RL Vol. 10.2 2012

Research in Language

edited by
EWA WANIEK-KLIMCZAK
IWONA WITCZAK-PLISIECKA
JAN MAJER

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

ŁÓDŹ 2012

REVIEWER
Linda Shockey

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. 6043/2012

ISSN 1731-7533

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

Contents

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0045-6

NEW WAYS OF ANALYSING THE HISTORY OF VARIETIES OF ENGLISH - AN ACOUSTIC ANALYSIS OF EARLY POP MUSIC RECORDINGS FROM GHANA

SEBASTIAN SCHMIDT
Justus Liebig University, Giessen

Abstract
Focusing on English in Ghana, this paper explores some ways in which early popular music recordings might be used to reconstruct the phonology of colonial and post-colonial Englishes in a situation where other recordings are (mostly) absent.

While the history of standard and, to a certain degree, non-standard varieties of "Inner Circle Englishes" (Kachru 1986) has received linguistic attention, diachronic investigations of Outer Circle varieties are still the exception. For the most part, descriptions of the history of post-colonial Englishes are restricted to sociohistorical outlines from a macro-sociolinguistic perspective with little if any reference to the linguistic structure of earlier stages of the varieties. One main reason for this lack of diachronic studies is the limited availability of authentic historical data. In contrast to spoken material, written sources are more readily available, since early travel accounts, diaries or memoirs of missionaries, traders and administrators often contain quotes and at times there are even documents produced by speakers of colonial Englishes themselves (cf. the diary of Antera Duke, a late 18th century Nigerian slave trader; Behrendt et al. 2010). Such material provides insights into the morphology, syntax and the lexicon of earlier stages of varieties of English (cf. Hickey 2010), but it is inadequate for the reconstruction of phonological systems. Obtaining spoken material, which permits phonological investigation, is far more difficult, since there are comparatively few early recordings of Outer Circle Englishes. In such cases, popular music recordings can fill the gap.

I will present first results of an acoustic analysis of Ghanaian "Highlife" songs from the 1950s to 1960s. My results show that vowel subsystems in the 1950s and 1960s show a different kind of variation than in present-day Ghanaian English. Particularly the STRUT lexical set is realized as /a, ɔ/ in the Highlife-corpus. Today, it is realized with three different vowels in Ghanaian English, /a, ɛ, ɔ/ (Huber 2004: 849). A particular emphasis will also be on the way *Praat* (Boersma and Weenink 2011) can be used to analyze music recordings.

1. Introduction

The present study is concerned with the structural development of World Englishes. Focusing on English in Ghana, the former British Gold Coast colony, this paper explores ways in which early popular music recordings might be used to reconstruct the phonology of colonial and post-colonial Englishes in a situation where other recordings are (mostly) absent. The database consists of early 20th century music recordings from

the former Gold Coast colony and the emerging independent post-colonial nation of Ghana. The recordings contain lyrics that can be regarded as authentic historical near-spoken data. This source as a linguistic database is made accessible, because recordings of colonial or early post-colonial Englishes are rare. In several pilot studies (cf. Huber & Schmidt 2011a and 2011b, Schmidt 2011a) the potentials and methodological challenges of using early popular music lyrics for the analysis of earlier stages of Outer Circle (Kachru 1986) varieties of English have been explored.

For the present study, a pilot corpus of popular music lyrics from 1950s Gold Coast Colony and then, from 1957 onwards, post-independence Ghana has been compiled. The actual recordings and the transcribed lyrics have been subject to an auditory (Huber & Schmidt 2011a) and an acoustic analysis (Schmidt 2011b). In both studies, the focus is on sound change and on the differences between RP and Ghanaian English (GhE). The motivation of the acoustic analysis is first of all to investigate if early music recordings can be used within the context of an acoustic study at all and secondly to what extent the results fit in with the findings generated by the auditory analysis. The present paper brings the two analyses together. It is structured into three major sections, 1.) background information on the data, 2.) a report on the methods and tools that were applied, and 3.) a discussion of the findings and an outlook.

2. Background and Data

Kreyer and Mukherjee (2007) worked quantitatively and qualitatively on the style of pop song lyrics in general. They also investigated vocabulary and lexicogrammatical routines. In order to do so, the authors compiled the *Giessen-Bonn Corpus of Popular Music* (GBoP). The GBoP consists of transcripts of popular music lyrics of various heterogenous genres, such as rock music and rap. Kreyer and Mukherjee's study is based on the GBoP and focuses on written language. They call for a systematic, corpus-based approach to popular music lyrics as linguistic data on all descriptive levels. The authors show that it is worthwhile working with popular music lyrics as a linguistic database. Furthermore, they suggest that a corpus-based approach to the study of the language in popular music should be preferred.

Miethaner (2005) uses the *BLUR*-corpus (*Blues Lyrics collected at the University of Regensburg*) to reconstruct earlier stages of African American English (AAE). By applying corpus linguistic methodology, *BLUR* turns out to be an appropriate and valid representation of earlier AAE. Miethaner demonstrates that blues lyrics can be used to reconstruct the morphology, morphosyntax and syntax of earlier AAE.

Trudgill (1983) diachronically investigates English pop-singers' pronunciation. Among others, rhoticity serves as one linguistic variable. He observes a trend to sing in an Americanized way in the 1950s and 60s but this trend is weakened at the latest with the advent of punk-rock in the late 1970s in favour of a local English pronunciation. By comparing several records of The Beatles and The Rolling Stones from 1963 until 1969, Trudgill emphasizes the diachronic perspective of his study. As a result, the author shows the importance of linguistic models and of identity in the context of the language used in popular music.

In this tradition, Brato and Jansen (2008) focus on both southern and northern English indie rock bands, such as The Arctic Monkeys and The Kooks. Conducting an auditory analysis of a selection of songs, they find that the bands they looked at are generally English in their pronunciation and even exhibit regional accent features, such as typically marked Sheffield English.

West African popular music lyrics have also been subject to linguistic analysis. Both Coester (1998) and Culver (2007 and 2008) show an interest in the language of the late Nigerian musician Fela Anikulapo Kuti. As Coester (1998) points out, Kuti's lyrics are characterized by an "intermingling of languages" (Coester 1998). The author shows that the language in Kuti's 1970s and 1980s lyrics alternates between Nigerian Pidgin (NigP), Standard English (StE) and Yoruba, sometimes even within a single song. These alternations are frequent and give distinction to Kuti's style and the genre Afro-Beat of which Kuti is regarded to be the founding-father.

The present study is a corpus-based, diachronic analysis of phonological details in early Ghanaian popular music. The selected lyrics stem from a genre called Highlife. Highlife is a form of dance-music of West African origin which was popular both with the white minority and the local population (Bender 1985 and 2007, Collins 1986 and 1989, Oti 2009).

The musicologist Collins has worked extensively on Highlife music (cf. Collins 1986 and 1989), which he considers an umbrella term for West African popular music that had its heyday around the time of Ghana's independence in 1957 (cf. Collins 1986 and 1989). According to Collins, Highlife is characterized by "fusion" on the levels of musical styles, cultures and languages (Collins 1989: 221). English and Pidgin lyrics represent only a fraction of Highlife songs which were recorded in a variety of languages. Some of the songs contain both local Ghanaian languages and English or Pidgin English. There are also Pidgin elements within otherwise StE-oriented songs. Hybridity in terms of stylistic and cultural diversity as well as language fusion are, from a linguistic point of view, the central characteristics of early West African popular music.

Crucial for the acoustic analysis is the recording situation: According to Collins (personal communication), the singer stood near or in front of the recording microphone. The band was placed behind him and thus further away from the microphone. The music was originally distributed on gramophone records. These were digitalized and stored as .wav-data. It was particularly paid attention not to alter the voice in any way. To sum up, the voice of the singers is generally 'in front of the music' so that the music can be treated as background noise when vocals are measured.

3. Linguistic Context

Ghanaian English is an Outer Circle variety of English (Kachru 1982, 1986), which was brought to the territory of modern Ghana through trading contacts and colonisation (Huber 1999, 2008). Huber and Schmidt (2011a) locate modern GhE between nativization and the endonormative stabilization stage in Schneider's evolutionary model (Schneider 2003, 2007). Currently, GhE is the "*de facto* official language" (Huber 2008: 72) in Ghana, because the status of English in Ghana is not specified in the constitution

of the country. Nevertheless, GhE is spoken in most public domains such as schools, the media and in parliament. In contrast to local languages, it "has the advantage of ethnic neutrality" (Huber 2008: 73), which is an important aspect in a multi-ethnic and multilingual region.

Based on a structural investigation of GhE, as conducted by Huber (2008), "it should be kept in mind that on all descriptive levels, GhE is a system of tendencies rather than categorical differences from the British standard" (Huber 2008: 74). Especially in the public domain, the British standard has overt prestige. Ghanaian speakers of English often claim to sound RP, while in fact speakers often favour a distinct Ghanaian pronunciation to dissociate themselves from speakers of other West African varieties of English.

The present-day GhE vowel inventory is characterised by a reduction of the twelve RP monophthongs to the following seven: /i/, /e/, /ɛ/, /a/, /ɔ/, /o/, /u/ (Huber 2008: 75). Importantly, the RP central vowel /ʌ/ is not part of the GhE vowel system (Huber 2008: 76).

4. Pilot Study I: Auditory Analysis

The vowel quality of RP /ʌ/ (STRUT; Wells 1982 and 2010) varies considerably in present day Ghanaian English (cf. Huber 2008). This is why the standard lexical set STRUT was chosen as the variable for an auditive study of the pronunciation in early Ghanaian popular music lyrics (cf. Huber & Schmidt 2011a). The STRUT vowel is here defined as the central monophthong lower than schwa (cf. Ladefoged 2006). Huber and Schmidt (2011a) compared the /ʌ/ vowel sub-system in the corpus of early Highlife lyrics with Huber's (2008) report on contemporary GhE. The song lyrics were transcribed orthographically by Schmidt and proof-read by Huber and students from the University of Ghana. Word lists containing all RP STRUT words were extracted from the Highlife corpus. Then, both authors coded RP STRUT variants as follows:

open vowel = a
half-open back vowel = o
closed back vowel = u
undecided/between 'a' and 'o' = m

Depending on the actual realization, *love* would, for example, be coded as *love_o, love_*a or *love_m.* An inter-rater agreement of 97,2% was reached.

As expected, variation in the realisation of RP STRUT in the 1950s/1960s songs is clearly observable. The main variants, though, are /a/ and /ɔ/. /ɛ/, a current GhE variant of RP STRUT, was not found in the Highlife corpus. Surprisingly, *come* was consistently realized as /kʊm/ in the song "Apolonia" by The Builders Brigade Band. To date, not enough is known about the singer or the band to give a solid explanation, particularly, because *love* is realized throughout the song as /lav/.

5. Acoustic Analysis

For the acoustic analysis, the transcriptions of the songs were transferred to PRAAT text grids as required for most PRAAT scripts (Lennes 2003). The selection of songs had to be revised, though. In the auditory analysis, even rather damaged recordings could be included, because, after some training, human coders could work well with them. However, when analysed with PRAAT (Boersma and Weenink 2011), the formants in these songs could not be measured to a satisfactory degree.

My text grids consist of three tiers, a 'line'-tier, a 'word'-tier and a 'vowel'-tier. All variants of RP STRUT are marked on tier 3, the vowel-tier. Due to the relatively small number of data-points this was done manually. A modified version of Lennes' (2003) script was used to measure the marked sections on the vowel-tier. The generated output was normalised using the NORM vowel-normalisation suite by Thomas and Kendall (2011). The Bark Difference Metric was chosen for normalisation, because this method works well with vowel sub-systems (Thomas and Kendall 2011). Figure1 shows the plotted RP STRUT words in the selected songs. Z3-Z2 represents the front-back dimension, Z3-Z1 the height dimension in analogy to a standard vowel chart (cf. Ladefoged 2006). The STRUT words auditorily coded as 'a' by Huber and Schmidt (2011a) are plotted in red. They cluster in the lower half of the diagram whereas the 'o' words in blue gather in the upper part. *Love* coded as 'm' (green) falls in between. For current GhE, Huber (2008) shows that the open vowel /a/ and the half-open vowel /ɔ/ are typical realisations of RP STRUT. The acoustic analysis confirms that RP /ʌ/ was already realized as /a/ and /ɔ/ in 1950s/60s GhE.

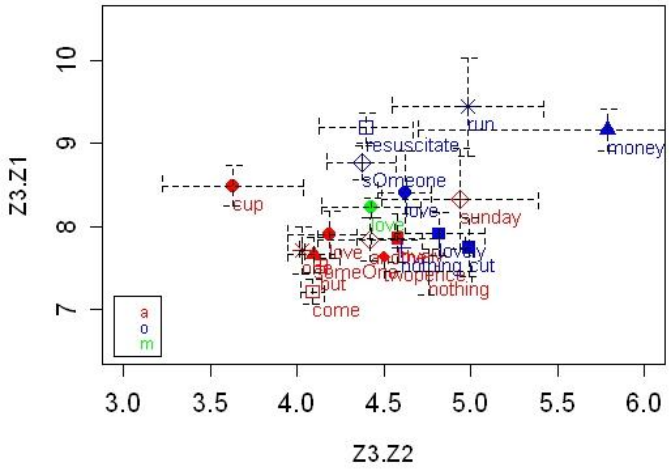


Figure 1: Vowel plot of RP STRUT words in a selection of early Ghanaian Highlife songs.

As the diagram shows, the acoustic pilot study confirms the findings from the auditory one (Huber and Schmidt 2011a). Apart from *cup*, all instances of RP STRUT in the Highlife corpus cluster towards the back-end of Z3-Z2. There is a tendency for words coded as 'o' to display a backer quality than tokens coded 'a'. There is also a tendency towards a height divide between 'a'-tokens and 'o'-tokens with only *Sunday* as an outlier. Figure1 also confirms the correct application of the 'm'-code, because the token *love* coded as 'm' is located between most 'a' and 'o' tokens.

The results for RP STRUT words in the Highlife corpus encourage further acoustic analyses aiming at a complete representation of the vowel system of early Highlife songs.

In conclusion, the lyrics of early popular music recordings can be analysed acoustically. Furthermore, the present study also shows that the results from the auditory study correspond to a large extend with the acoustic analysis. Through the application of both methods we get a glimpse of the English spoken in Ghana in the 1950s/60s.

6. Challenges

Early popular music in its original form is stored on various analogue records. Visits to archives, for example to the African Music Archive (AMA), Mainz, Germany, and experience from field work show that much depends on the condition of the actual record. Record here - since we are talking of the 1950s/60s - basically means shellac gramophone records and vinyl records. Depending on the frequency of use, the technology used for playing the records and the conditions of the respective archives or storerooms, the records deteriorate. Deterioration is inevitable due to the material characteristics of shellac and vinyl. Loss of data quality and sometimes of whole collections of music has to be taken into consideration. Apart from the purely physical aspect mentioned above, it is a challenge to contextualize the data.

Although ethno-musicologist Coester is currently working on the biographies of early Highlife singers, not much is known in detail about them. From a sociolinguistic or sociophonetic perspective, it would be helpful to know more about the singers, their L1s, educational background and if they had lived or toured extensively abroad for longer periods of time, for instance in Nigeria or the USA. Basic information about the singers can often be retrieved from the labels on the records. For example, in the case of the song "Awirehow" by E.T. Mensah and His Tempo's Band, the vocalist is identified as Dan Acquaye. In the case of "The Tree and the Monkey", also by E.T. Mensah and His Tempos Band, Julie Okine is mentioned, who is so far the only female singer in the corpus of early Highlife recordings. Due to typical regional and ethnic affiliation, though, it can be inferred from the names with some certainty to which ethnic group in Ghana a person belongs. *Okine*, for instance, is a Ga name (Anderson, personal correspondence).

Another challenge is the acquisition of data. Highlife recordings are scattered over various archives all over the world. The Gramophone Library of the Ghana Broadcasting Corporation (GBC) in Accra for instance holds a vast collection of shellac records from the 1950s and 1960s that is being digitalised. Recordings made by Decca West Africa

are held at the British Library. For the purpose of linguistic analyses, digital recordings in high .wav quality are indispensable and so an extensive database needs to be compiled. Technical issues prove less challenging than legal issues in this respect. It is often not clear who the copyright owners are and if digital copies of the recordings can be made available for research purposes.

7. Outlook

In order to generate a vowel system of early Ghanaian Highlife pronunciation, the methodology outlined above has to be repeated for other lexical standard sets as well, especially those which exhibit different realizations in RP and GhE (Huber 2008: 74, 81). Some Highlife songs are performed in a more 'spoken' way (performed somewhat similarly to talking blues). In these songs, vowel length merging can also be analysed.

Turning to consonants, /t/-affrication is a variable worth investigating. It is described by Wells (1982) as "a common allophone of /t/ in a London accent [which] is a heavily affricated [ts], thus [tsɑɪʔ ~ tsɑɪts] tight, [ˈpʰɑtsi] party" (Wells 1982: 31). As Huber (2008) observes for GhE, /t/-affrication has currency there, because "in the Fante dialect of Akan, /t/ has two allophones: [t] before back vowels and affricated [ts] before front vowels. Speakers of the dialect sometimes transfer this allophony to English and, for example, pronounce the name Martin [matsin]" (Huber 2008: 84).

Although current GhE is described as non-rhotic (Huber 2008: 87), the pronunciation of post-vocalic /r/ is a feature of a number of singers. An analysis of rhoticity in the Highlife corpus could provide empirical evidence of this phenomenon. An hypothesis which needs to be tested is whether rhoticity can be attributed to an orientation towards American popular music (cf. Trudgill 1983).

In the long run, early popular music recordings from Ghana will be compared to recordings from other colonial or post-colonial contexts. Nigeria with its extensive heritage of Highlife and Afro-Beat is an obvious contender for comparative studies. The same is true for Sierra Leone where Calypso culture brought forth an extensive number of recordings containing English or Krio lyrics in the 1950s and 1960s.

The advantage for linguists, who are interested in diachronic perspectives of popular music is that there is an ongoing, though not unproblematic (Hassold 2005), recording tradition in West Africa. From this rich source we are currently compiling a comprehensive corpus of music lyrics.

References

Behrendt, Stephen D., A.J.H. Latham and David Northrup (2010). *The Diary of Antera Duke, an Eighteenth-century African Slave Trader*. Oxford: OUP.

Bender, Wolfgang (1985). *Sweet Mother - Moderne Afrikanische Musik*. München: Trickster Verlag.

Bender, Wolfgang (2007). *Der Nigerianische Highlife. Musik und Kunst in der populären Kultur der 50er und 60er Jahre*. Wuppertal: Peter Hammer Verlag.

Boersma, Paul and David Weenink (2011). *Praat: Doing Phonetics by Computer*. Version 5.2.44. [Computer Programme].

Brato, Thorsten & Sandra Jansen (2008). "'You used to gerri' in yer fishnets, now you only gerri' in yer nightdress': Regional and supraregional accents in English rock songs". Presented at *The Thirteenth International Conference on Methods in Dialectology*. Leeds, 04 August. http://www.thorsten-brato.de/en/conferences-presentations/. Accessed: 20 February 2012.

Coester, M. (1998). "Language as a product of cultural contact". In: *ntama Journal of African Music and Popular Culture*. http://www.uni-hildesheim.de/ntama/. Accessed: 5. March 2012.

Collins, John (1986). *E.T. Mensah: King of Highlife*. London: Off the Record Press.

Collins, John (1989). "The early history of West African highlife music". *Popular Music*. Vol. 8, No. 3. 221-230.

Culver, Christopher (2007). "Fela's Nigerian English". http://www.christopherculver.com/linguistweblog/2007/10/felas-nigerian-english/.

Culver, Christopher (2008). "A linguistic approach to Fela Kuti's lyrics". http://www.christopherculver.com/linguistweblog/2008/01/a-linguistic-approach-to-fela-kuti%E2%80%99s-lyrics/. Accessed: 5. March 2012.

Hassold, Finn (2005). *Die Krise des Highlife - Zur Entwicklung der populären Musik in Ghana*. München: GRIN.

Hickey, Raymond (ed.) (2010): *Varieties of English in Writing. The Written Word as Linguistic Evidence*. Amsterdam: John Benjamins.

Huber, Magnus (2004). "Ghanaian English: Phonology." In Kortmann, Bernd and Edgar W. Schneider (eds.): *A Handbook of Varieties of English. A Multimedia Reference Tool. Volume 1: Phonology*. Berlin: Mouton de Gruyter. 842-865.

Huber, Magnus (2008). "Ghanaian English: phonology." In: Rajend Mesthrie (ed.). *Varieties of English 4 - Africa, South and Southeast Asia*. Berlin: Mouton de Gruyter. 67-92.

Huber, Magnus & Sebastian Schmidt (2011a). "New ways of analysing the history of varieties of English - Early Highlife recordings from Ghana". Presented at *ISLE 2*. Boston, 17-21 June 2011.

Huber, Magnus & Sebastian Schmidt (2011b). "Investigating the history of Pidgin English - Early Highlife Recordings from Ghana". Presented at *The 2011 Summer Conference of the Society for Pidgin and Creole Linguistics*. Accra, Ghana, 2-6 August 2011.

Kachru, Braj (ed.) (1982). *The Other Tongue: English across cultures*. Urbana: University of Illinois Press.

Kachru, Braj B. (1986): *The Alchemy of English. The Spread, Functions, and Models of Non-Native Englishes*. Chicago: The University of Illinois Press.

Kreyer, Rolf & Joybrato Mukherjee (2007). "The style of pop song lyrics: a corpus-linguistic pilot study". *Anglia*. 125 (1). 31-58.

Ladefoged, Peter (2006). *A course in Phonetics*. Boston: Thomson.

Lennes, Mietta (2003). "collect_formant_data_from_files.praat". [Computer Script]. Source: http://www.helsinki.fi/~lennes/praat-scripts/public/collect_formant_data_from_files.praat. Accessed: 5. March 2012.

Oti, Sonny (2009). *Highlife Music in West Africa: Down Memory Lane*. Lagos, Nigeria: Malthouse.

Schmidt, Sebastian (2011a). "Tracing the Lyrics – Early Highlife Recordings from Ghana as Linguistic Data and Cultural Artifacts". Presented at *the GCSC-Workshop Korpus Kommunikation Kultur: Linguistik als Kulturwissenschaft*. Gießen, 4. November 2011.

Schmidt, Sebastian (2011b). "New Ways of Analysing the History of Varieties of English - An Acoustic Analysis of Early Pop Music Recordings from Ghana". Presented at *Accents 2011*. Lodz, 15.-17. Dezember 2011.

Schneider, Edgar W (2007). *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge University Press.

Schneider, Edgar W. (2003). "The dynamics of New Englishes: From identity construction to dialect birth". *Language* 79 (2). 233–281.

Thomas, Erik R. and Tyler Kendall (2007). *NORM: The vowel normalization and plotting suite*. [Online Resource: http://ncslaap.lib.ncsu.edu/tools/norm/]. Accessed: 19 February 2012.

Trudgill, Peter (1983). "Acts of conflicting Identity - The sociolinguistics of British pop-song pronunciation". In: Peter Trudgill. *On Dialect - Social and Geographical Perspectives*. Oxford: Blackwell. 141-160.

Wells, J.C. (1982). *Accents of English 1 - An Introduction*. Cambridge: CUP.

Wells, J.C. (2010). *Standard lexical sets*. http://www.phon.ucl.ac.uk/home/wells/stanlexsets.htm. Accessed: 19 February 2012.

Discography

E.T. Mensah and his Tempos Band (1950s). "Awirehow". Decca West Africa.

E.T. Mensah and his Tempos Band (1950s). "Day by Day". Decca West Africa.

E.T. Mensah and his Tempos Band (1950s). "Don't Mind Your Wife". Decca West Africa.

E.T. Mensah and his Tempos Band (1950s). "I Want to be Happy". Decca West Africa.

E.T. Mensah and his Tempos Band (1950s). "Inflation Calypso" Decca West Africa.

E.T. Mensah and his Tempos Band (1950s). "Sunday Mirror". Decca West Africa.

E.T. Mensah and his Tempos Band (1950s). "Tea Samba". Decca West Africa.

Note on the discography: RetroAfric, London, offers reissues of E.T. Mensah's most famous recordings.

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0037-6

THE DOUBLE-EDGED SWORD OF RP: THE CONTRASTING ROLES OF A PRONUNCIATION MODEL IN BOTH NATIVE AND NON-NATIVE ENVIRONMENTS

MIROSLAV JEŽEK
Masaryk University, Brno

Abstract
Received Pronunciation (RP) is often studied as the pronunciation model in Great Britain and non-English-speaking countries separately. What my paper focuses on is the duality with which RP is essentially endowed: the role(s) in which it has to satisfy the needs of *both* native and non-native speakers of English.

Whilst the claim that RP has changed recently goes unchallenged, the issue of reflecting these changes in the preferred transcription models is hotly debated. Upton's model of RP is one that does include several new symbols, motivated by an attempt to 'ensure that the description of a late twentieth century version the accent […] looks forward to the new millennium rather than back at increasingly outmoded forms' (2001:352). I discuss the feasibility of adopting Upton's model of RP as the pronunciation model in non-English speaking countries, where it is desirable to resolve the paradox that 'most of our teaching is aimed at young people, but the model we provide is that of middle-aged or old speakers' (Roach 2005: 394).

The observations I make are largely based on my MA research, which is now being modified for the purposes of my Ph.D. I asked undergraduate students of English in England and the Czech Republic to evaluate seven voices ranging from the clearly regional to the unquestionably RP. The objective was to discover which sounds are considered to fall within the scope of RP by students in both countries, which approach avoids treating RP as though it were to include only the sounds 'allowed by a preconceived model' (Upton 2000: 78). Further, the respondents were asked to comment on the most salient features in the recordings: what they opted to comment on reveals a marked difference in the role of RP as a model accent in the given countries. Societies which lack a prestigious non-regional accent are often oblivious to the social connotations RP carries. Whilst it seems technically impossible to replace the model accent in all teaching materials all over the world, creating awareness of the fact that a rather outmoded model of RP found in many textbooks may not always be the best option is a necessary step towards ensuring that non-English speaking students are not only understood but that their speech will attract no adverse judgements.

1. Introduction

RP, like any other accent, is subject to constant change. However, the transcription model found in materials for ELT purposes has changed little since Jones's transcription, first used in the English Pronouncing Dictionary published in 1917. The reasons are manifold. Upton (2001: 355) mentions the following as the most prominent ones:

- in the world of lexicography, phonological matters are not usually given priority (this is presumably brought about by the fact that most lexicographers are not phoneticians, hence they do not pay as much attention to the matters of pronunciation as they do to semantics and grammar)
- there is strong conservative pressure in the ELT divisions of publishing houses
- phonological redescription in ELT dictionaries would also entail the revision of a great number of other non-dictionary texts in which pronunciation is discussed — this would be rather impractical and, above all, too costly

For the aforementioned reasons it might seem to an outsider (in particular to someone who does not reside in the UK and whose first language is not English) that RP is an accent with little, if any, variation. The best testimony to prove that the opposite is true is the number of labels often attached to RP. The basic division phoneticians make is into an older, rather conservative, variety and a younger, modern one. The former is labelled 'traditional RP' (Upton 2008: 239), 'U-RP' (Wells 1982: 279), 'Refined RP' (Cruttended 1994: 80) or 'marked RP' (Honey 1991: 38). The latter is called 'mainstream RP' (Wells 1982: 279), 'General RP' (Cruttended 1994: 80), 'unmarked RP' (Honey 1991: 38), or there might not be any label at all: Upton (2000: 76) decided to call this modern variety simply 'RP' on the grounds that it is the mainstream variety and it can therefore 'legitimately lay claim to the RP label without qualification'.

RP is an accent endowed with both advantages and disadvantages. This has been well-documented in a wealth of research; cf. for example Giles (1990) and, more recently, Beal (2008). RP is viewed as competent, persuasive and intelligent, but, at the same time, as rather unfriendly and dishonest (Beal 2008: 29). This is the reason why I call RP a 'double-edged sword': it may open some doors for you but it may also close others.

Prof. Clive Upton, currently based at Leeds University, is the only linguist who has radically altered the transcription model of RP with the aim of providing a transcription model which avoids 'slavish imitation of the dictates of self-appointed arbiters of taste or style in language' (Upton 2003: viii). Instead, Upton only includes those sounds 'heard to be used by educated, non-regionally marked speakers rather than [sounds] "allowed" by a preconceived model' (Upton 2000: 78). Ramsaran shrewdly observes that '[i]f one excludes certain non-traditional forms from one's data, how can one discover the ways in which the accent is changing?'. In other words, one cannot use the same sieve, metaphorically speaking, over and over again to see who falls through and who does not. This is hardly a successful way of detecting linguistic change.

It is now time to turn our attention to the actual description of the model in question.

2. Upton's model of RP

Upton's model has been in use for about two decades now and the most notable publications where this model can be found include the world-famous *Oxford English Dictionary* (OED). Other dictionaries using Upton's model of RP are, for example, *The New Shorter Oxford English Dictionary* (from 1993 onwards), *The Concise Oxford Dictionary* (from 1995 onwards), and *The New Oxford Dictionary of English* (1998,

2003). Last but not least, *The Oxford Dictionary of Pronunciation for Current English* (2001) is also on the list, this being the only dictionary focusing solely on pronunciation.

The call for an updated version of the RP model had been around for some time before Upton decided to undertake the task of providing one. Gimson, in particular, insisted that a new set of criteria for redefining RP be found. These 'will result in a somewhat diluted form of the traditional standard' (1984: 53). In the same article Gimson adds his hope that

> the re-defined RP may be expected to fulfil a new and more extensive role in present-day British society. Its primary function will be that of the most widely understood and generally acceptable form of speech within Britain […] and more importantly for the future, this standard form of British speech can function as one of the principal models for users of English throughout the world
>
> (1984: 53)

2.1. RP Vowels

While most of the vowels employed by Upton are the same as in other (older) models of RP, there are several salient changes which have made his model a contentious issue. The following table taken from Upton (2008: 241-2) neatly summarises the differences between RP and traditional RP:

vowel	RP	shared RP/trad-RP	trad-RP
KIT		ɪ	
DRESS	ɛ		e
TRAP	a		æ
LOT		ɒ	
STRUT		ʌ	
FOOT		ʊ	
BATH	ɑː ~ a		ɑː
CLOTH	ɒ		ɒ ~ ɔː
NURSE	əː		ɜː
FLEECE		iː	
FACE		eɪ	
PALM		ɑː	
THOUGHT		ɔː	
GOAT	əʊ		əʊ ~ oʊ
GOOSE		uː	
PRICE	ʌɪ		aɪ
CHOICE		ɔɪ	
MOUTH		aʊ	
NEAR		ɪə	
SQUARE	ɛː		ɛə
START		ɑː	

vowel	RP	shared RP/trad-RP	trad-RP
NORTH		ɔː	
FORCE		ɔː	
CURE	ʊə ~ ɔː		ʊə
happY		i	
lettER		ə	
commA		ə	

Table 1: The vowels of RP and traditional RP (Upton 2008: 241-2)

Whilst some changes seem to be mere transcriptional preferences (e.g. DRESS or NURSE), others have raised a few eyebrows because they essentially alter the way RP is perceived and interpreted. Namely it is the TRAP, BATH and PRICE lexical sets that are discussed here in detail.

Firstly, the TRAP vowel is lowered so that the appropriate symbol is no longer the ash symbol [æ], but the cardinal vowel no. 4 [a]. Wells (2001) insists that it is not necessary to make the change as it is enough to retain the original symbol and simply redefine it. This is, however, hardly possible due to the fact that phonetic symbols are absolutes, therefore 'their interpretation cannot be altered to suit the new development, so that if anything is to change in the interests of accuracy and clarity it must be the label that is applied to the sound' (Upton 2008: 240). Upton goes on to argue that because ELT texts are broadly phonemic 'their users […] need to be provided with transcriptions which correspond as honestly as possible to the sounds of the modern accent' (2008: 240).

Secondly, Upton introduces the short BATH vowel [a], typically associated with the North of England, as a possible RP alternative to the usual long BATH [ɑː]. The logic behind this decision is relatively simple: people in the North of England no longer adopt the southern long BATH vowel; as a result even those who would normally be perfect RP speakers cannot be labelled thus because they retain the short BATH vowel. If the older model is taken as the norm, there is not (or soon will not be) a single RP speaker in the North and, more importantly, RP ceases to be a non-regional accent. Instead, it is immediately associated with the South of England. Upton then introduces 'southern' and 'northern' varieties of RP, thereby adhering to the universally accepted principle that 'RP is not to be considered as exclusively a southern-British phenomenon' (Upton et al. 2003: xiii).

Thirdly, the PRICE diphthong, changed from trad-RP [aɪ] to RP [ʌɪ], has come in for a significant amount of criticism. Wells (2001) admits that there is a lot of variation in the starting point of the diphthong but strictly dismisses Upton's choice as 'very unsuitable [because it] accords with the habits neither of RP nor of southeastern speech'. It is interesting to ponder a little on why the second element (south-eastern speech) is added in the previous quote from Wells. I understand why Wells is unhappy about Upton's choice of [ʌɪ] if he cannot see it used in RP at all, but adding that it is not present in south-eastern speech either seems to go against the criterion that RP should not be associated with any particular region. Incidentally, this is exactly the reason why

Upton's model of RP comes in for a lot of criticism—his inclusion of the short BATH allegedly deprives RP of its non-regional basis. Surely, RP should only allow—where possible, of course—supraregional sounds not associated with any particular region. One notable exception is the short/long BATH vowel, where both regions stick to their own varieties. A linguist can then either dismiss one of the two variants as non-standard or allow both in their model of standard pronunciation.

This idea is far from modern: in 1942 Vilem Mathesius, the founding father of English Studies in Czechoslovakia, observed that people from the Bohemia region (centred on Prague) pronounce the initial consonant cluster in the Czech word 'shoda' (Eng. 'agreement') voicelessly while people in the Moravia region (centred on Brno) prefer the voiced variant. Although the former, i.e. voiceless, pronunciation had traditionally been regarded as standard, Mathesius noticed that people from Moravia, though otherwise perfectly conforming to the standard-speaker model, stick to the voiced variant. In a dilemma very similar to the one Upton found himself in, Mathesius opts to accept both variants as standard (1982: 149).

2.2. RP Consonants

RP consonants are nowhere near as variable as its vowels; hence they pose considerably fewer problems for phoneticians. Many variants found in Upton's model are RP universals and are thus not unique to his model. The only consonantal feature worth mentioning here is the presence of optional intrusive /r/, as in 'drawing' [dɾɔː*(r)*ɪŋ]. The italics mean that the /r/ sound is intrusive rather than linking, which is shown in normal font.

3. Research

I conducted the research in 2009 for the purposes of my MA thesis. Right now, it is being modified and, hopefully, improved at Ph.D. level. The whole idea formed in my mind during my year-long stay at Leeds University in 2006-2007. It was not until then that I started to realise certain differences in the perception of RP in the UK and the Czech Rep.

3.1. Research objectives

- to compare the roles RP fulfils in the UK and the Czech Rep.
- to test the extent to which undergraduate students of English in both countries are aware of recent innovations in RP
- to discover which sounds are considered to fall within the scope of RP by students in both countries

3.2. Methodology

I set up a simple website which can still be accessed here: www.received pronunciation.wz.cz. I asked respondents from both the UK and the Czech Rep. to evaluate seven recordings which ranged from clearly non-RP/regional to trad-RP. All the UK respondents were, incidentally, English (although I would certainly not have discarded data from, say, Scottish or Welsh people). They were all aged 19-25 and were either of working or of middle-class background. They were from all sorts of regions within England—if we take into account the two best-known criteria which separate the North from the South (namely the BATH and STRUT vowels), then I can say I had 17 southern and 13 northern respondents. The Czech respondents were also aged 19-25; furthermore, I chose only those who model their speech the British way. Five of the seven recordings were made by me; the remaining two (including the trad-RP recording) were taken from Collins et Mees (2003). Each recording was accompanied by a questionnaire. First, the respondents were asked to indicate, on a scale of 1 to 7 (1-highly regional, 7-RP), how close to RP the given recording sounded to them. I view RP, like any other accent, quantitatively (more or less) rather than qualitatively (either…or). This is something foreign students often seem oblivious to: they think that someone either speaks RP or they do not. But this is utterly mistaken as Wells' category of Near-RP (Wells 1982: 279) testifies. Then they went on to fill in several write-in questions. I deemed it extremely important not to ask about any particular sounds so as not to put ideas into my respondents' minds. The questions were thus rather vague such as 'What is your overall impression of this speaker?' or 'Can you comment on any particular details which helped you make up your mind in the RP score question?'. What the respondents opted to comment on — regardless of whether their comments were positive or negative — reveals a marked difference in the role of RP as a model accent in the given countries.

4. Results

It is perhaps not surprising that what I ended up with was just a hotchpotch of comments which were then classified into categories by the common topic. The most salient categories include the following: intelligibility, regionality, social status, education, poshness. There were admittedly some more categories, namely euphony, speed, authenticity, appropriateness, and rhythmicality, but these were found rather awkward to deal with or useless and will not be taken into account in the Ph.D. research.

A very simple table below illustrates the differences between GB and CZ respondents:

CATEGORY	GB respondents	CZ respondents
Intelligibility	3	26
Regionality	37	9
Social status	14	3
Education	12	6
Poshness	11	4

Table 2: observations by topic (measured in index points)

What is immediately observable is the fact that for Czech learners of English the crucial aspect when they assess English speakers is intelligibility. The remaining four categories are not nearly as important for them as they are for their British counterparts. This is obviously perfectly understandable and entirely predictable, but it shows without any doubt that the roles of RP in native and non-native environments are markedly different and should therefore be kept separate whenever transcription models are discussed.

Czech university students of English are, of course, told about the regional and social connotations RP carries but I argue there is a huge gap between knowing something and feeling it intuitively. Czech learners of English often see RP as the most intelligible accent and thus consent to learn it almost automatically. Unfortunately, the model of RP they find in teaching materials is outdated, which is rather startling, for the recordings found in the very same textbooks often do not correspond with the transcripts. One could argue that these recordings are not then RP (and unquestionably many of them really are not), but it would then mean that there are no RP recordings in modern textbooks of English. The next question then suggests itself: Why are these teaching materials full of phonetic transcriptions of an accent which does not appear in them at all?

The TRAP vowel is a case in point. While the transcriptions invariably insist on [æ], the recordings include voices with lowered [æ] for which it seems more appropriate to choose [a]. Specifically, I am now talking about Maturita Solutions textbooks used mainly in secondary schools—there are several pronunciation exercises which stress the importance of distinguishing such minimal pairs as 'pat' [pæt] and 'pet' [pet]. Sadly, the TRAP vowel is predominantly realised as [a] in the recordings (this might be so because of the fact that the majority of the voices, without any doubt, belong to people in their twenties, if not younger, which in itself is a very welcome step, of course). It then takes me a lot of time explaining to my students that there is no need to attempt [æ] and that [a] is perfectly acceptable. For many Czech learners of English, the adoption of [a] would certainly help to make the situation easier since they have [a], unlike [æ], in their repertoires.

The question in which respondents were asked to evaluate the recordings on a scale of 1 to 7 (1-highly regional, 7-RP) provided some intensely interesting data as well. Three speakers' scores are worth looking at in greater detail.

SPEAKER	GB respondents	CZ respondents
Speaker 3 (most regional)	2.7	3.45
Speaker 4 (modern RP)	5.1	3.45
Speaker 6 (trad-RP)	6.3	5

Table 3: RP scores for three selected speakers

I have decided to retain the original numbers the speakers had been assigned in the RP Test in order that the readers could visit the website and listen to the recordings for themselves.

As we can see, the most regional Speaker 3 (the accent is, by the way, not a particularly strong one, the voice belongs to a Ph.D. student of the English language from Middlesbrough) received exactly the same score from Czech respondents as modern-RP Speaker 4 did. There are two possible explanations: either students in the

Czech Republic failed to spot those regional features which clearly are not RP (e.g. lowered STRUT and monophthongised GOAT) or their perception of RP is rather outdated and what is considered modern RP now in the UK is still perceived as non-RP in the Czech Republic. The latter explanation, however, is made somewhat doubtful in the light of the next observation: Czech respondents failed to assign the highest RP score to the trad-RP speaker. Although the score of 5 might appear to be high, it must be kept in mind that Speaker 6 sits roughly in the middle with the fifth highest score of all. British respondents, on the other hand, unmistakably and unanimously placed Speaker 6 at the very top of the rank.

The comments Czech respondents made about Speaker 6 reveal that the accent is not only 'weird' but also, according to a number of them, regional, too. Crucially though, the accent was ranked fourth in the intelligibility question for Czech respondents. Generally speaking, the accent was not popular with either set of respondents. For British respondents the overwhelming perception of the accent was that of sounding extremely posh.

The comments from both sets of respondents have also shown that while lowered TRAP and short BATH vowels are RP sounds for English respondents, they are not so for their Czech counterparts. Intrusive /r/ is most assuredly an RP sound for both sets, as is, in fact, the glottal stop replacing /t/ in other than intervocalic positions. /t/-glottaling is not treated here for it has been covered extensively elsewhere (cf. Hannisdal 2006).
The last RP sound I want to discuss here in greater detail is the PRICE diphthong. It is one of the most contentious issues in Upton's model of RP and the one for which Wells (2001) finds the least sympathy. This diphthong did draw some comments from British respondents, many of whom noticed the backed first element. The decision as to whether or not this falls within the scope of RP was, however, far from unanimous—about 60% of those who did comment on it considered [ʌɪ] to be an RP sound.

Most revealing is the conspicuous lack of any comments on the part of Czech respondents. The reason why they failed to spot any variation here is quite simple: in the Czech phonological system there only are five monophthongal vowels /a/, /e/, /i/, /o/, and /u/ and three diphthongs /au/, /eu/, and /ou/ (Dankovicova 1999: 72). As far as the /a/ vowel is concerned, its realisation varies to a large extent ranging from [a~ɑ~ʌ]. The front vowel is common in Bohemia whereas the back one is typical of Moravia. This variation is merely allophonic; as a consequence, Czech learners of English have trouble distinguishing minimal pairs such as *fun*/*fan* when these are pronounced by a native speaker of English whose *fan* vowel is realized as [a] and not as [æ]. It is then far from surprising that Czech respondents did not comment on the PRICE diphthong in the RP Test at all.

5. Conclusion

The results of my research seem to suggest that trad-RP is a now such a rarity it has lost its function in the ELT field. It appears to be so obsolete that some Czech respondents mistook it for a regional accent; moreover it is not the most intelligible dialect any more. This might have been brought about by far greater exposure to a higher number of native

British accents in the past two decades. Learners of English in the Czech Republic rely less and less on textbook CDs and turn to some more natural/authentic sources (TV programmes of all sorts are immensely influential in this respect) when trying to improve their pronunciation.

Upton's model of RP seems highly suitable for Great Britain since it reconciles the two opposing tendencies still present in British society—namely the desire to speak better but, at the same time, to avoid sounding posh and elitist. This is well documented in Beal who comes to the conclusion that 'British society today is every bit as hierarchical as that which spawned the elocution movement of the 18th century, but […] the models of good pronunciation are no longer the aristocracy but the professional and entrepreneurial classes who can provide employment' (2008: 38). But RP is no longer the automatically preferred accent. Call centres are a case in point—their workers 'avoid both the unfriendly connotations of RP, and the uneducated associations of broad regional accents, and so are acceptable to a wide range of callers' (Beal 2008: 30-1). Surely Upton's model of RP is a step towards a less elitist perception of the accent.

Wells (2001) objects to Upton's model of RP because he sees it as an unnecessary threat to the 'hard-won uniformity' which had been achieved in the transcription of RP. He believes that 'supposed gains did not make up for the sacrifice of an agreed standard' (2001). What should we do, though, if the agreed standard, albeit so laboriously gained, does not reflect the true state of affairs any longer?

Introducing Upton's model to the Czech Republic, however, appears to face many obstacles.

The first and seemingly insuperable obstacle is money. Re-editing and republishing the vast numbers of teaching materials in which pronunciation is discussed would not only be highly impractical but also too expensive.

Secondly, for the reasons mentioned in the Introduction there is not enough support to carry out these changes anyway.

Thirdly, I fear some of the changes would only bring about more confusion for the overwhelming majority of learners (in particular for those who do not study English at university, which is the lowest level where phonetic symbols are dealt with properly in the Czech Republic) for whom phonetic symbols are abstruse and who learn pronunciation by way of imitation rather than by way of pronouncing dictionaries.

Last but not least, RP in the Czech Republic lacks the social and regional connotations it has for native speakers in Great Britain. The roles of RP in the two countries in question are markedly different. What seems necessary in Britain might not be so in the Czech Republic: whilst updating the model in Britain makes sure that the accent is rid of the redolence of social privilege, there is no such problem in the Czech Republic.

It seems, nonetheless, important for Czech learners of English to be aware of the incessant change RP is subject to (it is not a petrified accent, although it is for obvious reasons more resistant to change than other accents). Likewise they should take into account the wealth of connotations this accent is endowed with. They should know that for many people in Britain RP (particularly in the traditional form) is not the preferred accent and the reaction to it may not always be positive.

RP is the accent used in the Czech Republic as the model accent. This seems extremely unlikely to change in the foreseeable future (if a completely radical change is not undertaken, e.g. replacing RP with the General American accent). I am convinced that it is eminently desirable to resolve the unhappy situation in which the accent often heard from CDs is in certain particulars considerably different from the transcription provided. It is true that CDs often contain recordings with a wide variety of accents; many of them are (slightly) regional and are also different from the phonetic symbols used in the teaching materials. These, however, are not presented as the model students should imitate.

References

Beal, Joan. 2008. Shamed by Your English? In Joan Beal et al. (eds) *Perspectives on Prescriptivism*. Bern: Peter Lang, 21-40.

Collins, Beverley et Inger Mees. 2003. *Practical Phonetics and Phonology*. London: Routledge.

Cruttenden, Alan. 1994. *Gimson's Pronunciation of English*, 5th ed. London: Arnold.

Dankovicova, Jana. 1999. Czech. In *Handbook of the International Phonetic Association*, Cambridge University Press, pp. 70–74.

Giles, Howard et al. 1990. The Social Meaning of RP. In Susan Ramsaran (ed) *Studies in the Pronunciation of English: A Commemorative Volume in Honour of A. C. Gimson*. London: Routledge, 191-211.

Gimson, A. C. 1984. The RP Accent. In Peter Trudgill (ed) *Language in the British Isles*. Cambridge: Cambridge University Press, 45-54.

Hannisdal, Bente Rebecca. 2006. *Variability and change in Received Pronunciation*, Ph.D. dissertation. Bergen: University of Bergen.

Honey, John. 1991. *Does Accent Matter?*. London: Faber and Faber.

Mathesius, Vilem. 1982 [1942]. *Jazyk, kultura a slovesnost*. Praha: Odeon.

Roach, Peter. 2005. Representing the English Model. In Katarzyna Dzubialska-Kolczyk & Joanna Przedlacka (eds) *English Pronunciation Models: A Changing Scene*. Bern: Peter Lang, 393-9.

Upton, Clive. 2000. Maintaining the Standard. In Robert Penhallurick (ed) *Debating Dialect: Essays on the Philosophy of Dialect Study*. Cardiff: University of Wales Press, 66-83.

Upton, Clive. 2001. Revisiting RP. In Malcolm Jones (ed) Essays in Lore and Language: Presented to John Widdowson on the Occasion of His Retirement. *Sheffield: National Centre for English*, 351-68.

Upton, Clive, William Kretzschamr and Rafal Konopka. 2003. The Oxford Dictionary of Pronunciation for Current English**.** Oxford: Oxford University Press.

Upton Clive. 2008. Received Pronunciation. In Clive Upton & Bernd Kortmann (eds) *Varieties of English: The British Isles*. New York: Mouton de Gruyter, 237-52.

Wells, J. C. 1982. *Accents of English* (3 vols). Cambridge: Cambridge University Press.

Wells, J. C. 2001. IPA Transcription Systems for English.
http://www.phon.ucl.ac.uk/home/wells/ipa-english-uni.htm [accessed February 2012]

USING NIGERIAN ENGLISH IN AN INTERNATIONAL ACADEMIC SETTING

UNA CUNNINGHAM
Stockholm University

Abstract
This study examines the English pronunciation of a group of Nigerian students at a university in Sweden from the point of view of their intelligibility to two groups of listeners: 1) native speakers of English who are teachers at the university; 2) nonnative speakers of English who are teachers at the university. It is found that listeners who are accustomed to interacting with international students do better than those who are not, and that native speakers of English do no better or worse than non-native listeners. The conclusion is drawn that locally useful varieties of Nigerian English may not easily be used as for wider communication and that students preparing to study abroad would find it useful to gain access to a more widely intelligible variety.

1. Background

Many students from all around the world find their way to universities in Sweden. There are a number of reasons why Sweden is attractive to international students. The standard of living is high, and so is the standard of education. In addition, it is well known that many Swedish people speak English. Swedish universities offer a fair number of Master's programs and a few undergraduate programmes taught through the medium of English. It is not difficult for international students to study in Sweden, even if they have no knowledge of the Swedish language. Another, quite compelling, reason for the interest Swedish universities have attracted from international students is the fact that Swedish higher education had until recently no tuition fees, not even for students from outside of Europe. Many students have realised that in Sweden they have the chance of getting a world-class education without paying fees.

Many Nigerian students who come to Sweden to study English have received most or all of their previous education through the medium of English. It comes as a shock in many cases for these students to find that they are not viewed by their teachers in Sweden as native or near-native speakers of English. They may fail language proficiency courses and find that their English does not work as well as they expect it to in communication with their teachers and with other international students, in particular those who are non-native speakers of English.

The influx of students from other parts of the world has not been entirely without problems. There are a number of inconsistencies between the Swedish education system and its counterparts in other countries. One problem we have had is with the way foreign qualifications are judged by the Swedish National Agency for Service to Universities

and University Colleges who centrally administer admissions to Swedish universities. Each year the Agency produces a handbook where the national qualifications in many countries are listed, explained and compared to the Swedish qualifications upon which the admissions system is based. Unfortunately it appears that in a number of cases the Swedish system is overly generous in its conversion of foreign grades. For example, it is necessary for students who have attended school in Nigeria to achieve a grade 8 (the lowest pass grade) in English O-level (SSCE/WASSCE). This is deemed as equivalent to the Swedish upper secondary course English B, which is in turn deemed equivalent to IELTS level 6.0 with at least 5.5 in each section of the test. This fulfils the English language prerequisites for any programme of study in any faculty at any Swedish university. Ironically, it appears that at least in the past, Nigerian universities do not accept students to any faculty with less than a credit (grade 6) in O-level English (SSCE/WASSCE) (Ufomata 1996).

The Swedish system of higher education was designed to cater for the needs and expectations of Swedish school leavers. For many years this was adequate. When Sweden entered the EU in 1995 the number of international students increased with exchange schemes such as the Socrates-Erasmus programme which funds and facilitates the exchange of students and staff between universities in Europe. Such students stay for a semester or a year and return to their home universities with their credits to take their degree there. The influx of students from beyond the EU coincided with the introduction of degree programmes (as opposed to short courses which the student collects until the appropriate number of credits and a degree thesis have been achieved allowing the student to apply for a degree). In the global higher education market, degree programmes are much more transparent and attractive than the loosely bound selection of courses which leads to a degree that has been usual, at least in the humanities, in Sweden.

The EU has, through what is known as the Bologna process, attempted to impose a degree of uniformity on European higher education. It is, in theory, possible for students to wander from one European university to another, taking their credits where they may. Of course, in practice, things are not always that simple, but there is at least a level of understanding of the way the system works in other parts of Europe. When students from other parts of the world apply to Swedish courses and programmes they may find that their qualifications are not well regarded. Students from Pakistan, for example, may find that they need to have completed both a BA and an MA to be deemed to have a qualification equivalent to a Swedish bachelor's degree. Students from Russia may find, to their dismay, that only three of their five years of university education will be considered.

These circumstances lead to a situation where many students are admitted to study at too high a level due to the prerequisites being inappropriately low. In fact such students often have a primary problem with insufficient proficiency in English language. The University provides an English language needs analysis to discover such cases early on (in the first week of study) so that students can be offered courses in English for academic purposes during one or sometimes two semesters, before proceeding to their planned programme of study. This preparatory study improves students' proficiency levels while simultaneously introducing them to the means, methods and models of learning which shape the student experience at a Swedish university.

English has no official status in Sweden, which means that Sweden is part of Kachru's expanding circle (Kachru 1992). In Sweden, the requirement of English language proficiency for the study of English at university level is set to match that held by Swedish school-leavers who have taken two years of English at upper secondary school. Previously such students will have studied English for at least 7 years at primary and lower secondary school. In addition, they are bombarded by English from TV, cinema, music and computer games. They have ample opportunity to hear English and most young people can switch to English with minimal inconvenience when they need to, which is fairly often, given that Swedes travel extensively and cannot expect to meet only Swedish speakers outside Sweden and that Sweden has many international visitors. Swedish young people possess considerable communicative competence in English. While their speech may be accented and their grasp of the vagaries of English grammar tenuous, they speak and understand English adeptly. Consequently, university English courses in Sweden are generally designed to teach grammatical accuracy and academic reading and writing skills at the initial level, rather than pronunciation and communication, moving quickly on to the kind of courses in English language, linguistics, literature and culture that can be found at universities anywhere in the English-speaking world.

What then do we require of students who are to take part in our courses in terms of English language proficiency? One criterion for the required level in the needs analysis is that students need to be able to understand native and non-native speakers of English speaking clearly and at normal tempo, which is what they need to be able to do if they are to take part in classes. Another is that they need to be able to express themselves coherently orally and in a free writing task. Yet another part of the needs analysis is a grammar test, corresponding to the IELTS levels 5.0 and 6.0 which are the levels required for our preparatory courses and ordinary undergraduate and graduate courses respectively.

To address the needs of students who are admitted to the university with less English proficiency than we require as shown by the needs analysis, we have designed our preparatory courses in English for academic purposes. In fact we have found that some Swedish students also appreciate these courses, either because they have only one year of upper secondary English, or because they have been away from education for some years and feel that their English needs refreshing before they continue. Even the occasional native speaker of English turns up on our courses for a number of reasons, often involving limited educational opportunities. Our needs analysis will pick up these speakers through their lack of certainty regarding the grammaticality of Standard English constructions and their written disfluency. For students who have learned English as a foreign language in their home country in the so-called expanding circle, it may be disappointing to find that the level achieved is not adequate for study in Sweden, and some do insist on disregarding the advice of their teachers and continuing on the programme or course to which they have been admitted, generally with disappointing results. For Swedish students and for the occasional inner circle speaker from the UK or the US, the preparatory courses offer a chance to remediate the gaps in their English in a context which is not face-threatening. The students who find it hardest to accept a disappointing result in the needs analysis are those who come from the outer circle, those for whom English is a second language, often the language of their education.

There is a serious problem here which is faced by all international academic environments. While, on the one hand, speakers of what McArthur (2002) calls New Englishes rightly demand full respect and recognition of these as legitimate varieties of English, they, like some of the Old Englishes such as my own Northern Irish English are not always ideally suited to international communication. As an educated speaker of Northern Irish English, when I left Northern Ireland to study in Britain, I quickly learned to modify my pronunciation to facilitate communication with non-Northern Irish interlocutors. Significantly, this can be done without compromising speaker identity as, in my case, a person from Northern Ireland. Initially I was not easily understood and my pronunciation was the object of comment. Failure to change my more "extreme" pronunciations might eventually have led to those I interacted most with getting used to my way of speaking, but the social and educational cost would be considerable. The result is that I, like many speakers of non-standard accents and dialects, switch between accents depending on my interlocutor and the communicative situation.

There is a significant distinction to be made at this point between English and Englishes. One of the definitions of a language as opposed to a dialect refers to the criterion of mutual intelligibility. While I would not like to suggest that the less widely intelligible accents of English are not English – we are after all talking about accents rather than syntactic or lexical variation – English is a very special case. We ask more of English than has ever been asked of any language in our history. Not only is it an important lingua franca, allowing genuine international communication, it is also a local living language for millions of everyday speakers in many different countries. But we are fooling ourselves if we claim that a speaker can wander from one communicative situation to another without modifying his or her English according to the communicative situation. Two speakers of any variety of English will be able to speak together in a different way than if one of them were to converse with a speaker of another variety in another place, and in yet another way if speaking to an EFL or ESL learner (even one who has the same variety of English as a target for their learning) whose proficiency may well be limited.

The problem may arise as a consequence of postcolonial insecurities and defensiveness regarding the status of New Englishes. If there is indeed a Standard Nigerian English pronunciation, which seems relatively problematic given the variation described in e.g. Banjo (1971), Bamgbose (1995) and Ufomata (1996), it may not be very useful for international communication, just as can be said of Glasgwegian and various Northern Ireland accents, not to mention some kinds of southern US accents, or broad Australian, or Scouse, Geordie or any other well-defined accent of English. No linguist would question the legitimacy of any of these inner or outer circle accents. What happens is that these are not adequate for international or even interregional settings. To say this is not in any way to denigrate these accents–they are obviously linguistically adequate and important carriers of sociolinguistic markers. But it is important to separate the functions of English in local and international communication. English as a language of international communication is not the same as speaking to your neighbour in Glasgow, Birmingham, Hong Kong or Lagos. There is little point in insisting on the right to use the particular forms and phones that mark a speaker as a speaker of a particular variety if there is a failure to communicate.

It does, of course, take two to communicate. A good deal has been written about the need for native speakers of inner speaker varieties to become more informed and tolerant listeners such that they might be better prepared to perceive and interpret unfamiliar accents of English (Phillipson 1992). There is a good deal of individual variation in how flexible listeners are in their attempts to understand what they are hearing, and probably the personal language history of the listener will be relevant in how easily they understand other accents (Cunningham 2009). In addition, experience of the accent in question will also be significant in how easily an individual can understand a particular accent (Kirkpatrick, Deterding et al. 2008; Rooy 2009).

There is a difference, however, between attempting to understand a speaker of English as a foreign language (EFL) who has a foreign accent and attempting to understand a speaker of a New English who is a speaker of English as a second language (ESL) and who has an accent associated with that variety of English. The two situations are similar, but there is a difference in speaker and listener expectations. Both speakers will have had the experience of being a learner of English and presumably of having instruction in pronunciation of English. Both will often have learned English from a teacher with their own language background in a class of others with the same background. However, the EFL speaker will often have had British or American English as a model for their learning, while the ESL speaker may well have had the New English variety in question as their model. Where there is a breakdown in communication these speakers may behave differently. The speaker of a New English has different expectations of his or her variety being met with respect and may be extremely reluctant or unable to offer alternative pronunciations, finding that an intolerable infringement of their speaker integrity. The EFL speaker may be better prepared to try different pronunciations and formulations.

Jenkins has led the way in the description of English as a language of international communication (EIL) e.g. Jenkins (2002, 2005) and where at least one party is not a native speaker of English, English as a Lingua Franca (ELF) e.g. Jenkins (2006, 2009), Berns (2008), Seidlhofer (2009) and Watterson (2008). In this context, it is unproblematic to discuss matters such as an ELF core phonology (Jenkins 2000). Crystal (2003,124) and Graddol (1997, 56) discuss the possibility that English might develop into a number of mutually unintelligible varieties, but that this would be mitigated by a parallel competence being built in a globally standard English for international communication, leading to a diglossic situation which is reminiscient of that currently operating in countries like Sweden where English is used as soon as a non-Swedish participant is involved while Swedish is used between Swedes. The data presented in this paper suggests that this may already be a necessity.

Smith and Nelson (1985) teased out the distinction between intelligibility, comprehensibility and interpretability. Intelligibility is the concern of this paper, and deals with word or utterance recognition, such that a listener would be able to transcribe an utterance which he or she finds intelligible.

Intelligibility is not an absolute. Intelligibility is a factor related to a specific speaker-listener communicative event. An utterance or a speaker cannot be said to be intelligible or not intelligible in any absolute sense. A speaker can be more or less intelligible to different speakers in different situations.

A lack of intelligibility is a problem for speaker and listener alike, and a good deal of work has been done on various aspects of intelligibility, e.g. Smith and Rafiqzad (1979), Smith & Nelson (1985), Jenkins (2002) and Berns (2008). Smith and Nelson (1985) point out that there is general agreement that it is unnecessary for every speaker of English to be intelligible to every other speaker of English, but that we do need to be intelligible to those with whom we are likely to communicate in English.

Naturally, the time is long past when native inner circle speakers are the only legitimate judges of what is intelligible, and few would maintain that native speakers are automatically more intelligible than non-native speakers e.g. Smith and Rafiqzad (1979). As the number of speakers for whom English is one of a number of languages grows and has long ago exceeded the number of so-called native (monolingual?) speakers of English, the imagined native speaker is not often the implied interlocutor for learners of English in either EFL or ESL situations.

This study uses data from Nigerian students and thus it is relevant to consider the role and status of English in Nigeria. A good deal has been written on this topic which is confounded by the multitude of languages spoken in the country (some 400 in some sources e.g. Gut and Milde 2002). The colonial history of countries such as Nigeria have led to a situation where English is retained as a language of business, education and media as well as interethnic communication (Gut 2007), although, Nigerian Pidgin English also serves for interethnic communication. Due to a complex mesh of factors including linguistic attitudes and language policies in the outer circle countries in general and Nigeria in particular, these speakers may not appreciate their first languages, sometimes referring to them disparagingly as dialects, vernaculars or local languages. A good deal has been written and will continue to be written about the need for African languages to take a more prominent role in the lives of the people of Africa, e.g. (Prah 2002). The role of English in Nigeria, as elsewhere in Africa, and the attitudes of Nigerians to English and other Nigerian languages are sensitive topics.

The distinction between second language varieties of English such as Standard Nigerian English and learner varieties of those with Standard Nigerian English as their target variety is far from clear cut. The nature of the relationship between English-based varieties in Nigeria has not, to my knowledge, been fully explored. In other comparable postcolonial contexts a continuum has been described which spans from a basilect, perhaps represented here by Nigerian Pidgin English to an acrolect which would be close to the British English which was the variety once imposed upon Nigeria, as suggested by Ufomata (1996).

Adamo (2007) writes that "English has itself (to a certain extent at least) become a Nigerian language". She points to nativization of English as indexical of its integration into the culture of the community. Like the Nigerian author, Achebe, she sees Nigerian English as having "communion with its ancestral home but is altered to suit its new surrounding" (Achebe 1975). She writes further that "When a people are alienated from their language(s), as is the case in Nigeria today, they gradually become alienated from their culture" She argues that English, however nativized, will not serve as a national language, and calls for an indigenous language to take that place. At the same time she is realistic and points to the efforts made to standardize, nativize and codify Nigerian English to enable it as a carrier of Nigerian culture.

The status of Nigerian English as a variety of English has been questioned (Ajani 2007). This is certainly a central question if we are to be able to decide whether the English spoken in Nigeria is a variety of English which can carry a culture or if we are to regard it as a learner variety. In the words of Kachru "what is 'deficit linguistics in one context may be a matter of 'difference' which is based on vital sociolinguistic realities of identity, creativity and linguistic and cultural contact in another context" (Kachru 1991). Ajani (2007) sets the position of a standardised Nigerian English against the early position of English teachers in Nigeria who refuse to accept any model but the native British model. Ajani relates this debate to the US Ebonics debate, rejecting AAVE as a legitimate variety for use in education. He further questions whether speakers of one of the 400 languages of Nigeria, e.g. Hausa, will sound the same when speaking English as will a speaker of another language, e.g. Yoruba or Igbo.

Bamgbose (1982) views the emergence of a Nigerian English as a natural outcome of the language contact situation in the country. He accounts for three mechanisms at work in generating usages in the Nigerian English: the interference, deviation and creativity approaches involving "interference" from the mother tongue (or possibly Nigerian Pidgin English), "deviation" from the native British norm and the creative inclusion of elements of local languages as well as English to create new items respectively. Bamgbose rejects the native model for Nigerian learners and suggests that the educated speaker of Standard English be the model. This standard has not, however been well described.

Schneider (2003) compares the evolution of postcolonial Englishes in language contact situations to the acquisition of a second language such that the phonology of such new varieties will display features that resemble transfer from the phonology of "indigenous languages". This view is shared by Hickey (2004:519) who writes on cluster simplification in Asian and African Englishes that "this is determined largely by the phonotactics of the background language(s). In the case of Nigeria, there are a multitude of such background, or substrate languages. It is estimated that almost 400 languages are spoken in Nigeria (Bamgbose 1971, Agheyisi 1984). This does, of course, depend on how the languages are defined. Prah (2009) claims that the number of languages, as defined by criteria of mutual unintelligibility might be far fewer. He states,

> "What is not easily recognized by many observers is that most of what in the literature, and classificatory schemes, on African languages passes as separate languages in an overwhelming number of cases are actually dialectal variants of "core languages." In other words, most African languages can be regarded as mutually intelligible variants within large clusters (core languages)."

Ufomata (1996) offers an account of the continuum that exists with native-like accents at one end (deemed essential for a career as newsreader) and "other varieties which can be defined negatively in relation to these standard accents". Ufomata goes on to say that the Nigerian standard is socially accepted and internationally intelligible. Bamgbose (1995) suggests that this accent should be taught in schools. Ufomata accounts for some of the main features of Educated Spoken Nigerian English, describing them with reference to RP phonemes. These are:

- The vowels of *ship and sheep* are both pronounced [i]
- *Food* and *foot* are both pronounced with [u]

- *Bath* and *bag* are both pronounced with [a]
- The vowels of *play* and *plough* are monophthongized to [e] and [o] respectively
- The initial consonants of *thin* and *then* are pronounced [t] and [d] respectively
- Heavy nasalization of vowels preceding nasals and the dropping of word-final nasals.

Previous work on the intelligibility of Nigerian English has indicated that rhythm and intonation are the biggest problem (Stevenson 1965). Syllables that would be unstressed in other varieties of English may not be reduced in any way in this variety. This study will add to our knowledge about the intelligibility problems experienced by Nigerian English speakers and their non-Nigerian interlocutors.

Banjo (1971:169-70) in an often cited account describes four discrete varieties of Nigerian ranging from what he calls Variety 1 which is marked by wholesale transfer of phonological, syntactic, and lexical features of Kwa or Niger Congo to English, spoken by those whose knowledge of English is very imperfect and neither socially acceptable in Nigerian nor internationally intelligible, through Variety 2 and Variety 3 which are described as progressively closer to standard British English in syntax, semantics and lexis, though still different in phonetic features with increasing international intelligibility to Variety 4 which he describes as identical to standard British English. This last may correspond to the "newsreader variety" described by Ufomata (1996). It seems likely that there is in fact a continuum ranging perhaps even from a basilect represented by Nigerian Pidgin English through Standard Nigerian English to the British-like acrolect.

2. Material and Methods

The three students who have provided the stimuli for this study are young men aged between 23 and 34 from Nigeria. They came to Sweden to study a bachelor's programme in English language, literature and culture. When they arrived to take up their studies they took part in the needs analysis mentioned above, and all three of them were found to be have an inadequate level of English proficiency on both their oral skills (receptive and productive) and their mastery of standard English grammar. The students involved in this study have been educated in English-medium schools since primary school. When asked which is their first or native language, all three indicated that English was their first language. This is in spite of the fact that further enquiry revealed a) that they did not encounter English until they began primary school, b) that English was not the language they used to speak to each other, choosing the Nigerian language Igbo for that purpose in the case of two speakers (the third speaker did not speak or understand Igbo) and Nigerian Pidgin English otherwise, c) that English was not the language they used to talk to their families and d) that their English was not a language they mastered in terms of grammatical consistency, vocabulary size and written or spoken fluency according to the results of our needs analysis. Their English appears to all intents and purposes to be a learner variety. The distinction between learner varieties and New Englishes is, of course, not always easy to draw, and these young men have presumably had Nigerian English as a model and target for their English learning.

The 21 listeners were recruited from among students and staff at a Swedish university, both those who regularly come into contact with international students and those who do not. Seven of the listeners were native speakers of English from the England, the US, Scotland, Ireland and Australia and 14 were non-native speakers of English with French, Swedish, Russian, Italian, Finnish and German as their first languages. Six of the native English speakers and five of the non-native English speakers had extensive experience hearing international Englishes of many kinds through contact with our extremely international student body. Others had less such contact and experience.

The three speakers each recorded a set of material including a text, a wordlist, a set of words contrasting high front vowels and postvocalic consonant voicing embedded in carrier phrases in phrase final and non phrase final position, a set of semantically meaningful sentences and a set of semantically unpredictable (but still grammatical) sentences and a set of true/false questions. The last three items on this list are the same material as used in another study reported in Munro and Derwing (1995). The stimuli used in this study were selected from the semantically meaningful sentences. These sentences were designed to include some sounds and sound combinations that are generally challenging for many ESL and EFL speakers in sentences where the contexts is not especially helpful to the listener. In other words, comprehension will not be an aid to intelligibility, while the sentences are still considerably more natural than the test words in carrier phrases that were also recorded.

Eight sentences were used in this study, uttered by speaker N1 apart from sentences 5 and 8. Sentences 4 and 8 are the same, but were spoken by two different speakers.

1 A big farmer lifts a large load.
2 A confident guy viewed a natural scene
3 A fair judge gives a second chance
4 A hundred sheep took a dangerous trip
5 My girl climbed a red car (speaker N2)
6 A pool is better than seventeen orange trees
7 A thin lady taught a musical language
8 A hundred sheep took a dangerous trip (speaker N3)

Speakers were presented with the stimuli using an online test facility built into the learning platform used at the university. Listeners heard the utterances individually through headphones and they could listen as many times as they wanted to the utterance and were then asked to write what they heard. They could take the test online at a time convenient to themselves.

3. Results

Table 1 shows the results provided by the listeners for the first sentence, *A big farmer lifts a large load* as uttered by speaker N1. The listener responses are divided into those obtained from native vs. non-native speakers of English, and those used or not used to international Englishes. As can be seen, the responses were very varied, from the imaginative *It is summer, live the blue life* to two cases, one native and one non-native listener who heard the utterance as intended by the speaker.

Used to International Englishes?	NS	NNS
yes	It is summer, live the blue life A big farmer lifts a large loot A big farmer leaves a large lodge The big farmer lives large loge A big farmer lives a large looge? A big farmer lifts a large load.	A big farmer leaves a large Luke. The big farmer lives in large lu??? A big farmer lifts a large load. A big farmer lives in a large luge A big farmer lives a large luuk?? A big farmer lives in large louge
no	A big farmer lives a large look.	A big farmer lives a large loot The big farmer lives in a large loot A big farmer lives a large... A big farmer lives in large ? A big farmer leaves a large look... A big farmer lives in a large loudge The large farmer lives a large lodge The big farmer leaves/lives a large?

Table 1. Responses from native and non-native English speakers used and unused to international Englishes listening to speaker N1 saying *A big farmer lifts a large load.*

What we see here is that the listeners have difficulty reconstructing the elided /t/ in *lifts*; they are unsure whether the intended vowel gives *leaves* or lives. They are interpreting the word *load* produced with a [u] *loot, look or Luke* to name just a few, and the speaker's slightly affricated /d/ in *load* is interpreted as *lodge or large.* The listeners are doing their best to listen with an open mind as they try to make sense of the utterance. This leads to incomprehensibility as well as unintelligibility in Smith and Nelson's (1985) terminology.

Other stimuli sentences produce similarly creative reconstructions as listeners do their best to comprehend the only sporadically intelligible speech of the speakers. Table 2 summarizes the responses, with the intended word at the left of each row and the listener perceptions in subsequent columns.

Speaker N1					
a	a 5	the15			Other 1
lifts	lifts 3	leaves 4	lives 13		Other 1
load	load 2	loot 3	luke/luuk/look 4	lodge etc. 7	Other 5
confident	confident 16	competent 1	coffee 3		Other 1
guy	guy 4	car 3	guard 4	girl 3	Other 7
viewed	viewed 17	filled 1			Other 3

scene	scene 9	sin 2			Other 10
fair	fair 10	friend 2			Other 9
judge	judge 5	choice 4	church 1	george 6	Other 5
gives	gives/give 15				Other 6
sheep	sheep(s)14	ship(s) 3			Other 4
dangerous	dangerous 18				Other 3
pool	pool 17	poo 2			Other 2
orange trees	orange tree(s) 16				Other 5
thin	thin 6	teen/team 8	tin 2		Other 5
lady	lady/ladies 16				Other 5
Speaker N2					
girl	girl 5	gate/gay 5			Other 11
climbed	climbed 13				Other 8
red	red 17				Other 4
car	car 14				Other 7
Speaker N3					
sheep	sheep(s) 14	ship(s) 5			Other 2

Table 2. Summary of intelligibility issues in all eight stimuli sentences spoken by speakers N1, N2 and N3 showing numbers of responses

So what we see here is that speaker N1 (like N3) does not distinguish between the vowels in e.g. *sheep* and *ship* as evidenced by the confusion experienced by listeners in these words as well as *lifts*, *scene* and *thin*. As mentioned above, his reduction of consonant clusters or affrication of consonants in the coda lead to misperception of the words *lifts*, *competent* and *judge*. We can further note that his realisation of post vocalic nasals as nasalised vowels misleads or confuses the listeners in the words *confident*, *scene* and *thin*. His monophthong pronunciation of the vowels in *guy* and *fair* causes many listeners to guess wildly at the speaker's intention. For speaker N2, the very open [a] pronunciation of the vowel in *car* confuses a third of the listeners, while only less than a quarter of the listeners could reconstruct *girl* from what they actually heard.

The listeners who came closest to hearing the speakers' intended words were both native and non-native speakers of English, but they were both quite used to hearing international Englishes. The listeners who did least well were in one case a native speaker who does in fact have experience of international Englishes, and the non-native inexperienced listeners.

4. Discussion

There is nothing unexpected about the results reported above. Jenkins (2000, 2002) has posited that certain parameters need to be upheld if speech is to be internationally intelligible. These speakers of Nigerian English, perhaps even Standard Nigerian English, as described by Ufomata (1996) and Bamgbose (1982) do not maintain the distinctions outlined by Jenkins, and their speech as elicited for this study is patently not

intelligible to the non-Nigerian native and non-native speakers of English who are listening to them.

Some descriptions of Nigerian English compare the variety to RP as a target variety, e.g. Ufomata (1990, 1996). But the question of the status of Nigerian English as a variety of English or a New English is very relevant here. If Nigerian English is a legitimate variety of English, there is no reason why it should not be used as a model for Nigerian learners of English. Eka (2003:35) writes that this is "the variety of world Englishes spoken and written by Nigerians within the Nigerian environment". So the question of whether or not the features of Nigerian English are to be viewed prescriptively as errors or descriptively as features of Nigerian English depends of the speakers' intentions. If they are intending to speak Nigerian English, they are not making errors – they are succeeding in their intention. But if they are aiming at a more internationally intelligible variety, then the features of their pronunciation can be seen as errors and may be corrected if the students take part in classes in English pronunciation (which the speakers in this study actually did as part of their course in Sweden. This Nigerian English is not a language of wider communication as defined by Bamgbose (1991).

Smith and Nelson (1985) suggest that if a listener expects to understand a speaker it is more likely that this will indeed be the case. Nonetheless, the listeners in this study do appear to expect certain things of the utterances they hear. In line with the ideas expressed in Jenkins' Lingua Franca Core (Jenkins 2002), there are some sounds that should not be elided and some vowel distinctions that should not be neutralised if intelligibility is to be maintained.

It is not only the pronunciation that is affected by the first language. Listeners will listen according to the salient cues to vowel and consonant identity, voicing, etc. that operate in the languages they speak, particularly in their first languages. Native speakers of English will identify postvocalic voicing in words like *bat* vs *bad* according to the length of the vowel rather than the vocal fold vibration (voicing) during the stop phase of the postvocalic consonant. In fact, in the speech of many individuals, the stop will be devoiced, though still lenis (Cruttenden 2008). If a speaker of another variety is transmitting other cues to postvocalic voicing but failing to shorten the vowel before a voiceless consonant, the native speaking listener may fail to pick up on the intended voicing. In any kind of communication involving speakers of different varieties, listeners need to be as flexible as they are able to be, although, unless they have considerable experience of listening to a particular speaker they may not be able to read the cues transmitted by the speakers.

Levis (2005) explicates the difference between nativeness and intelligibility as learner targets (see also Cunningham (2009)). Hung (2002) questions the need to "improve" non native pronunciation of English. He asks why teachers should modify learners' naturally acquired phonology of English and when it is worth the learners' and their teachers' efforts to do so. The answer Kirkpatrick, Deterding et al. (2008) offer to the question is that intelligibility criteria must be decisive here. The research of Kirkpatrick, Deterding et al has taken place in the Hong Kong context. In Nigeria too, we are dealing with learners of English as a second, not a foreign language and Nigerian English is a Nigerian language and is used to convey speaker identity. International intelligibility may not, however, be high on speakers' lists of priorities. Failure to speak in a way that is intelligible to a wider circle of listeners than that found in a local

Nigerian context is only problematic if the speech is indeed directed to non-Nigerian listeners. Even then, it is no more acceptable to insist that a Nigerian English speaker change his or her pronunciation to suit the listener than it would be to require the same of a Welsh, Australian or Northern Irish speaker.

There are two ways to go here. The Nigerian (Welsh / Australian / Irish) English speaker can adjust his or her pronunciation, moving along the continuum to a less regionally marked pronunciation, if he or she has access to such a variety, or the listener can learn more about Nigerian (Welsh / Australian / Irish) English in order to become a more experienced and "in tune" listener, what Catford would have described as "lowering one's intelligibility threshold" (Catford 1950). Now in the case of a non-Nigerian listener who is in Nigeria, the latter alternative is reasonable and realistic, but in the case of a Nigerian English speaker in the diaspora, it is not realistic to expect one's listeners to be prepared for perceiving Nigerian English. The speaker must adjust his or her speech or face having interlocutors miss a good deal of what is said.

In discussion of the use of English as a language of international communication, or English as a lingua franca, mutual intelligibility is a major concern (Cunningham 2009; Rooy 2009). Without intelligibility, communication is severely hampered. If speakers of Nigerian English mean to use their English as a language of wider, or international communication, they need to move along the continuum that is Nigerian English to a point where they avoid those features that are least helpful to their listeners such as the realisation of postvocalic nasals as vowel nasalisation, the elision of postvocalic /l/ and the mapping of English vowels onto a severely reduced set of vowels. This does not in any way mean that they need to speak Standard Southern English, or even to sound anything but Nigerian. It is fully possible to signal one's identity in accent without impairing intelligibility. The educated Nigerian speaker, just as the educated Northern Irish, Scottish or Indian speaker, needs to have access to more than one register. There are situations when such speakers will want to move in the other direction, back along the Nigerian English continuum, when for reasons of credibility, integrity, solidarity and identity it is necessary and desirable to enhance the very pronunciation features that impair international intelligibility.

To conclude then, it would seem that whatever legitimacy this variety might have in a national Nigerian context, it is not particularly useful for communication outside the Nigerian context. If speakers intend to make themselves understood in a pan-African context or further afield such as is the position of the students who come to Europe to study, they will need to modify their pronunciation. This is true of all peripheral varieties, or indeed perhaps all varieties where Jenkins' Lingua Franca Core features are not a part of the phonology. Certainly speakers of some Scottish or Northern Irish varieties of English also need to modify their pronunciation when interregional or international intelligibility is at stake. Efficient communication is a two-way affair. It relies upon speakers and listeners meeting in their expectations, and there will usually be an accommodation of interlocutors to each other (Coupland 1984).

However, it is necessary to balance the phonetic integrity of the speaker with the needs of the listener. Nigerian English is a member of the family of English languages (McArthur 2002). But the speaker needs to have access to a point high enough on the basilect-acrolect continuum that is Nigerian English if international intelligibility is to be achieved. There is a clear need for teaching in English for international communication

alongside teaching of Standard Nigerian English if Nigerians are not to cut themselves off from international discourse and the wider international community.

In many parts of Africa parents are reported to be enthusiastically seeking English medium schooling for their children from an early age, even from preschool in many cases. A number of African nations have implemented legislation stipulating that children will be educated through the medium of English either from the start or from a certain age. This is far from uncontroversial, as both political opinion and research in bilingual education suggest that children might learn better in the language or languages they actually speak than in a foreign language (Prah 2002; Garcia, Skutnabb-Kangas et al. 2006). The empowerment of the languages of Africa is an important issue and the use of indigenous language in African schools is held by Prah and others to be the only way forward if more than a small English-speaking elite are to have access to academic success. One of the reasons why English-medium schooling is sought after by parents is that they believe it will give the children access to a language of wider communication. While this is the case in many African nations, it may not be the case in Nigeria. In Nigeria, children are schooled in English from an early age, but the variety of English used is naturally Nigerian English. Nigerian English speakers who do not gain access to a more acrolectal variety of Nigerian English as part of their education will not be intelligible to either their fellow Africans or to the wider international community. While the English that is needed as a language of wider communication need not be restricted to the Lingua France Core, Seidlhofer (2009: 243) points out that "ELF and postcolonial Englishes are very different realities on the ground.". The political desire to view all varieties of English as mutually intelligible must not be allowed to stand in the way of speakers of Nigerian English from acquiring a more widely understood pronunciation.

References

Achebe, Chinua. (1975). *Morning yet on creation day*. London: Heinemann.

Adamo, Grace. E. (2007). Nigerian English. *English Today* 231: 42-47.

Agheyisi, Rebecca. N. (1984). Minor languages in the Nigerian context. *Word* 35: 235-253.

Ajani, Timothy T. (2007). Is there indeed a Nigerian English? *Journal of Humanities and Social Sciences* 1:1. Available at http://www.scientificjournals.org/journals2007/j_of_hum.htm Accessed 27 June 2009.

Bamgbose, Ayo. (1971). The English language in Nigeria. In *The English language in West Africa*, ed. J. Spencer. London: Longmans.

Bamgbose, Ayo. (1982). "Standard Nigerian English: Issues of identification".In *The other tongue: English across cultures*, Ed.Braj. B. Kachru. Urbana: University of Illinois Press.

Bamgbose, Ayo. (1991). *Language and the nation*. Edinburgh: Edinburgh University Press.

Bamgbose, Ayo. (1995). English in the Nigerian environment. In *New Englishes: A West African perspective*, ed. Ayo Bambgose, Ayo Banjo and Andrew Thomas, 9-26. Ibadan: Mosuro.

Banjo, Ayo. (1971). On codifying Nigerian English: research so far. In *New Englishes: A West African perspective*, ed. Ayo Bambgose, Ayo Banjo and Andrew Thomas, 1995, 203-231. Ibadan: Mosuro.

Berns, Margie. (2008). *World Englishes, English as a lingua franca, and intelligibility*, London: Blackwell Publishing.

Bern, Margie. (2009). English as a lingua franca and English in Europe. *World Englishes* 28 (2): 192-199.

Catford, John C. (1950). Intelligibility. *English Language Teaching* 1: 7-15.

Coupland, Nikolas. (1984). Accommodation at work. *International Journal of Sociolinguistics* 4-6: 49-70.

Cruttenden, Alan. (2008). *Gimson's Pronunciation of English*. London, Hodder Arnold.

Crystal, David. (2003). *English as a Global Language*. Cambridge: Cambridge University Press.

Cunningham, Una. (2009) Models and targets for English pronunciation in Vietnam and Sweden. *Research in Language*, 7, 113-128.

Cunningham, Una. (2009). Quality, quantity and intelligibility of vowels in Vietnamese-accented English. In *Issues in Accents of English II*: *Variability and Norm*, ed. Ewa Waniek-Klimczak. Newcastle: Cambridge Scholars Publishing Ltd. 3-22.

Eka, David. (2003). The English language: changes and chances within the Nigerian environment. *Journal of Nigerian English and Literature* 4: 32-41.

Garcia, Ofelia, Tove Skutnabb-Kangas, and Maria E. Torres-Guzma. (2006). *Imagining Multilingual Schools: Languages in Education and Glocalisation.* Clevedon: Multilingual Matters.

Graddol, David. (1997). *The future of English*. London: British Council.

Gut, Ulrike. (2007). First language influence and final consonant clusters in the new Englishes of Singapore and Nigeria. *World Englishes* 26: 346-359.

Gut, Ulrike and Jan-Torsten Milde. (2002). The prosody of Nigerian English. Available at http://www.lpl.univ-aix.fr/sp2002/pdf/gut-milde.pdf . Accessed 27 June 2009

Hickey, Raymond. (2004). Englishes in Asia and Africa: origin and structure. In *Legacies of colonial English*, ed. Raymond Hickey. Cambridge: Cambridge University Press: 503-535.

Hung, Tony. T. N. (2002). English as a global language and the issue of international intelligibility. *Asian Englishes* 5: 4-17.

Jenkins, Jennifer. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Jenkins, Jennifer. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics* 231: 83-103.

Jenkins, Jennifer. (2005). Implementing an international approach to English pronunciation: The role of teacher attitudes and identity. *Tesol Quarterly* 393: 535-543.

Jenkins, Jennifer. (2006). Current perspectives on teaching world Englishes and English as a lingua franca. *Tesol Quarterly* 401: 157-181.

Jenkins, Jennifer. (2009). English as a lingua franca: interpretations and attitudes. *World Englishes* 28 (2): 200-207.

Kachru, Braj. (1992). Teaching world Englishes. In *The Other Tongue, English across Cultures*, ed. Braj B. Kachru. Urbana, University Illinois Press.

Kachru, Braj. B. (1991). Liberation linguistics and the Quirk concern. *English Today* 25: 3-13.

Kirkpatrick, Andy., David Deterding, and Jennie Wong. (2008). The international intelligibility of Hong Kong English. *World Englishes* 273-4: 359-377.

Levis, John. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *Tesol Quarterly* 393: 369-377.

McArthur, Tom. (2002). *The Oxford Guide to World English*. Oxford, Oxford University Press.

Munro, Murray J. and Tracey M. Derwing. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech* 38: 289-306.

Phillipson, Robert. (1992). *Linguistic Imperialism*. Oxford, Oxford University Press.

Prah, Kwesi K., Ed. (2002). *Rehabilitating African languages : language use, language policy and literacy in Africa: selected case studies* Cape Town, South Africa Centre for Advanced Studies of African Society CASAS.

Prah, Kwesi K. (2009). The burden of English in Africa: From colonialism to neo-colonialism. Lecture presented at *Mapping Africa in the English-Speaking World.* University of Botswana.

Rooy, Susan. C. V. (2009). Intelligibility and perceptions of English proficiency. *World Englishes* 28 (1): 15-34.

Schneider, Edgar. (2003). The dynamics of new Englishes: from identity construction to dialect birth. *Language* 79: 233-81.

Seidlhofer, Barbara (2009). Common ground and different realities: world Englishes and English as a lingua franca. *World Englishes* 28 (2): 236-245.

Smith, Larry E. and Cecil. L. Nelson (1985). International intelligibility of English: directions and resources. *World Englishes* 4(3): 333-342.

Smith, Larry E. and Khalilullah Rafiqzad (1979). English for cross-cultural communication - question of intelligibility. *Tesol Quarterly* 133: 371-380.

Stevenson, K. J. (1965). Reflections on the teaching of English in West Africa. *Journal of Nigeria English Studies Association* 32: 227-235.

Ufomata, Titi. (1990). Acceptable models for TEFL. In *Studies in the pronunciation of English: A commemorative volume in honour of A.C. Gimson*, 212-218, ed. S. Ramsaran. London & N.Y.: Routledge.

Ufomata, Titi. (1996). Setting priorities in teaching English pronunciation in ESL contexts. London, UCL. Available at http://www.phon.ucl.ac.uk/home/shl9/ufomata/titi.htm. Accessed 27 June 2009.

Watterson, Matthew. (2008). Repair of non-understanding in English in international communication. *World Englishes* 273-4: 378-406.

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0044-7

INITIAL GLOTTALIZATION AND FINAL DEVOICING IN POLISH ENGLISH

GEOFFREY SCHWARTZ
UAM Poznań

Abstract
This paper presents an acoustic study of the speech of Polish leaners of English. The experiment was concerned with English sequences of the type *George often*, in which a word-final voiced obstruent was followed by a word-initial vowel. Acoustic measurements indicated the degree to which learners transferred Polish-style glottalization on word-initial vowels into their L2 speech. Temporal parameters associated with the production of final voiced obstruents in English were also measured. The results suggest that initial glottalization may be a contributing factor to final devoicing errors. Adopting English-style 'liaison' in which the final obstruent is syllabified as an onset to the initial vowel is argued to be a useful goal for English pronunciation syllabi. The implications of the experiment for phonological theory are also discussed. A hierarchical view of syllabic structures proposed in the Onset Prominence environment allows for the non-arbitrary representation of word boundaries in both Polish and English.

1. Introduction

In the development of English pronunciation teaching materials, issues of phonological representation may lead to conflicting strategies with regard to given aspects of the target language phonology. For example, the *ship-sheep* contrast may be described in terms of a number of different phonological features, including [tense], [ATR] and [long]. This variety of description can confuse learners and teachers alike, and lead to undesirable results. I have heard many learners, presumably on the basis of the descriptions of "long /i:/", produce unnaturally long vowels in words such as *sheep*. In this connection, it may be worthwhile to go back and re-evaluate traditional phonological descriptions of target-language segments with the goal of increasing teacher and learner awareness of their most salient properties. For the *sheep-ship* contrast, studies such as Kaźmierski (2007) have suggested that the dynamic properties of English vowels (Strange 1989), in particular diphthongization, are worth focusing on. With regard to the nasal /ŋ/, Schwartz (2011) found that learners' tendency for stop insertion in words such as *singer* may be alleviated be revising the traditional 'velar' description of the sound. Briefly stated, it some cases it may be worth re-evaluating our descriptions and representations of the most difficult aspects of L2 speech.

The English voice contrast in word-final position may represent another candidate for this kind of representational refinement. Due to aerodynamic factors, final voiced obstruents present a phonetic challenge for foreign leaners in the acquisition of English phonology. Final devoicing (FD), a well-known feature of the phonology of many

languages, is one of the more frequently cited contributors to a Polish accent (e.g. Gonet and Pietroń 2004) in English. Its avoidance is a priority in ESL pronunciation teaching. Final devoicing frequently occurs in the native language as well, but without the neutralization of the laryngeal contrast, the preservation of which may be attributed to two phonetic parameters. The first, the relative duration of the final consonant and its preceding vowel, is an example of well-documented phonetic universal (Chen 1970, Maddieson 1997) that English has chosen to exaggerate (Port and Dalby 1982). Vowels are clipped before unvoiced consonants, which are longer in duration. Voiced consonants are shorter in duration, and the vowels preceding them are longer.

English speakers often employ an additional strategy in overcoming the phonetic challenge of final voiced obstruents. They have a tendency to 'liaise' the final consonants with the beginning of the following word, especially if that word begins with a vowel. In other words, phrases such as *hold on* and *tries it* are generally pronounced by native speakers as if they were written *whole Don* or *try zit*. As a result of this process, the final obstruent in question loses its 'final' status, and is no longer in the environment for FD to apply. Liaison is described in most teaching materials I have seen, but it is usually relegated to descriptions of 'connected speech phenomena' that do not comprise the main focus of textbooks. In the context of Polish instruction in English pronunciation, final devoicing is emphasized as an area of L1 interference that must be overcome. While the durational properties discussed above do get some mention, liaison is rarely mentioned in connection with final voicing.

With regard to liaison, the Polish and English phonological systems are diametrically opposed. Liaison in English results in a rearrangement of syllabic affiliation – the final consonant becomes an onset to the first syllable of the following word. This process may be motivated by an apparent universal preference for syllables with consonantal onsets. In Polish, resyllabification across word boundaries is impossible (Rubach and Booij 1990). The preference for consonantal onsets is satisfied by means of an alternative strategy: glottal stop insertion. Glottal stops may be claimed to fill an 'empty onset' position to repair a non-optimal syllabic structure. However, glottalization may have further prosodic implications, underlying the 'initial' status of vowels at the start of a word, and reinforcing the 'finality' of the preceding consonant, thereby preserving the context for final devoicing. As a consequence, although FD in Polish English is generally described as a simple segmental error, it may have far-reaching phonological consequences. In particular, the study presented here touches on the question of how Polish and English differ with respect to the representation of word boundaries.

These phonological considerations suggest that in Polish English we might look for a correlation between FD and glottal stop insertion - we would predict that speakers who glottalize initial vowels in English should be more likely to devoice final obstruents. A preliminary study (Rejniak 2011) of a corpus of Polish English speech suggests that such a correlation indeed exists. An impressionistic analysis found that the number of devoicing errors rose in accordance with the number of glottal stop insertions. This paper will present the results of an acoustic study of Polish English speech that seeks to investigate this correlation. After some discussion of the phonetic parameters under study in Section 2, the experimental procedure and results are described in Section 3. Section 4 offers a new phonological perspective on these issues, and Section 5 concludes.

2. Phonetic aspects of (de)voicing and glottalization

The phenomenon of final obstruent devoicing is a well-known feature of a large number of languages, and is particularly prevalent in the languages of Europe, including most members of the Slavic family, German, Dutch, and Catalan. It may be seen as a necessary aerodynamic consequence of the final portion of a sequence of speech sounds, during which airflow through the vocal tract has a natural tendency to diminish. Since airflow through the glottis is what makes voicing possible, the decrease of airflow is expected to be accompanied by a lack of vocal fold vibration. As a result of this challenge, languages that maintain laryngeal contrasts in final position often employ additional strategies to produce a distinction. For example, in French one may often observe a short vowel after the release of a final voiced consonant, suggesting that extra effort has been made to maintain the airflow required for voicing. Vowel intrusion may be seen as a process that is parallel with the classic liaisons in phrases like *les hommes* [le zɔm] 'the men'. The result is a syllable-initial consonant during which it is easier to maintain voicing.

Before pauses and consonant-initial words, final obstruent devoicing often occurs in English, particularly in the case of fricatives. However, the "voice" contrast is preserved through exploitation of a known phonetic universal (Maddieson 1997): vowels are longer before voiced consonants than before voiceless ones. The magnitude of this difference is much greater in English than in other languages (Chen 1970), so we may assume that English has exaggerated this phonetic property in order to keep the laryngeal contrast readily perceptible in the face of weak or absent vocal fold vibration in final consonants. Alongside this difference in vowel duration, we find that consonants too are also marked by universal voice-related durational properties: voiced consonants are shorter than voiceless ones. English may be claimed to exaggerate this property as well. While this fact is widely noted with regard to aspirated initial stops, the extended duration of final voiceless consonants in English, a feature described in experimental phonetic studies (e.g. Port and Dalby 1982), is not a priority of most English pronunciation materials.

While Polish is if course known for final obstruent devoicing, the Southern and Western regions of the country have been observed to exhibit voicing between vowels across word boundaries. This process, known as Poznań-Cracow voicing, has been found to neutralize the laryngeal contrast in favor of the voiced variant, so the phrase *brat Ewy* 'Eve's brother' is pronounced as [bradevɨ]. This voicing process may conceivably be interpreted as a form of 'liaison' that blocks the insertion of glottal stops, which are voiceless. Our acoustic study includes four speakers from the Wielkopolska region where this voicing process is attested.

The term glottalization may be associated with two different phonological phenomena. In the study of English accents and pronunciation, the terms glottalization and glottaling are frequently associated with a process by which /t/ is replaced by glottal stops. As an allophone of /t/, the glottal stop is commonly assumed to be the result of a lenition process in casual speech. By contrast, the glottalization of word-initial vowels serves as a marker to a prosodic boundary. It represents a form of strengthening, making the syllable boundary more robust for listeners. Our focus in this paper will be on the glottalization of word-initial vowels.

In English, initial glottalization has been found to be largely dependent on higher-level prosodic structures. That is, it most frequently appears on word-initial vowels at phrase boundaries, but not within a phrase. For example, in a study based on a corpus of radio announcers' speech, Dilley et al (1996) found that glottalization rates for phrase-initial vowels were around 60%, while word-initial vowels within phrases were glottalized around 20% of the time. In Polish, glottalization appears to be a syllable-level process, motivated by the preference for consonantal onsets. The process has been reported to be present on word-initial vowels (Dukiewicz and Sawicka 1995, Gussmann 2007) without reference to phrase-level structures. It may even be found within words on morpheme-initial vowels: *nauka* 'science' may surface as [na^{ʔ}uka]. As a result, although there is little published data that is comparable to the studies describing English, it is reasonable to assume that glottalization in Polish is more widespread than in English, which largely limits the process to phrase-initial position.

One important aspect of glottalization that may be observed in both initial vowels as well as glottalized allophones of /t/ is phonetic variability. While a canonical glottal stop is characterized by a full closure, this feature often fails to surface in natural speech. This is especially true in the case of intervocalic glottal stops, which may be perceived on the basis of drops in pitch and small irregularities in the periodicity of the vocal wave. The various irregularities have been described for English in Redi and Shattuck-Hufnagel (2001) and Ashby and Przedlacka (2011). As it turns out, in the study described in this paper it will be possible to describe glottalization in terms of the duration of full closure. This is due to two factors: (1) we will not analyze glottalization at vowel hiatus where full closure is not often achieved, and (2) we will analyze second language speech, in which casual speech processes such as the reduction of glottal closure should be relatively infrequent.

3. Experimental method

This section describes an acoustic study of Polish English speech. Our experiment will address the following questions.

1. To what extent do Polish speakers transfer initial glottalization into their English speech?
2. What effect does initial glottalization have on the realization of final voiced obstruents in Polish English?
3. Do speakers from dialect regions associated with Poznań-Cracow voicing show different behavior in their L2 with regard to these parameters?

3.1. Subjects and Data

10 first year students of English at the School of English at Adam Mickiewicz University in Poznań, Poland participated in the acoustic study. The students were recorded in a soundproof recording booth. The linguistic materials were comprised of a list of English sentences containing sequences of word-final voiced obstruent(s) + word-initial vowel, such as *George often*, *today's express train*, *Fred's aunt*. The data set

included 20 such sequences, as well as additional sentences to control for list reading effects. The sentence list was read twice by each subject, producing 40 sequences for analysis per speaker*10 speakers = 400 tokens for analysis. A native speaker of American English also read the sentence list.

3.2. Acoustic measurements

The acoustic measurements were performed by hand using the Praat program. The following measurements and calculations were made.

1. Duration in milliseconds of vowel preceding final consonant (VD)
2. Duration in milliseconds of final consonant (CD). For stops and affricates this measurement combined both closure and noise bursts/frication.
3. V/C ratio: (VD/CD)
4. Duration in milliseconds of glottal closure (GC) from end of consonant noise to onset of voicing on the vowel.
5. Duration in seconds of each sentence (RATE), allowing for the control of speech rate.

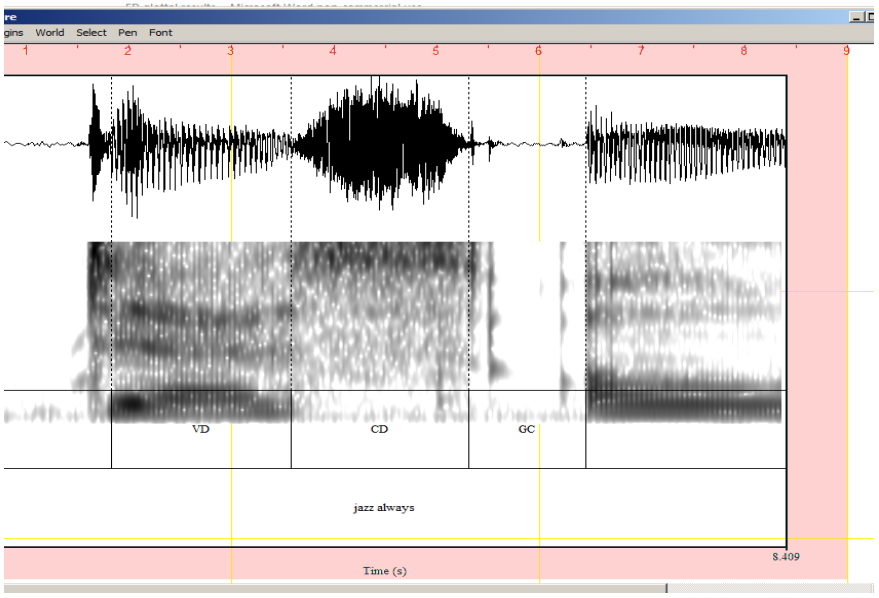


Figure 1 – Illustration of acoustic measurements for the sequence *jazz always*. VD represents vowel duration (142ms), CD consonant duration (140ms), and GC indicates glottal closure duration (92 ms).

Figure 1 presents an illustration of the acoustic measurements on the sequence *jazz always*. The following measurements (Rate is not included in this illustration) were made on this token: VD=142 ms, CD=140 ms, V/C=1.01, GC=92 ms.

The GC measurement was complicated somewhat by irregularities in the vocal wave associated with glottalization (Redi and Shattuck-Hufnagel 2001, Ashby and Przedlacka 2011). Figure 2 shows an example of this difficulty in the obstruent-vowel sequence in the phrase *George often*. Note that at vowel onset there are two pulses of highly laryngealized voicing. Since this type of irregularity is associated with the perception of glottal stops, in such cases the GC measurement was extended to the onset of modal voicing, characterized by a regular periodic pattern in the waveform. The glottal pulse trackers in Praat were of assistance in identifying the onset of modal voicing.

For the purposes of the research questions, the V/C ratio and the GC measurements allow us to characterize the degree of final devoicing and the extent of initial glottalization. A higher V/C ratio is associated with error-free final voiced obstruents. A shorter GC measurement indicates that the consonant and vowel have been liaised, while longer glottal closure is of course associated with glottal stops.

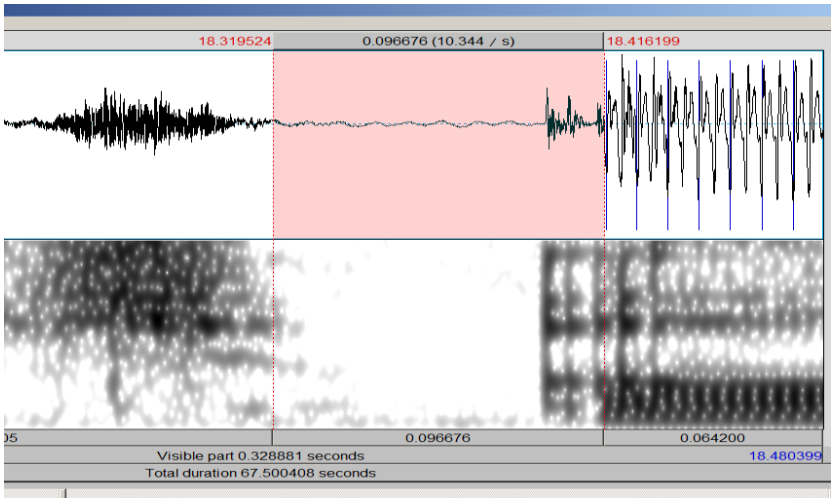


Figure 2 – Glottal closure duration measurement in *George often*. Measurement includes two cycles of highly laryngealized glottal pulsing.

3.3. Results - Individual data

The mean results for each individual speaker are provided in Table 1. Three speakers had Rate measurements that fell outside of the standard deviation for the entire group, indicated by shading in the appropriate cell. These speakers were excluded from the group analysis. Note that the GC measurements for the non-native subjects exhibited a wide range, from just under 9 to over 100 milliseconds. The native speaker showed measurable glottal closure in just one of 40 tokens, for an average GC measurement of less than 2 milliseconds. The native speaker's V/C ratio was 2.9, while that of the non-natives ranged from 1.33 to 2.72

Speaker	VD	CD	GC	V/C	RATE
Native	148.1	57.9	1.33	2.90	*1.96*
pm1	118.5	102.1	85.9	1.33	2.26
kp1	116.2	68.7	30.06	2.07	2.07
kp2	114.4	79.1	58.8	1.55	2.06
kp3	169.3	76.6	57.05	2.72	2.62
pm2	164.0	88.9	103.5	2.05	*2.74*
wm1	120.7	98.8	59.4	1.32	2.23
wlkp1	157.4	94.36	60.6	1.77	*2.67*
wlkp2	138.2	94.8	67.2	1.63	2.55
wlkp3	145.6	75.28	8.9	2.02	2.55
wlkp4	144.4	99	26.1	1.65	2.41

Table 1 – Mean results for each individual speaker. Shaded cells denote speakers whose Rate average fell outside of the group standard deviation. These speakers were excluded from the group analysis.

3.3 Group Data and interaction of GC with voicing parameters

To investigate the possible effects of GC duration on the voicing parameters, each individual measurement was placed into one of three categories depending on the value of the GC measurement. Type 1 was comprised of GC measurements of less than 40 ms, and may be described partially or completely liaised. Type 2 included GC measurements between 40-79 milliseconds, while Type 3 covered glottal closures of over 80 ms. From the 8 speakers analyzed in the group data, there were 112 tokens of type 1, 119 tokens of type 2, and 89 tokens of type 3.

Parameter	Type 1; 0-39ms; n=112	Type 2; 40-79 ms; n=119	Type 3; >80 ms; n=89
V/C	2.25	1.65	1.42
VD	142.3	134.7	131.7
CD	71.7	88.9	104.9

Table 2 – Voicing parameter means sorted according to three types of glottal closure duration.

The mean results of the measurements sorted according to glottal closure duration are presented in Table 2. A one-way analysis of variance (ANOVA) was performed to

establish the effects of Glottal Closure token type on the voicing parameters. Significant effects ($p<0.01$) were found for both V/C ration and Consonant Duration. No significant effect was found for Vowel Duration ($p=0.17$). Post-hoc Tukey tests were performed on the pairs of means. For V/C ratio and Consonant Duration all pairwise comparisons were significant. For Vowel Duration, Type 1 and Type 3 were significantly different, while the other pairwise comparisons were insignificant

3.4 Effects of dialect

Our study may also raise the question of whether intervocalic voicing across word boundaries, a feature associated with certain dialect regions, may be found in these speakers' L2 English, and if so, what if any effect does it have on the parameters of final voicing. Of the 10 subjects recorded for this experiment, 4 of them reported that they were raised in Wielkopolska, an area of Poland associated with intervocalic voicing. The results of the acoustic study were thus sorted according to dialect background to investigate any possible effects on the parameters under study. The dialect results are given in Table 3, which shows the mean values of the voicing and glottal closure parameters, as well as the percentage of Type 1 (liaised) tokens. A one-way ANOVA revealed a significant effect of dialect on Glottal Closure duration, which was shorter for the Wielkopolska speakers. In addition a chi-square test on the percentage of liaised (Type 1) tokens was significant: Wielkopolska speakers were more likely than the others to produced 'liaised' sequences. No significant effect of dialect was found for the voicing parameters

Dialect	VD	CD	GC	V/C	%Type 1
Wielkopolska	139.5	84.1	40.9	1.74	47.5
Other	133.8	90.2	67.4	1.77	29.5

Table 3 – Acoustic measurements sorted for dialect background.

3.5 Discussion

The results of the acoustic study support the hypothesis that glottalization of initial vowels may contribute to final obstruent devoicing in the speech of Polish leaners of English. In this connection it is interesting to observe the results obtained from the native speaker, who showed almost across-the-board liaison, as well as the highest V/C ratio of all the recorded subjects. Table 4 shows a comparison of the native speaker with group mean values of the non-natives. For the Polish speakers, the average glottal closure duration of 60.8 ms fell within the Type2 range, while the V/C ratio was 1.76, notably lower than that of the native speaker. Thus, it is reasonable to claim that liaison is a clear

aspect of native-like speech that contributes to the production of 'final' voiced obstruents.

	V/C	GC	%Type 1
Native control	2.9	1.33	97.5
Non-native	1.76	60.8	38

Table 4 – Comparison of native speaker control with group data for Polish learners.

When the analyzed tokens were divided into three types of the basis of Glottal Closure duration, a clear effect was found for token type on both the V/C ratio and the Consonant Duration. Importantly, the effect of token type on Vowel Duration was not significant. This fact suggests that we may rule out speech rate as a factor in the group results. While one may be inclined to attribute initial glottalization to the fact that the subjects were speaking more slowly in a foreign language, such rate effects should have been equivalent for each of the parameters involved. This was not the case – only the final consonant was affected.

The effect of Glottal Closure duration on the duration of final consonants may be attributable to a process of final lengthening by which segments are lengthened at the end of prosodic constituents (Beckman and Edwards 1990). Final devoicing and final lengthening should be expected to co-occur. The longer a consonant is, the more likely it is to be unvoiced, since more effort is required to sustain the glottal airflow required for voicing over the course of a lengthened consonantal constriction. In other words, we are witnessing the manifestation of a phonetic universal by which unvoiced consonants are longer than voiced ones.

When liaison occurs, the context for final lengthening (and final devoicing) is eliminated; the consonant is no longer final. Thus, although final lengthening does occur in English (Beckman and Edwards 1990), the native speaker produced liaised consonants instead of longer final ones that would be more susceptible to devoicing. These results suggest that Polish and English have somewhat different representations of "final" and "initial" positions. We will take up this issue in detail in the following section.

The data from the dialect groups may complicate the picture. The results indicated that speakers from Wielkopolska produced more 'liaised' tokens, but this did not seem to have a significant effect on the durational patterns associated with voicing. That is, more liaison did not necessarily imply less devoicing, at least in terms of the temporal parameters. One possible clue in explaining the discrepancy associated with the Wielkopolska speakers may be found in the behavior of one speaker, who in many instances showed an intrusive vowel before a glottalized initial vowel. For example, in the phrase *Today's express train*, the speaker produced a short vocoid after the final /z/, and then showed full glottal closure on the initial vowel of express, resulting in a sequence [zᵊʔɛ]. Since a full glottal stop is produced, we may not claim that liaison has been acquired. The speaker appears to have adopted a vowel-insertion strategy to produce fully voiced final obstruents, perhaps diminishing the significance of the temporal parameters.

4. The phonology of boundaries

The acoustic study described in this paper reflects a fundamental difference in the phonology of English and Polish with regard to the behavior of speech sounds at word boundaries. Stated briefly, word boundaries in Polish seem to block many common phonological processes that might be expected to accompany the concatenation of two sounds. In English, on the other hand many such processes are common.

For example, Polish morphology shows a number of palatalization processes that turn coronal stops into alveolo-palatal affricates before certain grammatical endings. Thus, the locative of the form /kot/ 'cat' is /kotɕe/. The traditional assumption is that it is the frontness of the vowel in the ending that conditions the alternation – the /t/ is said to be 'palatalized'. In a sequence *kot jest* 'the cat is', one might expect the palatal glide in *jest* to cause palatalization of the /t/. It does not, so we may assume that the concatenation process that results in the alternation at the morpheme boundary does not apply at the word boundary. Conversely we frequently observe palatalization in an analogous sequence *got you* in English, which is often pronounced as *gotcha*.

These facts are connected with the notion of resyllabification across word-boundaries, by which a word-final coda consonant is reinterpreted as the onset to the following syllable. Thus, for English we may make a generalizing statement that a sequence /tj/ in a syllable onset results in a post-alveolar affricate. Resyllabification is impossible in Polish (Rubach and Booij 1990), so the /t/ and the /j/ in *kot jest* must be analyzed as belonging to two separate syllabic constituents. Liaison in English may be interpreted as a form of this type of resyllabification.

The Onset Prominence representational environment (OP; Schwartz 2010) offers a useful set of materials for analyzing the different behavior in Polish and English at word boundaries. OP builds on recent insights into the structural nature of segmental phonology, in particular manner of articulation (Golston and Hulst 1999, Pöchtrager 2006). The basic building block of the OP representational environment, which may be seen as the functional equivalent of a universal CV structure, is given in Figure 3. The tree represents the acoustic signal as a hierarchical structure, from which both segmental representations and prosodic categories such as syllables are derived. Manner is defined by the layers of structure present in the segmental representation. Figure 3 represents a stop-vowel sequence.

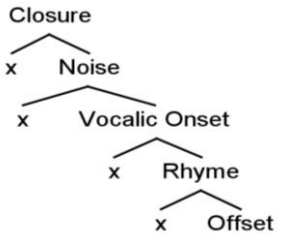


Figure 3 – Basic building block of syllabic structures in the OP environment.

In the OP environment, syllabic constituents such as the one in Figure 3 are formed from the concatenation of individual segmental structures. Consider Figure 4, in which the representation of the stop /k/ contains the top two layers of the structure in Figure 3, the /w/ contains the Vocalic Onset node, and the vowel and final /k/ represent the Rhyme. These structures combine to form the English word *quick*. Such a sequence, since it proceeds down the hierarchy is predicted to be contained a single syllabic constituent[1]. The basic principle for syllabification is that a tree may be "absorbed" into a higher level structure to its left, so the three structures in Figure 4 may combine into a single constituent.

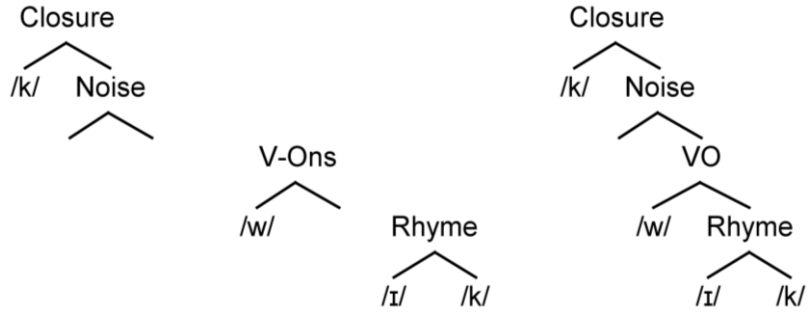


Figure 4 – Syllabification of English *quick* in the OP environment

For the representation of initial vowels in Polish, we claim that they contain an additional layer of structure, namely the Closure node associated with stops. Since it is not a lower-level structure than the preceding consonant, the vowel may not be absorbed into the tree to its left. Resyllabification does not occur, and the "final" status of the consonant is reinforced. This is illustrated in Figure 5, which shows a word-final /d/ followed by an initial /e/ as they would be represented in Polish using OP structures.

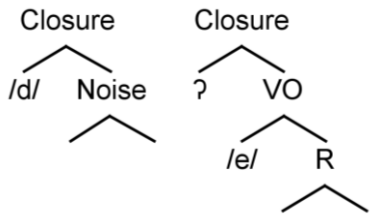


Figure 5 – Sequence of word-final /d/ and word-initial /e/ in Polish. The active Closure node on the vowel blocks the merging of the two trees.

1 The presence of the /k/ in the rhyme is the result of a submersion process for codas that will not be relevant for the present paper.

In Figure 6 we see an analogous sequence in English. Note that the English vowel does not contain the extra structure, and the tree on the right may be absorbed into the one on the left, reflecting liaison and resyllabification. The difference between Polish and English is captured in terms of the structural properties of the initial vowel. Initial vowels in Polish are larger structures than they are in English. They might be thought to contain a "built-in" glottal stop, which blocks resyllabification across word boundaries[2].

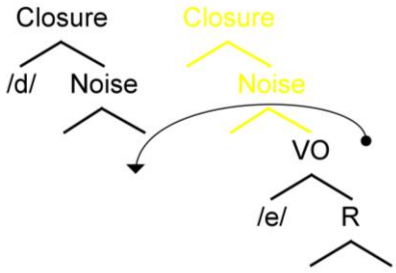


Figure 6 – English sequence of final /d/ and initial /e/ producing liaison.

This representational approach comes with benefits for both phonological theory and comparative descriptions of Polish and English phonology upon which we may base teaching materials. The advantage concerns the representation of phonological boundaries. In phonology, this has been a recurring problem. Symbols (such as + and #) traditionally used to represent such boundaries are inherently arbitrary in nature (e.g. Scheer 2008). By contrast, in the Onset Prominence environment, such boundaries may be constructed using the structure of segments themselves – they are truly 'phonological' entities. With regard to teaching materials, the value of OP representations lies in the fact that they are hierarchical. Unlike a linear string of segmental symbols, this approach allows for a faithful model of what actually happens in speech.

5. Final remarks

This paper has described an acoustic study of the speech of Polish learners of English. The results, as well as the ensuing phonological discussion, suggest the need to establish principled representations of phonological boundaries. Languages appear to show systematic differences in the behavior of word-initial and word-final segments, which manifest themselves in a number of processes found in Second Language speech. The Onset Prominence environment offers a principled explanation of these differences, with benefits for both phonological theory and second language speech acquisition.

2 In the case of *kot jest*, resyllabification is prevented by the 'promotion' of the structure of the /j/.

References

Ashby, M. & J. Przedlacka. 2011. The stops that aren't. Ms. UCL.

Beckman, Mary & Jan Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In J. Kingston & M. Beckman (eds.), *Papers in Laboratory Phonology* I: 179-200. Cambridge, UK: Cambridge University Press.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22: 129–159.

Dilley, L., Shattuck-Hufnagel, S., and Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24: 423-444.

Dukiewicz, L. and I. Sawicka. 1995. *Gramatyka współczesnego języka polskiego – fonetyka i fonologia*. Kraków: Wydawnictwo Instytutu Języka Polskiego PAN.

Golston, C. and H. van der Hulst. 1999. Stricture is structure. In B. Hermans and M. van Oostendorp, (eds.) *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: John Benjamins: 153-173.

Gonet, W. & G. Pietroń. 2004. The Polish Tongue in the English Ear. *Zeszyty Naukowe PWSZ w Koninie*, nr 1/2004 (4), pp. 56-65.

Gussmann, E. 2007. *The Phonology of Polish*. Oxford: Oxford University Press.

Kaźmierski, K. 2007. Dynamic specification in English vowels – implications for Polish learners. Unpublished BA thesis. UAM Poznań.

Maddieson, I. (1997). Phonetic universals. In Hardcastle, W. and J. Laver (eds). *The Handbook of Phonetic Sciences*. Cambridge: Blackwell. 619-640.

Pöchtrager, M. 2006. The structure of length. Ph.D. dissertation. University of Vienna.

Port, R. & J. Dalby. 1982. C/V ratio as a cue for voicing in English. *Perception and Psychophysics 2*. 141-152.

Redi, L. and S. Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29: 407-429.

Rejniak, J. 2011. What is final? Factors affecting devoicing errors in the speech of Polish learners of English. Unpublished BA thesis. UAM Poznań.

Rubach, J. and G. Booij. 1990. Syllable structure assignment in Polish. *Phonology* 7: 121-158.

Scheer, T. 2008. Why the Prosodic Hierarchy is a diacritic and why the Interface must be Direct. *Sounds of Silence: Empty Elements in Syntax and Phonology*, edited by Jutta Hartmann, Veronika Hegedüs & Henk van Riemsdijk, 145-192. Amsterdam: Elsevier.

Schwartz, G. 2010. Auditory representations and the structures of GP 2.0. *Acta Linguistica Hungarica* 57: 381-397.

Schwartz, G. 2011. Avoiding stop insertion after English angma - a representational solution. In: Dziubalska-Kołaczyk, K.; Wrembel, M.; Kul,M. (eds.). *Achievements and perspectives in SLA of speech: New Sounds 2010.* Berlin: Peter Lang: 241-250.

Strange, W. 1989. Dynamic specification of coarticulated vowels spoken in sentence context. *JASA* 85: 2135-2153.

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0030-0

THE EFFECT OF WORD-INITIAL GLOTTALIZATION ON WORD MONITORING IN SLOVAK SPEAKERS OF ENGLISH

JAN VOLÍN
MÁRIA UHRINOVÁ
RADEK SKARNITZL
Institute of Phonetics, Faculty of Arts, Charles University in Prague

Abstract
The study investigates the impact of glottal elements before word-initial vowels on the speed of processing of the phrases taken from natural continuous speech. In many languages a word beginning with a vowel can be preceded by a glottal stop or a short period of creaky voice. However, languages differ in the extent of use and functions of this glottalization: it may be used to mark the word boundary, for instance, or to add special prominence to the word. The aim of the experiment was to find out whether the presence of the glottal element can influence reaction times in a word-monitoring paradigm. Users of different languages – Slovak and Czech learners of English, as well as native speakers of English – were participating in perception testing so that the influence of the mother tongue could be determined. The results confirm the effect of both glottalization and the L1 of the listeners. In addition, a significant effect of test item manipulations was found. Although the phrases with added or deleted glottal stops displayed no obvious acoustic artefacts, they produced longer reaction times than items with naturally present or absent glottalizations. We believe that this finding underlines the importance of inherent stress patterns, whose alterations lead to the increase in processing load.

1. Introduction

Linguists of most methodological backgrounds have a similar concern. Whether they are generativists, structuralists or constructionalists, they have to establish the inventories of items that are relevant for language communication. The research in sound patterns of languages of the past decades has shown that it is unproductive to remain stuck with narrowly defined phonemes and ignore rich symbolic structure provided by other speech phenomena. Descriptive units, whose distinctive power rightfully draws attention of language users, can change lexical meanings, but cannot explain on their own why some speakers communicate more effectively, are better accepted, and induce more cooperative behaviour than others (Local 2003; Hawkins 2003).

One of the elements that occur in most languages with non-phonemic status and still could influence intelligibility of speech and the smooth flow of communication with all its consequences is the glottalization of word-initial vowels. In this study, the term

glottalization will be used for glottal stops or perceptually equivalent glottal events, e.g., creaks, rapid drops in F0 or intensity, etc., which precede words beginning with onsetless syllables. Languages differ in the extent of use and roles or functions of such glottalizations (e.g., Przedlacka and Ashby 2011; Gordon and Ladefoged 2001; Redi and Shattuck-Hufnagel 2001; Kohler 1994).While in some they can be treated as external juncture signals that indicate an important autosemantic morpheme boundary, in others they may add special prominence to words with which they are used. In such cases the prosodic structure or the semantics of the utterance may be reflected. In phonological terms, the word-initial prevocalic glottalization can be viewed as a specific treatment of onsetless syllables in critical positions (Schwartz 2011).

While the production of glottal elements is often noted and explored, the perceptual aspect of the problem remains unclear. It might be hypothesized that speakers who regularly produce glottalizations would rely on their presence in the speech signal when they have to process it. By analogy, the greatest sensitivity can be expected in those rare languages where glottal stops act as phonemes. However, English is described as a language where word-initial prevocalic glottalization is facultative, and it is only used to emphasise a word if such an emphasis is contextually appropriate (Wells 1990: 327; Cruttenden 1994: 155). It is even recommended to foreigners to avoid production of glottal elements before most of the words beginning with vowels (especially frequently occurring grammatical *of*, *in*, *is*, *are*, *a*, *and*, etc.) to prevent unnatural "choppy" flow of speech (O'Connor 1980: 101). In such circumstances, inappropriate presence of glottal elements might even hinder mental processing of speech since it produces unnatural or unpredictable rhythmic configurations.

As our ultimate concern is English as a foreign language, the matter is even more complicated. Foreign speakers of English try to model the speech behaviour of native speakers, yet they struggle with production stereotypes from their own mother tongue. The extent to which they either benefit or suffer from the presence of glottal elements in speech can thus differ depending on their native situation.

In our previous study, we found significant differences between Czech and Spanish speakers of English (Bissiri et al. 2011). Spanish learners of English, in whose L1 glottalization is used infrequently and mostly as a marker of emphasis, benefited less from the presence of word-initial glottalizations than native speakers of Czech, which uses glottalization frequently as a signal of juncture. However, these results are difficult to interpret unambiguously since apart from differences in the general use of glottalizations, Spanish differs from Czech typologically. The phonotactic patterns and the prosodic plan of the two languages endow the learners of English with quite different predispositions. Moreover, the EFL teaching in the two countries seems to draw on different resources: both the general motivation of students and the teaching methods may not be comparable.

Therefore, we decided to examine the differences between reaction times to words with and without glottalization in Slovak speakers of English. Slovak is in many features similar to Czech (they both are Western Slavonic languages) and speakers of these languages are able to reasonably communicate even without special language instruction. Also, the EFL methodology is essentially the same in the two countries: the Czechs and Slovaks had lived under one central government until 1989 and they keep sharing many of their social and cultural traditions. On the other hand, the two languages

differ in the exploitation of word-initial glottal stops: the use of glottalization in Slovak is reportedly low and word-initial vowels regularly cause assimilation of voicing of the final consonant of the preceding word. This means that rather than providing the onsetless word-initial syllable with a glottal consonant-like element, the speakers of Slovak prefer to tie the word-initial vowel quite firmly to the preceding consonant. For instance, the word *tak* [tak] in the Slovak phrase

(1) *tak ale nie* [tag_ale ɲie] (in Engl. *but not this way*)

will be pronounced with [g] due to the tightly adhering [a] of the following word. The similarly sounding phrase in Czech, on the other hand, will contain glottalization and the preceding word-final [k] will remain voiceless:

(2a) *tak ale ne* [tak ʔale ne] (in Engl. *but not this way*)

Even in the case of less careful pronunciation where the glottal element might be missing, the assimilation of voicing will not happen (again, cf. Geoff Schwartz' concept of onsetless syllables).

The objective of our study is thus to investigate the influence of L1 on the perceptual impact of glottalizations in English while abstracting from profound differences in phonological systems (Spanish and Czech) and in language instruction. Slovak and Czech listeners will be compared mutually and against the benchmark performance of native English listeners.

We have stipulated two sets of hypotheses. The first set concerns the influence of glottalization, and the null hypothesis states that there is no effect of the presence or absence of a word-initial glottal element on reaction times when monitoring the speech signal for target words. An alternative hypothesis says that the presence of glottalization highlights the target word thus facilitating its perception. Reaction times in such a case should be shorter. Another alternative would argue that the presence of the glottal segment breaks the natural flow of English (as argued in some pronunciation textbooks) and produces the effect similar to that reported by Buxton (1983): rhythmically impaired utterances lead to longer reaction times in word-monitoring experiments.

The second set of hypotheses concerns the mother tongue of the EFL learners. The null hypothesis would deem it irrelevant. The first alternative would suggest that the Czech listeners will benefit more from the presence of glottal stops as they use them on a regular basis in their mother tongue. The second alternative would argue that the Slovak listeners, who only use glottal stops to highlight words (similarly to the English) will have shorter reaction times to words with glottal segments than the Czech listeners, to whom the glottalization of word-initial vowels does not signal anything special.

2. Method

The experiment was based on the word-monitoring paradigm (Kilborn & Moss, 1996). In this design, respondents are given a target (a word usually printed on a computer screen) and they listen to auditory stimuli for that target. Their task is to press a

designated key as soon as they detect the word. Their reaction time (or the so-called latency) is measured from the acoustic onset of the word to the moment when the key was pressed. We used the DMDX software – a package developed specifically for reaction time measurements (Forster & Forster, 2003).

Natural continuous speech provided the material for the stimuli. Five native speakers of southern variants of British English read news bulletins that were earlier broadcast at the BBC World Service. Forty-eight phrases were extracted from the recordings such that the target words could not be guessed from the semantic cues, i.e., all common collocations of the target words were avoided. For instance, in the phrase *Arafat last month as partial promised reforms* the conjunction *as* was the target. Clearly, the extraction of the sequence from a longer sentence does not help the listeners to guess when the target word might come. Similarly, in the phrase *with ten men after the striker Thiery Henry* the listeners were asked to react to the word *after*. The targets were placed anywhere between the second and the fifth stress-group. Distractors with the target in the first stress-group were only used to keep the listeners alert, but were not analyzed. Some more distractors were prepared with consonants in the word-initial position so that the listeners would not figure out the nature of the true targets.

One half of the true targets occurred naturally with glottal stops, the other half without them. These 48 items were processed in sound editing software to create artificial stimuli with the opposite value of glottalization, i.e., the naturally occurring glottal stops were deleted and the items without glottal stops were provided with an spliced one. Obviously, all possible care was taken to produce items that could not be recognized as artificial, i.e., the items were without clicks and other discontinuities, with smooth transitions of formants and the fundamental frequency track. These manipulations were carried out with the help of Praat, Sound Forge, and Matlab software packages.

Altogether, 96 targets and 36 fillers were used in the perceptual testing. The listeners were 90 adults in three equally-sized groups by their L1. Thirty were native English students and employees of a British university, 30 Czech and 30 Slovak learners of English. They were tested individually through headphones in a sound treated booth.

3. Results

The results confirm previous findings of the positive effect of glottalizations on the latencies: the words with pre-glottalized word-initial vowels are spotted faster than words linked to the preceding words. Repeated measure ANOVA returned highly significant effect of glottalization: F (1, 87) = 481.4; $p < 0.001$. Figure 1 indicates that the latencies were about 450 ms and items with glottal stops were spotted about 60 ms faster than the items without it.

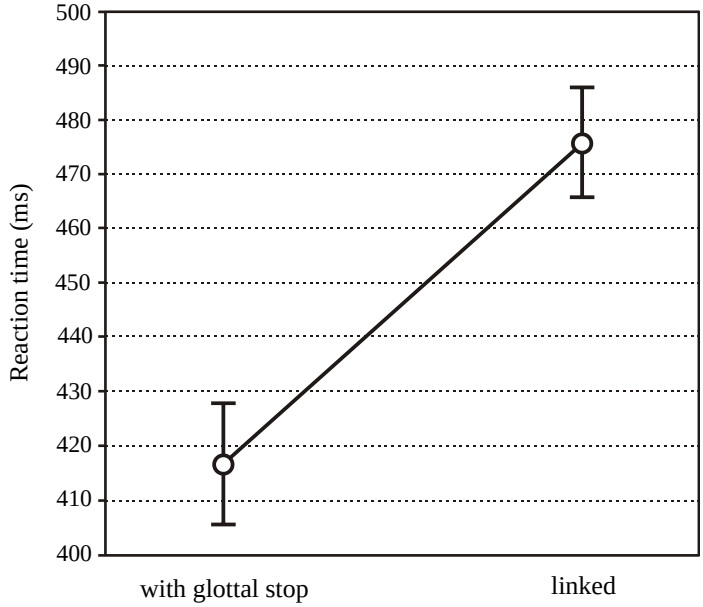


Figure 1: Mean reaction times of all listeners to words with (on the left) and without (on the right) the word-initial pre-glottalization.

The main effect of the mother tongue (the between-group factor) was also found highly significant: F (1, 87) = 11.96; $p < 0.001$. Post-hoc pairwise comparisons revealed that the English listeners were significantly faster than both the Slovak and Czech listeners, while Czechs and Slovaks did not differ significantly from each other. Although the difference between the latter groups was not statistically significant, Figure 2 shows that the Slovaks were on average faster than Czechs. That, however, does not address the hypotheses about the influence of glottalization and, therefore, the interaction between the variables is of interest. Analysis of variance returned significant interaction between the mother tongue of the listeners and the glottalization variable: F (1, 87) = 7.26; p = 0.0012 Figure 2 indicates that this result is again caused by the difference between the English on the one hand, and the Czech and Slovak on the other hand. Although there are allegedly differences in the production of the word-initial glottalization between Czech and Slovak, we found no difference in perceptual testing between the speakers of these two languages.

In addition to this main outcome, we carried out some further analyses to find out, whether the reaction times could have been influenced by any of our captured linguistic or other variables. These analyses were also based on ANOVA for repeated measures, but were calculated for individual test items rather than for individual subjects.

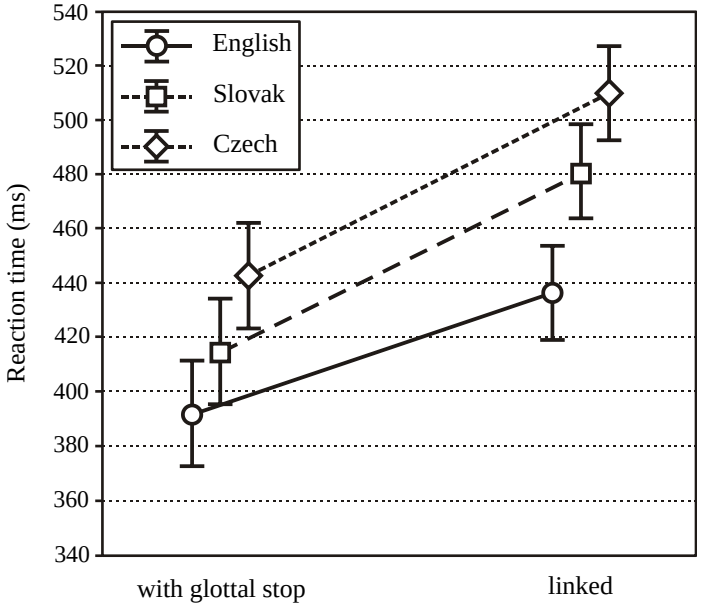


Figure 2: Interaction between the variable of the mother tongue and glottalization. Mean reaction times of the three listener groups to words with (on the left) and without (on the right) the word-initial pre-glottalization.

First of all, we found a significant effect of word stress. Reactions to words with stressed initial vowels were faster: F (1, 3740) = 25.1; $p < 0.001$. Figure 3 displays the mean reaction times which suggest that the English listeners benefited more from the presence of stress than the other two groups, whose behaviour with respect to word stress was again very similar. There was no significant interaction between stress, mother tongue and glottalization ($p = 0.86$).

We also decided to test the effect of the target position in the phrase. The factor of position had four levels: the items in the second stress-group were labelled *early* (no first stress-group targets were tested), the third stress-group was *mid*, the fourth was *late-mid*, and the remaining items were *late*. Unlike the findings in Buxton (1983), our results did not show any interesting trend. The *early*, *mid* and *late-mid* positions led to practically the same result and only the *late* position produced significantly longer reaction times.

Similarly, we did not find any significant difference between reactions to structural words (e.g., conjunctions, prepositions) and content words (e.g., nouns, adjectives). Semantic status obviously did not matter in the word-monitoring task. This may have been caused by the fact that the test items were extractions from longer sentences and their semantics was damaged: the price we had to pay to meet the requirement of unpredictability of the targets.

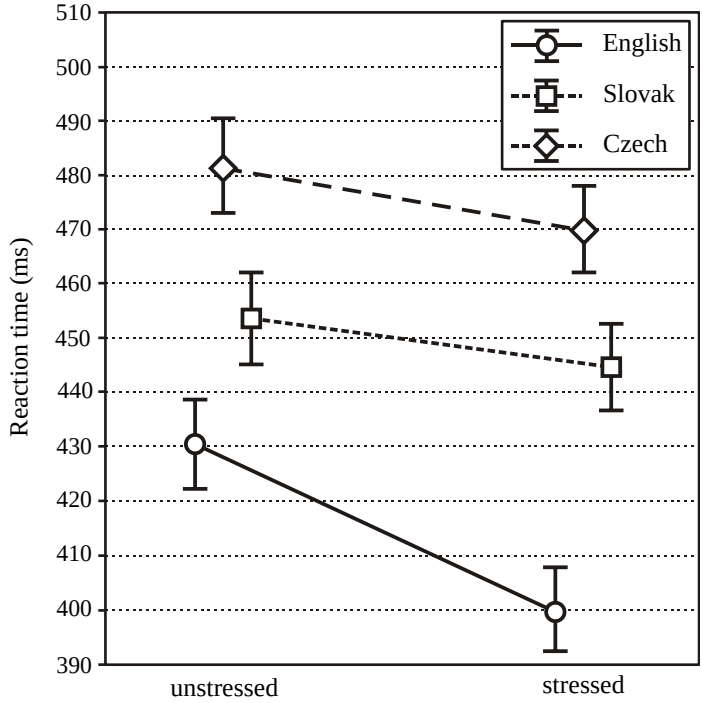


Figure 3: Mean reaction times of the three listener groups to words with stressed initial vowel (on the right) and with unstressed initial vowel (on the left).

The last variable we tested was that of *manipulation*. Our set of 96 items consisted of 48 instances of natural production of glottalization or natural linking (24+24). The other half of the test items had glottal stop either edited out or added (again 24+24). Although the manipulated items did not exhibit any consciously perceivable artefacts, we wanted to know whether there was any difference in reaction times to them. Figure 4 shows that manipulation indeed matters and there is even highly significant interaction between this variable and glottalization: F (1, 3734) = 144.6; $p < 0.001$. The items in which glottal stops were edited out behaved in the same way as the analogical natural items, but the items where the glottal stop was added led to slower reactions compared with items where glottal stop was naturally present. This result is discussed below.

4. Discussion

The presence or absence of the glottal element before a word-initial vowel influences the perceptual processes in all three language groups. However, our new group of listeners – the Slovaks – did not produce results similar to the Spanish sample we investigated previously. Although the Slovak listeners should differ from the Czech ones in the same direction as the Spanish, they did not produce a similar effect, they did not differ significantly from the Czech listeners. A possible explanation is that mutual contacts of

Czechs and Slovaks which are, for instance, reflected in the fact that they do not have to learn each other's language and still understand each other without difficulties, overrule the influence of the native language on the perception of a facultative prosodic marker like the glottal stop before a word-initial vowel. Perhaps the Spanish, who should be using glottalization similarly to the Slovaks, interact less with speakers of languages where glottalization is common. (The French, for example, are known to link words very consistently without glottalizing the onsetless syllables.) Another explanation could be that despite the traditional descriptions in grammar books the younger generation of Slovaks uses more glottal stops than the older generations used to. This possibility is supported by our informal observation but has to be verified empirically.

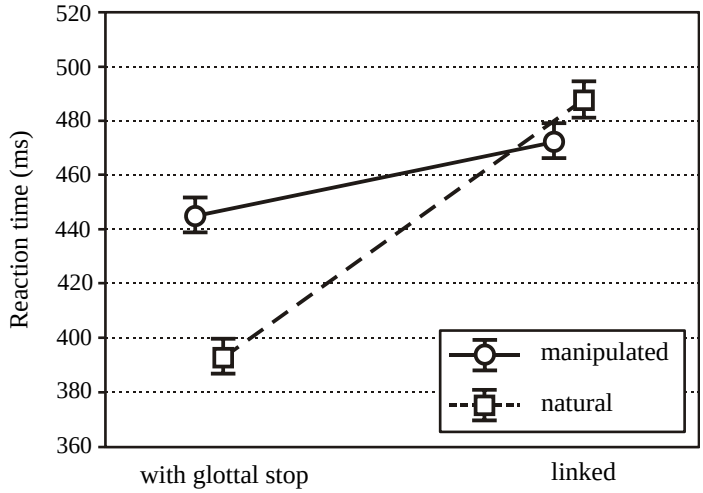


Figure 4: Mean reaction times of the listeners to words with (on the left) and without (on the right) the word-initial pre-glottalization according to the manipulation status of the item.

The general effect of stress confirms the expectations based on the earlier work of other researchers, but smaller impact of stress on Czech and Slovak listeners is, to our best knowledge, a new empirical finding. However, the effect of the target position in the phrase and the effect of the semantic status of the words were not confirmed. As stated above, we assume that the semantic unpredictability of the carrier phrases could have caused this result.

On the other hand, we found a significant effect of test item manipulations. Although the phrases with added glottal stops displayed no obvious acoustic artefacts, they produced longer reaction times than items with naturally present glottalizations. We believe that this finding underlines the importance of inherent stress patterns of a language, whose alterations leads to the increase in processing load (*cf.* Buxton, 1983).

Acknowledgment
The research was supported by the internal grant of the Faculty of Arts, Charles University in Prague. The authors would also like to thank to M-P. Bissiri who, as the

intern at the Institute of Phonetics in Prague at the time of the initial stages of the study, collected some of the data.

References

Bissiri, M. P., Lecumberri, M. L., Cooke, M. and Volín, J. 2011. The Role of Word-Initial Glottal Stops in Recognizing English Words. In: *Proceedings of the 12th Annual Conference of ISCA Interspeech*: 165-168. Florence: ISCA.

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

Cruttenden, A. 1994. *Gimson's Pronunciation of English*. London: Edward Arnold.

Dilley, L. Shattuck-Hufnagel, S. and Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24: 423-444.

Forster K.I. and Forster, J.C. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35/1: 116-124.

Gordon, M. and Ladefoged, P. 2001. Phonation types: a crosslinguistic overview. *Journal of Phonetics* 29: 383-406.

Hawkins, S. 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31: 373-405.

Kilborn, K. and Moss, H. 1996. Word Monitoring. *Language and Cognitive Processes* 11/6: 689-694.

Kohler, K. 1994. Glottal stops and glottalization in German. *Phonetica* 51: 38-51.

Local, J. 2003. Variable domains and variable relevance: interpreting phonetic exponents. *Journal of Phonetics* 31: 321-339.

O'Connor, J.D. 1980. *Better English Pronunciation*. 2nd Edition. Cambridge: CUP.

Przedlacka, J. and Ashby, M. 2011. Acoustic correlates of glottal articulations in Southern British English. In: Proccedings of ICPhS XVII: 1642-1645. Hong Kong: IPA.

Redi, L. and Shattuck-Hufnagel, S. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29: 407-429.

Schwartz, G. 2011. Final devoicing in Polish English: Segmental or prosodic error? *Presentation at Accents 2011*, Lodz: UoL.

Wells, J.C. 1990. *Longman Pronunciation Dictionary*. Harlow: Longman.

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0035-8

MORE ON THE VOICING OF ENGLISH OBSTRUENTS: VOICING RETENTION VS. VOICING LOSS

WIKTOR GONET
RADOSŁAW ŚWIĘCIŃSKI
UMCS Lublin

Abstract
In Gonet (2010), one of the present authors found out that English word-final phonologically voiced obstruents in the voicing-favouring environment exhibit asymmetrical, if not erratic, behaviour in that voicing in plosives is most often retained while in fricatives voicing retention concerns only about 1/3 of the cases, with the other possibilities (partial and complete devoicing) occurring in almost equal proportions. The present study is an attempt at exploring the intricacies of devoicing in English to examine to what extent the general tendency towards obstruent devoicing is overridden by voicing retention triggered by adjacent voiced segments both within words and across word boundaries. This study is based on a relatively large knowledge base obtained from recordings of spontaneous R. P. pronunciation.

1. Introduction

The present study is a follow-up on Gonet (2010), whose focus was on consonantal voicing in the word-final position. The paper presented the behaviour of English obstruents and indicated that the voicing of English word-final obstruents is best described by referring to the combination of word position and the voicing of the initial sound in the following word. These combinations fall into two major classes:

- **phonation-favouring** (if they are followed by a vowel or a voiced consonant),
- **phonation-impeding** (before a pause or before a voiceless sound).

The study reviewed a number of publications, including those by Ball and Rahilly (1999), Catford (1964, 1977, 1988), Clark and Yallop (1990), Davenport and Hannahs (1998), Fujimura and Erickson (1999), Gimson (1962, 2001), Gonet (1989, 2001), Gonet and Stadnicka. (2006), Jassem (1983), Ladefoged (1971, 1975), Lisker and Abramson (1964), Maddieson (1999), Ohala. (1999), Port and Rottuno (1979), Raphael et al. (1975), Roach (1983), Shockey (2003), Szpyra-Kozłowska (2003), Van den Berg (1958), and was based on a large body of recordings of spoken English by 6 native speakers. Yet the results exhibited asymmetrical, if not erratic, behaviour; the details are presented in Table 1 as well as Figures 1 and 2.

	BEFORE A PAUSE	BEFORE A VOICELESS CONSONANT	BEFORE A VOICED CONSONANT
PLOSIVES	---------- Partialy dev. Completely dev.	Fully voiced Partially dev ----------	Fully voiced ---------- ----------
FRICATIVES	---------- ---------- Completely dev.	---------- ---------- Completely dev.	Fully voiced Partially dev. Completely dev.

Table 1. Voicing in English word-final obstruents (Gonet 2010).

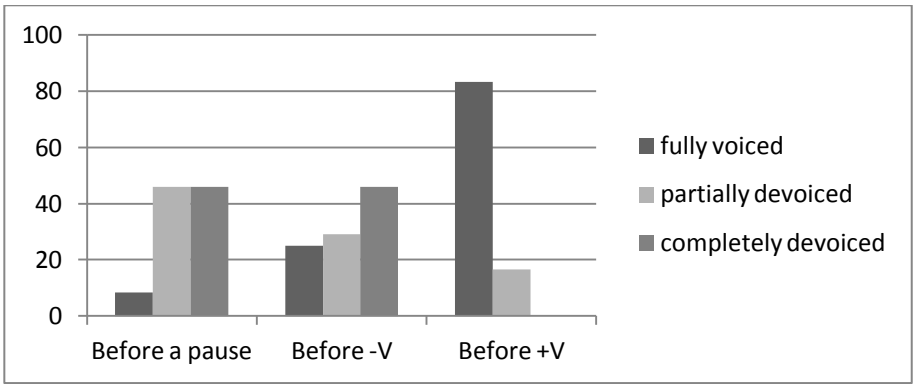


Figure 1 Distribution of voicing in word-final plosives (Gonet 2010).

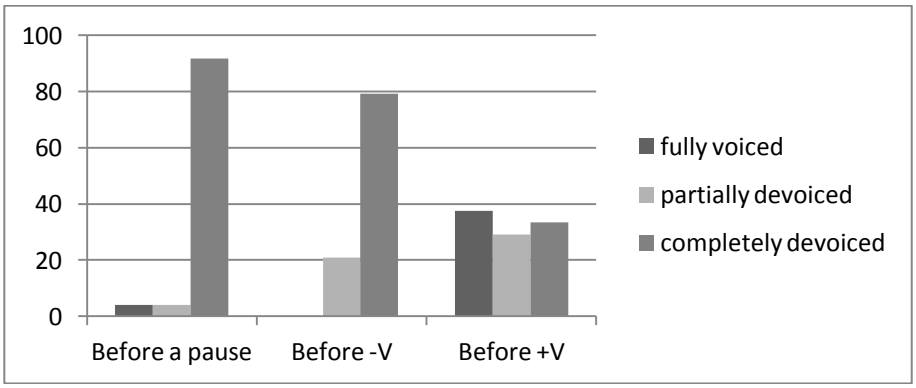


Figure 2 Distribution of voicing in word-final fricatives (Gonet 2010).

Many authors indicate that obstruents have a natural tendency to devoice, especially in voicing-impeding environments. Hence, for voiced obstruents, hypothetically there apply 2 opposing forces:

- Devoice an obstruent, especially in word-final position
- Retain voicing, especially before a voiced sound

In view of the above, the goal of the present study was to explore the question to what extent the general tendency towards obstruent devoicing is overridden by voicing retention triggered by adjacent voiced sounds both within words and across word boundaries.

2. Design of the experiment

As most of the studies on obstruent voicing in English are based on audio material elicited in the form of read wordlists or lexical items embedded in sentence-frames, it appeared imperative that this study should be based on spontaneous speech. For this reason, the authors extracted audio from 4 high definition video recordings of interviews with native speakers of English (2 male, 2 female), whose accent features were characteristic of broadly defined Received Pronunciation.

2.1 Method

The audio recordings were then analyzed with a view to extracting sequences of sounds, in which (phonologically) voiced obstruents were flanked by other voiced segments. From each of the recordings, 200 samples were taken out. The selection was not random; the samples were extracted one after another as they appeared in the recording. Thus obtained 800 tokens of obstruents (X) between voiced sounds (V) could generally be classified into three categories (word initial (V#XV), word medial (VXV), and word final (VX#V):

V#**X**V	*have **g**o, my **b**usiness, editor**s** of*
V**X**V	*e**d**itors, a**b**out, bu**dd**ing, su**gg**estion*
V**X**#V	*ha**v**e go, an**d** I, an**d** er, insi**d**e of*

The waveforms and spectrograms of the samples were then inspected and labelled as either ‘fully voiced’ or ‘devoiced.’ The analyzed tokens were assigned to the first category when voicing was maintained throughout the closure and release in the case of stops, and during the entire period of close approximation in spirants. The segments were classified as ‘devoiced’ whenever there was loss of voicing in the medial phase of the stop and/or VOT was positive, and in the period of close approximation in fricative segments.

Examples of both cases are shown below. Figures 3 and 4 present voicing maintained throughout all stages of the plosive’s articulation; a fully voiced fricative is exemplified in Figure 5, whereas Figures 6 and 7 show devoiced obstruents.

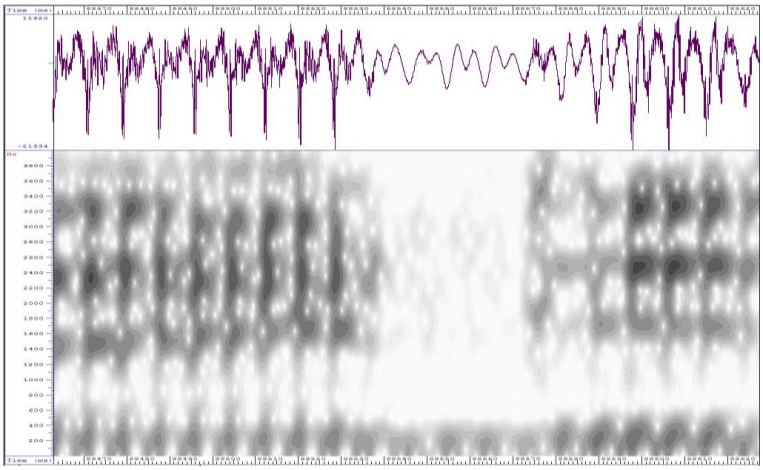


Figure 3 Full voicing of closure in [edɪ]*tors*

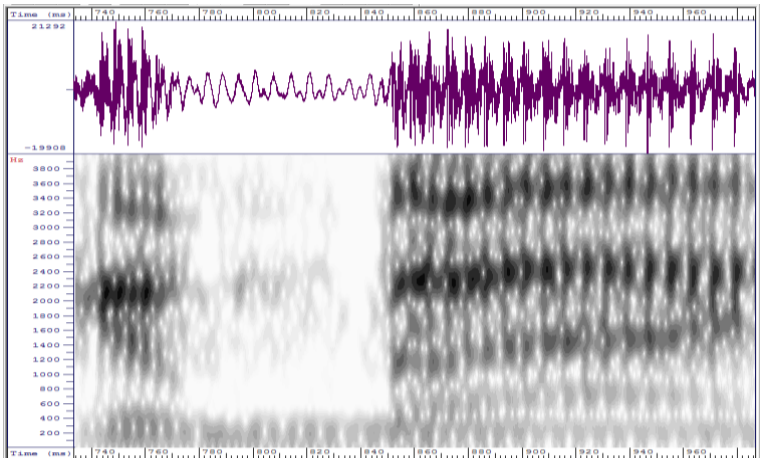


Figure 4 Full voicing in closure in welc[om#ba]ck

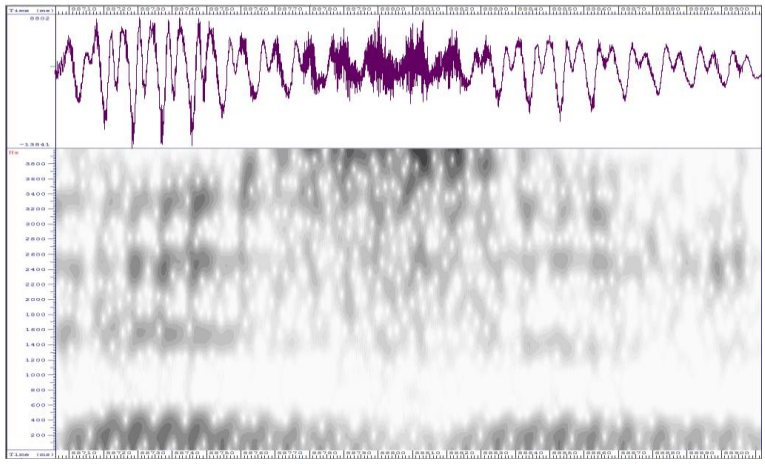


Figure 5 Full voicing od /z/ in edit[əz#ə]f

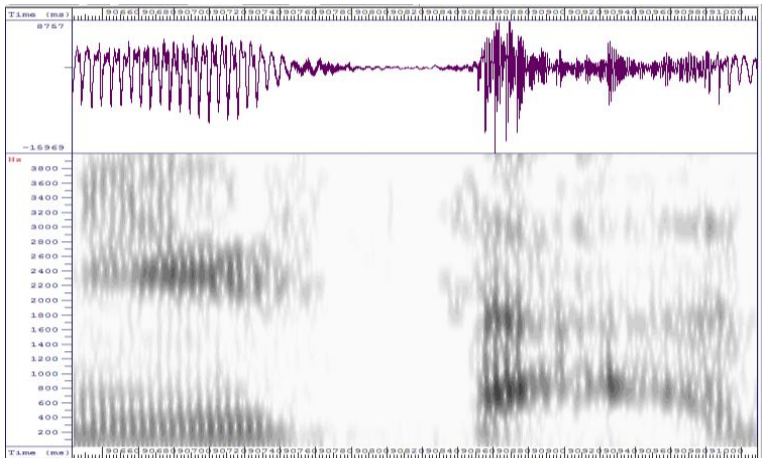


Figure 6 Devoicing of /z/ in u[s#ð]at

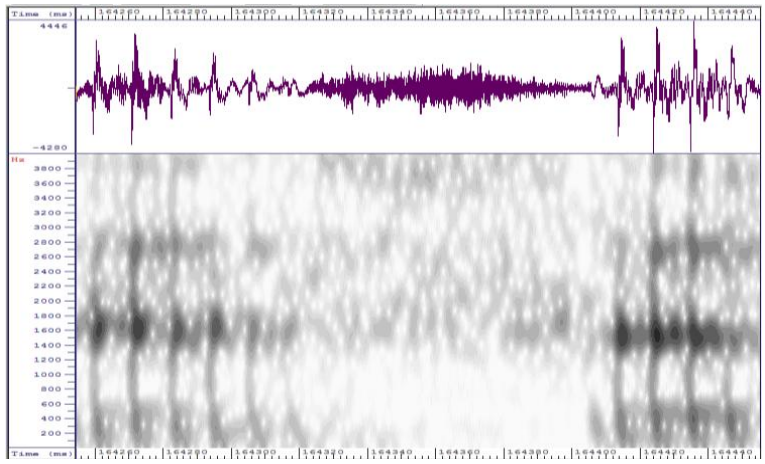


Figure 7 Devoicing of /z/ in character[s#ə]f

2.2 Results

Overall, 34 per cent of all the tokens were pronounced with voicing loss. The sections below present a detailed analysis of the results, taking into account the following factors:

- phonological category of the examined obstruents
- manner of articulation
- position in the word
- following and preceding context
- stress
- position in the syllable
- lexeme type

If we view the number of devoiced tokens in individual lenis obstruents, it appears that the differences between particular sound categories are more incremental than radical (cf. Fig. 8).

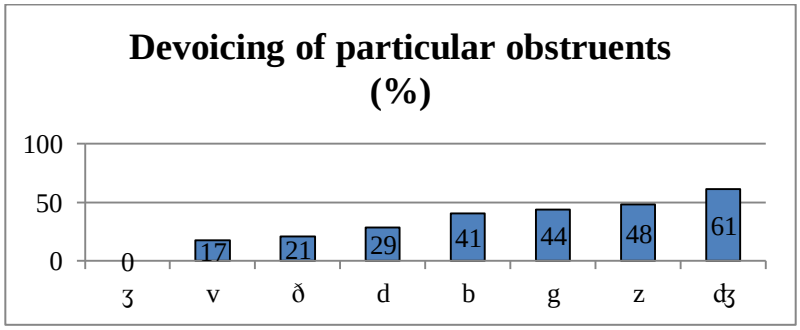


Figure 8 The percentage of devoiced tokens in particular sounds.

Although the arrangement of sounds in the sequence looks random and does not indicate any relationship with place or manner of articulation, there is a statistically significant difference ($p<0.001$) between the affricate which tends to be devoiced in more than 60% of the cases, and plosives and fricatives, in which devoicing occurs, respectively, in 35% and 30% of the cases (Figure 9). Moreover, the results obtained for obstruents containing fricative segments are in line with those presented in Haggard's study (1978) in that there appears a similar progression of devoiced sounds /v/ - /z/ - /ʤ/, with the palato-alveolar affricate becoming devoiced most often, and the labio-dental fricative most frequently retaining its voicing. It should also be noticed that the result for the palato-alveolar fricative /ʒ/ should not be regarded as valid for the whole category of lenis palato-alveolar fricatives due to the extremely low frequency of the sound; there occurred only one instance of this consonant in the analyzed material (*Asia*).

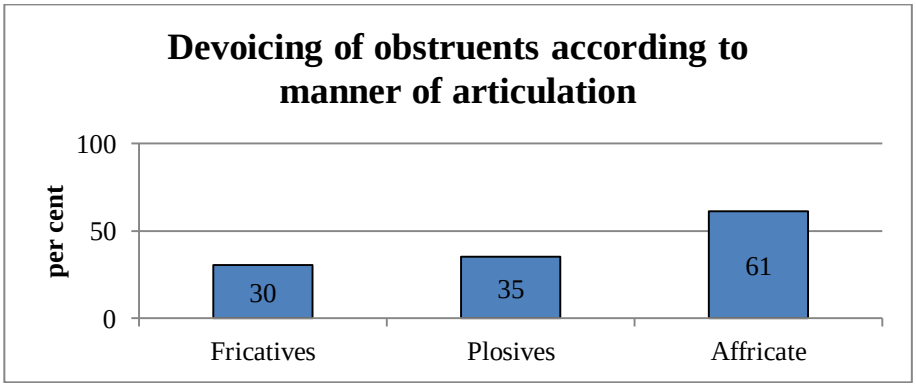


Figure 9 The percentage of devoiced tokens in particular manners of articulation.

In regard to the position in the word, voicing is retained most often word internally (80%), whereas most devoicing occurs word-initially (44%, Fig. 10), which shows the relevance of word boundaries in the implementation of voicing as pointed out by Docherty (1992:32). Similarly, in the case of plosives, the results (Fig. 11) match those in Flege and Brown (1982) and Westbury (1979) in that the sounds are least frequently devoiced in word-medial position, namely in 18% and 3.5%, respectively. The more frequent occurrence of word-medial devoicing in the present study, particularly in comparison to Westbury's result, could stem from the fact that the above mentioned analyses were carried out on elicited disyllabic words, not on spontaneous speech.

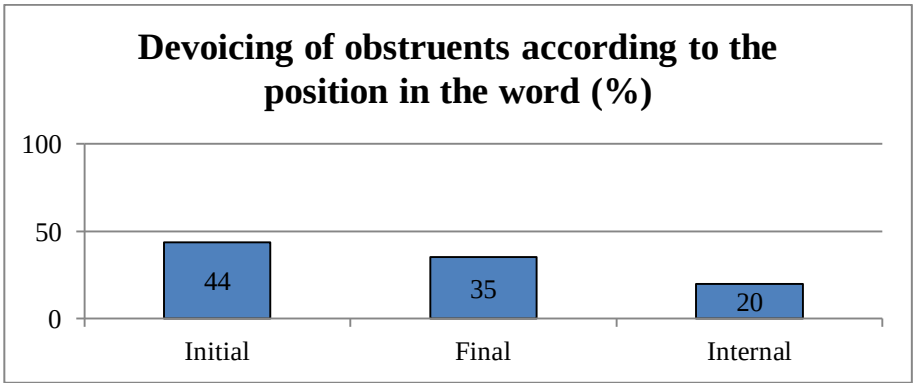


Figure 10 The percentage of devoiced tokens in different word positions.

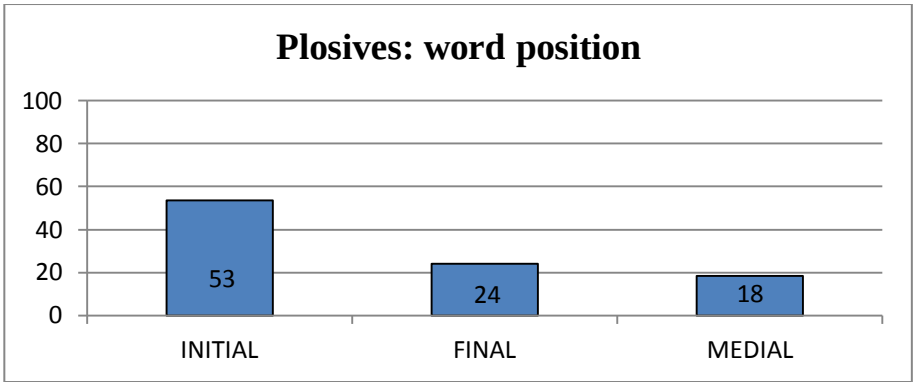


Figure 11 The percentage of devoiced plosives in different word positions.

Regarding the contexts in which obstruents occur, they are most often devoiced in the vicinity of an adjacent obstruent: 59% in the preceding, and 54% in the following context. In the context of preceding and following vowels and sonorants, devoicing is less frequent ($p<0.001$, cf. Figures 12 and 13). An analogous observation was made by Haggard (1978) in a study of words pronounced in isolation, which confirms that the neighbouring sounds are a relevant factor in the realization of voicing.

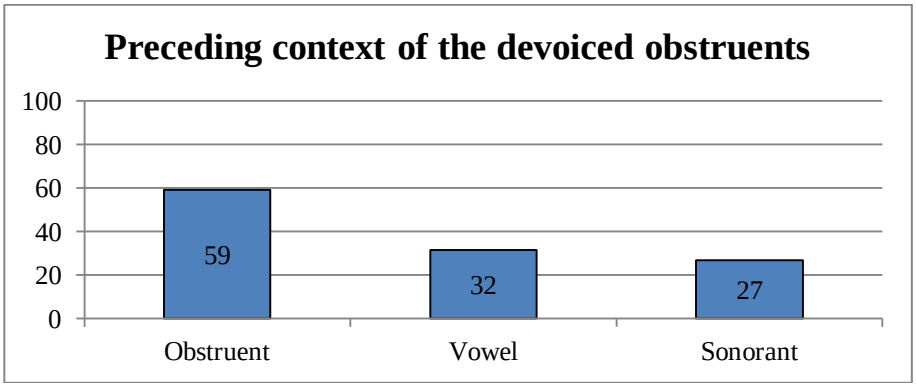


Figure 12 The percentage of devoiced tokens as preceded by specific sound categories.

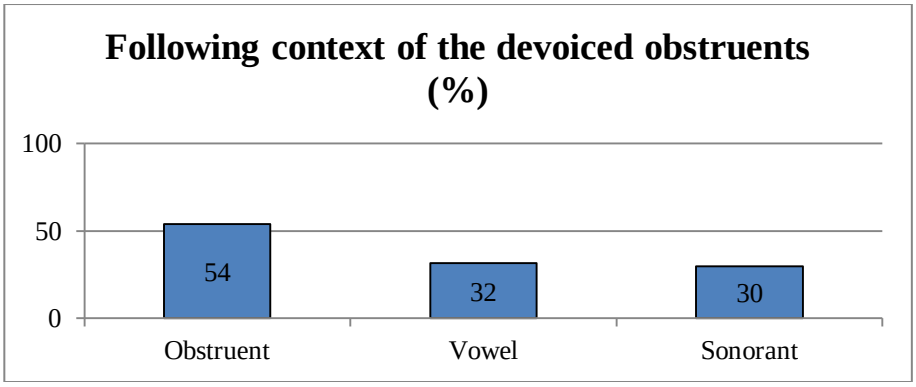


Figure 13 The percentage of devoiced tokens as followed by specific sound categories.

Considering the effect of stress on the voicing of intervocalic lenis obstruents, there is more devoicing ($p<0.001$) in stressed, than in unstressed, syllables (Fig. 14), while the position in the syllable does not exert a statistically significant effect on the whole (Fig. 15). Assigning word-medial obstruents to syllables was performed according to the Maximal Onset Principle (Goldsmith, 1990:128).

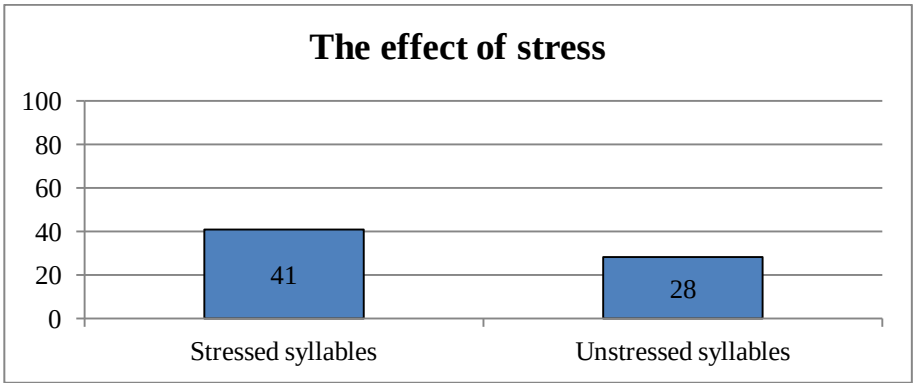


Figure 14 The effect of stress on the percentage of devoiced tokens.

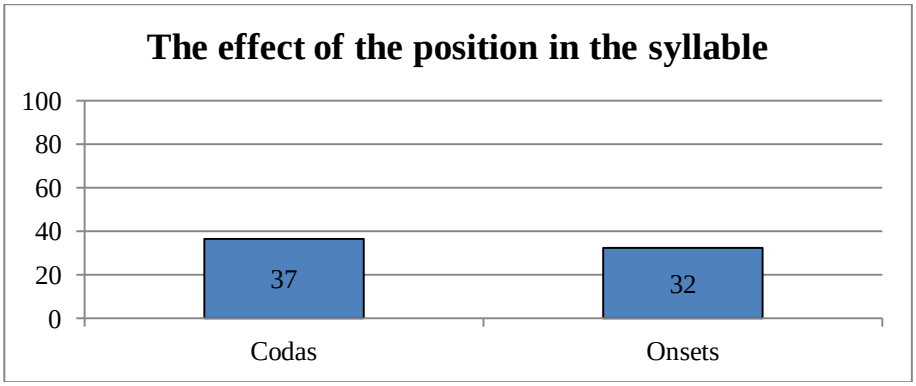


Figure 15 The effect of the position in the syllable on the percentage of devoiced obstruents.

When the interaction of stress and syllable position is taken into account, it appears that the greatest percentage of devoiced obstruents appears in stressed onsets. However, there is a similar amount of devoicing in the opposing environment, i.e. in unstressed codas, while significant differences concern the two previously mentioned contexts vs. stressed codas and vs. unstressed onsets (p=between 0.001 to 0.01, Fig. 16). Thus, it cannot be stated that a particular combination of the position in the syllable and the existence or lack of stress enhance of hinder devoicing.

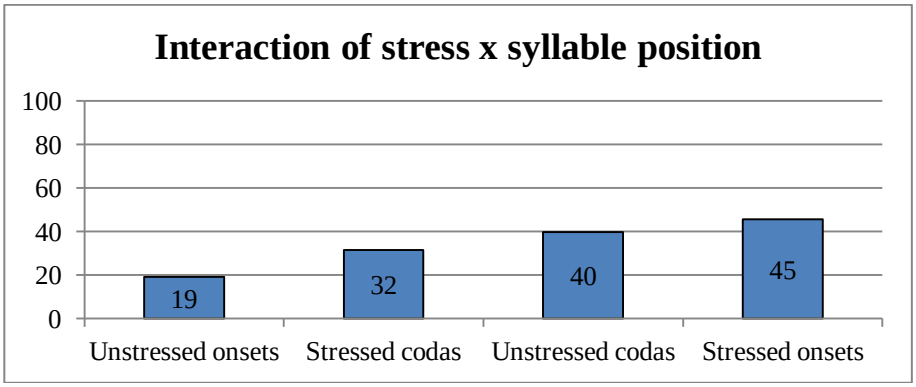


Figure 16 The effect of stress and the position in the syllable on the percentage of devoiced obstruents.

The distinction between function and content words has not found a reflection in the amount of devoicing, and was found in 31% and 36% of cases, respectively (Fig. 17).

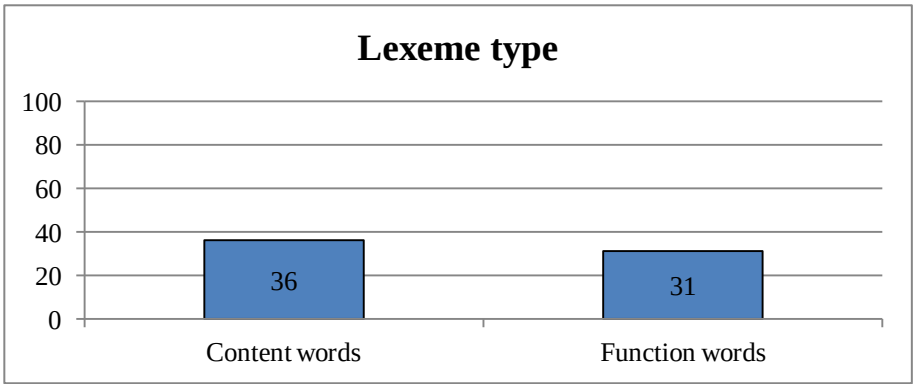


Figure 17 The percentage of devoiced obstruents in content and function words.

Let us now review the effect of stress in each manner of articulation (Figures 18-20).

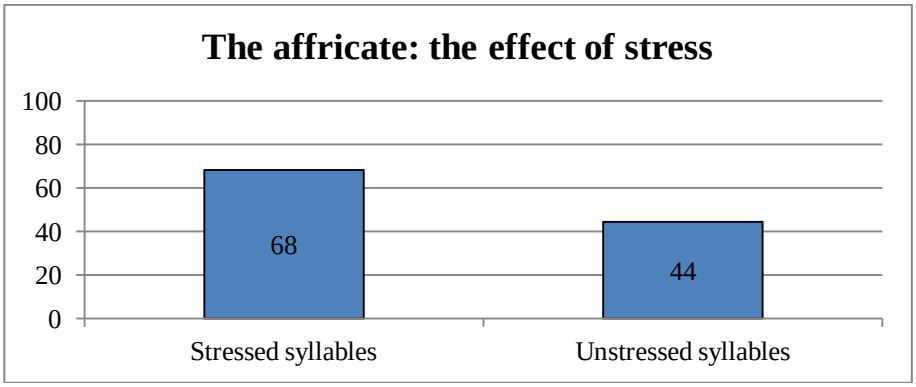


Figure 18 The percentage of devoiced affricates in stressed and unstressed syllables.

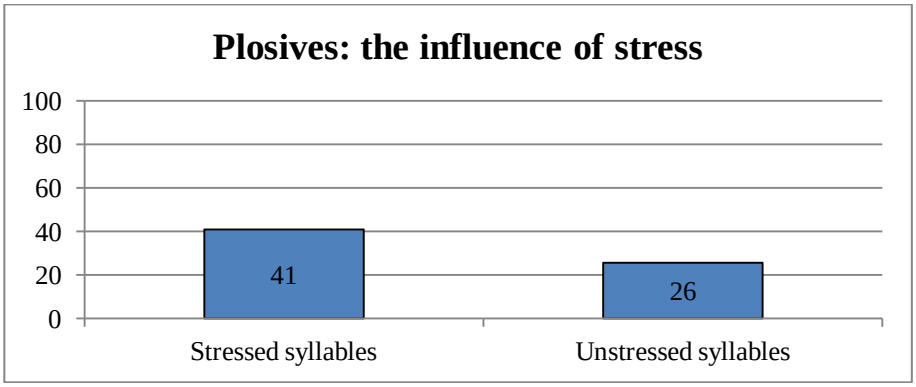


Figure 19 The percentage of devoiced plosives in stressed and unstressed syllables.

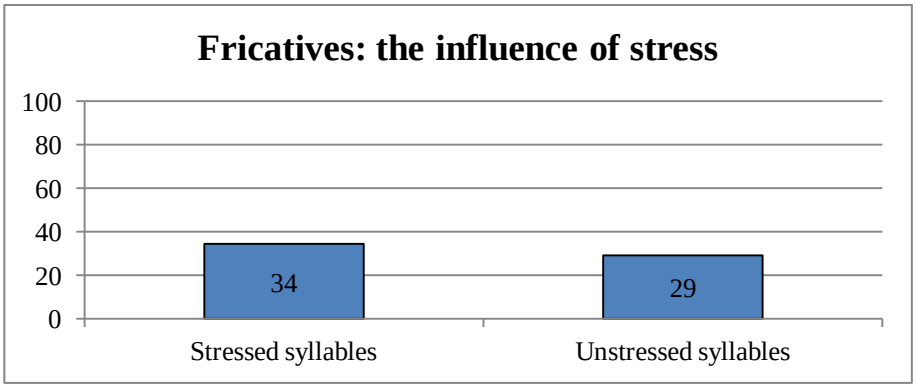


Figure 20 The percentage of devoiced fricatives in stressed and unstressed syllables.

Significant differences between the amount of devoicing in stressed vs. unstressed syllables were found were found in the affricate (Fig. 18) and in plosives (Fig. 19), while in fricatives the differences were not significant (Fig. 20).

Another comparison was done for the position in the syllable. As was observed in the effect of stress, here, too, the figures for affricates (Fig. 21) are markedly larger than those for fricatives (Fig. 21) and plosives (Fig. 22).

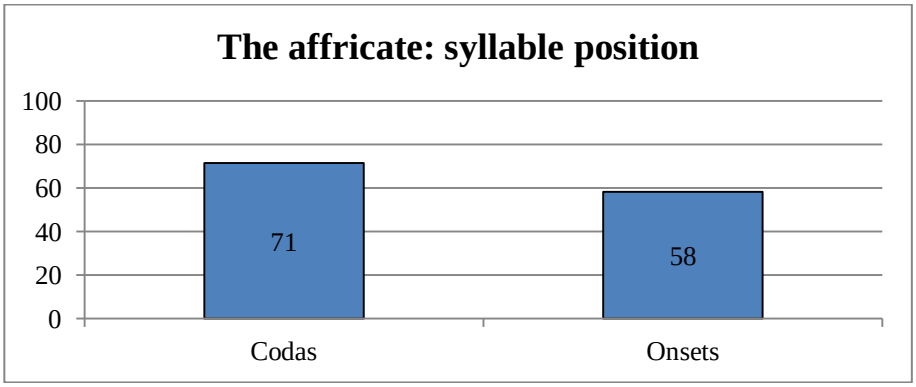


Figure 21 The percentage of devoiced affricates in the onset and coda of the syllable

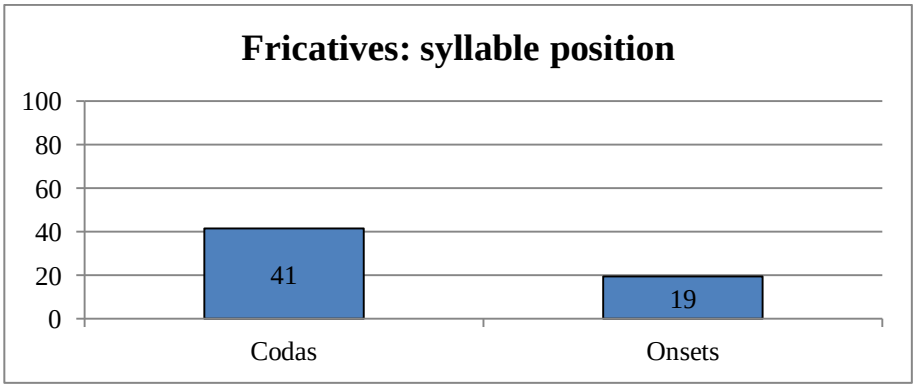


Figure 22 The percentage of devoiced fricatives in the onset and coda of the syllable

The relation of devoicing vs. position in the syllable is reversed in plosives, where more devoicing was noted in onsets than in codas (Fig. 23).

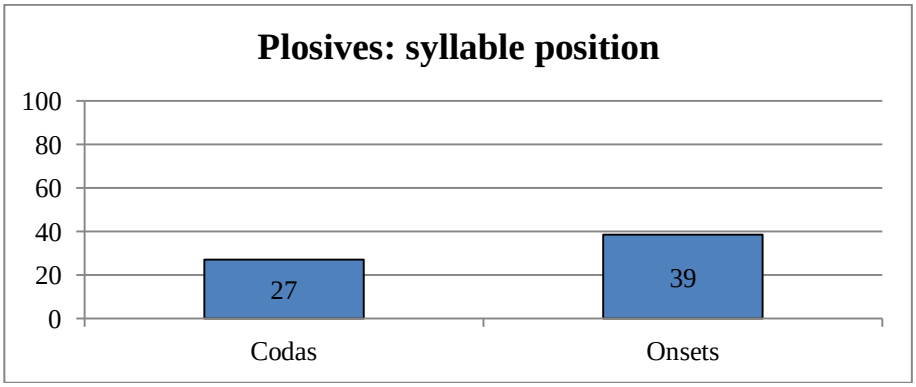


Figure 23 The percentage of devoiced plosives in the onset and coda of the syllable.

Finally, let us observe the interaction of devoicing with the position in the syllable x stress (cf. Fig. 15 averaged across manner of articulation).

As there appeared no token containing the palato-alveolar affricate in an unstressed coda, Figure 25 shows only three bars for the contexts available in the study.

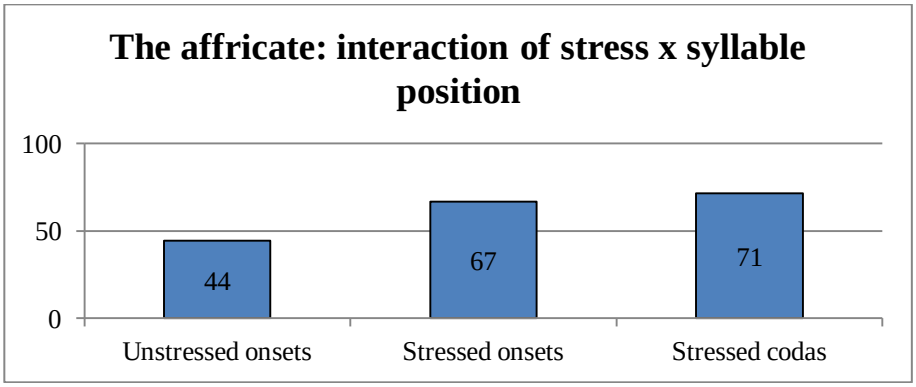


Figure 24 The percentage of devoiced affricates in stressed and unstressed codas and onsets

Thus in the affricate, devoicing is significantly stronger ($p<0.001$) when under stress. The results in plosives (Fig. 23) are similar to those in fricatives (Fig. 24), with unstressed onsets and stressed codas favouring devoicing more than the remaining two contexts ($p<0.001$).

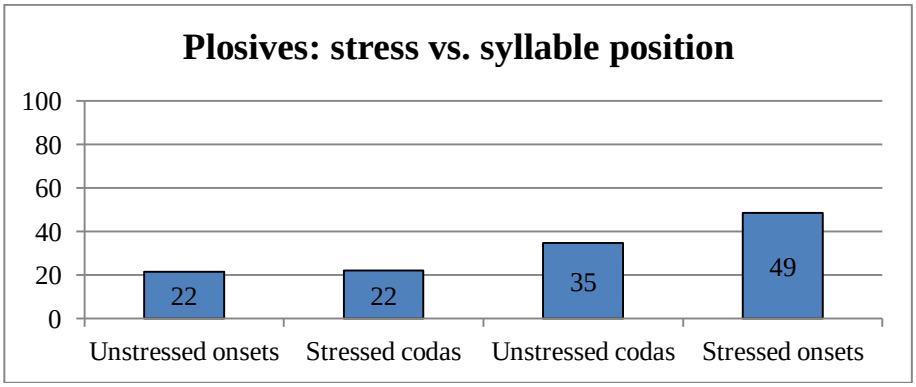


Figure 25 The percentage of devoiced plosives in stressed and unstressed codas and onsets.

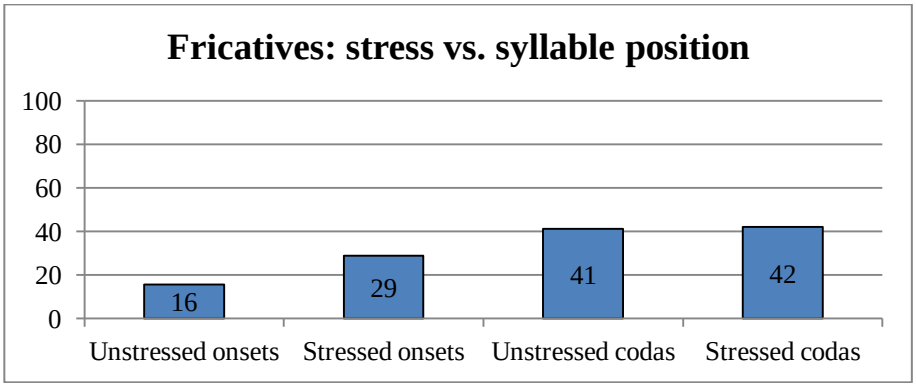


Figure 26 The percentage of devoiced fricatives in stressed and unstressed codas and onsets

3. Conslusions

Most of the factors considered in the present study appear to affect voicing in intervocalic obstruents. Regarding particular sound categories and manners of articulation, the affricate is devoiced twice as frequently as plosives and fricatives, and of other obstruents, /z/ is most frequently devoiced, probably because its voicing is often predictable morphologically and does not have to be manifested phonetically, while /v/ and /ð/ were devoiced rarely. Plosives are devoiced still less frequently than /z/.

Considering the position of analyzed sounds in the word, it is interesting to see that obstruents devoice more frequently when word-initial than when word-final. This shows that in English the tendency to prolong VOT in stressed syllables exerts a stronger effect than the reduction of Voicing-Into-Constriction.

Examining voicing in relation to adjacent sounds, it was noted that preceding and following voiced obstruents do not retain voicing as strongly as one would expect; vowels and sonorants exert a stronger voicing-retention effect.

Devoicing is also conditioned suprasegmentally, as most frequently devoicing takes place in stressed syllables.

References

Ball, M.J. / Joan Rahilly, J. (1999) *Phonetics The Science of Speech*, London: Arnold

Catford, J. C. (1988) *A Practical Introduction to Phonetics*, Oxford: Oxford University Press, (2001)

Catford, J.C. (1964) "Phonation types: the classification of some laryngeal components of speech production", (in:) Abercrombie, D. et al. (eds.) *In honour of Daniel Jones* (p. 26-37), London: Longman

Catford, J.C. (1977) *Fundamental Problems in Phonetics*, Edinburgh: Edinburgh University Press

Clark, J. / Yallop, C. (1990) *An Introduction to Phonetics and Phonology*, Oxford: Basil Blackwell

Davenport, M. / Hannahs, S.J. (1998) *Introducing Phonetics and Phonology*, London: Arnold

Docherty, G.J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin: Foris Publications.

Flege, J. & Brown, W., Jr. (1982). The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. *Journal of Phonetics*, 10, 335-345.

Fujimura, O. / Erickson, D. (1999) "Acoustic phonetics", (in:) Hardcastle, W.J. / Laver, J. (eds.) *The Handbook of Phonetic Sciences* (p. 65-115), Oxford: Blackwell Publishers

Gimson, A.C. (1962) *Gimson's Pronunciation of English*, London: Arnold, (2001)

Goldsmith, A. 1990. *Auto segmental and Metrical Phonology*, Massachusussetts: Basil Blackwell LTD

Gonet, W. (1989) *Factorial Analysis of the Duration of R.P. Monophthongs in Monosyllabic Words*, diss., Institute of English, Maria Curie-Skłodowska University, Lublin; Acoustic Phonetics Department, Polish Academy of Sciences, Poznań

Gonet, *W.* (2001) "Obstruent Voicing in English and Polish. Pedagogical Perspective", *International Journal of English Studies*, 1, nr 1, 73-92

Gonet, W. (2001) Voicing Control in English and Polish: A Pedagogical perspective. *International Journal of English Studies*. Murcia (Spain): Universidad de Murcia, pp. 73-92.

Gonet, W. 2010. Dispelling the Myth of Word-Final Obstruent Voicing in English: New Facts and Pedagogical Implications. In: E. Waniek-Klimczak (ed.), *Issues in Accents of English 2*. Cambridge: Cambridge Scholars Publishing pp. 361-376.

Gonet, W. and K. Różańska (2003). Voice Onset Time in Word Initial Lenis Plosives in the Speech of Four BBC Presenters. *Speech and Language Technology* Vol. 7, pp. 35-52. Poznań: Polish Phonetic Association.

Gonet, W. and L. Stadnicka. (2006). Vowel Clipping in English. *Speech and Language Technology*. Vol. 8. Poznań: Polish Phonetic Association.

Haggard, M. (1978). The devoicing of voiced fricatives. *Journal of Phonetics* 6. 95-102.

Jassem, W. (1983) The Phonology of Modern English, PWN: Warszawa

Ladefoged, P. (1971) Preliminaries to Linguistic Phonetics, Chicago: University of Chicago Press

Ladefoged, P. (1975) A Course in Phonetics, New York: Harcourt Brace Jovanovich

Lipowska, E.B. (1991) Voicing of word final obstruents in R.P. English, M.A. thesis, Maria Curie-Skłodowska University, Lublin

Lisker, L. and Abramson, A.S. (1964) "A cross-language study of voicing in initial stops: Acoustical measurements", (in:) *Language and Speech* 10, 1-28

Maddieson, I. (1999) "Phonetic universals" (in:) Hardcastle, W.J. / Laver, J. (eds.) *The Handbook of Phonetic Sciences* (p. 619-639), Oxford: Blackwell Publishers

Ohala, J.J. (1999) "The Relation between Phonetics and Phonology", (in:) Hardcastle, W.J. / Laver, J. (eds.) *The Handbook of Phonetic Sciences* (p. 674-694), Oxford: Blackwell Publishers

Port, R.F. and Rottuno, R. (1979). "Relation between voice-onset time and vowel duration", (in:) *Journal of the Acoustical Society of America*, 66, 654-662

Raphael, L.J. et al (1975) "Vowel duration as cues to voicing in word-final stop consonants: spectrographic and perceptual studies", *Journal of Speech and Hearing Research*, 18, 389-400

Roach, P. (1983) *English Phonetics and Phonology. A Practical Course*, Cambridge: Cambridge University Press, 1991

Schockey, L. (2003) *Sound Patterns of Spoken English*, Oxford: Blackwell Publishing

Szpyra-Kozłowska, J. (2003) "The Lingua Franca Core and the Polish learner", (in:) W. Sobkowiak, E. Waniek-Klimczak (eds.) *Dydaktyka fonetyki języka obcego. Neofilologia. Zeszyty Naukowe*, 5, 193-210, Płock: Państwowa Wyższa Szkoła Zawodowa

Szpyra-Kozłowska, J. et al. (2003) "Priorytety w nauczaniu języka angielskiego", (in:) W. Sobkowiak, E. Waniek-Klimczuk (eds.) *Dydaktyka fonetyki języka obcego. Neofilologii. Zeszyty Naukowe*, 5, 211-223, Płock: Państwowa Wyższa Szkoła Zawodowa

Van den Berg, J. (1958) "Myoelastic-aerodynamic theory of voice production", *Journal of Speech and Hearing Research* 3, 1, 227-244

Westbury, John (1979) Aspects of the Temporal Control of Voicing in Consonant Clusters in English, *Texas Linguistic Forum 14*. Department of Linguistics, University of Texas, Austin.

MEASURING VOWEL DURATION VARIABILITY IN NATIVE ENGLISH SPEAKERS AND POLISH LEARNERS[1]

ANDRZEJ PORZUCZEK
University of Silesia

Abstract
This paper presents a set of simple statistical measures that illustrate the difference between native English speakers and Polish learners of English in varying the length of vocalic segments in read speech. Relative vowel duration and vowel length variation are widely used as basic criteria for establishing rhythmic differences between languages and dialects of a language. The parameter of vocalic duration is employed in popular measures such as ΔV (Ramus et al. 1999), VarcoV (Dellwo 2006, White and Mattys 2007), and PVI (Low et al. 2000, Grabe and Low 2002). Apart from rhythm studies, the processing of data concerning vowel duration can be used to establish the level of discrepancy between native speech and learner speech in investigating other temporal aspects of FL pronunciation, such as tense-lax vowel distinction, accentual lengthening or the degree of unstressed vowel reduction, which are often pointed out as serious problems in the acquisition of English pronunciation by Polish learners. Using descriptive statistics (relations between personal mean vowel duration and standard deviation), the author calculates several indices that demonstrate individual learners' (13 subjects) scores in relation to the native speakers' (12 subjects) score ranges. In some tested aspects, the results of the two groups of speakers are almost cleanly separated, which suggests not only the existence of specific didactic problems but also their actual scale.

1. Introduction

Foreign language (FL) pronunciation is traditionally assessed by the teacher on the basis of immediate subjective impressions. Although in classroom teaching practice this will probably remain the basic approach, the recent development of PC-operated methods of speech analysis has made them available to people outside the circle of professional laboratory phoneticians, including FL teachers, who can now consider the use of acoustic analysis as an interesting accessory didactic aid.

Not all speech signal parameters can be easily employed for pedagogical purposes, but speech unit duration measurement is relatively reliable and informative. The segmentation of speech chain is not always an easy task even if clear and consistent criteria are applied, and it is time-consuming, but before the automatic methods are made fully reliable, manual segmentation gives the researcher a better insight into the data.

[1] Research supported by the Polish Ministry of Science and Higher Education via Grant No.:0576/B/H03/2010/38.

The duration of speech units provides a researcher with a lot of useful information. Vowel length appears to be a particularly interesting aspect of speech timing from the point of view of the Polish learner of English (cf. Waniek-Klimczak 2005). This is because relative vocalic duration in English can cue

- tense/lax vowel contrast (as an accessory cue)
- fortis/lenis contrast in coda
- prominence distribution
- prosodic domain boundaries
- rhythm patterns

Polish, however, is characterised by

- no tense/lax vowel distinction
- the voiced/voiceless contrast neutralised in coda
- very little unstressed vowel reduction
- allegedly weaker accentual lengthening.

Moreover, although final lengthening and initial strengthening are said to be universal phenomena, we may face cross-linguistic discrepancies in the scale of their effects on prosodic unit duration. Finally, Polish gives the listener more syllable-timing impression despite extremely complex consonant clusters.

All these discrepancies may lead to cross-linguistic interference in the process of FL learning. A number of researchers dealing with English phonetics pedagogy indeed report problems with insufficient intrinsic vowel length distinction (Sobkowiak 1996, Szpyra-Kozłowska 2003, Nowacka 2008, Bryła 2010), insufficient unstressed syllable reduction and too short prominent syllables in Polish learners (Avery and Ehrlich 1996, Hewings 2004, Dziubalska-Kołaczyk et al. 2006, Gonet et al. 2010) and especially insufficient vowel reduction in Polish learners of English (Luke and Richards 1982, Sobkowiak 1996, Hewings 2004, Nowacka 2008, Gonet et al. 2010, Porzuczek 2010). Most opinions, however, are formulated with reference to auditory assessment and pedagogical experience.

2. Objectives of the present study

There are two main objectives of the present study:

- to provide evidence for vocalic timing differences between native English speakers and Polish learners that will illustrate the scale of learners' problems with the 'short'/'long' and stressed/unstressed temporal vocalic contrasts,
- to illustrate the developmental tendencies in the learners' speech by repeating the testing procedure after 7 months of study including a course of practical phonetics.

The obtained evidence can also be used for further investigations into the rhythmic patterns of the Polish learner's English speech.

3. Method

The subjects were 13 Polish first-year students of English at a teacher training college. Their performance (2 recording sessions – October 2006, May 2007), originally recorded

for a more comprehensive study of EFL speech timing (Porzuczek, in press), was analysed in comparison to the performance of 12 English secondary school students in Cambridge, downloaded from the IViE database (Grabe et al. 2001). The participants read the Cinderella passage (Grabe et al. 2001, see Appendix). They had been given time to practise the reading prior to the recording.

The tested material included 46 vocalic syllable nuclei (see Appendix):

- 20 unstressed reduced vowels (17 non-phrase-final)
- 20 stressed monophthongs (10 non-phrase-final), (5 'long' vowels, 12 'short' vowels, 3 æ's)
- 6 stressed diphthongs (3 non-phrase-final)

Vowels adjacent to approximants and phrases showing significant interspeaker differences in prominence distribution were avoided. Stressed syllables were thus lexically and syntactically determined. This approach helps to reduce the problems which call for automatic segmentation (e.g. Loukina et al. 2011). The acoustic analysis for the purposes of the present research was based on manual segmentation and measurement (standard criteria) from the spectrograms and waveforms using the PRAAT software (Boersma 2001). The data analysis involved descriptive statistics including group and personal vowel duration medians, means and standard deviation. Raw measurements were normalised for speech rate by using proportions of vowel class mean durations and VarcoV (Dellwo 2006, White and Mattys 2007). VarcoV is calculated as the percentage proportion of standard deviation from mean vowel duration (SD) to mean vowel duration (VarcoV=SD*100%/meanV, where V=vowel duration).

Acoustic research tools based on duration, such as the recent rhythm measures, yield results marked with significant individual variation. As Loukina et al. (2011) notice, in cross-linguistic rhythm studies more variation is often found between individual speakers than between languages. The same problem may therefore appear in comparing native and non-native speech within one language. This poses a problem of data interpretation, especially for normative didactic purposes. It seems justified though to assume that results out of the range of native speakers' scores indicate non-native-like pronunciation features.

4. Results

Predictably, group means show significant differences between native and non-native English speech in both investigated aspects. Mean stressed vowel durations are presented in Table 1.

group\V class	D	L	A	S	text grand mean
PL1	199	147	120	98	133 (SD=65=48%)
PL2	176	137	115	91	122 (SD=58=48%)
ENG	203	147	137	85	130 (SD=72=55%)

Table 1: Mean durations (ms) of particular vowel classes (D=diphthong, L=long, A=ash, S=short) in stressed syllables and vowel length variability (Porzuczek, in press).

The general results suggest similar articulatory rates in both groups of subjects, as indicated by similar mean vowel durations. Stressed vowel duration variability is higher in native speakers (ENG). After the training (PL2), the learners noticeably accelerate, but the variability index (SD/mean duration) remains identical. There is also a larger temporal difference between particular vowel classes in the pronunciation of native speakers.

Table 2 presents more information concerning the performance of individual speakers, which is important in the context of teaching groups of learners and setting the norms.

group\V class	D:S	L:S	A:S
PL1	1.8-2.25 (**2.1**)	1.22-1.75 (**1.5**)	.92-1.51 (**1.25**)
PL2	1.57-2.33 (**1.9**)	1.21-1.78 (**1.5**)	.94-1.59 (**1.33**)
EN	1.95-2.82 (**2.4**)	1.47-2.29 (**1.7**)	1.12-1.85 (**1.69**)

Table 2: Vowel class mean length proportions in individual speakers' score ranges. Group medians in parentheses.

It turns out that the learners' group medians for L:S ratio (1.5) in both recordings approximate the native speakers' minimum (1.47). However, the ranges largely overlap and, despite significant group differences, most Polish learners fall within the norms of native-like performance. Individual speakers' scores are shown in Appendix B.

The results indicate that the duration contrasts between vowel classes are clearer in native speakers. Still, even though group scores differ significantly, there are a number of native speakers who show less vowel length variation. This may suggest that either many Polish learners make a proper distinction between the vowel classes, at least for the 'long'/'short' vowel contrast, or that the scale of this quantitative distinction is irrelevant as long as a minimum contrast level is reached, e.g. approximately a 1.5:1 ratio for the present text. In order to account for possible effects of extraneous variables, we tried to observe the impact of pre-fortis clipping and final lengthening. The relevant calculations showed 15% shorter vowels in pre-fortis positions in the native performance. The learners made such vowels 8% shorter in the first recording and 16% shorter in the second. There was more difference in final lengthening, however, which made the native vowels three times longer than in non-phrase-final syllables, while the Polish learners made their vowels in prepausal syllables twice as long (Table 3). The ratio, which we call FLQ (final lengthening quotient), is obtained by dividing a subject's mean vowel duration in phrase-final syllables by mean vowel duration in non-phrase-final syllables.

group	FLQ = mean final (N=7): mean non-final (N=19)
PL1	1.64-2.51 (1.95)
PL2	1.63-2.75 (2.09)
EN	2.28-3.32 (2.9)

Table 3: Personal final lengthening quotient (FLQ) ranges and group medians (in parentheses).

The same data, illustrating individual subjects' performance, are also presented in Fig. 1 below.

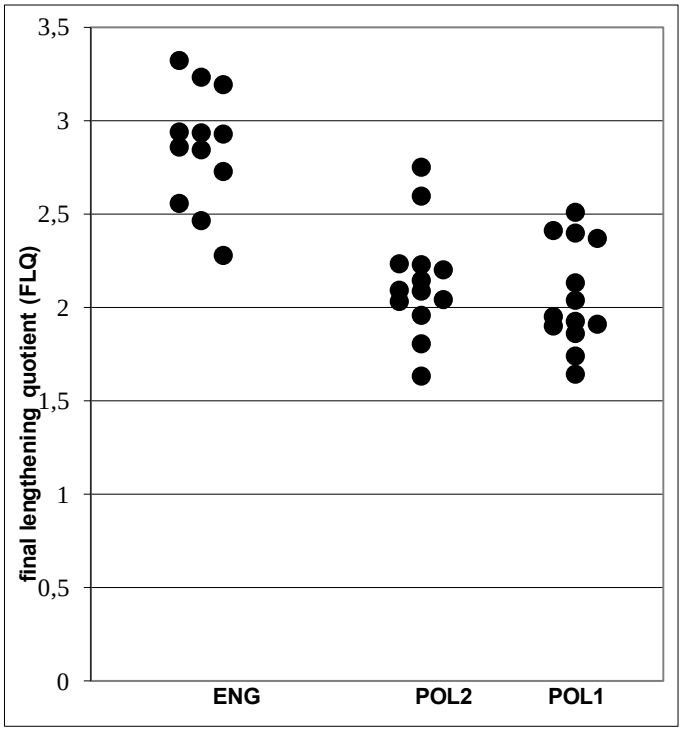


Fig. 1: Individual final lengthening quotient (FLQ) in English and Polish speakers.

The strong effect of final lengthening makes it advisable to present the results of the research with respect to non-phrase-final syllables as well as the overall scores, even though the process does not seem to have a very strong effect, for instance, on L:S ratios (Table 4) or general vowel length variability (Table 5), especially in terms of score ranges.

group\V class	L:S (non-final)	L:S (overall)
PL1	1.35-2.18 (**1.7**)	1.22-1.75 (**1.5**)
PL2	1.24-1.79 (**1.6**)	1.21-1.78 (**1.5**)
EN	1.5-2.32 (**1.7**)	1.47-2.29 (**1.7**)

Table 4: Personal 'long':'short' vowel ratio ranges and group medians.

1	2	3	4	5	6
group	overall (26)	non-final (19)	group mean VarcoV	overall (26) VarcoV	non-final (19) VarcoV
PL1	112-160 (132)	94-127 (108)	48	39-55 (49)	33-51 (39)
PL2	100-140 (127)	82-119 (106)	48	39-62 (47)	30-49 (36)
EN	106-155 (127)	87-121 (100)	55	44-63 (53)	30-51 (44)

Table 5: Personal mean vowel duration ranges and group medians (2-3). Personal vowel length variation (5-6).

Apart from final lengthening and pre-fortis clipping, there is yet another potential extraneous variable, viz. the complex and gradient nature of prominence. As was already mentioned earlier, because of the lack of a continuous scale that could be used to measure prominence taking into account all its components and their contribution, we can only try to control its effects on duration by careful selection of contexts where structural prominence is unambiguously distributed.

Generally, two conclusions can be formulated with respect to stressed vowel length variability. Firstly, all native speakers and a majority (2/3) of Polish speakers before training make the long vowels at least 50% longer than the short ones. Secondly, final lengthening appears much stronger in the pronunciation of native speakers.

Far more spectacular results are obtained if vowels in both stressed and unstressed syllables are taken into consideration. The differences can be captured by both VarcoV and vowel reduction quotient (VRQ), calculated for individuals by dividing their mean unstressed vowel duration by mean stressed vowel duration. Tables 6 and 7 show the relevant VarcoV (SD:M) results[2] for non-final contexts and all tested vowels. Native speakers' codes are shown in bold. Polish learners' codes are followed by "1" (1st recording) or "2" (second recording).

subject	**M**		subject	**SD:M**
CSM	**62**		**CSM**	**0.72**
AK2	**62**		**CER**	**0.71**
CHB	**63**		**CPT**	**0.68**
AS2	**63**		**CHB**	**0.68**
CTG	**64**		**CMF**	**0.67**
CMF	**68**		**CLP**	**0.66**
CER	**69**		**CTG**	**0.64**
CMA	**69**		**CLH**	**0.63**

subject	**M**		subject	**SD:M**
CTG	**81**		**CHB**	**0.78**
AK2	**81**		**CJE**	**0.77**
AS2	**82**		**CLH**	**0.77**
CSM	**82**		**CSM**	**0.75**
RM2	**83**		AK2	**0.75**
CHB	**87**		**CMF**	**0.74**
MG2	**88**		**CER**	**0.74**
CMA	**90**		**CLP**	**0.7**

[2] The figures are not multiplied by 100 as in the original VarcoV formula.

subject	**M**		subject	**SD:M**
RM2	**70**		**CJE**	**0.63**
CPT	**72**		AS2	**0.62**
CJE	**73**		AK2	**0.62**
AS1	**78**		**CJI**	**0.61**
CLP	**78**		**CMC**	**0.57**
MG2	**79**		AK1	**0.55**
PA1	**82**		PS2	**0.55**
AO2	**82**		**CMA**	**0.54**
RM1	**82**		PS1	**0.52**
PS2	**84**		AS1	**0.52**
CLH	**84**		JK2	**0.52**
AK1	**87**		PA1	**0.51**
CMC	**87**		RM2	**0.51**
PA2	**87**		MG2	**0.5**
CJI	**91**		AO2	**0.5**
PO2	**91**		LK1	**0.5**
AJ2	**92**		DK2	**0.49**
DK2	**92**		AO1	**0.48**
JK2	**93**		PA2	**0.48**
LK2	**93**		DK1	**0.48**
MG1	**95**		AJ1	**0.47**
AO1	**95**		MG1	**0.46**
LK1	**99**		PO1	**0.46**
JK1	**100**		MB1	**0.45**
AJ1	**101**		PO2	**0.44**
DK1	**101**		LK2	**0.43**
PS1	**101**		JK1	**0.43**
MB2	**102**		AJ2	**0.42**
PO1	**104**		RM1	**0.41**
MB1	**112**		MB2	**0.36**

Table 6: Non-final mean vowel duration (M) and duration variability (SD:M) (19 stressed vowels + 17 schwas)

subject	**M**		subject	**SD:M**
CPT	**91**		**CPT**	**0.7**
CER	**91**		**CMA**	**0.69**
CMF	**92**		AS2	**0.68**
RM1	**98**		**CTG**	**0.68**
AO2	**99**		**CJI**	**0.66**
AS1	**99**		AK1	**0.64**
PS2	**103**		**CMC**	**0.63**
CJE	**103**		AS1	**0.61**
CLP	**104**		PA1	**0.61**
PA2	**104**		PS2	**0.59**
LK2	**104**		PO1	**0.58**
PA1	**105**		PA2	**0.57**
CMC	**105**		RM2	**0.57**
PO2	**107**		JK2	**0.56**
DK2	**108**		PS1	**0.56**
JK2	**108**		DK1	**0.55**
MG1	**109**		AO2	**0.54**
AK1	**110**		DK2	**0.54**
AJ2	**111**		MG2	**0.52**
AJ1	**111**		AO1	**0.51**
AO1	**112**		MB2	**0.51**
LK1	**114**		LK1	**0.51**
JK1	**115**		AJ2	**0.5**
CLH	**116**		PO2	**0.5**
CJI	**118**		AJ1	**0.49**
DK1	**120**		RM1	**0.49**
PS1	**123**		JK1	**0.49**
MB2	**123**		MG1	**0.48**
PO1	**126**		MB1	**0.47**
MB1	**129**		LK2	**0.47**

Table 7: Overall mean vowel duration and duration variability (SD:M) (26 stressed vowels + 20 schwas)

The data from Tables 6 and 7 are also presented as a graph in Figure 2 for a clearer illustration of cross-group and individual differences.

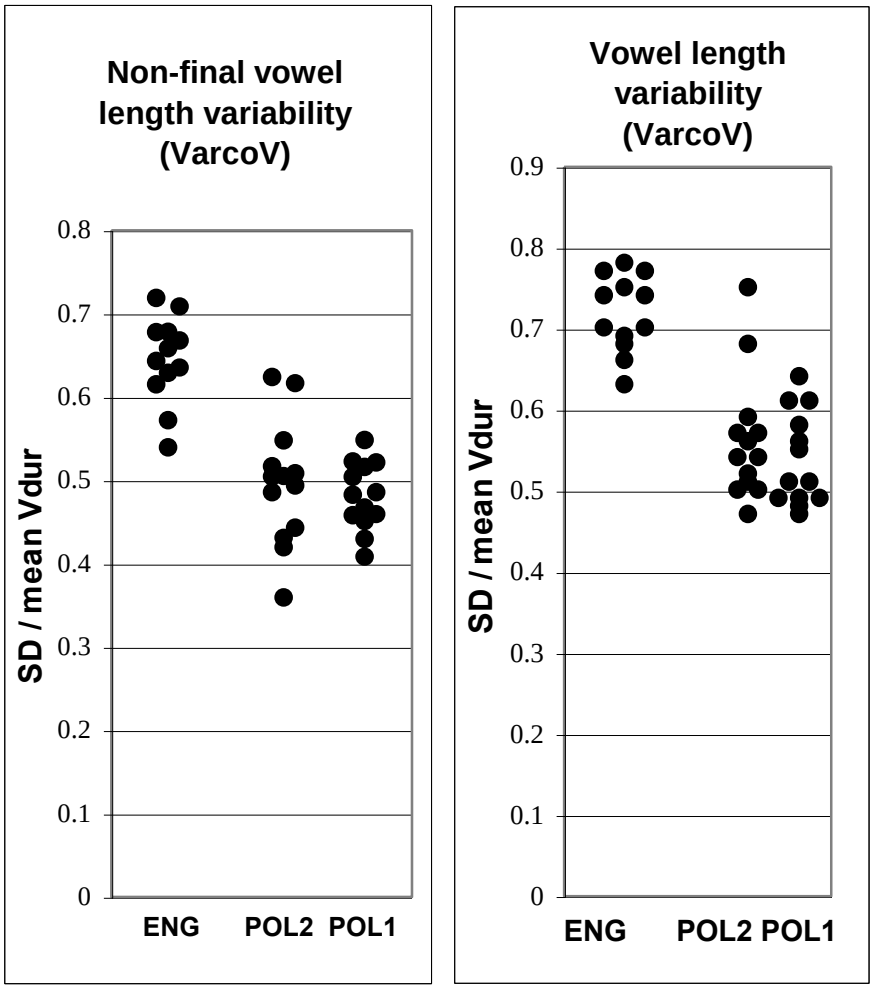


Figure 2: Vowel duration variability.

VarcoV shows the general vowel length variability, which may be influenced by other factors, while VRQ focuses on the stressed/unstressed distinction, and shows the scale of quantitative vowel reduction. It is presented in Table 8 and Figure 3.

S	**MstrV**		S	MstrV	**Mschwa**	S	**VRQ**
AK2	**82**		**CSM**	89.8	**31**	**CSM**	**0.34**
RM2[3]	**84**		**CHB**	90.2	**32**	**CHB**	**0.36**
AS2	**84**		**CMF**	98.1	**35**	**CMF**	**0.36**
CTG	**87**		**CER**	100	**36**	**CER**	**0.36**
MG2	**88**		**CTG**	87.1	**39**	**CMC**	**0.43**
CSM	**90**		AS2	84.1	**40**	**CLP**	**0.43**
CHB	**90**		AK2	82.2	**40**	**CPT**	**0.43**
CMA	**91**		**CPT**	98.8	**42**	**CJE**	**0.43**
RM1	**94**		**CJE**	100	**43**	**CTG**	**0.45**
AS1	**95**		**CMA**	91.1	**45**	**CLH**	**0.46**
CMF	**98**		**CLP**	107	**46**	**CJI**	**0.47**
AJ1	**99**		**CMC**	119	**51**	AS2	0.47
CPT	**99**		**CLH**	113	**51**	AK2	0.49
MG1	**99**		*RM2*	**83.5**	**55**	**CMA**	**0.5**
CER	**100**		AO2	106	**55**	AO2	0.52
CJE	**100**		PA1	106	**56**	PA1	0.53
PS2	**102**		**CJI**	121	**57**	DK2	0.54
AJ2	**105**		AS1	94.6	**59**	PA2	0.56
PA1	**106**		PA2	109	**61**	JK2	0.57
LK2	**106**		PS2	102	**64**	PO2	0.58
AO2	**106**		AK1	107	**64**	AK1	0.6
AK1	**107**		DK2	119	**64**	PO1	0.61
CLP	**107**		PO2	113	**66**	PS1	0.62
LK1	**108**		JK2	116	**67**	AS1	0.62
PA2	**109**		MG2	88.2	**68**	PS2	0.63
MB2	**109**		*RM1*	**93.8**	**69**	*RM2*	**0.65**
AO1	**112**		AO1	112	**77**	DK1	0.65
CLH	**113**		PS1	124	**77**	AO1	0.68
PO2	**113**		PO1	127	**77**	JK1	0.69
JK2	**116**		DK1	120	**79**	*RM1*	**0.74**
JK1	**118**		AJ2	105	**79**	LK2	0.74
DK2	**119**		LK2	106	**79**	MB1	0.75
CMC	**119**		JK1	118	**81**	AJ2	0.75

[3] The case of subject RM is an outstanding argument for the necessity to normalise the data for speech rate. Together with CMC, CLH and CLP it may also convince learners that high speed does not equal proficiency in FL speech performance.

S	**MstrV**		S	MstrV	**Mschwa**	S	**VRQ**
DK1	**120**		LK1	108	**87**	MG2	0.77
CJI	**121**		MG1	98.8	**91**	LK1	0.81
PS1	**124**		MB2	109	**93**	MB2	0.85
PO1	**127**		MB1	127	**95**	MG1	0.92
MB1	**127**		AJ1	98.5	**103**	AJ1	1.05

Table 8: Quantitative vowel reduction scale in native English speakers and Polish learners. S=subject, MstrV=mean stressed vowel duration, Mschwa=mean reduced vowel duration, VRQ=Mschwa:MstrV. Native speakers' codes in bold. Polish learners' codes followed by "1" (1st recording) or "2" (second recording).

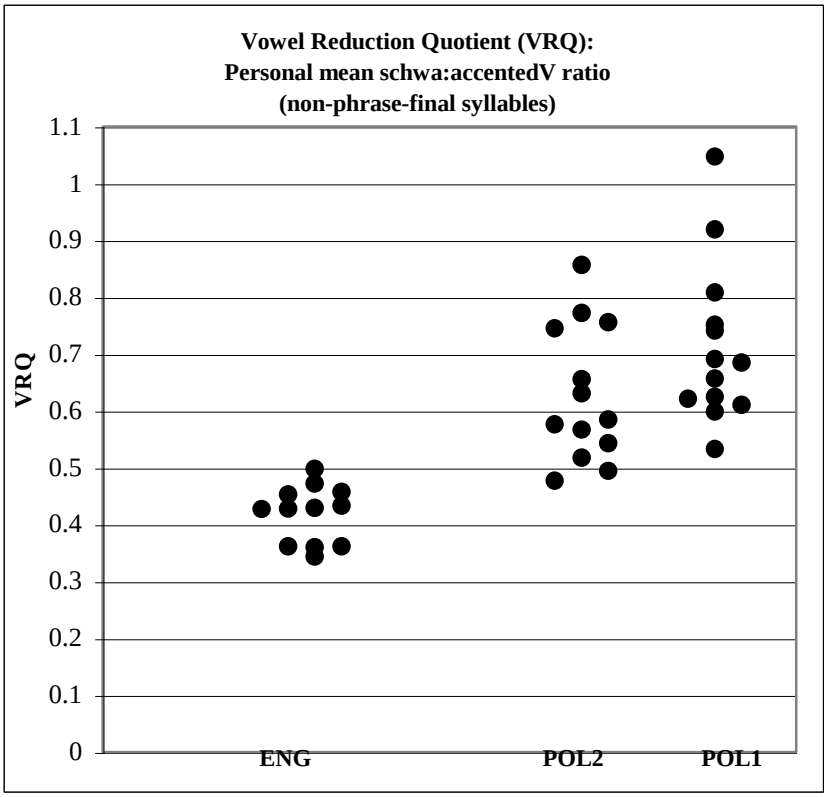


Figure.3: Vowel Reduction Quotient.

The VRQ scores suggest that in native English speech the unstressed vowels are at least 50% shorter than the stressed ones. Polish learners, even after pronunciation training, hardly ever reach this level of vowel reduction. The significant difference between the

groups is also reflected in group median differences. Table 9 presents both raw schwa durations and measures normalised for speech rate (VarcoV, VRQ).

measure\group	ENG	POL2	POL1
schwa median (ms)	42	64	77
VarcoV median	65	50	48
VRQ median	.43	.58	.68

Table 9. Group medians for vowel reduction and duration variability measures.

5. Conclusions

Simple descriptive statistics concerning vowel duration which were used in this study help to provide evidence supporting the following statements:

1. In Polish learners' read speech, there is less difference between 'long' and 'short' vowels than in native production (but the evidence is rather weak).
2. Final lengthening is considerably stronger in native speakers.
3. Vowel reduction is a serious problem for Polish learners, who produce too long unstressed vowels in terms of both absolute and relative durations. Despite some progress, this remains difficult even after training.
4. Considering all duration determinants combined, the Polish learners vary their vocalic length far less than do native English speakers, even though fluency problems, typical of learner speech, should probably contribute to more variation.
5. VarcoV and VRQ are efficient measures which show differences between native and Polish-accented English speech timing.
6. VRQ appears resistant to individual speech rate differences.
7. Because duration statistics are text-dependent, cross-linguistic studies are difficult to conduct. Useful data about native and non-native speakers can be gathered if standardised tests are introduced.

The measures presented in this paper show general differences between native English and Polish learner pronunciation but they can also serve as immediate didactic help in practical phonetics courses to enhance the learners' awareness of cross-linguistic differences and similarities and may help set concrete targets for practical pronunciation training.

References

Avery, P. and S. Ehrlich. 1992. *Teaching American English Pronunciation.* Oxford: Oxford University Press.

Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5 (9/10): 341-345.

Bryła, A. 2010. Phonetic properties of Euro-English – empirical evidence. In *Issues in accents of English 2: Variability and norm*, ed. E. Waniek-Klimczak, 37-60. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Dellwo, V. 2006. Rhythm and Speech Rate: A Variation Coefficient for ΔC. In: *Language and Language-processing*, eds. P. Karnowski and I. Szigeti, 231-241. Frankfurt am Main: Peter Lang.

Dziubalska-Kołaczyk, K., A. Bogacka, D.Pietrala, M. Wypych and G. Krynicki. 2006. PELT: an English language tutorial system for Polish speakers. *MULTILING-2006*, paper 012.

Gonet, W., J. Szpyra-Kozłowska, and R. Święciński. 2010. The acquisition of Vowel Reduction by Polish students of English. In *Issues in accents of English 2: Variability and norm*, ed. E. Waniek-Klimczak, 291-308. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Grabe, E., B. Post and F. Nolan. 2001. *The IViE Corpus*. Department of Linguistics, University of Cambridge. http://www.phon.ox.ac.uk/IViE. [Retrieved 7 September 2006].

Grabe E. and E. L. Low. 2002. Durational variability in speech and the rhythm class hypothesis. In *Laboratory Phonology* 7, eds. C. Gussenhoven and N. Warner, 515-546. Berlin, New York: Mouton de Gruyter.

Hewings, M. 2004. *Pronunciation Practice Activities.* Cambridge: Cambridge University Press.

Loukina, A., G. Kochanski, B. Rosner, C. Shih and E. Keane. 2011. Rhythm measures and dimensions of durational variation in speech. *Journal of the Acoustal Society of America* 129/5: 3258-3270. DOI: 10.1121/1.3559709

Low E. L., E. Grabe and Nolan F. 2000. Quantitative characterisations of speech rhythm: syllable-timing in Singapore English. *Language and Speech* 43: 377-401.

Luke, K.-K. and J. C. Richards. 1982. English in Hong-Kong: Functions and status. *English World-Wide* 3: 147-164.

Nowacka, M. 2008. *The Phonetic Attainment in Polish University and College Students of English. A Study in the Productive and ReceptivePronunciation Skills.* Unpublished Ph.D. dissertation. Maria Curie-Skłodowska University, Lublin.

Porzuczek, A. 2010. The weak forms of TO in the pronunciation of Polish learners of English. In *Issues in accents of English 2: Variability and norm*, ed. E. Waniek-Klimczak, 309-324. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Porzuczek, A. (in press). *The timing of tone group constituents in the advanced Polish learner's English pronunciation.* Katowice: Wydawnictwo Uniwersytetu Śląskiego

Ramus, F., M. Nespor and J. Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 72: 1-28.

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

Szpyra-Kozłowska, J. 2003. The Lingua Franca Core and the Polish Learner. In *Dydaktyka fonetyki języka obcego*, eds. W. Sobkowiak and E. Waniek-Klimczak, 193-210. Płock: Wydawnictwo Naukowe PWSZ w Płocku,.

Waniek-Klimczak, E. 2005. *Temporal Parameters in Second Language Speech.* Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

White, L. and S. L. Mattys. 2007. Calibrating rhythm: First language and second language studies. *Journal of Phonetics* 35: 501-522.

Appendix A

The read text and tested vowels. Unstressed reduced vowels in *italics*, stressed vowels in **bold**.

Once upon a time there was *a* girl called Cinderella. B*u*t everyone called h*er* C**i**nd*er*s. C**i**nd*er*s lived with her mother and two stepsisters called Lily and Rosa. Lily and Rosa were very unfriendly and they were lazy girls. They spent all their time buying new clothes and going t*o* p**ar**ties. Poor Cinders had to wear all their old hand-me-downs! And she h**a**d t*o* do the cleaning!
One day, a royal messenger came to announce a ball. The ball would be held at the Royal Palace, in h**o**nour of the Queen's only s**o**n, Prince William. Lily and Rosa th**ough**t this was div**i**ne. Prince William was g**or**ge*ou*s, and he was l**oo**king for a bride! They dreamed *o*f wedding bells!
When the evening of the ball arrived, Cinders had to help her sisters get ready. They were in *a* b**a**d m**oo**d. They'd want*e*d t*o* b**uy** some new g**ow**ns, but their m**o**ther said that they had en**ou**gh g**ow**ns. So they st**ar**t*e*d sh**ou**ting *a*t C**i**nd*er*s. 'Find my jewels!' yelled one. 'F**i**nd my h**a**t!' howled the other. They wanted hairbrushes, hairpins and hair spray.
When h*er* s**i**st*er*s h*a*d g**o**ne, Cinders felt very down, and she cried. Suddenly, a voice said: 'Why are you crying, my dear?'. It was h*er* fairy g**o**dm**o**ther!

Appendix B

Individual speakers' vowel class length ratios. Native speakers' codes in bold. Polish learners codes followed by "1" (1st recording) or "2" (second recording)

subject	**D:S**
CJE	**2.82**
CLH	**2.71**
CMA	**2.52**
CTG	**2.51**
CPT	**2.44**
CMF	**2.43**
CLP	**2.39**
AK2	2.33
CER	**2.32**
CHB	**2.3**
CJI	**2.3**
AS2	2.27
PS1	2.25
DK1	2.23
CSM	**2.14**
PO1	2.14
LK1	2.14
PA1	2.13
PS2	2.12
MB2	2.08
AK1	2.06
AS1	2.05
RM1	2.03
RM2	2.02
PA2	1.98
AO1	1.97
MG2	1.95
CMC	**1.95**
AJ1	1.93
JK2	1.92
JK1	1.91
LK2	1.9
PO2	1.88
MG1	1.87
AJ2	1.83
MB1	1.8
DK2	1.72
AO2	1.57

subject	**L:S**
CJE	**2.29**
CLH	**2**
CPT	**1.89**
CLP	**1.88**
CTG	**1.81**
AK2	1.78
MG1	1.75
PS2	1.73
CMA	**1.72**
CJI	**1.7**
DK1	1.7
PA1	1.67
PO1	1.67
RM2	1.64
PS1	1.59
MG2	1.59
AS2	1.59
PO2	1.59
CMF	**1.58**
DK2	1.55
RM1	1.53
CER	**1.52**
MB2	1.52
CMC	**1.52**
AK1	1.5
CSM	**1.48**
LK2	1.47
CHB	**1.47**
JK2	1.46
AJ2	1.45
LK1	1.44
AJ1	1.42
MB1	1.39
AS1	1.37
JK1	1.34
PA2	1.33
AO1	1.22
AO2	1.21

subject	**A:S**
CPT	**1.85**
CHB	**1.79**
CMA	**1.78**
CTG	**1.76**
CSM	**1.74**
CER	**1.7**
CJE	**1.68**
CJI	**1.64**
PO2	1.59
AS2	1.55
CLP	**1.52**
PS1	1.51
CLH	**1.5**
PS2	1.5
CMF	**1.48**
LK2	1.45
MB2	1.4
PO1	1.37
DK1	1.35
AS1	1.35
JK1	1.34
MG2	1.33
AK2	1.33
JK2	1.29
AK1	1.28
MG1	1.25
MB1	1.23
DK2	1.2
AO1	1.16
RM1	1.13
PA1	1.12
CMC	**1.12**
PA2	1.05
RM2	1
AO2	0.96
AJ2	0.94
AJ1	0.92
LK1	0.92

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0049-2

STRESSED VOWEL DURATION AND PHONEMIC LENGTH CONTRAST

TOMASZ CISZEWSKI
University of Gdańsk

Abstract
It has been generally accepted that greater vowel/syllable duration is a reliable correlate of stress and that absolute durational differences between vowels underlie phonemic length contrasts. In this paper we shall demonstrate that duration is not an independent stress correlate, but rather it is derivative of another stress correlate, namely pitch. Phonemic contrast, on the other hand, is qualitative rather than quantitative.

These findings are based on the results of an experiment in which four speakers of SBrE read 162 mono-, di- and trisyllabic target items (made of CV sequences) both in isolation and in carrier phrases. In the stressed syllables all Southern British English vowels and diphthongs were represented and each vowel was placed in 3 consonantal contexts: (a) followed by a voiced obstruent, (b) voiceless obstruent and (c) a sonorant. Then, all vowels (both stressed and unstressed) were extracted from target items and measured with PRAAT.

The results indicate that stressed vowels *may* be longer than unstressed ones. Their durational superiority, however, is not stress-related, but follows mainly from vowel-intrinsic durational characteristics and, to some extent, from the prosodic context (i.e. the number of following unstressed vowels) in which it is placed. In CV_1CV_2 disyllables, when V_1 is phonemically short, the following word-final unstressed vowel is almost always longer. It is only when V_1 is a phonemically long vowel that V_2 may be shorter. As far as diphthongal V_1 is concerned, the durational V_1~V_2 relation is variable. Interestingly, the V_1~V_3 relation in trisyllables follows the same durational pattern. In both types of items the rare cases when a phonemically short V_1 is indeed longer than the word-final vowel involve a stressed vowel which is open, e.g. [æ,ɒ], and whose minimal execution time is longer due to a more extensive jaw movement. These observations imply that both in acoustic and perceptual terms the realisation of word stress is not based on the durational superiority of stressed vowels over unstressed ones. When it is, it is only an epiphenomenon of intrinsic duration of the stressed vowel and extra shortness of non-final unstressed vowel.

As far as phonemic length contrast is concerned, we observe a high degree of durational overlap between phonemically long and short vowels in monosyllabic CVC words (which is enforced by a greater pitch excursion), whereas in polysyllables the differences seem to be perceptually non-salient (>40 ms, cf. Lehiste 1970). This suggests that the differences in vowel duration are not significant enough to underlie phonological length contrasts.

1. Introduction

Vowel duration has been given an enormous amount of research attention, both phonetic and phonological. It has also been generally accepted that duration is one of the major phonetic correlates of stress (cf. Fry 1955, 1958). In this paper we will concentrate on how phonemic length contrasts are curtailed by the operation of pre-fortis clipping (PFC) and the prosodic context (i.e. the number of the following unstressed syllables, or foot structure) in which the stressed vowel is placed. We will argue that PFC and the size of the foot obliterate quantitative vowel contrasts.

2. Experiment design

Four male speakers of Southern British English took part in a controlled experiment. Each subject read 162 target items (54 monosyllables, 54 disyllables and 54 trisyllables). All items were presented in two contexts: in isolation and phrase-finally (*Say the word...*). Target items were selected according to the following criteria: (i) all monosyllables were of the CVC type, (ii) all di- and trisyllables terminated in [i] (incidentally schwa), (iii) in the stressed vowel position all RP vowels an diphthongs were represented, (iv) the post-stress consonants were of three types: voiceless obstruents, voiced obstruents and sonorants (each vowel and diphthong was placed in all three consonantal contexts), (v) where possible, the initial C was a voiced obstruent. Only vowels were measured in the present study. The total number of observations amounts to 652 (162 vowels x 2 contexts x 4 subjects). The significance of the durational differences between stressed vowels in isolated vs. phrase-final context was tested for all vowels in all three groups of target items (mono-, di- and trisyllables) separately. We hypothesised that both isolated and phrase-final pronunciations are in fact identical by virtue of being followed by silence. Thus, if the phrase-final lengthening effects occur (for individual vowels or globally for all vowels within an item in terms of their total duration), they should be observed in both contexts. One-way Anova (with an alpha of .05) confirms that there is no significant effect of the context on both stressed and unstressed vowel duration ($p>.05$). Thus, the two sets of data were combined which increased the sensitivity of further statistical tests ($n=104$ for an individual subject in each group of items, i.e. 1-, 2- and 3-syllables).

Vowel duration was measured with PRAAT (Boersma and Weenink 2005) using waveforms and spectrograms. For vowels followed by consonants, vowel onset was identified as the point where the target vowel full formant structure was reached and the end of the vowel corresponded to the beginning of the closure phase. The termination of word-final vowels was assumed to coincide with the end of periodic wave accompanied by dispersion of F2/F3.

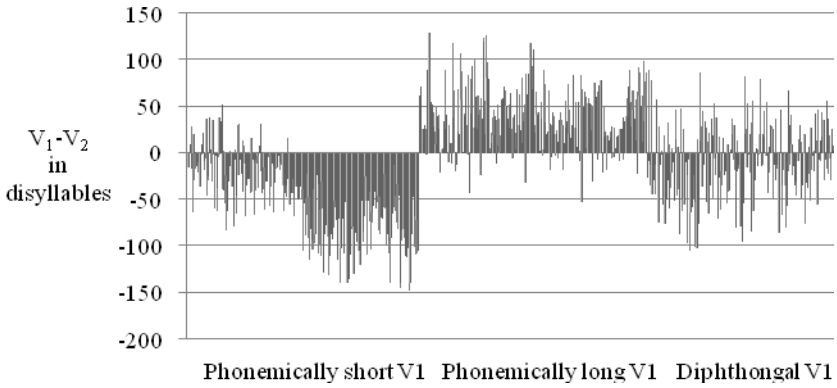
3. Vowels duration: a problematic stress correlate

Earlier studies have shown that there exist three acoustic correlates of stress, i.e. *f*0, duration and intensity. According to Fry (1955, 1958), Bolinger (1958) and Morton and

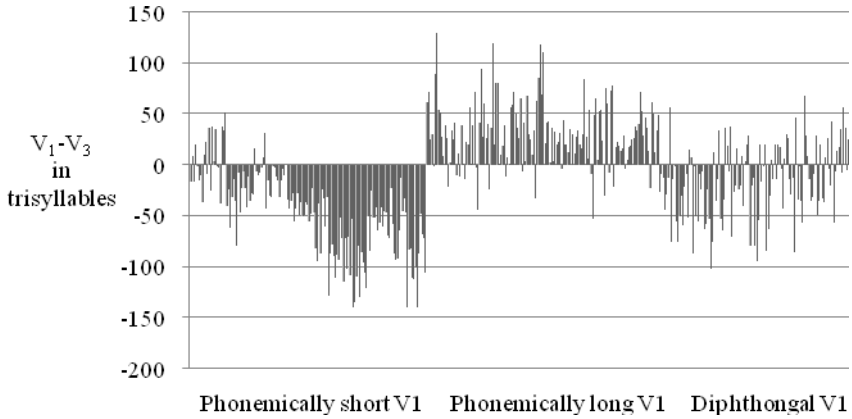
Jassem (1965) the correlates differ in their contribution to stress perception: *f*0 provides the strongest cue, increased duration has a slightly lesser perceptual value and intensity is the weakest correlate. As argued by Lieberman (1960), however, vowel duration is the weakest correlate. A different point of view is presented by Cutler, Dahan and Donselaar (1997: 154) who argue that there is "peculiar redundancy of stress cues in English" and it is also segmental structure that provides robust information about stress.

In essence, the null hypothesis tested in the present study assumes that there exists a fixed $V_{STRESSED}>V_{UNSTRESSED}$ relation that holds for all three phonetic correlates of stress, duration being one of them. Thus, V_1 in polysyllabic items should invariably be longer than the following unstressed vowels (V_2 in 2- and 3-syllable words and V_3 in 3-syllable words). The durational superiority of the stressed vowel over the unstressed ones within a lexical item, however, is not as obvious as it may seem. Admittedly, in trisyllabic words V_1 was found to be generally longer than the following unstressed vowel (V_2). The mean differences between the two vowels for each subject were as follows: S1=61 ms; S2=69 ms; S3=52 ms and S4=77 ms. However, not in all cases was the difference between V_1 and V_2 positive. V_1 did happen to be shorter than V_2 (S1=5.5%; S2=4.6%; S3=14.5% and S4=0.6% of items in the sample). Although such instances were relatively infrequent in each sample, the very fact that they did occur raises doubts about the validity of $V_{STRESSED}>V_{UNSTRESSED}$ relation. We do not think, however, that this provides sufficient arguments for rejecting it. It has to be mentioned that V_2 was longer than V_1 only very specific contexts: (i) when V_1 was followed by a coda consonant (e.g. *density, dignity*) and/or the consonant following V_2 was a stop (e.g. *Kennedy, Canada*). The former context accounts for the extra shortness of V_1 and the latter one for the lengthening of V_2 due to a slightly longer closure phase before the following stop. Furthermore, since the coda consonant is generally assumed to contribute to the phonological weight of the syllable rhyme, its duration should also be taken into account. If added to the pre-coda vowel, the total duration of the CV rhyme would have certainly eliminated all instances in which V_1 alone was shorter than V_2 in trisyllables.

Much stronger doubts concerning the durational domination of the stressed vowels over the unstressed ones appear when V_1 duration in di- and trisyllables is compared with that of word-final unstressed vowels (e.g. ***biddy***, ***bigamy***). In disyllables, when V_1 is phonemically short, the following word-final unstressed syllable is almost always longer. It is only when V_1 is a phonemically long vowel that V_2 is shorter. As far as diphthongal V_1 is concerned, the durational V_1~V_2 relation is variable. Interestingly, the V_1~V_3 relation in trisyllables follows the same durational pattern. In both types of items the rare cases when a phonemically short V_1 is indeed longer than the word-final vowel involve a stressed vowel which is open, e.g. [æ,ɒ] and whose minimal execution time (Klatt 1986) is longer due to a more extensive jaw movement. These observations imply that both in acoustic and perceptual terms the realisation of word stress is not based on the durational superiority of stressed vowels over unstressed ones. When it is, it is only an epiphenomenon of intrinsic duration of the stressed vowel and extra shortness of non-final unstressed vowels, as illustrated in the graphs (1) and (2). Hence, to a large extent it is accidental.



Graph 1: V_1-V_2 difference in duration (ms) in 2-syllable items (all subjects)



Graph 2: V_1-V_3 difference in duration (ms) in 3-syllable items (all subjects)

In consideration of the above, we have to reject the idea that stressed vowels are longer than the unstressed ones within the same item. In terms of duration the $V_{STRESSED}>V_{UNSTRESSED}$ relation is neither stable nor does it seem to be stress-related.

4. Pre-fortis clipping effects, or how phonemic contrast gets neutralised

In principle the PFC effects should be observed in all vowels which are followed by a voiceless obstruent regardless of the vowel position and the prosodic context. Thus, it should affect stressed and unstressed vowels alike as it is a stress-independent

phenomenon. As observed by Kim and Cole (2005), there exists an inversely proportionate relation between the duration of the stressed syllable and the number of syllables that follow. However, this regularity is also contextually independent of PFC. While on average the duration of the stressed vowels is expected to decrease in longer items, the compression effect may not suspend the operation of PFC. Thus, the mean difference in milliseconds between the duration of stressed vowels followed by a voiceless obstruent and those followed by a voiced one is expected to diminish without threatening the significance of the difference itself.

Thus, according to the null hypothesis, regardless of the durational differences between the stressed vowels in shorter vs. longer items and stressed vs. unstressed vowels, the PFC effects, which are merely related to the voicing of the following consonant, should be constant. If this claim is falsified, i.e. the PFC effects turn out to be insignificant for some group of items or some prosodic context, the conditioning factor must be singled out which is responsible for the PFC suspension. An alternative hypothesis, in our view, must assume that it is caused by the intervocalic durational relations within polysyllabic items. The existence of such interdependences entails a postulation of a higher-level constituent which controls the interactions between the total number of syllables and the degree of stressed vowel shortening before a fortis consonant. We assume that this constituent is the metrical foot.

First, let us consider the durational differences relating to the PFC in the group of monosyllables ending in voiced vs. voiceless obstruent. Rather unsurprisingly, the one-way Anova test (alpha .05) confirms that PFC has a highly significant effect (S1 p=1.28E-14; S2 p=2.59E-08; S3 p=.0007; S4 p=1.6E-15) on vowel duration for all subjects regardless of the phonemic length of the vowel.

In disyllables the pre-voiced/pre-voiceless durational difference between stressed vowels (V_1) remains statistically significant, although it has to emphasised that the *p*-values are generally higher and the mean differences are smaller (S1 p=.0002; S2 p=.004; S3=.0007; S4 p=.02).

As far as trisyllabic items are concerned, however, for all subjects the PFC effects on V_1 duration turn out to be non-significant (S1 p=.02; S2 p=.07; S3 p=.56; S4 p=.23). Moreover, the mean differences in duration between $V_1+C_{VOICELESS}$ and V_1+C_{VOICED} are further reduced, both generally and for an individual subject. Noteworthy is also the fact that while the mean difference in the duration of pre-voiced vs. pre-voiceless vowels in monosyllables (53.6 ms~113.7 ms) may be safely assumed to be perceptually salient, this is not so obvious in the case of di- and trisyllables, where the difference range is 27~33.9 ms and 20.9~26.6 ms, respectively.

In conclusion, PFC affects stressed vowels to a different degree depending on the number of syllables that follow. Thus, the probability of its occurrence is inversely proportionate to the overall vowel duration of the word.

Let us now pay attention to another surprising fact, namely that phonemically identical vowels followed by a voiceless obstruent are not necessarily shorter than that those followed by a voiced one. The percentage of cases when the vowel in $VC_{VOICELESS}$ is longer than VC_{VOICED} is presented in Table 1 below. For each speaker, the left-hand column shows the number of items where the pre-voiced vowel was actually longer than the phonemically identical pre-voiceless one and the right-hand one the percentage of such occurrences in the data sample (n=36).

	S1		S2		S3		S4	
Monosyllables	0	0%	1	3%	10	28%	0	0%
Disyllables	1	3%	7	19%	13	36%	8	22%
Trisyllables	12	33%	11	30%	12	33%	14	39%

Table 1: Number of instances in which a stressed vowel is longer than a phonemically identical vowel despite the PFC context.

This seems to undermine the very relation between the duration of a vowel and the voicing of the following consonant. This observation does not necessarily falsify PFC. As argued by Kingston and Diehl (1994), PFC is a feature which enhances phonemic contrast and as such it facilitates speech perception. As Gussenhoven (2007: 146) puts it "the implementation of pre-fortis clipping [...] is a concession to the hearer by way of compensation for the frequent devoicing of the voiced obstruent." Thus, this compensation is more likely to occur when the phonemic distinctiveness is threatened. Its degree observed in experimental conditions will then depend upon the organisation of the input. Since in our experiment the order of target items was randomised (i.e. items like *bit* and *bid* were never placed consecutively), there was no (or very little) necessity of contrast enhancement.

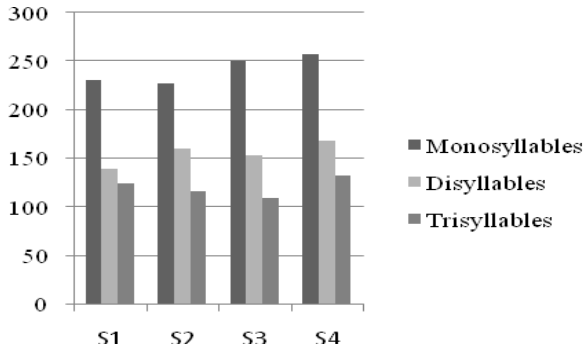
Since on the other hand, PFC is aerodynamically conditioned 'because the transglottal pressure difference creating the airflow driving vocal fold vibration is hard to maintain in the face of the impedence by the oral constriction of obstruents' (Gussenhoven: ibid.), its effect on vowel duration is likely to be observed even if distinctiveness is not threatened (e.g. in a randomised experimental input). This does not mean, however, that it *must* occur as the aerodynamic conditioning may be successfully counterbalanced by the prosodic one (which may also be aerodynamic in nature). Pre-fortis clipping, then, is both an articulatorily motivated and speaker-controllable parameter which may be latent (i.e. producing statistically and perceptually insignificant differences in vowel duration) when the vowel contrast is safe.[1] In terms of speech processing, considering the fact that the perceptual information load is directly proportionate to the number of the syllables within an item (cf. the cohort theory by Marlsen-Wilson and Tyler (1980)), in monosyllables the number of instances in which a vowel followed by $C_{VOICELESS}$ is longer than the phonemically identical vowel followed by C_{VOICED} is the lowest.

To sum up, PFC has been shown to have the greatest effect on vowel duration in monosyllabic items. The degree of durational difference between pre-voiced and pre-voiceless vowels in the stressed position is inversely proportionate to the overall length of an item, i.e. the effect is lesser on the stressed vowels in disyllables than on those in monosyllables and it becomes insignificant in trisyllabic items. Pre-fortis clipping appears to be both an articulatorily motivated and speaker-controllable process which may be latent (i.e. producing statistically and perceptually insignificant differences in vowel duration) when the vowel contrast is safe.

[1] A typical context for its activation is the presentation of length contrast in minimal pairs (*beat~bead*), e.g. in the process of phonetic instruction.

5. Durational overlap between phonemically long and short vowels

We observed that (i) mean stressed vowel durations systematically decrease as the number of following unstressed syllables increases and (ii) the differences between stressed vowel durations in mono- and disyllables are significantly greater (67-97 ms) than those between di- and trisyllables (15-43 ms).

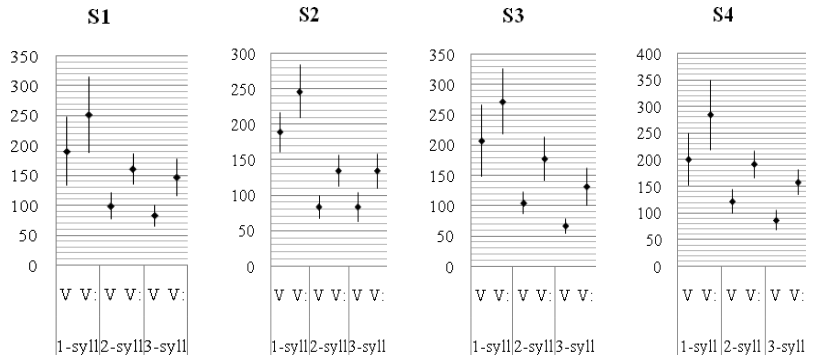


Graph 3: Mean stressed vowel durations (ms) in mono-, di- and trisyllabic items

Theoretically, one would expect that the systematic decrease in V_1 duration in 2- and 3-syllable words should result in a simultaneous obliteration of phonemic length distinctions and, consequently, pose a potential threat to their perception. However, the danger of eliminating phonemic length distinction in polysyllabic items is not as serious as it may seem. Recall that the inter-speaker variation ranges from 18ms to 43.9 ms, which does fit neatly in the non-distinguishable window (10~40 ms) established by Lehiste (1970: 13). The durational deficiency of V_1 in polysyllabic items may also be successfully compensated for by a more robust segmental context. Note that, paradoxically, due to the fact that as the number of the syllables grows, the number of potential vowel-consonant permutations increases rapidly, which reduces the chances of generating, for instance, a trisyllabic minimal pair (whose semantic contrast relies entirely upon V_1 quantity) virtually to zero. Thus, the substantially reduced V_1 recognition time in di- and trisyllables can hardly impede the process of the whole word recognition. Language economy should, therefore, allow to loosen the length contrast requirements where intelligibility is not threatened, i.e. in polysyllabic forms, and strengthen it if the recognition of an item is largely dependent on the recognition of the vowel, i.e. in monosyllables. So much of the theory. What emerges from our data, however, is a completely opposite regularity. It is in the monosyllabic items where the stressed long and short vowels display durational convergence rather than in di- and trisyllables. This conclusion was arrived at by mapping the mean durations of phonemically long and short vowels onto the corresponding standard deviation values. Thus, we have calculated the span of a durational window for the two classes of stressed vowels in 1-, 2- and 3-syllable words by adding the standard deviation for each group to its mean duration on the one hand and subtracting the standard deviation from the

corresponding mean duration on the other. The resulting windows for phonemically short and long vowel durations in each group of items were then compared for each subject with a view to extracting the degree of overlap, which was calculated in the following way: ($V_{MEAN\ DUR.} + V_{STD\ DEV.}$) − ($V:_{MEAN\ DUR.} - V:_{STD\ DEV.}$). We assumed that there is an inversely proportionate relation between the degree of the durational overlap and the robustness of the phonemic length contrast in a particular group of items.

It turns out that for all subjects the durational overlap was observed only in monosyllabic items (S1=60.2 ms; S2=8.2 ms; S3=48 ms and S4=29.8 ms) and not in di- and trisyllables. This is graphically illustrated in (x) below. Mean duration values are represented by ♦.



Graph 4: Long/short durational overlap in 1-, 2- and 3-syllable items

Thus, despite the (misleading) fact that the differences in mean durations between long vs. short vowels remain constant for all three groups of target items (cf. the distances between ♦s in each V/V: pair), the durational overlap between long and short vowels in monosyllables indicates that the phonemic contrast is, at least to some extent, suspended in this particular context. Bearing in mind the doubtful perceptual value of the long-short V_1 contrast in polysyllables and a fair amount of durational long-short overlap in monosyllables, we have to conclude that in general the phonemic contrast, at least in the dialect of English investigated in this study, is qualitative rather than quantitative in nature. What follows is that the perception of phonemic length contrast and the production of phonemically conditioned differences in vowel durations may be two different phenomena. While the distinctions do have their articulatory manifestation, their perception (due to the fact that they are below just noticeable difference) are based on quality rather than quantity.

6. Conclusions

The present findings may be summarised as follows. Duration alone is not an independent stress correlate. It is rather a derivative of other correlates (pitch in particular). Stressed vowels may be longer than unstressed ones. Their durational

superiority, however, is not stress-related but follows mainly from vowel-intrinsic durational characteristics. The operation of PFC obliterates the durational contrasts. Phonemic contrast is qualitative rather than quantitative. In monosyllables there is a high degree of durational overlap between phonemically long and short vowels (which is enforced by a greater pitch excursion), whereas in polysyllables the differences do exist but are perceptually non-salient.

References

Bolinger, Dwight. 1958. A theory of pitch accent in English. *Word* 14: 109-149.

Cutler, Anne, Dahan Delphine and van Donsellar, Wilma. 1997. Prosody in the comprehension of spoken language: a literature review. *Language and Speech* 40: 141-202.

Fry, Denis B. 1955. Duration and intensity as acoustic correlates of linguistic stress. *JASA* 27: 765-768.

Fry, Denis B. 1958. Experiments in the perception of stress. *Language and Speech* 1: 126-152.

Gussenhoven, Carlos. 2007. A vowel height split explained. Compensatory listening and Speaker Control. In J. Cole and J. I. Hualde (eds.) *Laboratory Phonology* 9*:* Mouton de Gruyter: 145-172.

Kim, Heejin and Jennifer Cole. 2005. The stress foot as a unit of planned timing: evidence from shortening in the prosodic phrase. *Proceedings of Interspeech* 2005, Lisbon, Portugal: 2365-2368.

Kingston, John and Randy Diehl. 1994. Phonetic knowledge. *Language* 70(3): 419-454.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59: 1208-1221.

Lehiste, Ilse. 1970. *Suprasegmentals.* Cambridge, MA: The MIT press.

Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *JASA* 32: 451-454.

Marslen-Wilson. William D. and Loraine K. Tyler. 1980. The temporal structure of spoken language understanding *Cognition* 8: 1-71.

Morton, John and Wiktor Jassem 1965. Acoustic correlates of stress. *Language and Speech* 8*:* 159-181.

• **Research in Language, 2012, vol. 10.2** • DOI 10.2478/v10015-011-0029-6

TRANSFER, SIMILARITY OR LACK OF AWARENESS? INCONSISTENCIES OF GERMAN LEARNERS IN THE PRONUNCIATION OF LOT, THOUGHT, STRUT, PALM AND BATH[1]

ALEXANDER KAUTZSCH
University of Regenburg

Abstract
The current study presents acoustic analyses of non-high back vowels and low central vowels in the lexical sets LOT, THOUGHT, STRUT, PALM and BATH as pronounced by German learners of English. The main objective is to show that learners of English at university level are highly inconsistent in approximating the vowels of their self-chosen target accents British English (BrE) and American English (AmE). To that end, the acoustic qualities of the English vowels of learners are compared to their native German vowels and to the vowels of native speakers of BrE and AmE. In order to facilitate statements about the effect of increased experience, the study differentiates between students in their first year at university and in their third year or later. The results obtained are highly variable: In some cases the learners transfer their L1 vowels to English, other cases show clear approximations to the target vowels, while other cases again document the production of new vowels neither found in German nor in English. However, close approximation to the target vowels only sometimes correlates with higher proficiency. This might be an indicator of a low level of awareness of systematic differences between the BrE and AmE vowel systems. But the data also indicate that the more advanced learners produce more distinct AmE BATH vowels and BrE THOUGHT vowels than the less advanced learners, which points to a partial increase of awareness resulting from increased experience. All in all it seems that raising the awareness of differences between target accents in L2 instruction is necessary if the envisage goal is for learners to reach near-native pronunciation.

1. Introduction

In varieties of English around the world words of the lexical sets LOT, THOUGHT, and BATH are pronounced in different ways. This leads to different degrees of overlap with the lexical sets PALM and STRUT.

In the two major varieties of English which German learners aim at, namely BrE and AmE, these differences manifest themselves as shown in table 1.

[1] I would like to thank one anonymous reviewer for helpful comments on this paper. All remaining shortcomings are my own responsibility.

Lexical set	Example	BrE	AmE
LOT	body	['bɒdi]	['bɑ(ː)di]~['bɔ(ː)di]
THOUGHT	raw	[rɔː]	[rɑ(ː)]~ [rɔ(ː)]
BATH	dance	[dɑːns]	[dæns]
PALM	father	['fɑːðə]	['fɑ(ː)ðɚ]
STRUT	run	[rʌn]	[rʌn]

Table 1: Examples for BrE and AmE pronunciation of the lexical sets LOT, THOUGHT, BATH, PALM, and STRUT

The LOT and THOUGHT vowels are less rounded in AmE than in BrE and can be variable in quantity, as indicated by "(ː)" in table 1 (cf. Wells 1982: 120, 122, 124, 476). Many native speakers of AmE merge THOUGHT and LOT either to [ɔ(ː)] or to [ɑ(ː)] (cf. Wells 1982: 473-476). A short and low AmE pronunciation of LOT / THOUGHT is very similar to STRUT. The BATH vowel matches with PALM in BrE, with TRAP in AmE. Similar to LOT and THOUGHT, AmE PALM can be comparatively short. (cf. Wells 1982: 118-124).

It could be hypothesized that for learners of English who are not aware of these differences between and variability within varieties of English and perceive the language they are learning as a monolithic whole, these multiple pronunciations of mid and low back vowels and low central are likely to be interpreted as a highly variable model. As a result, learners are inconsistent in targeting their self-chosen accent, making use of a plethora of vowels from different models.[2]

Along these lines, the present paper studies German learners of English at university level, mostly students in teacher training. It sets out to describe and interpret their inconsistencies in the production of vowels in the lexical sets LOT, THOUGHT, STRUT, PALM and BATH with respect to the learners' self-chosen target accents, in this case either BrE and AmE. Such an interpretation needs to take into account the notions of interlanguage, L1 transfer, similarity, and awareness.

In the case of the English vowels in LOT, THOUGHT, STRUT, PALM and BATH as pronounced by German learners of English similarity to the German vowels SOCKEN, BOTEN, HATTEN, and BATEN might be expected to lead to transfer, especially since these vowels receive little attention in formal instruction.

However, the acoustical analyses presented here show that in many cases learners use sounds different from both L1 and L2, which can be seen as an empirical manifestation of interlanguage. The likely reason for the learners' inconsistencies, therefore, cannot be pinned down to transfer alone, but also to a lack of awareness of the highly heterogeneous nature of the input around them.

[2] Even if this might be, according to the anonymous reviewer, an "unwarranted assumption", it is a reasonable one. Unfortunately, no previous studies supporting this claim could be discovered.

2. English and German non-high back and low central vowels

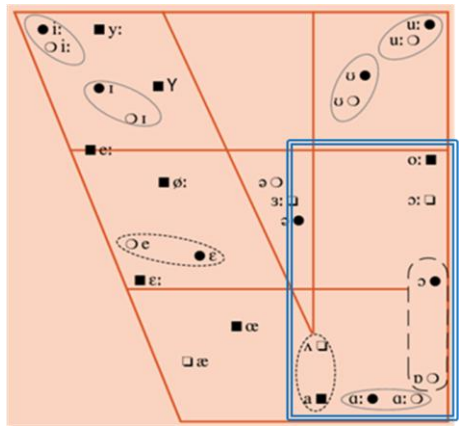


Figure 1: The vowels of English (RP) and German (taken from Kortmann 2005:182)

Figure 1 provides a contrastive overview of the vowel systems of English (RP) and German. The vowels to be dealt with in the present study, viz. the non-high back vowels and the low central vowels, are highlighted by the box.

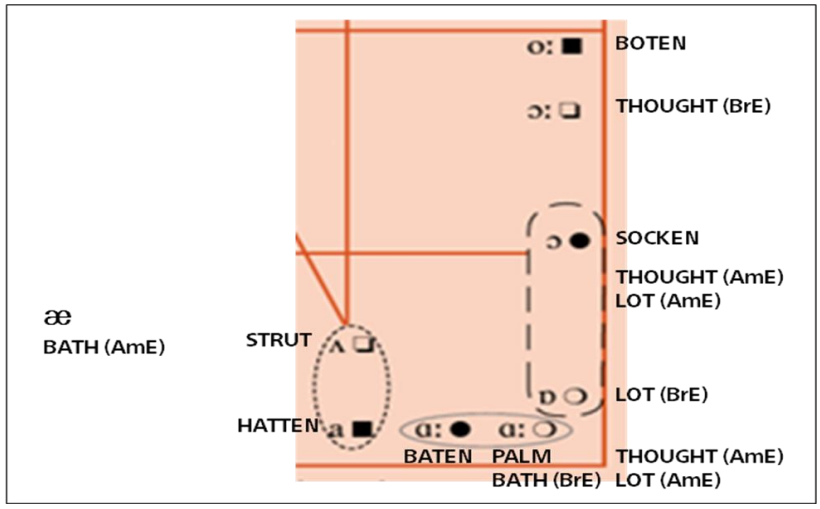


Figure 2: Non-high back vowels and low central vowels of English (RP) and (adapted from Kortmann 2005:182)

Figure 2 zooms into the relevant area and roughly places lexical sets for German, BrE and AmE at the traditional articulatory locations of vowels.

The four German vowels are represented by BATEN, HATTEN, BOTEN and SOCKEN. The BATEN and HATTEN vowels are long and short low central vowels,

respectively, with HATTEN being slightly fronter than BATEN. BOTEN and SOCKEN are long and short mid back vowels, respectively, with BOTEN being considerably closer than SOCKEN.

American and British STRUT and PALM are close to German HATTEN and BATEN. British THOUGHT is close to German BOTEN, the American versions are more open and less rounded, and can be as open as to match PALM. British LOT is, from an articulatory perspective, the rounded counterpart of [ɑ], American LOT is less rounded and its variants can be very similar to those of THOUGHT (LOT-THOUGHT merger, cf. Wells 1982: 473-476).

In BATH BrE uses the same vowel as in PALM, while the AmE BATH vowels equals TRAP and is realized as [æ].

The following section briefly surveys relevant concepts of SLA theory and makes some predictions of possible problems and routes of transfer in the acquisition of the vowel systems of BrE and AmE by German learners.

3. SLA theory: L1 transfer, similarity and awareness

On the basis of the differences between German and English vowel space mentioned above, the present section will briefly discuss the notion of interlanguage in connection with L1 transfer in L2 phonological acquisition and suggest that the outcome of L2 phonological acquisition is very likely to be connected to the level of awareness learners have for the details of the sound system of their target accent.

Interlanguage as introduced by Selinker (1972) entails the widely accepted notion that learners when acquiring a second or foreign language "create a language system", which is not seen as a "deficit system [...] but as a system of its own with its own structures" (Gass and Selinker 2008: 14). The elements of the interlanguage are either from the learner's L1 or from the L2. In addition there are so-called "new forms", elements that belong neither to the L1 nor to the L2 (*ibid.*).

The process responsible for L1 elements being present in the L2 is L1 transfer. Especially in L2 phonological acquisition L1 transfer is, despite its behaviourist roots, a well accepted concept (cf. Major 2008 for a detailed discussion).

What often goes hand in hand with L1 transfer is the notion of cross-linguistic similarity in that it addresses "the question which phenomena are more susceptible to transfer and which are not" (Major 2008: 71). Here most researchers agree that "[t]he more similar the phenomena the more likely transfer will operate; however, what constitutes similar is not always clear-cut" and "a more rigorous and universally agreed upon definition of similarity would seem necessary (Major 2008: 74).

Along these lines Bohn (2002) states that "[a]rmchair methods and acoustic and articulatory comparisons can, at best, serve as a starting point" (Bohn 2002:209) and according to Strange (2007) "[c]ross linguistic similarity is difficult to measure without perception data" (Strange 2007: 45). In other words, the only reliable way to define

similarity is through perception experiments (cf. Strange 2007, Bohn 2002, Strange and Shafer 2008).[3]

In this vein the acoustic data presented here will serve as a starting point for the description of L2 phonological acquisition of learners faced with more than one model. But they will also serve to support the claim that similarity is a highly relative concept. Equivalence classifications of sounds on the basis of assumed similarities are subconscious processes of which learners are not aware. It seems to be necessary to see similarity in Major's (2008: 75) terms as slowing down acquisition, but to different extents and on an individual basis.

Two examples from L1 German L2 English learners will serve to illustrate this. It has been shown elsewhere (Kautzsch 2010a) that in the case of the English mid and low front vowels [e] and [æ] in *bed* and *bad* (i.e in the lexical sets DRESS and TRAP) German has only one short vowel counterpart [ɛ] as in BETTEN, which is then - due to equivalence classification – used in both English contexts. The distinction between [e] and [æ] develops quite late in German learners. This is very likely due to the fact that this distinction does not feature prominently in German ESL/EFL classrooms, resulting in a low level of awareness for this difference.[4]

In the case of the dental fricatives /θ/ and /ð/, which due to their absence in German could be seen as dissimilar and therefore should be easier to acquire, the relativity of similarity becomes even more apparent. The success rate of German learners here is much higher, although there remain a considerable number of learners who do not manage to acquire these sounds and use the alveolar fricatives [s] and [z] instead. From a similarity perspective this means that for those learners who succeed in the acquisition, the dental fricatives are dissimilar enough from German sounds as not to be classified as equivalent. For those who fail, a perceived similarity with alveolar fricatives persists. Again, the different performances of these learners seem to be connected to awareness. As soon as one is aware of two sounds being different, they become more dissimilar and are thus acquired faster. And since much emphasis is placed on the dental fricatives in ELT in Germany, a higher level of awareness is created and the approximation to the target sound is on the whole more successful.

The notion that learners can be made aware of similar phenomena is not new in SLA. It is inherent, for example, in "Focus on Form" (presented for example in ch. 11.5 in Gass and Selinker 2008), or in the "Noticing Hypothesis" (developed by Schmidt in a series of articles on attention and awareness: Schmidt 1990, 1994, 1995, 2001, 2010). Making learners aware of certain structures (or in our case sounds) seems especially applicable in the classroom, less so in natural, immersive settings (cf. Krashen 1985, Gass and Selinker 2008).

As far as the vowels under scrutiny in the current study are concerned, they receive little attention in the ESL/EFL classroom of German learners of English. And when

---

[3] For three popular models which incorporate similarity as a central concept see the Speech Learning Model (Flege 1995), the Perceptual Assimilation Model (Best and Strange, 1992; Best 1994, 1995) and the Native Language Magnet model (Kuhl 1993, 1991).

[4] A similar situation is reported upon in Kautzsch (2010b) where German learners of English are very inconsistent in their realizations of non-prevocalic *r* when aiming at BrE or AmE.

considering Schmidt and Frota's (1986) claim that "a second language learner will begin to acquire the target like form if and only if it is present in comprehended input and 'noticed' in the normal sense of the word, that is consciously", it must be assumed that German learners will have difficulties in acquiring the vowels in LOT, THOUGHT, BATH, STRUT, and PALM; transfer will possibly be at work to some extent.

3.2 Predictions for German learners of English

Based on the cross-linguistic analysis above, the present section will make some predictions for the acquisition of a BrE and an AmE vowels system by German learners.

For German learners of English aiming at BrE the German and English non-high back vowels and low mid central vowels are similar in their articulatory properties and in their relative positions, i.e. the German system has the same contrasts as the BrE system, namely a pair of short and long mid back vowels, and a pair of short and long low central vowels. Thus it would be easy for German learners aiming at BrE to apply German BATEN, HATTEN, BOTEN and SOCKEN in English PALM / BATH, STRUT; THOUGHT, and LOT, respectively. In other words, L1 transfer can be expected, but at the same time few inconsistencies will arise since the two systems contain the same distinctions.

For students aiming at AmE there are several options to utilise their German vowels in English. HATTEN and BATEN may be matched with STRUT and PALM, but BATEN might also be used in THOUGHT and LOT, if pronounced as a very open vowel [ɑ(ː)]. Alternatively, when THOUGHT and LOT are pronounced as [ɔ(ː)], the SOCKEN or BOTEN vowel is likely to occur. However, BOTEN being a rounded close mid vowel, it is also possible that it is not employed at all. SOCKEN, on the other hand, may turn out to be too short to be used in LOT and THOUGHT. Thus, if LOT and THOUGHT are not pronounced similar to [ɑ(ː)], learners need to acquire a new sound. The same applies to the BATH vowel, which needs to be matched with TRAP and pronounced as [æ], a new sound that does not belong to the German vowel inventory.
In sum, it seems that the acquisition of the AmE system is more inconsistency-prone than the acquisition of the BrE system.

4. Data

The learners analyzed in this study are 20 students of English from the University of Regensburg. All have been chosen on the basis of a stable L1 background, i.e. they were born and raised in two adjacent regions of Bavaria, the south-eastern of the federal states of Germany: the Upper Palatinate (*Oberpfalz*) and Lower Bavaria (*Niederbayern*).

10 students each have AmE and BrE as their self-chosen target accent. Each of the target accent groups contains two proficiency levels: 5 learners each are "beginners" (Beg), i.e. students of English in their first year at university, and "advanced" students (Adv), i.e. learners in their 3rd year of later. What matches proficiency in this sample is the students' average time spent abroad in months: the beginners have spent 0.8 months

in an English speaking country, while the advanced students have been abroad for 8.3 months.

The analyses below will provide insights into how successful German learners of English are in approximating their self-chosen target accent and if they become more successful as proficiency and time spent abroad increases.

5. Method

The acoustical analysis to follow (section 6) will present the learners' English and German vowels and contrast them to BrE and AmE native speaker control groups.

The learners' vowels were elicited by means of reading two word lists, the one consisting of all English monophthongs, two from each lexical set, the other containing all German monophthongs. The present study picks out non-high back vowels and low central vowels as represented by the words below, the whole database thus totalling 200 English and 80 German monophthongs:

- *body, cot* (LOT)
- *bawd, caught* (THOUGHT)
- *bud, cut* (STRUT)
- *father, palm* (PALM)
- *bath, dance* (BATH)
- *Socken* (SOCKEN)
- *Boten* (BOTEN)
- *hatten* (HATTEN)
- *baten* (BATEN)

The wordlists educe the speakers' most monitored style and therefore provide access to their idealized targets. The recordings were made in a quite office setting, vowels were measured at the centre using Praat (Boersma and Weenink no date) and plotted by means of Kendall and Thomas's (2010) "vowels" package for "R" (The R Project for Statistical Computing no date), applying auditory-based Bark measure[5] for normalization to even out individual differences across speakers.

For the comparison with native speaker data, formant values published in two previous studies are utilised: The AmE vowels are taken from Hillenbrand et al. (1995), who analyse 45 men, 48 women[6] from Michigan (Great Lakes / Midland). The BrE vowels are taken from Deterding (1997, 1990), who provides the vowels of 8 men and 8 women from the South of England. In both studies, participants read lists of words in the context h_d[7].

---

[5] For further details on Bark normalization and the resulting Z values (cf. figures 3 to 10) see Thomas and Kendall (2007: "Methods") and Traunmüller (1997).

[6] Hillenbrand et al. (1995) also measure the vowels of children, but the present analysis only adopts the vowels of adults.

[7] Some scholars call for stable phonetic contexts when analysing vowels, because of variable coarticulation effects (cf. e.g. Bohn 2002:199). Others only avoid "tokens following the approximants [w], [j] and [r]" and tokens "before [ŋ] and dark [l], as all these sounds have severe coarticulatory effects on the vowel" (Deterding et al. 2008: 162).

6. Results

The results will be presented by means of two vowel plots for each of the learners' self-chosen target variety, for BrE in 6.2, for AmE in 6.3. The first plot in each section contains the average locations of the vowels as produced by the beginners and the advanced students to documents differences in the two proficiency groups. In addition these plots provide the average locations of the native speakers' vowels to illustrate the learners' degree of approximation to their target.

The second plot in each section adds the average locations of the learners' German vowels in order to obtain a visual idea of the degree of L1 transfer taking place, i.e. to show to what extent German learners use their native vowels in English.

6.1 British English Target

Figure 3 shows the results for beginners and advanced learners aiming at BrE in comparison to BrE native speakers. Advanced students are closer to native THOUGHT than the beginners (circle 1 in Figure 3). The LOT vowels (circle 2) are very close to native LOT for both groups. With the lexical sets STRUT, BATH and PALM, beginners are closer to native vowels (circle 3), while advanced students display a stronger – somewhat unnecessary – differentiation between these vowels (circle 4).

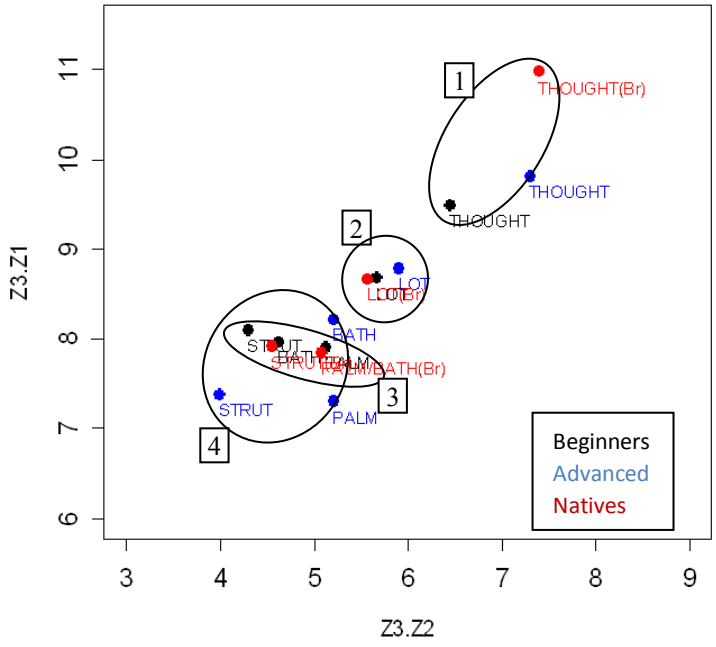


Figure 3: The BrE non high back vowels and low central vowels of German beginners and advanced learners and of native speakers of BrE

Adding the German vowels to the plot (Figure 4) results in the following picture: German SOCKEN (circle 1) is not used for LOT (circle 2), German BOTEN is very close to native THOUGHT but produce different vowels (circle 3), the pronunciation of STRUT and PALM is close to German HATTEN and BATEN for advanced students (circle 4), while the beginner's pronunciation of STRUT/PALM/BATH is closer to that of native speakers (circle 5). The advanced group's BATH vowel is considerably higher (circle 6).

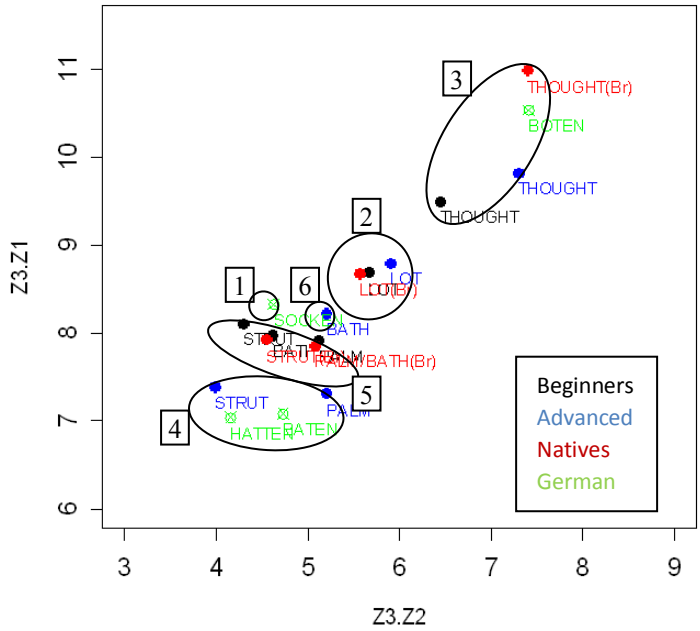


Figure 4: The BrE and German non high back vowels and low central vowels of German beginners and advanced learners, and the BrE non high back vowels and low central vowels of native speakers of BrE

Summing up, German high proficiency learners acquire the LOT vowel as a close approximation to native LOT and do not transfer German SOCKEN. Although German BOTEN is close to BrE THOUGHT, German learners use a different vowel, which is more open on average. In the case of PALM, BATH and STRUT, beginners are very close to the BrE target vowels, while advanced students' PALM and STRUT are closer to their native BATEN and HATTEN.

Thus, the predictions that German high proficiency learners aiming at BrE use their native vowels BATEN, HATTEN, BOTEN and SOCKEN in English cannot be confirmed; in other words expected L1 transfer take place to a limited extent only.

In addition, the increased experience of advanced students as opposed to beginners does not increase their approximation to target vowel sounds, in fact the beginners are closer to the target vowels in the case of PALM, BATH and STRUT.

6.2 American English Target

The results for German learners' non-high back and low central vowel with respect to an AmE target and in comparison to native speakers of AmE are shown in figure 5.

Both the beginners and the advanced students produce a close approximation to native THOUGHT and LOT, with the beginners being even closer (circles 1 and 2 in figure 5). The learners' BATH vowel is very different from native BATH; here the advanced students are closer to native BATH but still at considerable distance (circle 3). Moreover, German learners produce different vowels for LOT and PALM (circle 4). Finally, the learners' STRUT vowels are considerably lower than native STRUT (circle 5), with the beginners being close to native LOT. This mismatch between native and non-native STRUT, however, needs to be interpreted with caution. It cannot be seen as a failure to approximate a native target on the side of the learners. It rather results from the control groups' origin in the Greater Lakes region in the US. This is the area which is likely to have been in the initial stage of the Northern Cities Vowel Shift (cf. Labov et al. 2006: 187-208) at the time of recording and thus the speakers' pronunciation of STRUT does not represent a familiar target for learners. What this also illustrates is the theoretically and practically challenging situation of multiple and heterogeneous target accents.

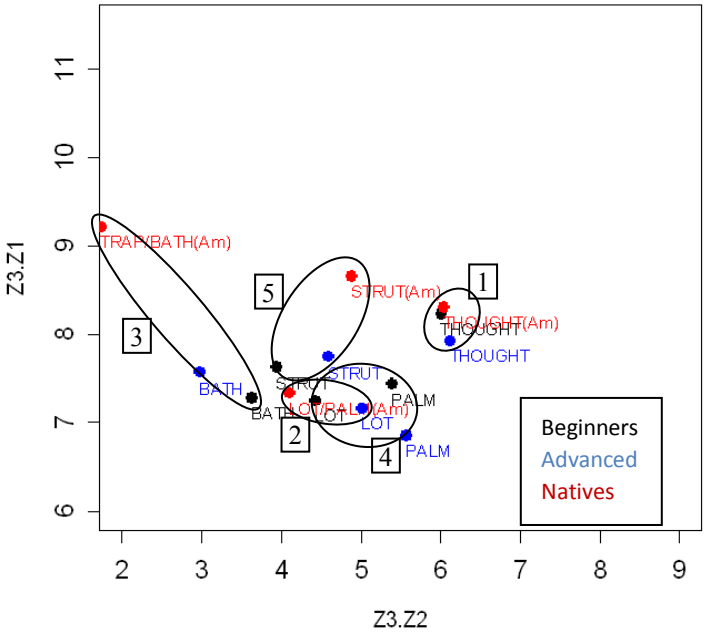


Figure 5: The AmE non high back vowels and low central vowels of German beginners and advanced learners and of native speakers of AmE

Adding the German vowels to the plot (figure 6) once more gives some insight into possible L1 transfer. BOTEN and SOCKEN are not used in the learners' English (circles

1 and 2). The learners' LOT vowels, as well as native LOT, are similar to BATEN/HATTEN (circle 3). THOUGHT, on the other hand, is a new native-like sound (circle 4), while BATH (circle 5), PALM (circle 6) and STRUT (circle 7) are new sounds which are neither German nor AmE.

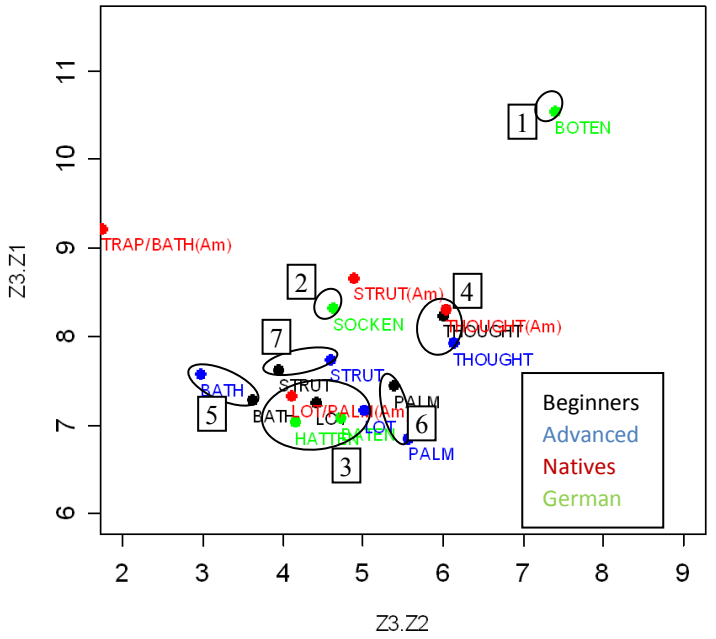


Figure 6: The AmE and German non high back vowels and low central vowels of German beginners and advanced learners, and the AmE non high back vowels and low central vowels of native speakers of AmE

Summing up, the only AmE vowel of German learners in which some degree of L1 transfer can be witnessed is the LOT vowel. It is close to German BATEN/HATTEN and learners make use of this proximity.

With AmE THOUGHT all learners use a vowel close to the target and different from German vowels, whereas in the cases of AmE BATH, PALM, and STRUT all learners use sounds different from German and AmE. In addition both learner groups maintain an (unnecessary) distinction of PALM and LOT. German BOTEN and SOCKEN, on the other hand, are not transferred. Similar to the learners aiming at BrE, increased experience on the side of the advanced students does not increase their approximation to the target.

6.3 Individual variation

In addition to the average locations of non-high back vowels and low central vowels as presented above (6.1 and 6.2), this section shows four vowel plots to illustrate variation

across speakers. The plots are again grouped by target accent and each accent group has one plot for beginners and one for advanced students. The ellipses around the mean values mark the acoustical ranges of the respective vowels.

Starting with the results for the BrE group, the beginners' vowels (figure 7) overlap to different extents than the advanced students' vowels (figure 8).

With the beginners, larger areas of STRUT and BATH overlap, PALM almost completely covers the area of STRUT, BATH and LOT, which results from some mispronunciations of PALM as [pɔːm]. Both BATH and STRUT overlap slightly with LOT, and so does THOUGHT with LOT.

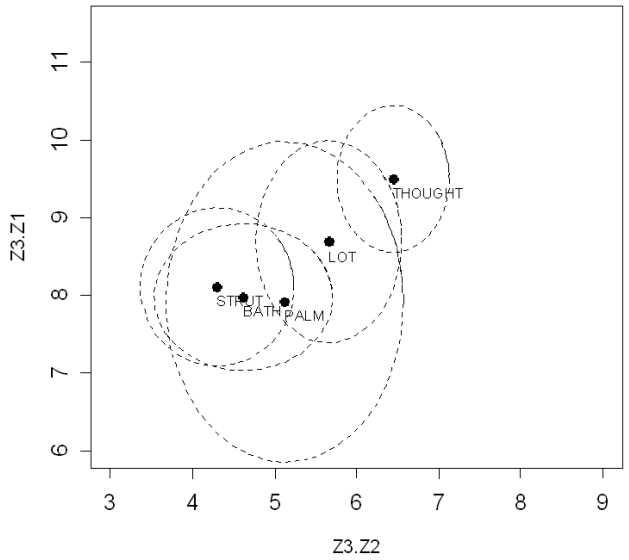


Figure 7: Beginners' individual variation in the pronunciation of BrE non-high back vowels and low central vowels.

The plot for the advanced students again shows the clearer distinction between PALM, STRUT and BATH mentioned above (6.1., figures 3 and 4). As a consequence, a wider area of vowel space is covered. This, however, does not result in a clear distinction between these vowels but rather leads to multiple overlaps of STRUT, PALM, BATH, and LOT. A noticeable difference between the beginners and the advanced students can be observed with respect to THOUGHT, which is almost completely distinct from LOT.

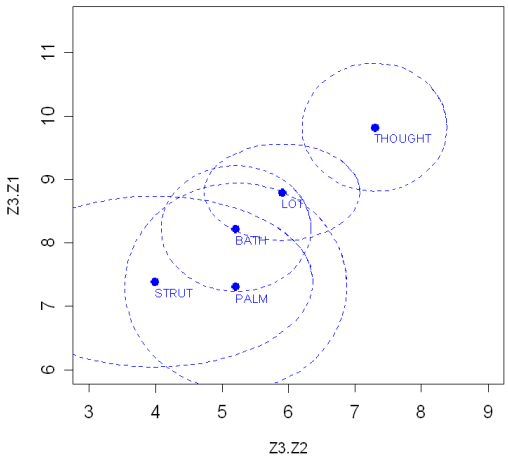


Figure 8: Advanced students' individual variation in the pronunciation of BrE non-high back vowels and low central vowels

Individual variation in the AmE target groups is shown in figures 9 (beginners) and 10 (advanced students). THOUGHT is almost fully distinct in both groups, overlapping to small extents with LOT in the advanced group and with PALM in both groups, the latter again being due to mispronunciations of PALM. In addition, both groups share a considerable overlap of LOT, STRUT and PALM. The evident difference between beginners and advanced students is that advanced students have a fully fronted version of BATH, fully distinct from LOT, STRUT, and PALM, whereas beginners' BATH strongly overlaps with these vowels.

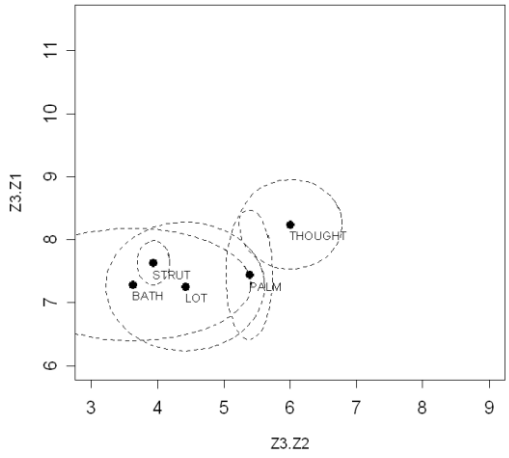


Figure 9: Beginners' individual variation in the pronunciation of AmE non-high back vowels and low central vowels

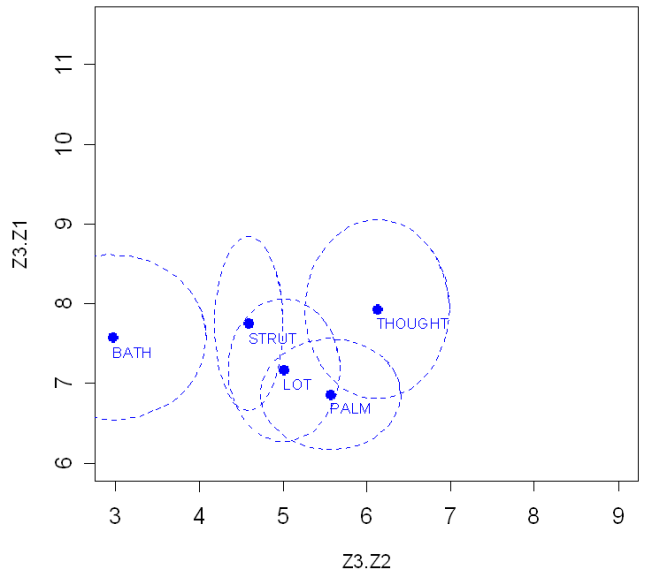


Figure 10: Advanced students' individual variation in the pronunciation of AmE non-high back vowels and low central vowels

In sum, this section has shown that the pronunciation of the vowels under scrutiny by German learners varies to a considerable extent, indicating that it seems difficult for learners to acquire a consistent and contrastive system, even after more than 10 year of instruction and some time spent abroad. In two cases, however, a differentiation of the vowel systems could be documented with the advanced students: THOUGHT becomes more distinct from LOT in the BrE system and BATH from STRUT/PALM/LOT in the AmE system. The overlaps and differentiations in the vowel systems under observation point to the fact that learners do make some progress in approximating a self-reported target accent with some vowels as proficiency increases, but fail to do so with others. A likely explanation might again be awareness. It is easy to picture that with greater experience in the foreign language, learners perceive a fronted version of BATH and a rounded and closer version of THOUGHT as symbols of AmE and BrE, respectively, and start to use these variants. Other characteristics, however, seem to go unnoticed.

7. Summary and Conclusions

The acoustical study of the mid and low back vowels and the low front voelws of 20 German learners of English at university level has yielded the following overarching results:

1. When targeting AmE or BrE non-high back vowels and low central vowels, German learners of English at university level make only little use of their native vowel systems. In other words, they are beyond a stage of strong L1 transfer.

2. The learners produce new vowels which are neither native German nor native English, which is a clear support for the reality of interlanguage as a system that, among other things, also contains "elements [...] that do not have their origin in either the NL or the TL" (Gass and Selinker 2008: 14).
3. Increasing experience in term of a closer approximation to the target is only reflected in two cases: BATH is more front in advanced learners aiming at AmE and THOUGHT is closer in advanced learners aiming at BrE. This might be due to an increased level of awareness of these vowels as a result of increases experience.
4. In general, however, experienced learners are not more native-like than less experienced learners with respect to the vowels under discussion after more than 10 years of learning.

All in all, the data presented here indicate that the learners of English analysed have not fully acquired an L2 sound system. Having demonstrated that only two very salient vowels start to be acquired at an advanced stage of proficiency, it seems that near-native pronunciation can only be acquired or learned – if at all - with attention to and awareness of the variability of the input[8]. Even experienced learners have no full awareness of the systematic differences between the two major accents of English. If near-nativeness in pronunciation is the envisaged goal of language learning, it is necessary to integrate awareness of varieties of English into the ESL/EFL classroom, and especially in teacher training at universities.

References

Best, C. and W. Strange. 1992. "Effects of phonological and phonetic factors on cross-language perception of approximants". *Journal of Phonetics* 20: 305-331.

Best, C. 1994. "The emergence of native-language phonological influence in infants: A perceptual assimilation model." In H. Nusbaum, J. Goodman and C. Howard, eds. *The Transition from Speech Sounds to Spoken Words: The Development of Speech Perception*. Cambridge, Mass.: MIT Press, 167-224.

Best, C. 1995. "A direct realist view of cross-language speech perception." In Strange, ed. 1995: 171-204.

Boersma, P. and D. Weenink. No date. "Praat: Doing phonetics by computer". [www.praat.org; accessed 31/06/2011] .

Bohn, O.-S. 2002. "On phonetic similarity". In P. Burmeister, T. Piske and A. Rohde, eds. *An Integrated View of Language Development*. Trier: WVT, Wissenschaftlicher Verlag Trier, 191–216.

Deterding, D. 1990. "Speaker Normalization for Automatic Speech Recognition". Ph.D.Thesis, Cambridge University.

[8] I fully agree with the anonymous reviewer that "[j]ust because less salient vowel qualities weren't learned it doesn't follow that near-native pronunciation would be possible even with attention to and awareness of the variability of the input". As usual in science, future research will shed more light on this issue.

Deterding, D. 1997. "The Formants of Monophthong Vowels in Standard Southern British English Pronunciation". *Journal of the International Phonetic Association* 27: 47-55.

Deterding, David, Jennie Wong and Andy Kirkpatrick. 2008. "The pronunciation of Hong Kong English". *English World-Wide* 29(2): 148-175.

Flege, J. E. 1995. "Second language speech learning: Theory, findings, and problems". In Strange, ed. 1995: 233–277.

Gass, S. M. and Selinker, L. 2008. *Second Language Acquisition: An Introductory Course* (3rd ed.). New York and London: Routledge.

Hansen Edwards, J.G. and M.L. Zampini, eds. 2008. *Phonology and Second Language Acquisition.* Amsterdam and Philadelphia: John Benjamins.

Hillenbrand, J., L. A. Getty, M. J. Clark and K. Wheeler. 1995. "Acoustic characteristics of American English vowels". *Journal of the Acoustical Society of America* 97(5):3099-3111.

Kautzsch, A. 2010a. "Exploring L1 transfer in German Learners of English: High front vowels, high back vowels and the BED/BAD distinction." *Research in Language* (Special Issue: *Proceedings of ACCENTS 2009, Lodz)* 8: 63-84.

Kautzsch , A. 2010b. Rhoticity in German Learners of English. Paper presented at World Englishes 2010 in Vancouver, Canada, August 2010.

Kendall, T. and E.R. Thomas. 2010. "Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package, version 1.1." [Software Resource: http://ncslaap.lib.ncsu.edu/tools/norm/]

Kortmann, B. 2005. "Chapter V: Contrastive Linguistics: English and German". In B. Kortmann. 2005. *Linguistics: Essentials*. Berlin: Cornelsen, 156-191.

Krashen, S. 1985. *The Input Hypothesis: Issues and Implications*. New York: Longman.

Kuhl, P.K. 1991. "Human adults and human infants exhibit a perceptual magnet effect for the prototypes of speech sounds, monkeys do not." *Perception and Psychophysics* 50: 93-107.

Kuhl, P.K. 1993. "Early linguistic experience and phonetic perception: implications for theories of developmental speech perception." *Journal of Phonetics* 21: 125-139.

Major, R. C. 2008. "Transfer in second language phonology. A review." In: Hansen Edwards and Zampini, eds. 2008: 63-94.

Labov, W., S. Ash and C. Boberg. 2006. *The Atlas of North American English*. Berlin: Mouton-de Gruyter.

Schmidt, R. 1990. "The role of consciousness in second language learning." *Applied Linguistics*, 11, 129-158.

Schmidt, R. 1994. "Implicit learning and the cognitive unconscious: Of artificial grammars and SLA." In N. Ellis, ed. *Implicit and Explicit Learning of Languages*. London: Academic Press, 165-209.

Schmidt, R. 1995. "Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning." In R. Schmidt, ed. *Attention and Awareness in Foreign Language Learning*. Honolulu, HI: University of Hawaii, Second Language Teaching & Curriculum Center, 1-63.

Schmidt, R. 2001. "Attention." In P. Robinson, ed. *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, 3-32.

Schmidt, R. 2010. "Attention, awareness, and individual differences in language learning". In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan and I. Walker. *Proceedings of CLaSIC 2010, Singapore*. Singapore: National University of Singapore, Centre for Language Studies, 721-737.

Schmidt, R. and S.N. Frota. 1986. "Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese." In R. R. Day, ed. *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, MA: Newbury House, 237-326.

Selinker, L. 1972. "Interlanguage". *International Review of Applied Linguistics* 10, 209–231.

Strange, W. 2007. "Cross-language phonetic similarity of vowels. Theoretical and methodological issues." In O.-S. Bohn and M.J. Munro, eds. 2007. *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege*. Amsterdam and Philadelphia: John Benjamins, 35-56.

Strange, W. and V.L. Shafer. 2008. "Speech perception in second language learners: The re-education of selective perception." In: Edwards and Zampini, eds. 2008: 153-192.

Strange, W. ed. 1995. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Baltimore: York Press.

The R Project for Statistical Computing. No date. "R". [www.r-project.org/, accessed 31/06/2011].

Thomas, E. R. and T. Kendall. 2007. NORM: The vowel normalization and plotting suite. [http://ncslaap.lib.ncsu.edu/tools/norm/, accessed: 07/06/2012].

Traunmüller, H. 1997. "Auditory scales of frequency representation". [http://www2.ling.su.se/staff/hartmut/bark.htm, accessed: 07/06/2012].

Wells, J.C. 1982. *Accents of English.* Cambridge, Cambridge University Press.

MISPRONOUNCED LEXICAL ITEMS IN POLISH ENGLISH OF ADVANCED LEARNERS

JOLANTA SZPYRA-KOZŁOWSKA
Maria Curie-Skłodowska University

Abstract
The paper is a continuation of the author's earlier studies in which she argues that it is the mispronunciation of whole words due to their incorrect phonological storage in the learners' phonetic memory that is more detrimental to successful communication via English than an inaccurate production of individual segments and suprasegmentals. Consequently, phonetically difficult words deserve to be thoroughly investigated and pedagogically prioritized.

The present study is a report on an experiment in which 20 English Department students, all advanced learners of English, were recorded having been asked to read a list of diagnostic sentences containing 80 words known to be problematic for Poles in terms of their pronunciation. This has been done in order to isolate and examine the major error types, to establish a hierarchy of difficulty among 8 sources of pronunciation errors, to compare the obtained results with the most common error types made by intermediate learners and to juxtapose the participants' subjective evaluation of the phonetic difficulty of words with their actual phonetic performance. The final goal is to draw pedagogical implications for the phonetic training of advanced students of English.

1. Introduction

A striking feature of foreign-accented English, including Polish-English, is a frequent occurrence of the so-called local errors, i.e. words stored in the learners' phonetic memory in an incorrect phonological shape. In a number of studies (Szpyra-Kozłowska 2011, Szpyra-Kozłowska and Stasiak 2010, Szpyra-Kozłowska in press) I argue that the use of such items is more detrimental to successful communication via English than inaccurately produced segments and suprasegmentals. In Szpyra-Kozłowska (in press) I present experimental evidence that local errors significantly decrease Polish learners' comprehensibility and intelligibility, create the impression of a heavy foreign accent and are irritating for native English listeners. Consequently, Szpyra-Kozłowska and Stasiak (2010) conclude that a shift is needed in phonetic instruction from the focus on sounds, sound contrasts and prosodies to the focus on the pronunciation of problematic words. To achieve this goal, however, a deeper insight is required into what types of items are phonetically difficult for learners of different L1 background and various levels of language proficiency. Szpyra-Kozłowska (2011) attempts to examine this issue in relation to intermediate Polish learners and identifies eight major sources of word pronunciation errors.

The present paper undertakes the problem of mispronounced words in the speech of advanced Polish learners of English. It is a report on an experiment in which 20 English Department students of Maria Curie-Skłodowska University in Lublin, Poland, were recorded having been asked to read a list of diagnostic sentences containing 80 words known to be problematic for Poles in terms of their pronunciation. This has been done in order

- to isolate and examine the major types of phonetically difficult words;
- to establish a hierarchy of difficulty among 8 chief sources of pronunciation errors in the speech of advanced learners;
- to compare the obtained results with those of intermediate learners;
- to examine the experimental results with the predictions of the PDI;
- to juxtapose the participants' subjective evaluation of the phonetic difficulty of words with their actual phonetic performance;
- to draw pedagogical implications for the phonetic training of advanced students of English.

It is hoped that although the study is carried out in the Polish context, many of the observations made here will be relevant for other types of foreign-accented English.

2. Sources of word mispronunciations

Many sources of word pronunciation errors commonly made by Polish learners are well-known and have been identified by previous research.

In this context Sobkowiak's (1999) work on the Phonetic Difficulty Index (PDI) should be pointed out as a valuable attempt to deal systematically with phonetically difficult words in Polish English. PDI (p. 214) "is a global numerical measure of the phonetic difficulty of the given English lexical item for Polish learners," meant to be included in machine-readable EFL dictionaries and thus having mainly lexicographic applications. It contains phonetic difficulty ratings of English words carried out by the author on the basis of his observations of Polish learners' pronunciation problems. The current list of error sources (personal communication) includes 61 issues which can, however, be grouped into more general categories. Thus, the largest set (26) concerns spelling-related problems while the next largest group (24) involves problems with the pronunciation of individual sounds and combinations of sounds (e.g. vowel hiatus, consonant clusters). A prominent position is also occupied by stress-related problems (5). The remaining error sources concern the incorrect application of Polish phonological rules (such as Word-Final Obstruent Devoicing and Voice Assimilation) to English, word length (more than 5 syllables) and several others.

It should be added that Sobkowiak's list and his PDI are of general nature and do not specify the relationship between the degree of words' phonetic difficulty and the learners' level of English proficiency. This is a serious drawback of his proposal since what is difficult for beginners might be fairly trivial for more advanced students. In other words, the phonetic difficulty of English words should be examined in relation to their proficiency level.

Taking this fact into account, in a recent study (Szpyra-Kozłowska 2011) I examined the sources of phonetic difficulty of English words in intermediate learners' speech. The following eight major types of issues have been isolated:

1. Spelling-related problems.
2. Phonetic 'false friends.'
3. Stress-related problems.
4. Pronunciation of consonant clusters involving interdentals.
5. Pronunciation of long words.
6. Pronunciation of words containing several liquids.
7. Pronunciation of words containing sequences of high front vowels.
8. Pronunciation of words with morphological alternations in related forms.

Since the types of problems listed above will be subject to experimental verification, some explanation of these issues is in order.

Spelling-related difficulties result from two kinds of interference. The first one involves interference from Polish spelling-to-pronunciation conventions incorrectly applied to English words. Thus, typical examples comprise silent letters pronounced in such items as <t> in *nestle* and <b> in *tomb*. Another problem stems from incorrect overgeneralizations of English letter-to-sound rules, for example, interpreting the digraph <ea> as the vowel [i:] in *steak* (as in *meat, leaf, teach*) or <ace> as [eɪs] in *surface* and *palace* (as in *face, lace*). As Polish learners have more access to written rather than to spoken English, spelling exerts a powerful effect on Polish English pronunciation.

Phonetic 'false friends' are numerous lexical items which occur in both languages in an identical or a similar orthographic form, but with different pronunciation. In the majority of cases they are cognates, e.g. E *chaos* / P *chaos*, borrowings from English, e.g. E *model* / P *model* or just accidental look-alikes, e.g. E *gnat* / P *gnat* 'bone.' A large group of such words are proper nouns which appear in both languages in rather different phonetic shapes, e.g. *Nepal* – E [nɪ'po:l] / P ['nepal] and *Sidney* – E ['sɪdnɪ] / P [s'idnej]. Similar or identical spelling suggests to the learners that their pronunciation must be similar as well.

For speakers of languages with fixed stress, such as Polish, learning the intricacies of the English stress system with its irregularities and exceptions is a genuine challenge. Thus, while Polish learners typically employ Polish penultimate stress to English words, e.g. *'Japan, in'dustry, demon'strated*, they frequently stress also other syllables, i.e. ultimate, e.g. *e'ffort, fe'male* and antepenultimate, e.g. *'successful, 'computer, ar'bitrary* (for a more detailed discussion of stress errors in Polish English see Waniek-Klimczak 2002).

The interdental fricatives, which are absent in Polish, belong to the most difficult sounds for many foreign learners. The degree of difficulty increases when they occur in combination with other consonants. Intermediate learners who participated in our previous experiment listed the following difficult words, all of which contain consonantal clusters with interdentals: *three, throw, birthday, maths, healthy, sixth*.

The next source of difficulty is the length of words. Longer words are problematic for the learners because of a variety of factors to be controlled: the placement of stress, the articulation of many different new sounds and complex sound sequences. The question that arises concerns the actual length of words which makes their pronunciation difficult. Below we list some examples, taken from Szpyra-Kozłowska (2011), supplied by intermediate learners as difficult because of their length,

(a) trisyllables: *excitement, adventure, Australia, picturesque*
(b) quadrisyllables: *relaxation, astonishing, surprisingly*
(c) quintisyllables: *encyclopaedia, occasionally, exaggeration*

According to these data, words marked as problematic because of their length contain three syllables or more. For intermediate learners the longer a word, the more difficult it is to pronounce.

One of the most interesting results of our study involving intermediate learners (Szpyra-Kozłowska 2011) was the discovery that the presence of several liquids, i.e. rhotics and laterals, contributes to the considerable pronunciation difficulty of a word. Here are some examples supplied by the participants: *appropriate, library, regularly, particularly, rarely, burglary*.

Many such items, apart from articulatory difficulty, are problematic because of their spelling since <r>, appearing in the word-final and preconsonantal position, is a silent letter in nonrhotic accents such as RP, generally taught in Poland. Since learners are often confused as to which r's to pronounce, many of them attempt to articulate all these letters, which creates several liquids in a single word.

The collected data include also words regarded as difficult by the respondents due to the fact that morphological alternations take place in the roots they contain. Since in English such changes are often highly irregular and idiosyncratic, this fact contributes to the perceived difficulty of the items in question. Some examples are presented below with related forms provided in parentheses.

(a) *so**c**iety* (so**c**ial), *nor**th**ern* (nor**th**), *sou**th**ern* (sou**th**), *an**x**iety* (an**x**ious)
(b) *c**a**n't* (c**a**n), *var**i**ety* (var**i**ous), *br**ea**the* (br**ea**th), *w**i**dth* (w**i**de),

In (a) segments subject to consonant alternations are underlined while in (b) vowel alternations are indicated. It is likely that pupils learn first more frequent words given in parentheses and when faced with less common related items, transfer the pronunciation from the former to the latter by analogy or due to preserving paradigm uniformity. The degree of difficulty increases due to the fact that in the above forms the alternating segments are spelt in the same way.

The last category of phonetically difficult words for intermediate Polish learners comprises items which contain two (a) or three (b) different high front vowels, i.e. [ɪ] and [i:], e.g.

(a) *reading, sleeping, cheating, speedy, greedy, sleepy*
(b) *believing, receiving, preceding, repeating,*

In such instances they tend to employ some kind of vowel harmony and pronounce two [i:] vowels (or rather its shorter and less tense Polish counterpart [i]).

Yet another problem with [ɪ] is created by the following words: *innocent*, *image*, *impression*, *important*, *industry*.

In these items the initial vowel is difficult for Polish learners to pronounce and usually replaced with Polish [i]. Apart from the powerful influence of English spelling, another active factor here seems to be a phonotactic constraint of Polish banning in the word initial position the occurrence of the Polish front centralized vowel [ɨ], very close to English [ɪ].

It should be added that, as demonstrated earlier, in many cases more factors than one contribute to the phonetic difficulty of words. For instance, long words are often problematic not only due to their length, but also because of the stress placement or combinations of sounds they contain.

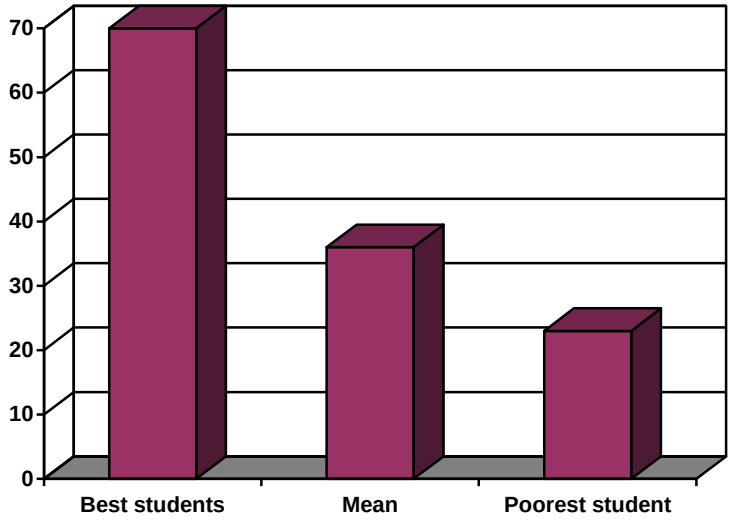
3. Experimental design

In October 2010 twenty randomly selected 4th year students of the English Department of Maria Curie-Skłodowska University, Lublin, Poland, all advanced learners of English, took part in the experiment in which they were asked to read aloud a list of sentences (see Appendix 1) containing 80 phonetically difficult words, with 10 items representing each of the 8 types of error sources discussed in the preceding section. The students were then individually recorded. After the recording they were given a short questionnaire to complete (see Appendix 2). They were provided with a list of 24 words which appeared in the diagnostic sentences (with 3 items representing each of the 8 categories) and asked to evaluate the degree of phonetic difficulty they posed for them (easy, medium and difficult to pronounce). They were also requested to select three particularly difficult words and comment on the source of the problem. The recordings were next auditorily assessed by the researcher.

4. Results and discussion

4.1. General results

The experiment yielded 1600 tokens, of which only 655 (41%) were pronounced correctly and 945 (59%) incorrectly. Graph 1 shows that the results of individual participants range from 22.6% of correctly pronounced experimental items by the poorest student to 70% by the best student in this group. The mean result is 36%.



Graph 1. Performance of best and poorest students

The above figures indicate that even for advanced students the experimental items constitute a serious learning problem which has to be approached and remedied.

4.2. Hierarchy of word pronunciation difficulty factors

The experimental data allowed us to establish the following hierarchy of difficulty of the 8 factors presented in section 2 for advanced students:
Relatively easy types (over 50% of correct responses):
Clusters of 'th' and consonants different than /s/ - 70%
Liquids – 66% of correctly pronounced tokens (with two words being considerably more difficult, i.e. *particularly* and *regularly* – only 7%)
Stress – 52% (particularly difficult: *caricature*)
Long words – 51% (particularly difficult: *artificiality*, *congratulatory*, *authoritarian*, *unintelligibility*)
Medium difficulty (25%-40% of correct responses)
Spelling – 40% (particularly difficult: *hideous*, *haven*, *thoroughly*, *Graham*)
Alternating forms – 33.5% (particularly difficult: *courteous*, *advantageous*, *managerial*, *infamous*)
Considerable difficulty (below 25% of correct responses):
Phonetic 'false friends' – 24% (particularly difficult: *algebra*, *gigantic*, *Disney*)
Clusters of 'th' and /s/ - 12% (particularly difficult: *strengths*, *lengths*)
High front vowels – 7.5% (all difficult)

The above data require some comments. First of all, within most categories there are words of a different degree of pronunciation difficulty for the participants. Only the set containing sequences of high front vowels is homogeneous in this respect in that all of them proved to be equally problematic. The same is true in the case of clusters of interdentals and other fricatives. Thus, in these two instances we can talk of truly global errors, not restricted to any particular lexical items. In the remaining cases there were both easier and more difficult words, which means that other factors, apart from the ones discussed here, are also relevant here. For example, while the words containing sequences of liquids do not pose any major difficulty for advanced learners, two items, i.e. *particularly* and *regularly* are commonly mispronounced by them.

Pedagogical implications of the established hierarchy of word difficulty are obvious. Advanced students should receive additional training in the pronunciation of words with sequences of high front vowels, items with clusters of interdentals and /s/ and forms which are subject to irregular morphological alternations. The next major source of errors is the existence of phonetic 'false friends', whose number runs into hundreds, not only in Polish, but also in many other languages as well. We would like to suggest that in phonetic practice use should be made of such 'minimal pairs,' employed in, for instance, Szpyra-Kozłowska and Sobkowiak (2011) . They should include both common and proper nouns, e.g.

English	**Polish**	**English**	**Polish**
atom	*atom*	*Adam*	*Adam*
boa	*boa*	*Nepal*	*Nepal*
safari	*safari*	*Madrid*	*Madryt*
marketing	*marketing*	*Cameron*	*Cameron*
model	*model*	*Clinton*	*Clinton*

4.3. A comparison of word difficulty for intermediate and advanced learners

A comparison of factors contributing to the difficulty of word pronunciation for intermediate and advanced learners shows that, statistically, the latter group has learnt to deal better with the issue of liquids, word stress, long words and spelling-related problems. Advanced learners have also fewer problems with consonantal clusters involving interdentals, with the exception of 'th' followed by /s/. The most difficult items turned out to be *lengths* and *strengths*, both containing clusters of three consonants. The issues which are problematic for both groups involve sequences of high front vowels within single words, phonetic 'false friends' and forms displaying irregular morphological alternations. This means that such difficulties should be given special attention in the phonetic training of all learners. These observations are summarized below.

Hierarchy of word pronunciation difficulty factors

Intermediate learners	**Advanced learners**
All factors of similar difficulty	Most difficult: high front vowels <th + s> phonetic 'false friends' less difficult: morphological alternations spelling-related problems long words stress-related problems sequences of liquids

4.4. Difficult and easy words

Let us now examine in some detail those words which, among the 80 diagnostic lexical items, proved to be particularly difficult or easy for the participants. The easiest words to pronounce, with over 85% of proper realizations are the following:

(a) *rural literally burglary barely*
(b) *monthly birthday hundredth*
(c) *various* (variety), *anxiety* (anxious), *sincerity* (sincere)

The examples in (a) contain sequences of liquids while those in (b) clusters in which the interdentals are combined with nonfricatives. The words in (c) participate in morphological alternations, as seen in related forms provided in parentheses.

The most problematic items can also be divided into several sets.

(a) *cheating ceiling greedy repeating deceiving*
(b) *courteous* (court) *advantageous* (advantage) *managerial* (manager)
(c) *algebra* (P algebra) *caricature* (P karykatura)
(d) *strengths lengths*

The first and the largest of them in (a) contains sequences of high front vowels The next one in (b) involves irregular morphological alternations. The third group in (c) has cognates in Polish. Finally, the last one in (d) comprises clusters of velar nasals, followed by interdentals and /s/.

Moreover, some words might be claimed to cause difficulty because of the complex relationship between spelling and pronunciation, for instance those with the suffix *–ous* added to stems ending in <e>, e.g. *courteous, hideous, advantageous*.

It should also be noted that some of the most problematic items are fairly long and contain at least four syllables: *advantageous, caricature, managerial*.

To sum up, the items provided in this section support the observations made earlier concerning the major sources of word mispronunciations. Yet another question that naturally arises in connection with the experimental items concerns the relationship between the phonetic difficulty of these words for the learners and their frequency of occurrence. We have found no meaningful relationship between these two issues. Thus,

the frequency figures, based on the British National Corpus of Spoken English, are identical or almost identical for many words belonging both to the category of difficult and easy words, e.g.

Frequency of easy words		Frequency of difficult words	
hundredth	*2*	*courteous*	*2*
burglary	*3*	*haven*	*3*
sincerity	*1*	*managerial*	*1*

The problem is that word frequency in the British National Corpus of Spoken English does not have to be the same as word frequency in foreign students' English for which no data are available. This means that the former source is of limited usefulness in predicting the degree of difficulty involved in word pronunciation.

4.5. Experimental results versus Phonetic Difficulty Index

In this section we examine the accuracy of Sobkowiak's Phonetic Difficulty Index, with its 10-point scale, in predicting the degree of difficulty of the experimental items.

It appears that in some cases the PDI values do coincide with the easy/difficult dichotomy established in this study, e.g.[1]

Easy words vs PDI value	Difficult words vs PDI value
sincerity – 0	*courteous* – 8
criticizing – 1	*caricature* - 7
literally – 1	*advantageous* – 7

In the majority of cases, however, no significant correlation between the two evaluations can be found. Thus, frequently our easy and difficult words are given the same PDI values.

Easy words vs PDI value	Difficult words vs PDI value
variety – 2	*haven* – 2
barely – 2	*ceiling* – 2
monthly – 3	*cheating* – 3

In some instances easy words have a higher PDI value than the difficult ones, e.g.

Easy words vs PDI value	Difficult words vs PDI value
rural – 4	*strengths* – 3
burglary – 5	*receiving* – 2
hundredth – 5	*algebra* – 3

[1] I am grateful to W. Sobkowiak for providing me with the PDI values of the experimental items. The difficulty scale ranges from 0 to 10, where the higher the score, the greater the phonetic difficulty of words.

An analysis of 24 easy and difficult experimental words with their PDI values shows that a correlation is found in about 50% of cases only. We can conclude that the PDI, in its present shape, is rather inaccurate as a measure of the phonetic difficulty of words for the advanced learners who took part in our experiment.

4.6. Students' evaluation of word difficulty

In the second part of the experiment the participants were asked to evaluate the degree of difficulty involved in the pronunciation of 24 experimental items representing eight types of factors isolated in section 2.

According to the subjects, the following factors make words difficult for them to pronounce:

- length (e.g. *unintelligibility, satisfactorily*)
- low frequency of occurrence (e.g. *unintelligibility, satisfactorily*)
- the presence of <th+s> clusters (e.g. *sixths, lengths, maths*)
- spelling and specific word endings (such as *–eous*) (e.g. *thouroughly*, *courteous, southern*)

Thus, the most frequently mentioned source of word difficulty was their considerable length (left unspecified), with two words judged as the most problematic of the tested items, i.e. *unintelligibility* and *satisfactorily*. The low frequency of the same words was also indicated as a cause of pronunciation problems. *Sixth, lengths* and *maths*, all containing interdental fricatives followed by /s/, were listed as difficult because of such clusters. Irregular spelling and sound correspondences in *thouroughly*, *courteous, southern* and *anxiety* were blamed for pronunciation problems with these items. The participants also enumerated some troublesome word endings, e.g. *-eous* (*courteous*), *-rily* (*satisfactorily*), *-rely* (*rarely*). Interestingly, problems with the correct placement of stress were not mentioned.

Of the 24 items subject to students' evaluation, the following five were judged the most difficult (the figures in parentheses indicate the number of participants who evaluated these words as such): *unintelligibility* (15), *satisfactorily* (10), *courteous* (5), *thoroughly* (5), *sixth* (5).

It is interesting to examine whether these judgements find confirmation in the students' actual performance. Taking into account the 24 items under consideration, the most difficult words for them to pronounce were as follows (with the percentage of correct realizations provided in parentheses): *courteous* (0%), *greedy* (0%), *innocent* (8%), *ceiling* (8%), *thoroughly* (16%).

These data show that the participants' opinions on word difficulty coincide with their phonetic performance only in the case of two items, i.e. *courteous* and *thoroughly*. In the remaining instances there is no such correlation. Thus, two words claimed to be the most difficult received the following scores: *unintelligibility* was pronounced correctly by 30% of the students and *satisfactorily* by 58%, which makes them items of medium difficulty. Interestingly, none of the respondents mentioned as problematic the following words with fairly low scores for correctness: *Murphy* (20%), *lengths* (20%), *southerners* (25%), *Nepal* (25%).

The observations reported here indicate that even advanced learners who undergo formal phonetic training are only partly aware of their pronunciation problems.

Let us examine in more detail the relationship between the students' phonetic performance and their assessment of word difficulty. We compared 8 evaluations of 24 words as easy or difficult to pronounce by 4 students with good pronunciation and 4 students with poor pronunciation with their actual realization of these items. We counted the number of matches and mismatches between the questionnaire answers[2] and the participants ' production of a given item. The results are shown below.

	items judged easy and pronounced correctly	**items judged easy and mispronounced**
good students	26	11
poor students	9	24

It turned out that good students were more accurate in their assessment of word difficulty than poor students. This means that good students more often consider words as easy when they can actually pronounce them correctly than poor students who often mark words as easy and yet mispronounce them.

5. Conclusions

It is hoped that the presented study has provided some insight into the issue of phonetically difficult words in the speech of advanced Polish learners. It has allowed us to make several observations which carry important pedagogical implications.

1. Phonetically difficult words abound in Polish-accented English of learners of different levels of proficiency, including intermediate and advanced students. Consequently, this issue should be given due attention in the course of their phonetic training.
2. The most important sources of word mispronunciations for advanced learners involve sequences of high front vowels, clusters of interdentals with other fricatives and phonetic 'false friends.'
3. Advanced students, when compared with intermediate learners, have fewer problems with spelling and stress-related issues, sequences of liquids, longer words, and clusters of interdentals with nonfricatives.
4. Since sets of phonetically difficult words for intermediate and advanced learners overlap only partially, there can be no one PDI common for all of them.
5. The comparison of students' evaluation of word difficulty with the experimental results indicates that even advanced learners are only partly aware of their pronunciation problems and cannot assess them objectively. Thus, more care should be taken to develop their skill of self-evaluation.

[2] A match was declared when an item was pronounced correctly and marked as easy to pronounce or when an item was mispronounced and marked as difficult to pronounce.

References

Sobkowiak, W. 1999. *Pronunciation in EFL Machine-Readable Dictionaries*. Poznań: Motivex.

Szpyra-Kozłowska, J. 2011. Phonetically difficult words in intermediate learners' English. In Pawlak, M,. E. Waniek-Klimczak & J. Majer (eds). *Speaking in contexts of instructed foreign language acquisition.*. Bristol: Multilingual Matters. 286-299.

Szpyra-Kozłowska, J. & W. Sobkowiak. 2011. *Workbook in English Phonetics*. 2nd edition. Lublin: Wydawnictwo UMCS.

Szpyra-Kozłowska, J & S. Stasiak. 2010. From focus on sounds to focus on words in English pronunciation instruction. *Research in Language* vol. 8., 163-174..

Szpyra-Kozłowska, J. in press On the irrelevance of sounds and prosody in foreign-accented English.

Waniek-Klimczak, E. 2002. Akcent wyrazowy w nauczaniu języka angielskiego. In Sobkowiak, W. & E. Waniek-Klimczak (eds) *Dydaktyka Fonetyki Języka Obcego. Zeszyty Naukowe PWSZ w Płocku,* t. III. Płock: Wydawnictwo Naukowe PWSZ w Płocku. 101-114.

Appendix 1

A list of diagnostic sentences. The tested items are in boldface.

1. His **mania** for watching **Disney** cartoons and **horror** films all night made the hotel **management increasingly uncomfortable**.
2. **Graham** and **Murphy** went straight from **Madrid** to **Nepal**, where they joined the **demonstrators** fighting for freedom in **Tibet.**
3. It is frequently claimed that at school **Einstein** showed no **enthusiasm** or **appreciation** for **algebra,** and his **maths** teacher **regularly** accused him of **cheating.**
4. He went to great **lengths** to **characterize appropriately** all his **strengths** and weaknesses to prove he was **innocent** and did not commit this **burglary.**
5. With mounting **curiosity** he examined the whole **area thoroughly** for the **sixth** time and decided that the evidence that the **infamous murderer** was there was not **satisfactorily** established and was purely **circumstantial.**
6. The **artificiality** and **unintelligibility** of his explanations created much **anxiety particularly** in this **rural area** were politicians are **rarely** trusted and their **sincerity** is frequently questioned.
7. The **rivalry** between the **southerners** and the **northerners** in this **sleepy** town was **closely** watched by the **managerial** staff of this **industry** who thought it **advantageous** for a **variety** of reasons.
8. In this **monthly** I saw a **caricature** of this **admirable, courteous** man whom the media keep **simultaneously** praising and **criticizing**.
9. It was her **hundredth birthday** and the **organizers** of the party made every **effort** to control the **chaos** and provide various attractions: a **gigantic neon** with a **congratulatory** message, a champagne **geyser** and a band of **robots** playing very **rhythmical** music.
10. They considered this place their **ultimate haven**, with **chestnut** trees, **thyme** and **heather** in the garden decorated with **hideous** dwarfs.
11. This **greedy, unsophisticated** person who kept **repeating** that he was **capable** of **deceiving literally** everybody held a prestigious **administrative** post.
12. He could **barely** wait for **receiving** an **explanatory** statement from this **authoritarian** official.
13. The **width** of this **ceiling** was truly **impressive**, but the paintings on it were fairly **imitative.**

Appendix 2

The questionnaire used in the experiment

Evaluate the difficulty of pronunciation of the following words by marking them as

E – easy to pronounce
M – medium
D – difficult to pronounce

1. thoroughly	2. demonstrators
3. lengths	4. Murphy
5. administrative -	6. particularly
7. greedy	8. anxiety
9. curiosity	10. geyser
11. sixth	12. unintelligibility
13. burglary	14. ceiling
15. southern	16. explanatory
17. chestnut	18. Nepal
19. criticizing	20. satisfactorily
21. rarely	22. courteous
23. innocent	24. maths

Now choose three words from the above list which you consider difficult to pronounce and comment on why you find them difficult.