

(5) “Kuuntelen sitä ja siinä on semmosia pieniä taukoja et sen aikana voi ääntää niitä.”
(“I listen to it, and there are pauses during which you can pronounce the words.”)
(Maria, primary level)

Based on Luukka et al. (2008) and Tergujeff (2012b), textbooks are the most widely used teaching materials in foreign language teaching in Finland: almost all teachers use textbooks. Another proof of the major role of textbooks in foreign language teaching is offered here, as many of the learners stated that during the lessons they do not often skip things in the textbook but cover all of it during the term.

(6) "Kyllä me ollaan noita tehty. Että melkein kaikki asiat täältä kirjasta on käyty. Ettei kauheasti hypitä kyllä."
("Yes, we have done those. We have covered pretty much all of the book. We seldom skip stuff.")
(Selma, lower secondary level)

3.3 Pronunciation teaching practices

When the learners talked about the pronunciation teaching they were receiving at school, they mentioned very traditional teaching techniques: mostly imitation and reading aloud. This suggests that despite the recommendations in the literature on the subject (e.g. Morley 1991; Celce-Murcia et al. 2010, 44–45), pronunciation teaching relies heavily on mechanical production without moving on to controlled practice and, finally, communicative tasks, as recommended at the stage when the learner has already learnt to produce the segments of the target language. However, it may also be that the teaching practices include more general oral skills (conversational) tasks which the learners do not label as pronunciation activities, since in their minds these consist of segmental-level mechanical production. The typical classroom practices reported are well exemplified in the following excerpt from the interview with Valtteri (lower secondary level):

(7) Valtteri: Se perinteinen on se että opettaja sanoo sanan oikein ja oppilaat sanoo perässä. Yrittää ääntää samalla tavalla. No mitenkähän sitä nyt yleensäkään… opetellaan. Aika lailla sillä tavalla.
("The traditional way is that the teacher says the word correctly and the pupils repeat it. Try to pronounce it the same way. Let me think how do we usually… study. Well, pretty much like that.")
Interviewer: Tuleeko muita harjoituksia mieleen?
("Can you think of any other tasks?")
Valtteri: No niitä sellaisia kai että pitää kuunnella nauhalta niitä sanoja ja pitää siinä kohtaa merkata missä se on se paino siinä sanassa.
("Well I guess those in which you have to listen to words and mark where the stress falls in that word.")
Interviewer: Mitä muuta opetetaan kuin painoa?
("What else do they teach, in addition to stress?")
Valtteri: Ei niitä enää sillä lailla opeteta kun nehän on tullu jo ala-asteella ne hommat että miten mitkä kirjaimet ääntyy missäkin tilanteessa minäkin ja tämmöset. Tämmöset hankalat sanat käydään erikseen. Aika lailla keskitytään niihin yksittäisiin äänteisiin että koko sana menee oikein.
("They don't teach that much anymore because it's all covered in primary school. The stuff about how letters are pronounced in different positions and that stuff. Difficult words are dealt with separately. We pretty much focus on individual sounds to get the whole word right.")

It also seems common for teachers to deal spontaneously with pronunciation when difficulties appear; e.g. when a pupil is unable to pronounce something, or there is a recurrent mispronunciation. This aspect has been discussed by Burgess & Spencer (2000), and by Macdonald (2002), who interviewed Australian ESL teachers reluctant to teach pronunciation. Macdonald mentions that pronunciation teaching is not always systematic but incidental in nature and that pronunciation is dealt with in class "as it comes up".

The present study focused in particular on the use of phonemic script in pronunciation teaching. This derives from the fact that the orthography of the learners' L1, Finnish, follows a principle of close letter-to-sound correspondence (Suomi et al. 2008, 141), and thus phonemic transcription can be seen as a beneficial learning tool for them, helping them to tackle their difficulties with the sometimes ambiguous spelling of English (cf. Wells 1996). What is more, it has been suggested that transcription skills and English pronunciation skills correlate in advanced Finnish learners of English (Lintunen 2004). Based on the interviews in the present study, phonemic script is not very commonly used in pronunciation teaching. A similar de-emphasis was found in an earlier study, based on classroom observations of the teaching of Finnish EFL teachers (Tergujeff 2012a), and a retrospective learner survey in Lintunen (2004, 183–188). However, the participants of the present study often stated that even though their teaching did not at the moment make use of phonetic symbols, these had been used earlier in their education – typically already in primary school. This is also supported by the fact that the primary level pupils interviewed here reported receiving teaching of phonetic symbols. The following excerpt is from an interview with Emma who is currently in lower secondary school.

(8) Interviewer: Ne on tuttuja sulle?
("You are familiar with them?")
Emma: Joo mutta mä en oikein osaa niitä. Tai siis silleen en oo koskaan osannu näitä kovin hyvin.
("Yes but I don't really know them. I mean I have never known them that well.")
Interviewer: Niitä ei ole varmaan paljon sitten opetettukaan?
("So have they not been taught thoroughly?")
Emma: No ku ala-asteella mä en ollu todellakaan tosi hyvä niinku englannissa -- niin mä en oikein keskittyny enkä halunnukaan oikeen oppia sitä nii vasta motivaatio nousi ku halus lähtee ulkomaille, nii en mä muista näistä kauheesti.
("Well I wasn't very good at English in primary school -- so I didn't concentrate and didn't even want to learn English. I had no motivation until I wanted to go abroad, so I don't remember much about them.")
Interviewer: Eli niitä on opeteltu ala-asteella mutta nytkö ei enää?
("So the symbols were taught in primary school but not anymore?")
Emma: ei yläasteella ole minun mielestä paljoo opetettu näitä.
("No they haven't been taught much in lower secondary school in my opinion.")

It seems likely that the interviewees had been taught phonetic symbols at some stage, even if they were not used in their current teaching, as despite the learners' tendency to downplay their skills, most of them were able to read phonemic transcriptions of single words presented to them in the interview. The idea was not to test their skills but to use

transcription reading as a stimulus for the discussion. The transcriptions included words such as *cat*, *fish*, *robot*, *anyone*, *religion*, *nothing* and *allergic*.

The learners were divided on the usefulness of knowing phonemic script. This division was not linked to age or level of proficiency, and may perhaps have more to do with personal preference or learner type. The same division was found among respondents to a teacher survey (Tergujeff 2012b). Even the reasoning behind the opinions was similar in teachers and learners: those who saw it as beneficial referred to checking the correct pronunciation of words, and those who did not referred to the actual spelling of words, feeling that phonemic script would interfere with the learners' spelling, as in the following quote:

(9) "-- en mä oikeen ymmärtäny että miksi noita pitäs tolleen kattoa. Ku ei niitä oikeastikaan noin kirjoteta."
("I didn't understand why they should be seen like that. After all that's not how you write English anyway.")
(Emma, lower secondary level)

3.4 Pronunciation models

As stated in 3.1, the learners do not seem to have great ambitions towards achieving a native-like pronunciation, and no strong preferences for a specific accent. A recent survey suggests that the pronunciation models most commonly used English pronunciation teaching across Europe are British Received Pronunciation (RP) and General American (GA) (Henderson et al. 2012). This finding is supported by the view of the learners interviewed for the present study. According to learners, the varieties used in the teaching they receive are British and American. Most of the learners stated that both are used, and that one is typically the main variety whereas the other is introduced on the side.

(10) "Britti. Sitä on. Ollaan me käyty vähän tota amerikanenglantiakin ja niitä eroavaisuuksia katottu."
("It's British, that's what it is. We have also explored American English a little, looked at the differences.")
(Selma, lower secondary level)

The majority of the learners reported that the British variety was the main pronunciation model taught, while for some learners it was American English. Other varieties, or introductions to these, were seldom mentioned. Valtteri, however, mentioned Canadian English:

(11) "Kyllä se nyt ollu vähän kumpaakin [britti- ja amerikanenglantia] nytte niinku viimesinä vuosina. Että sehän on alkanu ala-asteelta ja seiskaluokalle saakka brittienglantina mutta sitten meillä on ollut justiinsa tämä kirja missä se korosti niitä eroja ja täällon paljo tehtäviä alussa niistä. Täällon näitä sanaeroja ja kaikkee ja ääntämiseroja. Sitten täällä on kanadanenglantiakin. Tai täällä on pari kappaletta missä on vaan tehtäviä näistä."

> *("It's been both [British and American] these past few years. It began as British English in primary school until seventh grade but now we've had this book that emphasises the differences and there are lots of exercises on them. Differences in words and everything, and pronunciation. And there's Canadian English even. I mean there are a couple of texts with exercises.")*
> (Valtteri, lower secondary level)

The use of this variety has already come up in a previous study: in a recent survey (Tergujeff 2012b), 21.1% of the respondents (EFL teachers working in Finland, n=76) reported that they use Canadian English for receptive pronunciation tasks. It was suggested that this is due to the current EFL textbooks used in Finland, which also include audio material in Canadian English (ibid.). The use of different varieties offers opportunities for raising accent awareness, and even for receptive *accent addition* (i.e. adding accents to one's receptive repertoire by means of perceptual training, as suggested by Jenkins 2000, 208–212). As there is great variation in the pronunciation of English worldwide, it is good for learners to be aware of this and prepared to encounter people who speak differently from the main pronunciation model offered to them in teaching; see e.g. Cunningham (2009).

3.5 Amount and success of pronunciation teaching

In recent years, it has repeatedly been claimed that pronunciation teaching is generally neglected, both in Finland (Iivonen 2005, 46; Lintunen 2004, 215) and internationally (e.g. Fraser 2000, Gilbert 2010). This claim can be seen as related to the rise of Communicative Language Teaching (CLT), which started at the end of the 1970s. The proponents of CLT largely rejected traditional pronunciation teaching as incompatible with teaching language as communication (Celce-Murcia et al. 2010, 11), yet offered no comprehensible communicative set of methods in return (ibid., 9). Also, many teachers find pronunciation difficult to teach (Macdonald 2002), and are of the opinion that their training in how to teach pronunciation has been insufficient (Breitkreutz et al. 2001; Foote et al. 2011; Henderson et al. 2012; Tergujeff 2012b).

In the present study, learners were asked how much attention was paid to pronunciation in the teaching they were receiving, and what they thought of the amount they received. The primary level pupils reported receiving plenty of pronunciation teaching. Added to the findings on the teaching of phonemic script, it seems that more attention is paid to pronunciation teaching at the primary level than at the lower and upper secondary levels. This is in line with the national core curriculum for basic education (Finnish National Board of Education 2004), which emphasises the primacy of oral language skills in teaching foreign languages, and states that the weight given to written skills is to be added gradually (ibid., 139). The present study gives grounds to speculate that currently the weight given to written skills is added at the cost of pronunciation, as the majority of the learners expressed dissatisfaction with the amount of pronunciation teaching at the lower and upper secondary levels. They stated that this component had not been dealt with sufficiently, and they hoped for more teaching in this area. This view is exemplified by the following quote:

(12) "Aika vähän minun mielestä. Siis siihen keskitytään ihan liian vähän koska se olis paljon tärkeämpää kun mitä sitä nyt harjotellaan. Minusta se on aika vähästä mitä me sitä harjotellaan. Ja just että koko sen tunnin pitäs pystyä puhumaan englanniksi jotenkuten ettei aina menis siihen että puhuu sitte oikeesti suomeksi ja sanoo jotain vähän sinne päin. Tulis semmosta sujuvuutta."
("There's way too little focus on that, as pronunciation is much more important than you'd figure from the amount of practice at school. I think we practise pronunciation very little. And I think we should speak English the whole lesson and not slip into Finnish. It would bring that fluency.")
(Anna, lower secondary level)

While many of the learners expressed that pronunciation teaching is insufficient, some appeared to have taken action on their own initiative to develop their pronunciation skills. They had adopted an active role in the learning process, both in class and out of class. Emma, for example, said she reads texts aloud at home to practise pronunciation, and that she regularly asks her teacher how words are pronounced, as exemplified in the following quote. It also appears that she has identified a way of learning that suits her, and that she is aware of her own learning.

(13) "Olen ihan hyvin kyllä oppinu. Siis minä ite aina tykkään kysyä että miten tämä äännetään kun en tiiä. Ja silleen oon ihan hyvin oppinu."
("I have learnt pretty well. I like to ask how something is pronounced if I don't know it. I've learnt well that way.")
(Emma, lower secondary level)

Despite the fact that most learners interviewed for the present study claimed that pronunciation teaching was insufficient, they nevertheless considered that they had learnt English pronunciation at school. "Learning by doing" – presumably meaning exposure to English and practising it by speaking – was mentioned frequently by the learners. The learners seemed to be saying that it would be impossible *not* to learn pronunciation at school, which gives an interesting addition to the discussion on whether pronunciation is a teachable skill in the first place. The positive effects of formal pronunciation instruction have been challenged by Suter (1976) and Purcell & Suter (1980), but most studies conducted in this area have reported on developed pronunciation skills after teaching experiments (for a synthesis see Saito 2012). The following quote from Anna is an example of how obvious the learning of English pronunciation in class is to the interviewees.

(14) "Joo, kyllähän sitä väkisinkin oppii ja kun kuuntelee niin, tekemällä oppii."
("Oh yes, you just learn, and when you listen, yes, you learn by doing.")
(Anna, lower secondary level)

The learners reported they had also learnt English pronunciation outside school. It seems that the majority of them engage in various leisure activities that include the use of English. When asked whether he had learnt English pronunciation at school, Valtteri was of the opinion that leisure activities had taught him more about English pronunciation than formal teaching:

(15) "No, jaa-a. Enpä nyt sanois. Ehkä tietenki jonkun verran. Sehän on että ääntämään oppii aina ku puhuu. Koulussa tulee aika paljon kuitenkin käytettyä englantia kun nää on englannin tunteja. Kyllä sitä on jonkun verran tullu opittua mutta suurin osa tulee vapaa-ajalta."
("Well, I wouldn't say so. Perhaps a little. You learn to pronounce whenever you speak, that's how it is. At school we use quite a lot of English, as these are English lessons. I've learnt some but mostly in my free time.")
(Valtteri, lower secondary level)

The learners' descriptions of their pronunciation learning outside of the classroom included various contexts. However, many of them had to do with media. The learners seemed to feel that listening to music, watching television and films, and playing online games is beneficial to their learning of English pronunciation. Playing online games includes talking to other players using English as a lingua franca.

(16) "Aika paljo sillä lailla ku jotaki pelejä pelaa. Onlainina. Nii siinä käytetään aika paljon Skypea ja TeamSpeakiä ja näitä, että että voi kommunikoida. Se on helpompaa ku ruveta kirjottelemaan siinä kesken kaiken. Siinä oppii aika paljo."
("Quite a lot by playing games online. We use Skype and TeamSpeak and so on, to communicate. It's easier than typing in the middle of the game. You learn a lot like that.")
(Valtteri, lower secondary level)

Foreign contacts in person were also mentioned by the learners as learning situations. According to the interviewees, these encounters typically take place with tourists and foreign seasonal workers in Finland (as in the case of Lassi, lower secondary level).

(17) Lassi: Ääntämään? No en. Onhan se mitä kesällä ulkomaalaisten kans. Siinäki jotaki.
("To pronounce? No. Or well, a little with the foreigners in summertime. I guess that's something.")
Interviewer: Niin sä kuitenkin puhut joittenkin kanssa englantia koulun ulkopuolella? Keitä ne ovat?
("So you do speak English with people outside school? Who are they?")
Lassi: No yleensä thaimaalaisten kanssa. Ja joskus kun pelaa netissä ni niitten ulkomaalaisten kanssa.
("Usually Thais. And sometimes when I play online I speak with the foreigners there.")

Overall, the learners seemed able to identify their own learning of English pronunciation in both formal and informal contexts. The English language is strongly present in Finnish society, and it is known from a nation-wide survey that young people in particular use English also in their leisure activities (see Leppänen et al. 2011). The present study sheds light on the language learning involved in these activities, as many of the interviewees were of the opinion that they have learnt English pronunciation while playing online games, listening to music, watching television and films, and encountering foreigners.

4. Conclusions

The aim of the present paper was to explore learners' views on English pronunciation teaching in the context of Finnish schools from primary to upper secondary level. This interview study yielded the following main results in answer to the research questions. The learners considered intelligible and fluent speech to be their main goal in English pronunciation. They did not express aspirations for a native-like accent, and did not commonly have an accent preference. According to the learners, English pronunciation teaching mainly relies on traditional methods such as imitation and reading aloud. Tasks on word stress placement were also frequently mentioned. Phonemic script (on the usefulness of which the learners were divided) is used more in the teaching at the primary than lower or upper secondary level. The results also suggest that pronunciation teaching is not necessarily very systematic but rather is spontaneous in nature. Textbooks seem to play an important role in English pronunciation teaching. A British or American standard variety is generally used as the pronunciation model. Pronunciation is taught extensively at the primary level, and the learners expressed satisfaction with this. The learners at the lower and upper secondary level expressed the opinion that pronunciation is not paid enough attention to in teaching, and would like more pronunciation teaching. Notwithstanding, all the interviewees stated that their pronunciation skills had developed because of classroom activities. In addition, many reported learning pronunciation outside of the classroom, e.g. through media and personal encounters.

The results of the present study imply that more attention could be paid to pronunciation at the lower and upper secondary levels. According to the learners, it is not sufficient to focus on pronunciation at the primary level only; instead, they would like to see a continuation of pronunciation teaching at the later stages. The major role of textbooks in teaching imposes pressure on them, as there is a risk of language items being left out of the teaching if they are not dealt with in the textbook. In this connection, it is worth keeping the old proverb about "good servants but poor masters" in mind: textbooks are valuable tools for the teacher, but it is the curriculum that defines the objectives of teaching and the teacher who uses his or her expertise in planning and teaching the lessons (cf. Cunningsworth 1984, 1). Teachers could also pay more attention to opportunities for learning pronunciation outside of the classroom, and try to build bridges between learners' leisure and classroom activities. After all, many of the learners interviewed for the present study indicated that they had learnt English pronunciation outside school.

Acknowledgements

The author is grateful to the interviewees for sharing their views and experiences, and would also like to thank two additional experts for their contribution to the researcher triangulation. In addition, a warm thank you to Professor Hannele Dufva for her most helpful comments on the manuscript.

References

Breitkreutz, J.A., T.M. Derwing & M.J. Rossiter 2001. Pronunciation teaching practices in Canada. *TESL Canada Journal* 19 (1), 51–61.

Burgess, J. & S. Spencer 2000. Phonology and pronunciation in integrated language teaching and teacher education. *System* 28 (2), 191–215.

Celce-Murcia, M., D.M. Brinton, J.M. Goodwin & B. Griner 2010. *Teaching Pronunciation. A Course Book and Reference Guide.* Second edition. Cambridge: Cambridge University Press.

Cenoz, J.M. & L. Garcia Lecumberri 1999. The Acquisition of English Pronunciation: Learners' Views. *The International Journal of Applied Linguistics* 9 (1), 3–17.

Cunningham, U. 2009. Models and Targets for Pronunciation of English in Vietnam and Sweden. *Research in Language* 7 (1), 113–128.

Cunningsworth, A. 1984. *Evaluating and selecting EFL teaching materials.* London: Heineman Educational Books.

Dalton-Puffer, C., G. Kaltenboeck & U. Smit 1997. Learner Attitudes and L2 Pronunciation in Austria. *World Englishes* 16, 115–127.

Denzin, N.K. 1978. *The research act: A theoretical introduction to sociological methods.* Second edition. New York: McGraw-Hill.

Derwing, T.M. 2003. What do ESL students say about their accents? *Canadian Modern Language Review* 59, 547–567.

Finnish National Board of Education. 2004. *National Core Curriculum for Basic Education.* Helsinki: Author. Available at http://www.oph.fi/english/sources_of_information/core_curricula_and_qualification_requirements/basic_education. [Retrieved 4 Feb 2013]

Fontana, A. & A.H. Prokos 2007. *The Interview: From Formal to Postmodern.* Walnut Creek, CA: Left Coast Press.

Foote, J.A., A.K. Holtby & T.M. Derwing 2011. Survey of the Teaching of Pronunciation in Adult ESL Programs in Canada, 2010. *TESL Canada* 29 (1), 1–22.

Fraser, H. 2000. *Coordinating Improvements in Pronunciation Teaching for Adult Learners of English as a Second Language.* Canberra: DETYA.

Gilbert, J.B. 2010. Pronunciation as orphan: What can be done? *As We Speak, Newsletter of SPLIS, 2010.*

Henderson, A., D. Frost, E. Tergujeff, A. Kautzsch, D. Murphy, A. Kirkova-Naskova, E. Waniek-Klimczak, D. Levey, U. Cunningham & L. Curnick 2012. The English Pronunciation Teaching in Europe Survey: Selected results. *Research in Language* 10.1, 5–27.

Iivonen, A. 2005. Fonetiikan merkitys kielenomaksumisessa ja -opetuksessa. In A. Iivonen, R. Aulanko & M. Vainio (eds.), *Monikäyttöinen fonetiikka.* Mimeographed series of the Department of Phonetics 21, University of Helsinki. 3 edition.

Janicka, K., M. Kul & J. Weckwerth 2005. Polish Students' Attitudes to Native English Accents as Models for EFL Pronunciation. In K. Dziubalska-Kołaczyk & J. Przedlacka (eds.), *English Pronunciation Models: A Changing Scene.* Bern: Peter Lang. 251–292.

Jenkins, J. 2000. *Phonology of English as an international language.* Oxford: Oxford University Press.

Jones, K.W. 2001. "I've Called 'em Tom-ah-toes All My Life and I'm Not Going to Change!": Maintaining Linguistic Control over English Identity in the U.S. *Social Forces* 79 (3), 1061–1094.

Kachru, B.B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk and H. Widdowson (eds.), *English in the world: Teaching and learning the language and literatures*. Cambridge: Cambridge University Press. 11–36.

Kumpulainen, T. (ed.) 2010. *Koulutuksen määrälliset indikaattorit 2010.* Helsinki: Finnish National Board of Education. Available at http://www.oph.fi/download/130716_Koulutuksen_maaralliset_indikaattorit_2010.pdf. [Retrieved 4 Feb 2013]

Kvale, S. 2007. *Doing interviews.* London: Sage.

Leppänen, S., A. Pitkänen-Huhta, T. Nikula, S. Kytölä, T. Törmäkangas, K. Nissinen, L. Kääntä, T. Räisänen, M. Laitinen, P. Pahta, H. Koskela, S. Lähdesmäki & H. Jousmäki 2011. *National Survey on the English Language in Finland: Uses, meanings and attitudes.* Studies in Variation, Contacts and Change in English, Vol. 5. Available at http://www.helsinki.fi/varieng/journal/volumes/05. [Retrieved 4 Feb 2013]

Lintunen, P. 2004. *Pronunciation and Phonemic Transcription: A study of advanced Finnish learners of English.* Turku: University of Turku.

Luukka, M-R., S. Pöyhönen, A. Huhta, P. Taalas, M. Tarnanen & A. Keränen 2008. *Maailma muuttuu – mitä tekee koulu? Äidinkielen ja vieraiden kielten tekstikäytänteet koulussa ja vapaa-ajalla.* University of Jyväskylä: Centre for Applied Language Studies.

Macdonald, S. 2002. Pronunciation – views and practices of reluctant teachers. *Prospect* 17 (3), 3–18.

Morley, J. 1991. The Pronunciation Component in Teaching English to Speakers of Other Languages. *TESOL Quarterly* 25 (3), 481–520.

Pihko, M-K. 1997. *His English Sounded Strange. The Intelligibility of Native and Non-Native English Pronunciation to Finnish Learners of English.* University of Jyväskylä: Centre for Applied Language Studies.

Purcell, E.T. & R. W. Suter 1980. Predictors of pronunciation accuracy: A reexamination. *Language Learning* 30 (2), 271–287.

Saito, K. 2012. Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies. *TESOL Quarterly* 46 (4), 842–854.

Suomi, K., J. Toivanen & R. Ylitalo 2008. *Finnish sound structure: Phonetics, phonology, phonotactics and prosody.* Oulu: University of Oulu.

Suter, R.W. 1976. Predictors of pronunciation accuracy in second language learning. *Language Learning* 26 (2), 233–253.

Tergujeff, E. 2010. Pronunciation teaching materials in Finnish EFL textbooks. In A. Henderson (ed.), *English Pronunciation: Issues and Practices (EPIP): Proceedings of the First International Conference. June 3–5 2009, Université de Savoie, Chambéry, France.* Université de Savoie: Laboratoire LLS. 189–205.

Tergujeff, E. 2012a. English pronunciation teaching: Four case studies from Finland. *Journal of Language Teaching and Research* 3 (4), 599–607.

Tergujeff, E. 2012b. The English Pronunciation Teaching in Europe Survey: Finland. *Apples – Journal of Applied Language Studies* 6 (1), 29–45.

Waniek-Klimczak, E. 1997. Context for Teaching English Phonetics and Phonology at Polish Universities and Colleges: A Survey. In Waniek-Klimczak, E. (ed.), *Teaching English Phonetics and Phonology II. Accents '97.* Łódź: Wydawnictwo Uniwersytetu Łódzkiego. 5–17.

Waniek-Klimczak, E. & K. Klimczak 2005. Target in Speech Development: Learners' Views. In K. Dziubalska-Kołaczyk & J. Przedlacka (eds.), *English Pronunciation Models: A Changing Scene.* Bern: Peter Lang. 229–249.

Wells, J.C. 1996. Why phonetic transcription is important. *Malsori* 31–32, 239–242.

• **Research in Language, 2013, vol. 11.1** • DOI 10.2478/v10015-012-0012-x

ROADRUNNERS AND EAGLES

LINDA SHOCKEY
University of Reading
secondclarinet@gmail.com

MAŁGORZATA E. ĆAVAR
Eastern Michigan University
malgosia@cavar.me

Abstract
Our previous research on perception of gated casual English by university students suggests that *ceteris paribus*, Polish students are much more accurate than Greeks. A recent pilot study of casually-spoken Polish leads us to the conclusion that many shortcuts found in English are also common in Polish, so that similar perceptual strategies can be used in both languages, though differing in detail. Based on these preliminary results, it seems likely that perceptual strategies across languages tend towards the "eagle" approach - where a birds-eye view of the acoustic terrain without too much emphasis on detail is found - or the "roadrunner" approach, where phonetic detail is followed closely. In the former case, perceivers adjust easily to alternation caused by casual speech phonology while in the latter, perceivers expect little variation and possibly even find it confusing. Native speakers of Greek are "roadrunners", since there is little phonological reduction in their language there is little difference, for example, between stressed and unstressed syllables. We suggest that native speakers of Polish join English speakers as "eagles", which gives them a natural perceptual advantage in English. There is a conceptual similarity between this idea and that of the stress- or syllable-timed language, and we hypothesise that as in this case, there is a cline rather than a sharp division between eagles and roadrunners. As usual, more research is called for.

Keywords: perception, casual speech processes, prosody.

1. Introduction

It has long been axiomatic in foreign language instruction that the phonetics and phonology of one's L1 has a strong influence on production and perception of subsequent languages. Here, we will address an aspect which has received little attention: the contribution to perception of casual speech phonology. We will suggest that the study of "shortcuts" is more crucial for some students of English than for others.

Unselfconscious, conversational English is known to employ a variety of processes which increase contrast between stressed and unstressed syllables and reduce the size of consonant clusters, especially syllable-finally (see Shockey, 2003 for a summary). A few examples are:

thousand	[ˈθaʊz̩n̩]
can't	[kɑ̃ʔ]
surprise	[s̩ˈpraɪz̩]

These reductions or shortcuts are especially common in connected speech when consonants build up across a word or syllable boundary:

last night	[læsˈnaɪʔ]
mountain	[ˈmɑ̃ʊ̃ntn̩ ˈmɑ̃ʊ̃ʔn̩ ˈmɑ̃ʊ̃pm̩]
run the race	[rʌn̪ːə˞ɛɪs]
next week	[nɛksˈwik]
weakest link	[wikɨsˈlɪŋk]

We assume that one consequence of these very common reductions in spoken English is that native speakers learn to accept a variety of approximations to canonical pronunciation, based on an unconscious knowledge of what can be underachieved and what cannot. This knowledge forms part of their phonological competence and arguably involves recognition of the general phonetic profile of an utterance rather than an expectation of strict congruence with the most formal representation. We suggest the term "eagle" for this perceptual style because it involves recognising features of the landscape rather than precise detail.

When asked to recognise a gated[1] English sentence in which several conversational shortcuts are featured, native speakers of English generally achieve very high accuracy, with some delay. This has been reported in the literature for some time (cf. Bard et al, 1988). Typically, English native speakers can reinterpret a phonetic sequence as a reduced phonological string at the point when the conditioning factors are revealed. For example, as reported in Shockey, (2003, 97), when they hear "The screen [skrim] play", they interpret the second word as "scream" until they hear the "p", whereupon they usually reinterpret it as "screen".

2. The study

In an experiment reported elsewhere (Shockey and Bond 2012) we tested the perception of gated conversational English by speakers of two other languages, Greek and Polish. The groups tested were matched for age and experience with English. The stimuli were presented in 50-msec gates in quiet conditions.

[1] Gating is a process by which an utterance is presented in small incremental time units, building up from the beginning (Grosjean, 1980). Subjects are asked to judge what they have heard after each gate, and the percept builds up as more information becomes available.

The gated sentence was:
So it was quite good fun, actually, on the wedding, though...
[səʷɪʷsʷˡkwaɪʔgʊfʌnætʃʊɪn̪ːəˡwɛɗɨŋ..d̪ɵʉ]
/səʊ ɪt wɒz .....ækʃʊəli.....ðəʊ/

There was no 't' in 'it'
The [w] in 'was' was represented by rounding in the first syllable
The 't' in 'quite' was a glottal stop, there was no 'd' in 'good'
'actually' was significantly reduced
There was no separate dental fricative in 'the'
The fricative at the beginning of 'though' was pronounced as a dental stop

The surprising result was that Poles were much better at recognition of this phrase than the Greeks, nearly equalling the performance of English native speakers. To explain this discrepancy, we reasoned that Polish could have phonological strategies in common with English, because like English it is a language with a potential for complex consonant clusters, even though it differs prosodically. We postulated that, in accordance with the principles of Natural Phonology (Stampe, 1972), there would be a tendency to reduce complexity. Despite assurances from Polish speakers (not linguists) that they always pronounced their language exactly as written, we embarked on a pilot study of Polish casual speech.

Approximately 3 minutes of speech were recorded from three Polish radio talk programmes. The speakers included both males and females. The speech was casual and unguarded.

The two authors LS and MĆ looked at the excerpts independently; LS did a relatively fine-grained phonetic transcription, MĆ (a native speaker of Polish) produced a phonemic transcription. Both LS and MĆ looked at acoustic displays (amplitude waveforms and spectrograms) while transcribing.

Several notable casual speech shortcuts found in both recordings:

Vowel compression

słowa od	/swova od/	[swovod]
Polsce oni	/polstse oni/	[polstsoni]

Approximant compression

czy już	/tʃɨ juʒ/	[tʃuʒ]
dawno ja studio	/davno ja studjo/	[davnestudjo]

Vowel devoicing

klaps	/klɑps/	[kl̥ɑ̥ps]
przeprowadzam	/pʃeprovɑdzɑm/	[pʃ̥ęp....]

Vowel loss

teraz	/terɑz/	[terz]
tradycyjnej	/trɑdɨtsɨjnej/	[trɑdzɨnej] (twice)
to na tym	/to nɑ tɨm/	[tnɑtɨm]

Consonant Loss

wszystkiego	/fʂɨstcego/	[fʂɨscego]
tradycyjnej	/trɑdɨtsɨjnej/	[trɑdzinej] (twice)

Epenthesis

dwóch	/dvux/	[dəvux]
dni	/dɲi/	[dəɲi]
też w Polsce	/tɛʃ f polstse/	[tɛʒ ɛf ...]

While these reductions are not identical to those found in English, they result in an equivalent degree of phonetic variability.

With the caveat that this is only a preliminary study, we suggest that due to having to cope with phonological reduction, native speakers of Polish develop a perceptual strategy similar to that of native speakers of English: both groups are "eagles".

A study by Barry and Andreeva (2001) cites one example of cluster simplification in Greek, but as few clusters arise in the phonotaxis of the language there is understandingly no mention of other similar reductions. Nicolaidis (2001) describes the articulation of casual Greek based on electropalatography. She notes considerable variation in degree of achievement of canonical articulations for consonants and cites a small number of cases where consonants show no contact (mostly in intervocalic position) and two where entire syllables appear to be lost.

It is possible that native speakers of Greek do not cope as well with phonological reduction in English because their language does not incorporate many shortcuts. They therefore "hug the phonetic ground" more closely as patterned variation is not anticipated in the input: they are "roadrunners".

Of course, there could easily be a cline between the two extremes if, indeed, they prove to be valid at all.

3. Conclusion

Obviously, much more research along these lines is called for before firmer ground can be reached, we are building a theory on a very small amount of data. But based on our results so far, we suggest that it is not just native phonetic inventory and canonical phonotactics which aid or hinder perception in subsequent languages learned, it is also phonological **strategies**. Casual speech phonology is a crucial part of these strategies for learners of English, and we suggest that it is even more important for students whose L1 is low in this type of variability, such as Chinese. Support for this notion comes from gating results for 16 native speakers of Hong Kong Cantonese, all young women studying to be teachers of English who had achieved a high score on an English proficiency test, and none of whom correctly parsed a reduced English utterance, largely due to lack of knowledge of shortcuts (Shockey 2003, 121). Complex consonant clusters may be challenging and reducible (unstable) consonant clusters even more so. This may point to the conclusion that the study of casual English phonology is more important for speakers of languages with a marked tendency towards CV syllables than for native speakers of Germanic or Slavic languages.

Based on this notion, we are engaged in further research to see whether perception of gated casual speech bears out our categorisation of "eagles" and "roadrunners". Among our predictions are:

1) L1 Polish speakers will perform well at perceiving gated Polish with casual speech reductions (equivalent to L1 English speakers perceiving gated English casual speech).
2) L1 speakers of Spanish will be equivalent to L1 speakers of Greek at perceiving gated English casual speech.
3) L1 speakers of Catalan will be better than speakers of Spanish at perceiving gated casual English.
4) L1 speakers of Latvian (a Balto-Slavic language) will perform at the same level as speakers of Polish at perceiving gated casual English.

2,3, and 4 assume that subjects have achieved an equal level of instruction in English, which may prove the most difficult variable to control.

Acknowledgments
Thanks to Professor Andrzej Porzuczek and Professor Jolanta Szpyra-Kozłowska for comments on our Polish phonology

References

Bard, E. G., Shillcock, R. C. and Altmann, G. T. M. 1988. The recognition of words after their acoustic offsets in spontaneous speech: effects of subsequent context. *Perception and Psychophysics*, 44, 395–408.

Barry, W. J. and Bistra, A. 2001. Cross-language similarities and differences in spontaneous speech patterns, *Journal of the International Phonetic Association* 31/1, 51-66.

Grosjean, F. 1980. Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics,*11, 267-283.

Nicolaidis, K. 2001. An electropalatographic study of Greek spontaneous speech, *Journal of the International Phonetic Association* 31/1, 67-86.

Shockey, L. 2003. *Sound Patterns of Spoken English*, Blackwell.

Shockey L. and Bond, Z. 2012. Holistic perception of phonological variants (Holistyczna percepcja wariantów fonologicznych) in *Rhythm, melody and harmony in speech. Studies in honour of Wiktor Jassem.* Speech and Language Technology. Vol 14/15. Poznań, Poland, 199-209.

Stampe, D. 1979. *A Dissertation on Natural Phonology*. Chicago University Press.