The Pronunciation of English by Speakers of Other Languages

The Pronunciation of English by Speakers of Other Languages

Edited by

Jan Volín and Radek Skarnitzl

The Pronunciation of English by Speakers of Other Languages

Edited by Jan Volín and Radek Skarnitzl

This book first published 2018

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2018 by Jan Volín, Radek Skarnitzl and contributors

TABLE OF CONTENTS

INTRODUCTION

FOREIGN ACCENTS AND ENGLISH IN INTERNATIONAL CONTEXTS

JAN VOLÍN AND RADEK SKARNITZL

The title of the book in which you are presently taking interest comprises notions of "pronunciation", "English", and "speakers of other languages". All three concepts deserve a little comment that can make it easier to understand the contents and the general objective of the book.

The narrow meaning of the word "pronunciation" refers to the articulation of speech sounds like [iː, ɒ, s, ɡ, m]. This popular use of the word can be a bit misleading in the field of our research. The small speech segments are actually only descriptive tools reflecting what we recognize when we consciously observe and analyze spoken texts. They are most probably neither the true building blocks of the phrases that we utter, nor the templates we use for decoding spoken messages. Our understanding of the word pronunciation encompasses the production of stronger and weaker syllables (including their segmental features), melodies and temporal or amplitudinal attributes that make speech real in the psychological and neurophysiological sense.

An interesting justification of this stance is the term *accent* itself. It is primarily motivated by the prosodic phenomenon referring to the manifestation and distribution of prominences in the speech continuum. Groups of people share certain specific speech production features that are recognized by other groups of people as typical of the observed group. The fact that these features fall under an umbrella term of *accent* and not *vowelism or *phonemia acknowledge the importance of the wider approach to speech.

In the same vein, the terms phonetics or the adjective phonetic will refer to the entire complex sound structure of speech. We find expressions like, for instance, *phonetics and intonation* ridiculous (a simple Google search for this exact phrase returned 41,600 results). For a thoughtful

phonetician, intonation belongs to the domain of phonetics. (Coordination of the two terms is analogous to *food and apples* or *animals and rabbits*.)

English is currently the language of international communication and there are various theories why this has happened. Instead of speculating about languages that might take over, we build on the fact that hundreds of millions of people learn English as something serviceable, something they would like to master. To many, English is not the mother tongue. Those are the *speakers of other languages* in our title. The sound of their English is influenced by the sound structures of languages they have learned beforehand. We find these variations in the sound of English fascinating and for many practical reasons beneficial to explore and exhibit.

The work on this book started during the final stages of the 4th international conference *English Pronunciation – Issues and Practices*, which we organized in Prague in May 2015. More than seventy participants from four continents with 52 presentations manifested unusual dedication to research in the field. We realized that besides the proceedings on a CD we should invite some of the most dedicated researchers to expand on their topics and write a book chapter that would allow for sharing their findings with wider audiences. The peer-review process eliminated a few contributions and helped to improve the rest of them. The result is enclosed in this book.

It consists of four parts. In Part 1 broader, more general considerations of foreign-accented speech are exposed together with analyses of learner beliefs and attitudes to pronunciation instruction. Polish and Finnish learning environment is used to demonstrate certain issues. The second part brings several accounts of consonantal and vocalic phenomena demonstrated on Czech, German, Korean and Portuguese accents of English. Part 3 complements the preceding chapters with questions of speech prosody and adds Vietnamese and French-accented English. The fourth and final part considers methodological aspects of English pronunciation learning and offers inspiring tips for classroom management, testing pronunciation skills and enhancement of the learning process.

We would like to thank all the authors for their disciplined and responsible attitude to the book preparation and the staff of the publishing house for their professional, friendly and helpful approach.

Jan Volín & Radek Skarnitzl

PART I:

PERSPECTIVES ON 'ACCENTED' SPEECH

CHAPTER ONE

FOREIGN ACCENTS
AND RESPONSIBLE RESEARCH

JAN VOLÍN[1]

1.1 Introduction

In this chapter I am going to argue that the impact of foreign accents is far from trivial and when dealing with them, whether in discussions or in actions, we should do our utmost to avoid approaches based on ideologies or wishful thinking. Current science offers an alternative to these futile approaches. It provides adequate equipment to allow for a thorough exploration of the true nature of psychological and social consequences of accented speech. We are capable of searching for information that will substantiate useful practices both in classrooms and outside school settings. It is still prudent to acknowledge, however, that without genuine interest and impartial stances we will hardly succeed.

1.2 Controversies of variation and standard

It is a well-known fact that no two people in the world speak exactly the same way. Just as every individual exhibits a unique appearance, he or she also displays a unique manner of speaking. However, once we start studying these unique speech production patterns systematically, we realize that it is not a disarray of unpredictable idiosyncrasies. The pronunciation patterns can be grouped according to various similarities into accents. (Pronunciation is understood in its wider sense, i.e., including prosodic features – see Introduction to this book.)

In the past, people would recognize speech production peculiarities of the neighbouring villagers and they would commonly express such awareness in teasing, which often took form of fixed mocking phrases. The differing element would be exaggerated in these phrases or, metaphorically, magnified so as to create a caricature. The ancient idea that "we" speak correctly and whoever differs is incorrect is a deep-rooted one. Even open-minded individuals with the gift of tolerance, who embrace variation in the world and do not link strange to bad, still need to belong somewhere. The psychological need to be part of a community is innately human (e.g., Maslow, 1943), and speech provides quite a handy way to manifest belonging to or distancing oneself from various human communities (more about this in the following section).

The Anglophone world is quite rich in accents – but this is not necessarily due to its geographical vastness. Australia with its 7 692 000 square kilometres is more than thirty times larger than the United Kingdom, yet the accent variation is much richer in England, Scotland, Wales and Northern Ireland. For accents to develop, specific suitable social conditions must exist, and history does not provide those quickly and easily. The current increase in the mobility of human population makes geographical factors relatively weaker, but other factors of variation such as socio-economic status, ethnicity, gender or age still lead to differentiation among speech patterns of specific social groups.

The very special situation of English among other languages stems from its status of a widespread international language. It is often pointed out that the Earth now accommodates more speakers of English whose mother tongues are different from English than those who were born into English speaking families, i.e., native speakers of English. A natural consequence of this is that there are more speakers who produce foreign-accented English than those who display native pronunciation patterns. Obviously, this division builds on the standpoint of those who might want to claim the language, i.e., people who heard and learnt their first words from their English speaking parents. They somehow feel a stronger bond to the language and some of them even the right to fashion its fate. Although at times this feeling is labelled as undemocratic, it might prove difficult to change from the orators' pulpits.

Another current controversy that emerges relatively often when people discuss pronunciation variants of a language is the concept of the *standard*. In a wider sense the adjective *standard* can mean *typical* or *generally accepted* and the noun usually stands for a set of rules that describe some sort of a typical or generally accepted effect of human activities. (The gold standard in economy differs, but as a system for

setting the value of currency, it too has to be generally valid to be functional.)

The idea of the *standard form of a language* can be quite easily used to harass users of forms that are classified as non-standard. Lippi-Green, for instance, speaks of standard language ideology (SLI) and complains that the set of rules for the standard language use is based on the speech of upper-middle classes (Lippi-Green, 2012: 67). She is understandably bitter about speakers being bullied due to the use of various subordinated language forms, yet she seems to be a little too passionate in rejecting the whole concept of standard. While it is possible that she is correct in suggesting that the language standard is a collective delusion and should be abandoned altogether, it might be worthwhile to spend some time thinking of how standards are generated and what roles they fulfil. It is clearly one thing to disagree with social injustice, but a completely different thing to refuse to discuss it impartially so that all its important aspects can be considered. (It will be stressed repeatedly in this chapter that impartiality provides better chances to discover useful facts than partisanship.)

It can be universally observed that communities of speakers have appreciation for certain ways of speaking, or with regard to the narrower focus of the present book, certain pronunciation patterns. People cannot be prevented from evaluating the sound of someone's speech. Whenever something is said, a process of assessment is triggered on the part of the listener not only as to the contents of the utterance, but also as to its form. This process cannot be disabled at will. We are constantly trained to pay attention to the form, as it may contain important signals which modify and sometimes even invert the representative meanings of the words spoken. Furthermore, phonetic forms also signal the membership of the speaker in a group and, potentially, his or her attitude towards this membership (the desire to enhance or subdue it). It should be remembered that there are as many standards as there are groups of speakers, and one speaker normally belongs to more than one group. Countless groupings of speakers unconsciously produce the norms (very often tacit ones) that are observed and served. Norms are an inseparable part of the human social conduct.

However, if a standard form of an entire language is considered, what is usually meant is a widely understood and accepted variant that tends to signal some sort of detachment. Such a variant may offer the speaker the option to stay non-aligned, not linked to fractions, but appear somehow above ordinary groupings that one enters when solving problems of daily life. And just as the entire language community could be a complex

dynamic aggregate of numerous social groups, so is the "language standard" an ever-changing construct that cannot be described in its entirety. Educators may be unhappy at times not to have a complete and reliable description of the standard, but it has to be stressed that the language community as a self-organizing system always has enough (if not a vast majority of) mature members who do not have to be guided as to what social and language norms match individual situations.

When speakers opt for standard forms they do not express their admiration or allegiances to upper-middle classes. In many cultures the alignment is with education, politeness, carefulness or formality. (In addition to that, Milroy and Milroy (1992) discuss the link between the standard and prestige.) Various social norms exist that guide the speaker as to when to use the standard. If a school teacher instructs the children to use standard forms *all the time*, then he or she is obviously denying the existence of social norms. However, if a propagandist declares a war on the concept of the standard, then he or she makes the same blunder as the afore-mentioned school teacher.

It is clear from the propositions above that the standard is a concept based on relatively vaguely shared ideas. (The vagueness refers to the difficulty to capture the system through conscious analytical thinking and reach consensus with others.) When language teachers want to offer their students the option of generally accepted pronunciation norms, they may either use a published prescription and hope it is in reasonable harmony with reality, or rely on their own intuitions. From a researcher's point of view, the attempts to capture the standard in English lead up and down slippery paths. Hypothetically, the guidance could be provided by a consensus about the level of oddity that the listeners perceive when they encounter a given form. It could be argued that if a certain pronunciation form passes unnoticed by an absolute majority of the speakers, then it is accepted in terms of standard pronunciation. Empirical validation of such consensus, however, would definitely require many well-designed projects.

1.3 Neurophysiological and psychological background

In its psychological nature speech is behaviour. In an individual instantiation it comprises a set of actions with a communicative purpose. The previous section mentions a strong psychological need to belong. Groups of people who belong together regulate behaviour of their members by social norms, and speech behaviour is no exception to that. Perhaps, social norms should not be viewed separately from the speech-

related norms. Even a very cautious view should include the notion that speech decoding norms are a subset of social norms, and a bold view would argue that both sets of norms are very tightly intertwined and often inseparable. This holds not only for speech production but also for speech perception norms, i.e., those that regulate the way speakers decipher spoken messages. Implicit learning of how to understand certain phonetic features of an utterance takes place with great intensity in the first years of our lives and continues throughout the lifespan.

Another strong psychological need of an individual is the need to be respected or relevant. Unless seriously affected by an autistic disorder, a human being talks primarily not to pass information, but passes information to be relevant to others. To achieve this, not only the surface contents of the utterance must resonate with the listener's interests, but the form must be such that the listener understands. It must be emphasised especially in the context of foreign-accented speech that messages which are barely intelligible will not lead to the desired outcome. We might go even further and suggest that to be respected or relevant the speaker needs to be comfortably intelligible. It is a common experience of many people that without comfortable intelligibility frustration or irritation takes place. The listener (sometimes without knowing why) may grow increasingly uneasy or impatient, and may wish to reduce social contact with the speaker whose speech is taxing. Below I am offering a neurophysiological explanation of this.

The third psychological need of healthy humans is the need to be free. However, social foundations of our lives make freedom a complicated concept. For centuries, philosophers have been analysing the links between freedom and accountability, and even without philosophical training many people will ask: "Free – at whose expense?" This question is strongly phonetically pertinent. If we decide to free ourselves from pronunciation norms, then it is the listener who will have to pay the dues. It is a well-known principle observable at various levels of speech behaviour. The speaker's economy of effort has to be counter-balanced by the labour on the part of the listener. This principle is reflected in many cultures in the link between careful pronunciation and politeness. In other words, to display a polite stance the speaker may choose to exert greater effort so that their speech is clear. Conversely, to signal disregard one may merely choose to reduce the articulatory gestures including the pitch range and loudness.

From the discussion of the three selected psychological needs in the preceding paragraphs it is clear that attention to appropriate norms is strongly recommended. The consequences of not doing so can be quite

damaging. Yet, there is another severely damaging phenomenon that has to be mentioned in connection with accented speech and its psychological aspects. It is an unfounded negative attitude to a group of people called a *prejudice*. As Gordon Allport, the icon in the field of prejudice research, showed more than half a century ago, it is an extremely powerful phenomenon affecting at some point virtually anyone's life (Allport, 1954).

Why is prejudice – an attitude formed without much (sometimes without any) experience – so prevalent in our lives? It is a product of a cognitive process called stereotyping. Human lives require constant decision making and under certain conditions the decisions are a crucial factor in terms of survival. If the decisions have to be quick (in dangerous situations), there is no time to gather arguments for this or that choice. The individual has to draw on the past experience even if it is miniscule or indirect, i.e., based on what someone else reported. It seems that it is evolutionarily advantageous to decide, even if the decision is wrong, rather than stay passive. In the greater scale of events not doing anything means a smaller probability to succeed than acting, albeit sometimes erroneously. The ability of stereotyping developed to help humans survive.

Unfortunately, as the dangerous situations requiring fast decisions become scarce with the economic progress of human society, the positive aspect of stereotyping loses its relative dominance, while the negative aspect stays and grows. Our everyday lives are nowadays organized in such a manner that we typically do have time to collect evidence and contemplate the substance of problems. Yet the old cognitive mechanism seems to be better disposed to perform. Thus, hundreds of studies show that foreigners or minor ethnic groups are perceived with suspicion or worse. Allport's classic experiment with accommodation quest, when the same male was sometimes welcomed, sometimes refused over the telephone, based only on the surname he used to introduce himself (typical English names like Jefferson or Whitney meant success, Jewish-sounding ones like Rosenfield or Silverstein led to refusal), has been adapted to the phonetic domain repeatedly (e.g., Lambert et al., 1965; Aronovitch, 1976; Brennan & Brennan, 1981; Purnell, Idsardi & Baugh, 1999; Campbell-Kibler, 2007).

Prejudice should not be confused with the assessments that are produced by our fast adaptive brain mechanisms. Ambady and Rosenthal (1993) report that after just a thirty-second observation of a videoed lecturer (with the sound track off), students were able to produce estimates of their teaching quality that significantly correlated with evaluations of other students who were actually taught by those lecturers for a semester.

In other words, a very brief and incomplete observation may not differ from a long-term and a relatively thorough one. The quick unconscious processes may produce outcomes that are useful or misleading. Listening to low-pass filtered speech (i.e., speech where words cannot be recognized but rhythm and intonation can be perceived) for less than a minute produced assessments of medical doctors that significantly correlated with the number of law suits filed against them. Again, some unknown detail in the tone of voice leads to non-random estimates of the individual's personality (Ambady et al., 2002).

In a sense related to both prejudice and fast adaptive thinking are the results of Rubin (1992) and Niedzielski (1999). The former showed that even if two groups of respondents listened to the same speech sample, their memory retention and comprehensibility judgements differed. The only manipulated element in the experimental situation was a photograph of the alleged speaker that the respondents watched while listening. The latter achieved a similar effect by just orally informing respondents about the origins of the speaker. (Niedzielski did not measure memory retention, though – she was interested in the evaluation of differences between vowels.)

To explain the neurophysiological foundation of the effects of foreign-accented speech, it is useful to consider a few long known facts about speech perception. First of all, as Liberman and his colleagues already proposed in their Motor Theory of Speech Perception, the listener is not passive when a chain of acoustic events is unfolding in front of his ears (Liberman et al., 1967 or a revised account in Liberman & Mattingly, 1985). There is evidence that quite intense cerebral processes are going on and they comprise much more than just observation of the incoming acoustic signal and decoding the meaning from it. Our brain performs some sort of hidden mimicking of the articulatory gestures that produced the speech which is coming in through the ears. While listening to someone, we, to some extent, imitate that person's articulatory gestures by our neural production circuitry, except the outcome is not sent to the muscles so no external movements are executed.

Apparently, our brain can easily imitate only familiar gestures. If unusual muscular manoeuvres are used, the listeners struggle with the incoming speech signal. The term unusual pertains to timing and temporal alignment as well. Huggins (1979) reported his perceptual experiments with temporal patterns in speech and their impact on intelligibility. Standard rhythmic configurations of syllables positively influenced recognition of words in comparison with non-standard ones, which generated a considerably high word error rate. (Huggins himself, however,

speculated that this effect could be explained by malfunctions of short term memory: ibid. p. 283). Similar experiments a few years later already explicitly refer to the rhythm of speech, i.e., distribution of prominence contrasts in time (Buxton, 1983). Reaction time measurements revealed that ordinary rhythmic patterns lead to easier cerebral processing while the distorted ones require greater cerebral effort. Since then, numerous modifications or replications of these experiments confirmed that unusual acoustic patterns in speech activate additional cognitive resources, which may take its toll in areas such as attention or working memory (recently, e.g., Van Engen & Peelle, 2014).

The physiological framework of the phenomena observed in the experiments cited above is outlined in the Adaptive Resonance Theory (Pollen, 1999; Carpenter & Grossberg, 2002; Grossberg, 2003; Amis & Carpenter, 2009). Adaptive resonance is a general cognitive mechanism concerning recognition of visual or auditory objects and learning. An object in our field of interest could be a spoken word, morpheme, syllable etc. The mechanism builds on powerful predictive activities performed by our neurons. In perceiving speech, the incoming acoustic signal pre-processed by the auditory cortex meets with the expectational neural representations. Those are generated by the predictive activities and they are based on our experience. Our brain knows the language and various contexts well enough to produce quite accurate expectational representations. When these meet with the incoming signal and they reasonably match with it, an act of neural resonance takes place. It is a moment of synchronized activities of the neurons involved at the "meeting point". That is also the instant of object recognition, i.e., the moment when a word (morpheme, syllable) is recognized.

A lay person might wonder why neurons waste energy on prediction if they could just wait for what is coming in and compare it with some stored templates. The answer is simple. There is no extra energy spent on predictions. They happen due to the very nature of neurons. These little cells can only switch on and off, but the fascinating thing is that they do it in dependence on each other. They form associations and these help them map the structure of the language (or any system) and compute probabilities that, for instance, with this sound that one is most probable to occur, and with this word that one is likely to collocate, and after this idea that one should follow. So even if we do not know the topic that the speaker is discussing, based on our knowledge of the language and the world we still manage to be slightly ahead with our perception. Just confirming the predictions appears to be a more economical (and practical) way of perception than waiting with an empty plate.

How does this mechanism enter the debate on foreign accents? There is one crucial condition for smooth perceptual flow (i.e., for shorter reaction times, fewer errors – see above). It is the synchronization of the two streams of neural activity. The incoming and the expectational signals must meet at the right time. The brain is sending the expectational representations so that the timing of their encounter with the incoming imprints of reality is optimal. However, the timing in foreign-accented speech is not necessarily predictable and the neurons that are involved may have to repeat their activities to engage the reciprocal assembly. Without neural resonance an object is not recognized and repair mechanisms make the process of speech perception less economical. It is quite possible that irritation or unease on the part of the listener unfamiliar with the given type of non-standard patterning in speech is the consequence of extra labour his or her neural assemblies have to perform. In connection with this it is also interesting to consider the results of the study of Volín, Poesová and Skarnitzl (2014 – see below in Section 1.5).

1.4 Social significance and current approaches

Allusions or explicit references to accented speech can be found in very old scripts. The Old Testament (specifically The Book of Judges, Chapter 12) comprises a testimony of undercover invading soldiers being recognized and punished thanks to non-native pronunciation. It was the confusion of alveolar and post-alveolar voiceless fricative – a feature found regularly in Dutch, Finnish or Spanish accents of English – that allegedly cost thousands of lives. The importance of clear speech and its impact on listeners is mentioned by many ancient philosophers, orators, politicians. References to coaches giving paid lessons in "good speaking" and to the legendary self-taught Demosthenes (4th cent. BC) are illustrations of the awareness of the value attributed to the sound of speech. Until today, certain speaking styles are in many cultures linked to education, which is viewed as a prestigious commodity.

The social significance of pronunciation is also reflected in the modern empirical approach to speech, which dates back several centuries (e.g., von Kempelen, 1791). A prominent signpost in the history of social attention to speech is the foundation of the International Phonetic Association in 1886, in which teachers of foreign languages were heavily involved. It is useful to remember that the teachers themselves demanded scientific foundation for their work.

Yet unfortunately, in today's social debate on foreign accented speech, two extreme stances can be heard more often than anything. On the one

hand, there is the authoritative prescriptionism, and on the other hand, there is the naïve liberalism. The prescriptionists argue that there are "ways thing should be" and learners of languages are obliged to sound "proper". The trouble with identifying the proper is mentioned in Section 1.1 above. A compelling claim voiced by the prescriptionist camp concerns the social advantage that competent speakers possess. This claim, however, easily translates into not succeeding in life without decent pronunciation, and leads to anxiety of failure in some learners or the feeling of exclusion from the society (Gluszek & Dovidio, 2010; Gluszek, Newheiser & Dovidio, 2011). It would be very helpful to be able to measure objectively to what extent such feelings really motivate the individuals, or, conversely, impede their true potentials.

The other extreme is represented by claims that the importance of one's sound is grossly inflated and if we recognize that, we will be liberated. The underlying belief could be expressed by an infamous quote of the wife of a former top U.S. politician. When asked about drug addiction, she advised the affected subjects: "Just say no". A similar solution to the impact of accented speech is proposed by naïve liberals. They offer a simplistic impractical stance arguing by an individual's right to choose. Although very appealing in the political sense, the claims of naïve liberals are not supported by any rigorous, representative research. They, too, are most probably hugely demotivating in the learning environment.

These two camps are similar to football fans or to fervent supporters of competing political parties. Although their members are mostly educated people, they too have a strong, and in some cases even blinding desire to win. Unfortunately, they do not realize that the ability to see clearly and objectively is not under the conscious control of an individual. We do not see objectively at will. Once we accept a certain opinion as correct, we find it very difficult not to overlook counter-evidence, even if we do not want to cheat. Our mental capacities cannot be ordered to see the entirety clearly and accurately. In the field of academic research, however, we can help them by refusing to decide ahead what we want to find (or in its softer version – what the truth most probably is).

1.5 Responsible research

The number of researchers who are ready to carry out respectable experiments and observations in the field of foreign accents is growing. An indirect indicator could be seen in the number of conferences and journals dedicated to the topic. Although they do not necessarily testify

about the reputable scientific methods used in the field, they hint that societies are willing to spend money on this research, i.e., they consider it useful.

More direct proofs of the scientific prowess of researchers in the field can be found inside the dedicated journals and other publications. The methods used are compatible (or even identical) with those used in psychology, sociology or experimental linguistics – more traditional disciplines with an acknowledged impact on the development of science as a whole.

Apart from constantly scrutinized methodology, one of the features of current research work is prudency when it comes to strong claims. The availability of information which is typical of the present era (connected with the internet and relatively low costs of printing) makes scientists realize more than ever before that one single individual can hardly apprehend the complex phenomena of the world. The capacity of human cognitive mechanisms is quite limited, and without the collective effort we cannot hope to appreciate the complexities of natural or social structures. In today's setting, strong claims are linked to immaturity rather than boldness. If, for instance, the pressure to standardize language forms across language communities is called unjust (or malicious, evil, etc.), then we can use the *personal social network analysis* (PSNA) to find out what it actually means in a less abstract, but more expedient manner.

PSNA was introduced by social anthropologists (see, e.g., Mitchell, 1986), but soon attracted the attention of sociolinguists since it allowed to explain specific choices of speaking individuals rather than cumulate metrics of large groups. One of the important findings was the link between the denseness and tightness of the social networks and the willingness to accept influences from outside. Hence, communities with looser and not so numerous interpersonal ties may adopt standard features more easily than communities with closely knit ties. However, what is also clear is that the freedom not to standardize is traded for strict obligations to the group. An individual *must* rather than *can* observe the norms of the community and may be constrained in multiple aspects of personal life. Apparently, gang behaviour is not a matter of simple choice. It follows that fighting too fiercely for or against standardization means a risk of great harm, while thorough research of the problem can only help to avoid unpleasant (and sometimes even disastrous) consequences. (Too many people acquired bitter experience due to promises that they will all live as one nice happy family if they accept "historical inevitability". Ignoring psychological and social dispositions of humans has brought immense suffering in the name of "justice and order".)

A popular belief in the domain of foreign accents links all the perceptual difficulties with prejudice. As I have demonstrated in Section 1.2 above, there is growing evidence that most of the effects are not prejudice induced. One more example deserves to be mentioned. Lev-Ari and Keysar tested the impact of accentedness on the perceived truthfulness of statements (Lev-Ari & Keysar, 2010). They found that people attribute lower truthfulness scores to propositions that are spoken with a foreign accent. Although it would be quite easy to blame the results on prejudice, they carried out a follow-up study in which they explicitly urged the respondents to guard themselves against the foreign accent effects. The results suggest that, indeed, it is not necessarily prejudice, but rather "fluency effects". (Fluency is used in a metaphorical sense as the potential ease of processing – see, e.g., Oppenheimer, 2008.)

Results that point in the similar direction were obtained in the study by Volín, Poesová & Skarnitzl (2014). They extracted a number of longer statements from various English radio programmes and for each of them created a duplicate in which stressed vowels were artificially shortened and unstressed reduced vowels lengthened. The change was barely noticeable. Adult listeners were asked to try to judge the personality of the speaker from his or her voice. All guiding questions were linked to the emotional stability of the speaker and the estimates were implemented on a seven-point scale. A statistically significant outcome indicated that speakers with "less usual" rhythmic flow of vowels are perceived as more neurotic, irascible or less emotionally stable. It has to be emphasised that the durational manipulations did not model any specific foreign accent so, again, it would be difficult to argue that the listeners were prejudiced. Yet equalizing durations of stressed and unstressed syllables has been reported from many foreign accents of English.

Even in the domain of speech production, where all effects tend to be habitually attributed to the influence of the mother tongue, examples can be found which invite a re-evaluation of our approaches. For instance, Volín, Poesová & Weingartová (2015) found that certain features of F0 tracks (correlates of intonation) in Czech-accented English do not lie midway between Czech and English. They were found outside the space dividing Czech and English speech melodies. Rather than to the interference of the two prosodic systems the results seem to point toward certain "dysfluency" in the target language. (The use of "dysfluency" is again in the sense of Oppenheimer, 2008.) It is not to say that speakers of Czech English stuttered, stopped abruptly or re-started their sentences – they read them from a script with which they familiarized themselves ahead of reading. Most probably it was their implicit uncertainty about the

language, some sort of subliminal hesitation or hidden feeling of lower competence that prevented them from producing more vivid forms (i.e., those that exhibit ease of speaking). The intonations disclosed this even though the affective charge of the situation was not explicit.

All the examples above suggest that rather than being overly concerned with correctness, our teaching methods should pay more attention to fluency. And again, the research has to show to what extent. It is only too easy to become a newly established educationalist and start a "reform" that introduces a new extreme against an old one. In my lengthy career in schools, I have witnessed many such expensive exercises of newly gained power by officials in education, but the outcome has always been meager if any. Top-down reforms in state run schools principally offer a good opportunity to spend money. Apparently, there are individuals who believe that their reforms will improve the educational effectiveness, but since they do not test the new methods rigorously on impartially selected samples of students, and they do not care to disseminate the research outcomes among teachers in a comprehensible and trustworthy manner, the reforms make hardly any impact.

Responsible researchers are not biased, they are patient, diligent, and spend time and energy thinking of how to speak to teachers at schools, how to make them interested. This is because only interested teachers will care to implement the new training procedures. Without that they may travel to special seminars dedicated to the new method implementation and yet come back home unconvinced and unwilling to divert from their routines. For an extra bonus they may be willing to pretend they are supporters of the novelties, but this all usually fades away only to recur when a new cohort of attention seeking reformers is installed in the offices.

It is not easy to withstand the pressure from laymen on researchers to provide an unequivocal answer as to whether accents matter or not. Nevertheless, not knowing does not necessarily mean people should stay inactive. Everyone is invited to take a stand against injustice in individual cases (particularly those that one understands well). However, sweeping political measures should be discouraged for the time being, and a greater support for responsible research should be made available.

References

Allport, G. W. (1954). *The Nature of Prejudice*. Cambridge, Mass.: Addison-Wesley Publ. Co.

Ambady, N. & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 64, pp. 256–274.

Ambady, N., Laplante, D., Nguyen, T., Rosenthal, R., Chaumeton, N. & Levinson, W. (2002). Surgeons' tone of voice: A clue to malpractice history. *Surgery*, 132/1, pp. 5–9.

Amis, G. & Carpenter, G. A. (2009). Self-supervised ARTMAP. *Neural Networks*, 23, pp. 265–282.

Aronovitch, Ch. D. (1976). The voice of personality: Stereotyped judgements and their relation to voice quality and sex of speaker. *Journal of Social Psychology*, 99, pp. 207–220.

Brennan, E. M. & Brennan, J .S. (1981). Measurements of accent and attitude toward Mexican-American speech. *Journal of Psycholinguistic Research*, 10/5, pp. 487–501.

Buxton, H. (1983). Temporal predictability in the perception of English speech. In: Cutler, A. & Ladd, D. R. (Eds.), *Prosody: Models and Measurements*, pp. 111–121. Berlin: Springer-Verlag.

Campbell-Kibler, K. (2007). Accent, (ing) and the social logic of the listeners perception. *American Speech*, 82/1, pp. 32–64.

Carpenter G. A. & Grossberg, S. (2002). Adaptive resonance theory. In: Arbib, M. A. (Ed.), *The Handbook of Brain Theory and Neural Networks*, 2nd edition, 87–90. Cambridge, Mass.: MIT Press.

Gluszek, A. & Dovidio, J. F. (2010). Speaking with a non-native accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology*, 29, pp. 224–234.

Gluszek, A., Newheiser, A. K. & Dovidio, J. F. (2011). Social psychological orientations and accent strength. *Journal of Language and Social Psychology*, 30, pp. 28–45.

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31, pp. 423–445.

Huggins, A. W. F. (1979). Some effects on intelligibility of inappropriate temporal relations within speech units. In: *Proceedings of 9th ICPhS*, 283–289. Copenhagen: University of Copenhagen.

Kempelen, J. W. von (1791). *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine.* Wien: J. B. Degen.

Lambert, W., Anisfeld, M. & Yeni-Komshian, G. (1965). Evaluational reactions of Jewish and Arab adolescents to dialect and language variations. *Journal of Personality and Social Psychology*, 2/1, pp. 84–90.

Lev-Ari, S. & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology 46*, pp. 1093–1096.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74/6, pp. 431–461.

Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21/1, pp. 1–36.

Lippi-Green, R. (2012). *English with an accent: Language, ideology and discrimination in the United States*. London, UK: Routledge.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50/4, pp. 370–396.

Milroy, L. & Milroy, J. (1992). Social network and social class: Toward an integrated sociolinguistic model. *Language in Society*, 21, pp. 1–26.

Mitchell, J. C. (1986). Network procedures. In: D. Frick & H.-W. Hoefert (Eds.), *The Quality of Urban Life*, 73–92. Berlin: de Gruyter.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18/1, pp. 62–85.

Oppenheimer, D. N. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, 12/6, pp. 237–241.

Pollen, D. A. (1999). On the neural correlates of visual perception. *Cerebral Cortex*, 9, pp. 4–19.

Purnell, T., Idsardi, W. & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18/1, pp. 10–30.

Rubin, D. L. (1992). Non-language factors affecting undergraduate's judgments of non-native English speaking teaching assistants. *Research in Higher Education*, 33/4, pp. 511–531.

Van Engen, K. J. & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers In Human Neuroscience*, 8, pp. 1–4.

Volín J., Poesová K. & Skarnitzl R. (2014). The impact of rhythmic distortions in speech on personality assessment. *Research in Language*, 12/3, pp. 209–216.

Volín, J., Poesová, K. & Weingartová, L. (2015). Speech melody properties in English, Czech and Czech English: Reference and interference. *Research in Language*, 13/1, pp. 107–123.

CHAPTER TWO

ENGLISH PRONUNCIATION INSTRUCTION THROUGH THE ACCENT LENS

EWA WANIEK-KLIMCZAK

2.1 Introduction

Accent is adopted in this chapter as a working perspective for the discussion of problems and possible solutions in pronunciation instruction. Referring to accents, the chapter explores the role of native and non-native accents in specifying the model and the target in teaching and learning English. The main aim is to show that while accents, not only native but also non-native, need to be accepted as an obvious reality of natural language use, the instructional setting requires clear guidelines as to which pronunciations used by the learners need to be corrected. Such guidelines can be offered by a linguistic phonetics approach which focuses on linguistically relevant categories as the basis for sound system formation and usage (Ladefoged, 1997). Extending the original view to non-native systems, pronunciation instruction can be seen as aiming at the formation of sound categories recognized as linguistically relevant across accents of English, or in a specific accent of English, depending on the choice and needs of the learners. However, the realization of these categories, i.e. the degree of 'native-like' pronunciation is expected to vary and to be accepted as characteristic for a given non-native accent. The specificity of this approach is in the use of phonetic parameters for the description and analysis of sound categories. Thus, unlike a traditional phoneme/allophone approach as represented by e.g., Gimson's *Pronunciation of English* (Gimson, 1962), rather than talking about abstract categories and their physical implementation, the linguistic phonetics based approach advocates the use of physically-real sound description as the basis of the model for learners. The model is defined then in terms of the range of values typical for specific native accent or accents; the range used by non-

native speakers is expected to vary, with the target defined as the production of parameter values sufficient to mark the contrast at the category level.

There are two major reasons for adopting an accent perspective: firstly, accents are of major importance in pronunciation instruction, and secondly, more importantly, all of us, native and non-native speakers of English, speak with an accent. In this sense, an accent, defined as a 'loose bundle of prosodic and segmental features distributed over geographic and/or social space' (Lippi-Green, 1997: 42) is seen as a natural consequence of language experience. The key difference between native and non-native accents results from the nature of previous language experience. Thus, a native language speaker's accent reflects their experience of the variability within what is defined as one language system (intra-language experience) while a non-native speaker's accent reflects the experience of the variability in more than one language (inter-language). The inter-language has been defined by Major (2001: 1) as 'An adult second language learner's linguistic system (…) or simply the language of a non-native speaker', which 'is a product of a combination of parts of the L1, parts of the L2, and universals' (ibid.: 4). Notice, however, that discussing accents we refer to the accents of a particular language. Consequently, while we can talk about the inter-language experience of a non-native speaker, the elements of language transfer and language universals, it is the system of the language the accent of which we describe or discuss that functions as the point of reference. Consequently, extending Lippi-Green's definition to non-native accents, in the analysis of the elements of those loose bundles of segmental and prosodic features found in the speech of non-native English speakers, we define and describe non-native accents from the perspective of the system of English adopted as a point of reference, typically RP or GA, not unlike Wells in his *Accents of English* (1982) or Collins and Mees (2008) in their pedagogical description of accents for learners. Notice also that all accents can be analyzed at different levels: segmental or suprasegmental, and what can be called sub-segmental, i.e., the level of phonetic parameters, or the level of sound-categories and their phonetic realization[1]. However, when we talk about accents, what is foregrounded is the use of language and a specific user or groups of users of the language, with the focus on their language experience and other characteristics, including language attitudes and individual differences.

---

[1] Presenting the most comprehensive description of English accents, Wells (1982) refers to systemic and realizational differences, supplemented by lexical-incidential and phonotactic.

Thus, at a descriptive level, accent is something we all have and something we can describe. It is an inevitable (expected, natural) element of language use. However, when an accent goes to a language school, it becomes not only 'good' or 'bad' but also 'right' or 'wrong', 'correct' or 'incorrect'. This is not surprising: in parallel to other elements of the language system, grammar and vocabulary, pronunciation instruction relies on a clear identification of what is expected from the learner and 'success' is measured by the degree to which correctness has been reached. This is true for native and non-native accents alike, however, in the case of non-native accents, it is not only the question of 'good' or 'bad' in terms of socially desirable or not, but also, or - as many teachers will agree, more importantly – whether it is intelligible or not. Consequently, it is for a long time now that intelligibility rather than correctness has been stressed in formulating aims for pronunciation instruction. In the discussion of pronunciation priorities for intelligibility, the issue of an accent seems to have been lost, or at least moved to the background; this paper hopes to bring it back to the foreground in order to demonstrate that a non-native accent perspective offers possible solutions to many problems and challenges that we face in pronunciation instruction.

The chapter divides into two major parts: in the first one, problems and contexts for pronunciation instruction and accent issues are explored in comparison to other elements of the system of English (Section 2.2), and then from the perspective of pronunciation teachers, pronunciation learners and finally, fully proficient pronunciation learners who become users of English (Section 2.3). Finally, a linguistic phonetic approach to non-native accents is proposed as a possible solution to the problem of accents in pronunciation instruction (Section 2.4). The data used for the discussion come predominantly from the Polish context, which is treated here as representative for problems and challenges shared across different places where English is taught as an additional language.

2.2 Problems

Over the years, while teaching the pronunciation of English to non-native speakers – in my case Polish learners and users of English – I have begun to realize that there are two major problems that I need to solve: putting it in simple words, the first one is what to teach, and the second one – what to reach. These two problems may seem trivial for anyone involved in a sister-system element of the language, i.e. grammar. From the perspective of someone trained in teaching English as a Foreign Language (but not a specialist in grammar instruction, who will have further insights into the

problem), it seems that when we teach grammar, we know that we need to teach grammatical structures, starting from simpler / more frequent to the more complex / less frequent ones (most of the time we follow the order of grammatical elements suggested by the textbook we happen to be using). And unless we start looking into language variability, we accept the descriptions of the grammatical system without much controversy about the need to use e.g. the verb 'to be' in progressive forms, 's' in third person singular verbs or single negation. In other words, we have a clearly defined point of reference with respect to 'what to teach'; but more than that, we seldom have any doubts as to 'what to reach'. As teachers then, whenever we hear a learner say 'she ride', 'I talking', or 'I didn't take nothing' we react to these forms as 'wrong' and tend to correct them at an appropriate point during the lesson. However, if the learner uses 't' rather than 'd' in 'ride', no aspiration in 'talking' or 'take', or a sound closer to 'f' than 'th' in the word 'nothing', our decision as to categorizing these as 'wrong' and correcting the learner may not be equally straightforward and obvious. In other words, even when we accept the description of Standard English as the reference point in pronunciation instruction, the decision as to the degree to which this model will be used in practice as the basis for correction is far from obvious. Thus, although there are many similarities between grammar and pronunciation instruction when discussed with reference to when to correct, the decision as to what to correct differs. There are many other parallels between grammar and pronunciation based on the crucial distinction between accuracy, system-based instruction and skill practice in teaching speaking. However, the questions 'what to teach' and 'what to reach' seem to be specifically true for pronunciation: it is in pronunciation instruction that the debate continues as to the model itself and the degree to which this model determines priorities and/or success in pronunciation learning. Problems and challenges related to these two crucial questions will be briefly considered in the following sections.

2.2.1 What to teach?

First the problem is 'what to teach'. On the face of it, the answer is easy: when teaching pronunciation, we teach the sound system of English. Well, if only we could stop there, we could accept a description of the system of English and teach our learners – or rather help them learn – this system. The native-speaker model used in pronunciation classes would correspond to the description and the learner would be expected to recognize sound categories and contrasts and to (re)produce them in her (re)production. The 're' element proves to be crucial here. Historically speaking, it was

with emphasis on communication and production rather than reproduction that this straightforward, well-structured approach to pronunciation started to be amended so as to incorporate the communicative needs of the learners. This led to a major shift, with 'comfortable intelligibility' (e.g. Kenworthy, 1987), and not native-like (or near-native) pronunciation treated as the goal in pronunciation instruction. Notice that unlike the previous approach, with the variability introduced by the term 'comfortable' as well as 'intelligibility' (with the question of intelligible for whom?), the communicative approach has not provided clear answers as to the 'what to teach' question, as it is clearly assumed that not all elements of the system are equally important and not all of them need to be taught or learned. However, which ones they are has not been experimentally determined and remains obscure. The only set of pronunciation priorities for non-native speakers formulated to-date was proposed by Jennifer Jenkins (2000, 2002) in her intuition and observation rather than research-based[2] Lingua Franca Core (LFC). One of the problems with LFC is that it sanctions, or indeed advocates 'mispronunciations', such as the lack of a dental fricative or weak forms, with the latter observed to hinder communication between non-native speakers (ibid.). The suggestion that certain elements of native English pronunciation can be abandoned in pronunciation teaching has divided teachers of English pronunciation into two camps: those who adhere to it and try to put it into practice (e.g. Walker, 2005) and those who oppose it on linguistic (e.g. Szpyra-Kozłowska, 2005) or logical (e.g. Sobkowiak, 2005) grounds. Not surprisingly, the lack of solid research and a somewhat patronizing approach behind LFC (interpreted as 'let them use bad English') has provoked a strong negative reaction among many non-native teachers of pronunciation, who feel that (1) LFC does not address problems of their learners (as voiced by e.g. Szpyra-Kozłowska, 2005) and, importantly, find it a frustrating way of rejecting their own experience of being able to overcome many pronunciation problems. Teachers of English pronunciation who are not native speakers of English tend to share with their learners not only L1, but also the learning experience; they may rightly feel that they know best how to teach pronunciation to the specific L1 learners.

As a compromise between ELF and a traditional native-speaker model, Szpyra-Kozłowska in her recent book (2015) advocates the idea of NELF, i.e. Native English as a Lingua Franca, originally proposed in a slightly

---

[2] I am grateful to an anonymous reviewer for noticing and bringing to my attention a possible misinterpretation of LFC for an experimentally determined set of priorities, which it definitely is not.

jocular manner[3] by van den Doel (2008). What they suggest is going back to a native speaker model, RP or GA (with preference for RP), while accepting the fact that non-native speakers may not reach the native-like pronunciation level. Thus the question 'what to teach' receives the answer 'native model' once again. However, the novelty of NELF is in the outcomes of pronunciation instruction, with learners not expected to master all aspects of a native phonetic model.

> The NELF approach is intuitively adopted by an overwhelming majority of EFL teachers who, being predominantly non-native speakers of English and speaking it with some degree of a foreign accent, in their pronunciation instruction take a native phonetic model as a reference point, but focus on selected features of English which they consider particularly important for successful communication. They do not require their students to master every phonetic detail as they know very well that in this respect perfection is impossible to achieve. (Szpyra-Kozłowska, 2015: 27).

The same sentiment seems to guide Trudgill, when he voices his cautious support for native English models:

> (…) the sensible, pragmatic course is to continue, as before, employing ENL models [….] with an understanding that in most cases phonetic accuracy is unlikely to be achieved. (Trudgill, 2005: 93).

The above mentioned understanding that 'accuracy' or 'perfection' is unlikely or even impossible to be achieved, sensible and practice-based as it may be, does not seem to help in deciding what it really is that phonetic instruction aims to achieve. It seems that adopting a native English model and then accepting (understanding) the fact that it will not be mastered leaves us with a huge grey area, where deciding what to teach, what to correct, and what to accept as 'correct' is left for teachers to decide. In other words, even if we solve the problem of 'what to teach', we are left with 'what to reach'.

2.2.2 What to reach?

Once again, the problem is true for pronunciation rather than grammar, where it is difficult to imagine a controversy as to whether to treat 'She ride' as less or more acceptable than 'I talking' or 'I didn't take nothing'. In the case of pronunciation, on the other hand, the reaction of the teacher

---

[3] van den Doel, personal communication.