Andrzej Porzuczek

The timing of tone group constituents in the advanced Polish learner's English pronunciation

Wydawnictwo Uniwersytetu Śląskiego • Katowice 2012

The timing of tone group constituents in the advanced Polish learner's English pronunciation

NR 2940

Andrzej Porzuczek

The timing of tone group constituents in the advanced Polish learner's English pronunciation

Wydawnictwo Uniwersytetu Śląskiego Katowice 2012

Editor ot the Series: Językoznawstwo Neofilologiczne
Maria Wysocka

Referee
Ewa Waniek-Klimczak

CONTENTS

Introduction

Foreign language (FL) learning involves the acquisition of the lexicon and grammatical patterns of that language. The process is believed to be influenced by the learner's native language (L1) competence so strongly that the developing FL system has been referred to as interlanguage (Selinker 1974), accommodating elements of the mother tongue and the target language within a dynamic, "approximative system" (Nemser 1971). L1 influence or cross-linguistic interference (cf. Weinreich 1953) has naturally been recognised as a crucial factor that shapes the interlanguage of FL speakers. Although interlanguage is a system developed by an individual speaker, the influence of L1 makes it possible to distinguish national interlanguage varieties. Polish learners of English share numerous largely predictable L1-related features, which at the phonetic level contribute to the general impression of Polish accent. Consequently, a Polish-based interlanguage where English is the target is often named "Polglish" (cf. Sobkowiak 2001).[1]

The fact that the influence of L1 results in predictable problems in FL learning is very important for FL pedagogy. It suggests a selection or even a hierarchy of potential problems on which to focus with respect to the learners' linguistic background. This knowledge offers an opportunity to devise more efficient syllabuses, better suited for the needs of particular groups of learners.

Discrepancies between native and non-native pronunciation pertain to phonological differences between the two systems and problems with conceptual acquisition of the system, but the learner also needs to develop purely physical speech production skills, necessary for communication. Human speech production is naturally anatomically and physiologically conditioned.

[1] The term "Polglish" is used in this book for convenient reference to the Polish subjects' English pronunciation despite certain controversies that may be raised especially in connection with the dynamic aspect of the learners' language under investigation.

The physiological mechanisms responsible for speech production are usually fully developed in Polish learners of English because a majority of them do not start learning a foreign language before primary school. Therefore they need to "separate" L1-specific motor control habits from universal ones and try to replace the former or modify them to suit the needs of FL pronunciation. Certainly the economy principle suggests relying on the already acquired patterns of behaviour as much as possible, which calls for negotiating the degree of necessary alterations in the articulatory patterns. The Critical Period Hypothesis (CPH — cf. Lenneberg 1967) does not only refer to L1 acquisition but also to FL acquisition ability (Major 2001). What is more, as Scovel (1988: 101) points out, the critical period particularly affects the acquisition of FL pronunciation, which is "the only aspect of language performance that has a neuromuscular basis."

The influence of critical period is questioned by, for instance, Burns et al. (2007), Hoonhorst et al. (2009) and Rivera-Gaxiola et al. (2005). Moreover, Birdsong (1992), Bongaerts (1999) and Moyer (1999) argue that adequate instruction and practice may lead to mastering native-like FL pronunciation, although Scovel (1988: 185) only admits that as few as one in a thousand late FL learners ("superexceptional" ones) can overcome the critical period constraints. Similarly, Markham (1997) estimates this proportion to reach not more than 0.1%—3%.

The effects of critical period may be debatable, but there is no doubt that L1 interference is an important factor in FL learning. The question then arises about a specific target for the learner.

It is a commonly accepted view that intelligibility ensuring successful communication is the substantial goal in FL production. It is often claimed that FL speech should be intelligible and that foreign accentedness does not necessarily affect intelligibility (cf. Neri et al. 2002, Pennington 1999, Munro and Derwing 1995). Consequently, in order to facilitate this aspect of foreign language teaching and learning, efforts have even been made to work out a simplified model of English pronunciation for international communication. Gimson (1978) proposed 'rudimentary international pronunciation', and Jenkins (1998, 2000) suggested the priorities of Lingua Franca Core (LFC) for international communication.

However, native models of English pronunciation for foreign learners are still more popular because they make a clear, consistent target for language learning. Therefore, if foreign-accented production is acceptable as long as it is intelligible, it must be developed through imitation of natural, native patterns.

The effects of foreign accentedness on communication may be a matter of debate, but there is certainly a relatively large proportion of learners who, for a variety of reasons, are determined to minimise traces of foreign accent in their speech. Their motivations range from purely "aesthetic" reasons to fears

of being "subjected to discriminatory attitudes and negative stereotypes," as Felps et al. (2009: 920) point out, concluding that "by achieving near-native pronunciation, L2 learners stand more to gain than just better intelligibility." Non-native pronunciation problems faced by L2 speakers in their new language communities are reported in studies carried out by Anisfeld et al. (1962), Arthur et al. (1974), Lippi-Green (1997), Ryan and Carranza (1975), and Schairer (1992). Moreover, Lev-Ari and Keysar (2010) found that people speaking with a strong non-native accent appear less credible to listeners than those who have a mild foreign accent and native speakers. Finally, there are the non-native language teachers and prospective teachers, who naturally desire to serve as reliable immediate pronunciation models for their students (Jenkins 2007).[2]

Pronunciation training traditionally embraces segmental and suprasegmental phonetics. A lot of argument and many publications have been devoted to the problem of establishing proper relations between the two general aspects in glottodidactics. Prosodic features of FL speech have been regarded as an important but difficult part of the learning process (e.g. Szpyra-Kozłowska et al. 2003, Roach 2002, Celce-Murcia et al. 1996, Jenner 1999, Bogle 1996, Nunan 1995, Kenworthy 1987, Wrembel 2002, Nowacka 2003). Among others, problems with FL timing patterns have been pointed out in a number of studies (e.g. Flege 1984, Cunningham 1986).

Among the prosodic features of speech, the duration of utterance and its constituents is one of the most objective, measurable parameters of speech. If observations of this parameter reveal clear differences between learners and native speakers, they do not only signal the existence of a didactic problem, but also lead to better understanding of the Polish learner's problems with English pronunciation and suggest further research directions in which to investigate the other prosodic factors, related to duration.

This study is a diagnostic, pedagogically oriented analysis of Polish learners' English pronunciation with the focus on timing. It is designed to describe timing differences in the English read speech of native speakers and advanced Polish learners, and the rate and efficiency of approaching native standards in the course of English studies including practical phonetics training. The results indicate the characteristics of English pronunciation such as intrinsic segmental length, prominence conditions and position in the utterance[3] that lead to timing differences between the two groups of speakers. This kind of evidence that apart from indicating the problems shows their actual scale and developmental tendencies may prove helpful in designing practical pronunciation courses for Polish learners of English.

[2] However, Jenkins stresses that non-native teachers rarely achieve native-like competence.
[3] Detailed discussion in Chapter 3.

The methods of evaluating the timing discrepancies between the learners and native speakers with reference to vowel duration (SD:M ratio — section 5.4.3), syllable duration (SVI — section 6.2.5), and foot duration (section 6.3), considering their simplicity, which makes them readily available to an average FL pronunciation teacher, can serve as a useful didactic aid in the aspects of language acquisition, the assessment of which is often criticised for an entirely impressionistic approach. Admittedly, the study analyses duration alone, but this parameter reflects numerous important features of speech, and the methods can be helpful in teaching quantitative vocalic contrast, vowel reduction, the marking and distribution of prominences, and the rhythmic patterns of the target language.

The differences in speech timing between Polish learners of English and native speakers are characterised on the basis of an empirical study of read speech of Polish students at a teacher training college and a group of twelve native British English speakers. Read speech is a form of test where it is possible to compare the performance of the same text by two groups of speakers in controlled conditions and the prosodic organisation is closer to natural communication than in word lists or sentences. Experiments with isolated words, schematic sentence patterns or reiterant speech run the risk of the subject's focussing unnaturally on the form of speech, which may exert a strong influence especially on prosody. For instance, Tatham and Morton (2001: 191) claim that short sentences and unnatural utterances within frames "tend to develop a rhythm of their own which might well approximate to isochronic repetition of stressed syllables."

On the other hand, although it is spontaneous speech that provides the most reliable information about natural language performance, the obtained content is usually insufficient for inter-speaker comparison with respect to many investigated areas. Moreover, considering the specificity of foreign language learning in classroom conditions, it is controlled speech rather than spontaneous production that the teacher handles, especially if pronunciation training constitutes a separate course within the programme of studies. If an explicit pronunciation course is applied, then spontaneous conversational performance appears as the ultimate phase, or rather the result, of training.

This book consists of seven chapters, including the conclusion. The theoretical part discusses the grounds for duration research. Chapter 1, on the basis of main prosodic structure theories, considers possible utterance constituents which will best serve the purpose of the study and which have been reported to influence speech timing. It also provides a review of particular factors determining the duration of prosodic units, referring to the domain-and-locus approach. Chapter 2 deals with the nature of prominence and its influence on the duration of speech units at various levels of the prosodic hierarchy. Chapter 3 is a review of language rhythm and timing studies, from

the birth of the idea of stress- and syllable-timed languages proposed by Pike (1945) to contemporary research and measures based on vocalic and consonantal length variability. It also discusses previous studies concerned with temporal relations in the English pronunciation of Polish learners and cross-linguistic interference tendencies that may cause non-native timing of Polish-accented English speech.

Chapter 4 begins the practical part of the book. It describes the empirical study of timing in native and non-native English speech that forms the core of the work, and the chapter includes the main assumptions, research method and procedures, and the hypotheses to be tested. Chapter 5 contains the results of the study at the segmental level, mainly concentrating on vowel duration, and proposes an interpretation of the obtained data, while Chapter 6 offers results and discussion concerning the timing of higher level units: syllables, feet and tone group constituents. Chapter 7, the conclusion, makes the work complete. It summarises the main observations, offers general conclusions, provides illustrations of typical differences between native English and Polish-accented read speech, and points out possible directions for further research and implications for pronunciation pedagogy.

CHAPTER 1

SPEECH UNITS AND THEIR DURATION

1.1 The structure of utterances for prosodic analysis

Timing is an important aspect of prosodic organisation of speech, strictly connected with intonation and rhythm (see Chapter 3). Because these aspects are interrelated, the description of utterance or phrase structures forming the object of analysis will naturally refer to the same units, even though the present study is not directly involved in intonation understood as pitch contours (e.g. 't Hart et al. 1990) or pitch targets and boundary tones (Pierrehumbert 1980). Therefore, in order to establish a prosodic structure and define prosodic domains best serving the purposes of the present dissertation, a brief review of prominent prosodic theories pertaining to intonation constitutes the first part of this chapter.

1.1.1 Higher-level utterance constituents

In line with most phonological approaches, we will regard the utterance as the highest-level domain, a relatively autonomous speech unit in terms of prosody, syntax and discourse, often corresponding to a sentence. The segmentation of utterances in the context of pronunciation teaching has traditionally been based on the Nuclear Tone Approach, associated with the British School (Kingdon 1958, Crystal 1969, O'Connor and Arnold 1961, 1973, Gimson 1980, Cruttenden 1986). This approach was indeed designed for foreign learners of English. The fundamental unit of an utterance has been called a sense group (O'Connor and Arnold 1961), tone-group (O'Connor and Arnold 1973) intonation-group (Cruttenden 1997) or a tone-unit (Crystal 1969, Halliday 1970). The central element of such a unit is the nucleus (Palmer 1922), defined as "the stressed syllable of the last prominent word in a sense group" (O'Connor and Arnold 1961: 271). The part of the utterance from

the first pitch-accented syllable to the nucleus is referred to as the head (Crystal 1969, O'Connor and Arnold 1961) and the accented syllables preceding the nucleus are known as prenuclear accents (Cruttenden 1997). Unaccented syllables preceding the head are called prehead (Crystal 1969, Halliday 1970, O'Connor and Arnold 1973, Cruttenden 1997). Finally, the nucleus is followed by an optional tail (O'Connor and Arnold 1961), also called nuclear tail by Crystal (1969). The tail, like the prehead, consists of unaccented syllables alone. The described parts of the tone unit are further divided into feet. Abercrombie (1967) defines the foot as a tone unit constituent comprising a stressed syllable and the following unstressed ones up to the next stressed syllable. According to this view, a word boundary does not automatically terminate the foot, which can thus become a unit larger than the word (a cross-word foot — cf. White 2002).

A similar model of prosodic structure, proposed by Jassem (1952), distinguishes narrow rhythm units (NRUs), which, combined with anacruses, constitute total rhythm units instead of the feet. The narrow rhythm unit is also initiated by a stressed syllable but, in contrast to the foot, it only includes syllables within the same lexical unit, which actually implies a claim that word boundaries in connected speech may influence the prosody. For instance, the phrases "summer dresses" and "some addresses" would receive identical description in Abercrombie's approach, whereas according to Jassem, the first syllable of "addresses" forms an anacrusis, a constituent outside the narrow rhythm unit (cf. Bouzon and Hirst 2004, Hirst and Bouzon 2005).[1]

Although the British School model of prosodic structure is still popular, a lot of recent research is based on the Autosegmental-Metrical (AM) approach, proposed by Pierrehumbert (1980) and developed by Beckman and Pierrehumbert (1986) and Ladd (1986, 1996). The AM theory differs from the British School not only in the approach to tones as discrete units, or pitch targets (H or L) rather than contours, but also in the approach to phrasing, which has been based on the studies by Selkirk (1980) and Nespor and Vogel (1986). These researchers postulated a hierarchy of prosodic units, distinguishing two basic levels of phrasing, viz. the full intonational phrase (IP) and the intermediate phrase (ip). The latter is further divided into prosodic words (or clitic groups) and feet (cf. Turk and White 1999). The prosodic domains, except for the lowest constituents, are made up of at least one constitutent of immediately lower level (Selkirk 1978, 1995b, Nespor and Vogel 1986, Hayes 1989). The slight differences between individual authors are presented in Table 1.1.

[1] The difference between Abercrombie's and Jassem's models, investigated by Bouzon and Hirst and closely related to other considerations of the role of word boundaries, will be taken into account in calculating the timing proportions in our empirical study, although insufficient amount of data will not allow to draw reliable conclusions in support of any of the two approaches.

Table 1.1. The prosodic hierarchy

Model / Level	Selkirk 1986	Nespor and Vogel 1986 (non-recursive)	Beckman and Pierrehumbert 1986	Hayes 1989 (non-recursive)	Jun 1998	Fougeron and Keating 1997	Keating et al. 2003
Higher	utterance	utterance	utterance	utterance	utterance	utterance	utterance
↑	IP	IP	IP	IP	IP	IP	IP
↑	—	—	intermediate phrase	accentual phrase	accentual phrase (for Korean and French)	intermediate phrase/ phonological phrase	smaller phrase (varies across languages)
↑	phonological phrase (major + minor) Selkirk and Tateishi 1988	phonological phrase (always corresponds with syntactic boundaries)	phonological phrase	—	—	—	—
↕	prosodic word	clitic group	prosodic word	clitic group	—	—	—
↓	foot	word	foot	foot	word	word	word
↓	syllable	syllable	syllable	—	syllable	syllable	syllable
Lower	—	—	mora	—	—	—	—

A hierarchic structure can display layeredness or recursivity. In the case of prosodic hierarchy, arguments are presented for both approaches. Strict Layer Hypothesis (SLH) postulates that each prosodic constituent is contained in a constituent of the adjacent higher level (Selkirk 1984). However, some linguistic evidence (pitch and boundary strength) questions SLH, pointing at the possibility of nesting constituents within constituents of the same rank in prosodic hierarchy, thus allowing recursive phrasing (Ladd 1986).

Prosodic domains are also to some extent related to syntactic constituents. Selkirk (2005), in her Syntactic Grounding Hypothesis, proposes the following correspondence:

Table 1.2. Syntactically determined prosodic constituents (Selkirk 2005)

Syntactic constituent	Prosodic constituent
comma phrase (CmmP)	intonational phrase (IP)
lexical maximal projection (lexP)	major phonological phrase (MaP)
branching syntactic constituent	minor phonological phrase (MiP)
lexical word (lex)	prosodic word (PWd)

However, although the correspondence between syntactic and prosodic constituents is fairly strong, syntactic boundaries are coded in prosodic structure in not more than 65—84% (Fach 1999). These syntactically determined constituents can be identified on the basis of segmental phonological rule application (Nespor and Vogel 1986, Jun 1993) and intonation (Beckman and Pierrehumbert 1986). An Intonational Phrase, for instance, bears a complete intonation contour. Because it often comprises a full sentence, it can be coextensive with the utterance. Nespor and Vogel (1986) found some differences in the application of certain phonological rules, but Wightman et al. (1992) observed no difference in final lengthening, a process depending on the domain's rank in prosodic hierarchy (to be discussed in section 1.3.1).

Differences in proposed prosodic hierarchy models are also caused by differences between languages. Certain languages might require specific phrasing, which must be reflected in the prosodic hierarchy. For instance, the Accentual Phrase — a unit between IP and Phonological Phrase is introduced for prosodic description of Korean and French (Jun 1998).

As mentioned before, apart from lexical and syntactic information, the phrasing of speech depends on stress and accent, which highlight important elements of the spoken message.[2] Both British School and AM theories recognise nuclear accent, but in Pierrehumbert's view, prenuclear accent

[2] The problems of prominence will be discussed in section 1.4.

("head" in British tradition) does not constitute a separate category, and the same accent inventory is applied in either case.

Autosegmental-Metrical theory has given rise to a specific transcription system called ToBI (Tone and Break Indices) (Silverman et al. 1992), where the prosodic structure and intonation pattern of a phrase is described by means of H (high) and L (low) tones and their combinations, and break indices referring to the degree of cohesiveness between words. ToBI has become a standard tool for describing the prosody of English, but also a wide variety of other languages (e.g. Jun 2005), which often require "national" modifications of the system.

Summing up the two relevant approaches, we receive a possible inventory of a number of speech units capable of forming a multilevel hierarchy. The hierarchic structure of utterances refers to the relatively independent and "complete" fundamental unit of intonational phrase, at the level of tone/sense/intonation group/unit. If an IP consists of more than one intermediate phrase (Beckman and Pierrehumbert 1986), then the tone group corresponds with the latter unit.

1.1.2 Relations between the word (lexical or functional) and the foot

Basically, the intonational phrase is segmented into feet in the British tradition and into words in the ToBI system. However, authors differ in their treatment of the two domains. The role of words in prosodic patterns depends on their category. The distinction between lexical and functional categories (or content and function words) is highly significant for the syntactic properties of sentences (e.g. Jackendoff 1977, Chomsky 1986), and, naturally, the two categories also display different phonological properties (e.g. Sweet 1891, Jones 1964, Gimson 1962, Selkirk 1972, 1984, 1986, 1999; Nespor and Vogel 1986 — see Selkirk 1995a for a review).

The phonological differences spring from the fact that function words in English may appear in stressless "weak" forms, while content words always have at least one stressed, unreduced syllable. Reduced function words are accommodated in the prosodic structure in different ways. A content word may be isomorphic with a syllable (e.g. "cat") or it may function as a single foot, (e.g. "honest"), but it may also consist of several within-word feet (Turk and White 1999), e.g. "anticipation." A weak form of a function word bears no stress that could form a foot, so it is combined with a content word (as in "talk to him") to form a cross-word foot (White 2002), stress foot (Kim and Cole 2005), prosodic word (e.g. Selkirk 1978), or clitic group (cf. Turk and White 1999: 171).

1.1.3 The status of lower level units — syllables and segments

Both words and feet are traditionally parsed into syllables. According to Tatham and Morton (2001: 192; referring to Gimson 1962), "[a] syllable is a phonological unit which forms the basis of the prosodic parameters of rhythm, stress and intonation — it is defined in terms of its hierarchically organised structure based on its segmental (consonantal and vocalic) composition. Syllables must have one vowel as their nucleus with margins where, in English, from zero to three consonants precede the nucleus and from zero to four consonants follow the nucleus."

The existence of consonant clusters in most languages, including English and Polish, raises the problem of syllable boundary location. Most researchers accept the Maximal Onset Rule (Pulgram 1970, Kahn 1976) which classifies consonants between vowels as syllable-initial if the phonotactic rules of the language are not violated. Maximal Onset Rule ensures consistent syllabification, but it does not solve all possible problems. For instance, the problem of syllable-final stressed short (lax) vowels, as in *manor* is either ignored (e.g. Halle and Vergnaud 1987), maintaining the consistency of the principle, or resolved by classifying the problematic consonant as the coda (e.g. Selkirk 1982). Originally, Pulgram and Kahn regarded such consonants ambisyllabic. The problem is discussed in more detail in Duanmu (2008).

Blevins (1995: 207) provides an acoustically motivated definition of the syllable, stating that "[e]ach sonority peak define[s] a unique syllable." This definition reflects language users' intuitive syllabification ability, including not only unit recognition but also boundary location (consequently at sonority troughs). However, acoustic evidence is not always unambiguous and does not always lead to universally accepted syllabic parsing either.

Syllables, as mentioned above, are composed of segments. A segment, regarded as a phonetic realisation of a phoneme, still remains a phonological entity, reflecting the contrastive elements of word structure, rather than physical reality, as it often comprises several distinct phases of articulation (e.g. occlusion, burst and VOT[3] in plosive articulation).

Many recent phonological approaches question the ontological status of the syllable (e.g. Dziubalska-Kołaczyk 2002), and segment as a realisation of phoneme, which itself became a controversial notion soon after its invention (e.g. Twaddell 1935) and still remains controversial (e.g. Port 2007). The notions in question indeed lack commonly accepted precise, uncontroversial definitions, but their intuitive reality for language users (and language learners) reflected in, for instance, the undeniable practical usefulness of phonemic transcription, which visualises the pronunciation of words as strings of discrete units and

[3] VOT — Voice Onset Time (see Lisker and Abramson 1964).

prominence relations, justifies recognition of the two units at least for pedagogical purposes. According to Ladefoged (2004: 8),

> [...] phonologists have the problem of deciding whether they are describing something that actually exists, or whether they are dealing with epiphenomena, constructs that are just the result of making a description. Phoneticians are seldom faced with this problem.

The development of acoustic phonetics has offered fairly objective methods of speech chain segmentation (e.g. Peterson and Lehiste 1960, Turk et al. 2006) and even if not all segment transitions allow reliable, unambiguous interpretation, there are a number of acoustic cues to segmental boundaries, which also indicate the edges of higher-rank prosodic domains. Their practical application for the purposes of our empirical study will be discussed in Chapter 5.

Another interpretation of speech production handling the problematic segmental transitions and overlapping is proposed by Articulatory Phonology. This approach decomposes utterances into gestures — dynamic units of articulators' constriction action (Browman and Goldstein 1986, 1989, 1992). The organs of speech (the lips, the tip, dorsum and back of the tongue, velum, larynx) are independently controllable and traditional segments and syllables are composed of atomic gestures forming gestural molecules, i.e. "temporally coordinated [...] assemblies of gestures" (Goldstein et al. 2007: 387, after Byrd 1996, Saltzman and Byrd 1999, Saltzman and Munhall 1989). By introducing dynamic speech units, the theory offers an interesting solution to the problem of speech segmentation, as well as such prosodic phenomena as final lengthening (cf. section 1.3.1), which is accounted for in terms of temporal articulatory π-gestures (Byrd and Saltzman 2003, Byrd et al. 2000, 2006).

As mentioned above, the pedagogical robustness of segments and syllables, deeply rooted in glottodidactics, makes them suitable candidates for the basic units of utterances for the purposes of this study. However, in order to overcome inevitable problems connected with those traditional notions, we will make use of selected acoustically salient articulatory gestures as landmarks for phonologically controversial but consistent identification of segments, syllables and consequently, feet/words[4] and intonational phrases. The details are discussed in the research description (Chapter 4).

[4] If cross-word feet are allowed, then in actual sentences a word may be smaller or larger than the corresponding foot.

1.1.4 Conclusion

The prosodic structure of utterances encompasses a hierarchy of domains the relations between which may be reflected on the temporal plane. When we take the above considerations into account, the empirical investigation of phrases selected for the study may require, depending on individual examples, considering the following relevant domains:
— intonational phrase
— phonological phrase
— prosodic word
— foot/word
— syllable
— segment.

Additionally, subsegmental temporal parameters (c.f. Waniek-Klimczak 2005) will be taken into account if they may cause speech timing differences between Polish and English speakers.

Furthermore, with respect to the prosodic hierarchy key notion of headedness, although we assume the idea of left-headed domains, in preheads (anacruses), analysed on the foot level, no prominent element is distinguished. Finally, the structure of tested fragments of speech used in our empirical study does not require judgements concerning constituent recursivity.

1.2 Segmental length and its determinants

The duration of complex speech units obviously depends on the duration of their constituents. The duration of units described in section 1.1 depends on a number of factors, some of them universal, others — language specific. This section points out the main determinants of speech unit duration focussing on those which show different influence on timing in English and Polish. These factors will be considered in accounting for the timing problems of Polish learners of English demonstrated by the results of the empirical study. Universal factors will remain outside the focus of this study as they are irrelevant in FL acquisition process.

1.2.1 Intrinsic segmental length

The length of individual segments depends on their intrinsic characteristics and extrinsic factors, operating on higher levels of the utterance structure. The latter will be discussed in the subsequent sections of this chapter, devoted to the higher-level units.

1.2.1.1 Universal segmental duration determinants

Among the universal principles, connected with the anatomy and physiology of speech, Waniek-Klimczak (2005) mentions the following:

1. Tongue height in vowel articulation (referring to studies by Lehiste 1970, cf. also House and Fairbanks 1953, Peterson and Lehiste 1960, Delattre 1962, Elert 1964) is inversely proportional to vowel length.
2. The place of articulation of consonants is an important duration determinant: Labial consonants are longer than alveolars and velars (Waniek-Klimczak 2005: 25, after Lehiste 1970, Klatt 1976, Maddieson 1997, Luce and Charles-Luce 1985), while VOT duration is proportional to the degree of backness (Lisker and Abramson 1964, Cho and Ladefoged 1999), which suggests duration balance, e.g. between /p/ (longer closure, shorter VOT) and /k/ (shorter closure, longer VOT — cf. Docherty 1992).
3. Voiced consonants are shorter than voiceless consonants (Chen 1970, Klatt 1976, Port 1979, Lisker 1986), also in languages which do not exploit tense/lax opposition, e.g. Polish (Keating 1985).

The above relations are claimed to be universal and as such they do not influence segmental length in the pronunciation of native speakers and FL learners in dissimilar ways. Still, different L1 backgrounds may result in duration discrepancies between the two groups as a consequence of the following interference processes:

1. Non-native vowel quality may lead to additional uncontrollable length variation.
2. VOT is generally shorter in Polish and, arguably, phonologically less relevant than in English, which may affect the length of prosodic units unless longer VOT is offset by vowel shortening.
3. The regularity comprises contexts where voicing contrast is neutralised in Polish, hence voicing inconsistencies may also affect timing relations within speech units.

Therefore, despite acknowledging the universal character of the intrinsic temporal characteristics of segments, we must allow for other L1-related processes, which make the conditions for the application of the above rules different in the performance of Polish and English speakers.

1.2.1.2 Language-specific intrinsic segmental length

English, unlike Polish, is a language that uses vowel quantity as a phonologically relevant feature. Although all English vowel contrasts are

reflected in the sound quality, most classifications mention intrinsically long and short vowels and even the most popular transcription systems used for pedagogical purposes mark that feature explicitly (e.g. Wells 2008, Hornby 2000). Moreover, studies have been carried out to establish regular duration relations between short and long vowels by Wiik (1965) (quoted by Cruttenden 2001: 96), Gonet (1997), Peterson and Lehiste (1960), Lehiste (1970). Waniek-Klimczak (2005: 25), referring to the latter two sources, mentions a mean of 50% difference in length between the two vocalic classes. This difference can be increased to 100% or even more if the shortest and longest vowels are compared (e.g. Umeda 1975, Van Santen 1992). This typological discrepancy between English and Polish must be treated as a potential source of non-native timing in the learners' English pronunciation.

Although consonant quantity contrasts are not used in English, studies of speech timing must take into account the length of consonants as well. Jassem et al. (1984) measured vocalic and consonantal segments and their duration variability in recorded study units of *A Course in Spoken English: Intonation* (Halliday 1970) and distinguished 18 phone classes, then grouped into 6 duration categories based on the principle of minimum within-group variation and maximum between group variation (Table 1.3).

Table 1.3. The duration of phone classes (Jassem et al. 1984)

Rank	Class	Mean (ms)	FOOT	NRU	ANA	FOOT (rel)	ANA (rel)
1.	F	16.6	16.8	18.5	14.5	.907	.785
2.	D	47.6	48.4	49.8	45.2	.972	.907
3.	G	56.7	56.2	57.9	43.2	.970	.755
4.	E	59.2	59.8	66	50.1	.907	.759
5.	B	60.2	60.4	60.5	59.6	.999	.985
6.	N	61.8	62.2	62.9	60.5	.989	.962
7.	H	64.9	65.8	69.3	35.6	.948	.514
8.	K	65.3	66.1	68.5	55.4	.965	.809
9.	Z	70.3	73.2	77.2	55.4	.949	.718
10.	SC	76.2	75.9	77.6	65	.978	.838
11.	KH	85.5	86.3	88.5	77.4	.975	.874
12.	S	87.9	90.4	93.4	65.5	.967	.701
13.	AFV	93.9	93.8	95.3	80.6	.985	.846
14.	O	96.3	97.3	100	78.8	.973	.788
15.	KHA	117.2	118.3	118	—	1	
16.	AF	126.9	127.7	127.7	—	1	
17.	A	132.1	134.2	139.9	100.2	.959	.717
18.	FTH	137.5	137.5	136.5	—	1.01	—

Abbreviations: F — flaps, lenis stops /ɾ/, /b/, /d/, /g/; D — weak-friction lenis fricatives /ð/, /v/; G — non-syllabic vocoids /w/, /j/, /r/; E — checked non-open vowels /e/, /ɪ/, /ʊ/, /ʌ/; B — non-initial lenis stops /b/, /d/, /g/; N — non-syllabic nasals /m/, /n/; H — the aspirate and initial voiceless fortis aspirated stops /pʰ/, /tʰ/, /kʰ/; K — fortis unaspirated stops /p/, /t/, /k/; Z — heavy-friction lenis fricatives /z/, /ʒ/; SC — syllabic contoids /l/, /n/, /m/; KH — aspirated unaccented fortis stops /pʰ/, /tʰ/, /kʰ/; S — fortis fricatives /f/, /θ/, /s/, /ʃ/; AFV — lenis affricates /dʒ/, /dr/; O — close unchecked and the open checked vowels /i:/, /u:/, /æ(:)/, /ɒ/; KHA — accented fortis stops /pʰ/, /tʰ/, /kʰ/; AF — fortis affricates /tʃ/, /tr/; A — mid and open unchecked monophthongs /ɜ:/, /ɑ:/, /ɔ:/ and diphthongs; FTH — aspirated final fortis stops /pʰ/, /tʰ/, /kʰ/.

The study shows considerable differences in duration means and variability depending on the type of a segment. Obviously in many cases the duration of particular phones might have been influenced by their phonetic environment since often particular positional allophones of one phoneme are classified in different groups. The general influence of phonetic context on segmental duration is discussed in the next section.

The results obtained by Jassem et al. (1984) will later be considered in discussing the results of the present research.

1.2.2 Phonetic context as a segmental duration determinant

The intrinsic duration properties of speech sounds are modified by a number of contextual factors addressed in this section, including the phonetic characteristics of adjacent segments and the position of the segment within utterance structure. Another important factor, prominence, is discussed separately in Chapter 2.

One important segmental factor influencing speech timing is consonant elision. It may strongly influence syllable duration because a whole segment is deleted. This is not always the case, though, because elision may invoke "compensatory" lengthening of the adjacent segments. The process is more widespread in unstressed positions (to be discussed in Chapter 2 as well), but Cruttenden (2001: 237) mentions a lot of examples of consonant (especially alveolar stop) elision typical mainly of colloquial speech, but independent of stress. These include segments occurring between other consonants, e.g. "facts," "exactly," "handsome," also before a word boundary, e.g. "next day," "old man," "looked fine," and in negations, e.g. "mustn't lose."

Otherwise, little significant length modification has been reported in consonants, although Lisker and Abramson (1967), Flege et al. (1998) report evidence for longer VOT duration before high vowels, a regularity that is physiologically (cf. Waniek-Klimczak 2005: 27) rather than phonologically motivated.

More contextual variability is visible in vocalic segments. Whereas the influence of the preceding consonant on vowel duration is considered negligible (Peterson and Lehiste 1960), the consonant that follows is significant. The fact that English vowels are longer before voiced than voiceless consonants has been documented in numerous publications, e.g. House and Fairbanks (1953), House (1961), Chen (1970), Klatt (1973), Lisker (1974). A difference of 50% was reported by Peterson and Lehiste (1960).

This regularity was also found in other languages (Chen 1970), even those which neutralise word-final voicing contrast, including Polish. Slowiaczek and Dinnsen (1985) observed a 10% vowel lengthening before underlyingly

voiceless consonants. However, Jassem and Richter (1989) only found a mean 4ms difference in vowel length. Likewise, House (1961) and Keating (1985) did not observe that process in Spanish, Czech or Polish. Considering the research by Delattre (1962), Chen (1970), Mack (1982), and Laeufer (1992), who argued for language-specific range of pre-voice lengthening, and studies of the acquisition of vowel duration by children (e.g. Krause 1982, Ko 2007), we may consider vowel length effect a universal tendency, manifested to varying degrees cross-linguistically.

The fact that English actually uses vowel length variation as the main cue for coda phonological voicing contrast perception (Raphael 1972, Gimson 1974, Hogan and Rozsypal 1980), especially word-finally (Docherty 1992: 120), while Polish shows little variation in that respect, may become a source of potential timing problems that Polish learners of English face.

1.2.3 The effect of syllable structure

The structure of syllable, the smallest prosodic unit, comprising an optional onset and an obligatory rhyme (Fudge 1969, Blevins 1995), can influence the duration of its segmental constituents. In open syllables, ones with no coda, vowels are reported to be longer (Maddieson 1985, Nowak 2006). This tendency has been observed in many languages, but it is not obvious in Polish. Jassem (1962) found no vowel lengthening in open syllables, unlike Klessa and Śledziński (2007), who reported considerable lengthening, magnified in prepausal positions.

Syllable nucleus duration is strongly modified by the type of coda, as indicated in the previous section, but also by the onset. The main cue to syllable onset voicing, VOT, is enhanced by length modification of the vowel. The reduced vowel portion is shorter than VOT (Allen and Miller 1999), which indicates that the process does not only involve vowel devoicing without changing length proportions. A similar compensatory process adjusts the temporal relations between the nucleus and coda (e.g. Jassem 1971, Waniek-Klimczak 2005). According to Cunningham (2008: 3), "[i]n English, for example, the durations of vowels and of the stops which follow them have an inverse durational relationship." If this statement, besides for the fortis/lenis contrast, also holds true for intrinsic vowel length effects, it broadens the scope of Lehiste's (1977) claim that longer duration of fortis consonants and their shortening effect on the preceding vowel reflects a tendency for syllable rhyme isochrony.

Vowel duration differences before fortis and lenis codas (cf. section 1.2.2) are most significant in stressed syllables and nonexistent or negligible in unstressed ones (de Jong 1991). With respect to speaking style, Crystal and

House (1988c) found more evident length differences depending on coda voicing in clear speech in comparison to normal speech.

The structure of higher level units influences the duration of its constituents in a more complex way. The lengthening or shortening processes seldom affect single segments, and their influence may exceed even higher-level unit boundaries. This problem will be discussed in section 1.3.

1.3 Temporal relations within and across utterance units

Segmental duration obviously determines the duration of higher-rank units, which was discussed in the previous section. Thus the main duration determinant of a basic prosodic domain, i.e. syllable, is its complexity, which can be expressed as the number of constituent elements, i.e. phonemes, and their articulatory complexity (place and manner of articulation). Consequently, the duration of a foot, a larger domain, depends on the number and structure of constituent syllables, and in the same way higher-level units depend on the structure of their immediate constituents. This section handles the temporal relations within and across suprasegmental units, which are best reflected through a domain and locus approach (White 2002).

Apart from intrinsic segmental properties and immediate phonetic context, the length of utterances and their parts is determined by a number of factors which affect whole prosodic domains and temporal relations between them. The review of these factors will be divided into sections corresponding to individual prosodic hierarchy levels. The length modification of complex units requires a specification of the span of the change within this unit, i.e. the locus of a given shortening or lengthening process.

According to White (2002), there are two types of processes with regard to their locus (scope of operation):

— **domain-span** processes, influencing all elements of the unit, and
— **domain-edge** processes, affecting only the segments near the unit boundaries.

The present section is a review of such processes operating at various levels of the prosodic hierarchy reported in literature.

1.3.1 Universal domain-edge processes

A number of lengthening and shortening processes operate at the edges of all prosodic domains. Often the length modification magnitude is proportional to the domain's rank in prosodic hierarchy, while other processes influencing the duration of speech units are claimed to operate in a similar way at the edges of all recognised levels of prosodic hierarchy. Final lengthening and

initial strengthening/lengthening belong to the former category since they are reported to exert a stronger influence at higher-level domain edges (Cooper and Paccia-Cooper 1980, Gee and Grosjean 1983, Wightman et al. 1992, Fougeron and Keating 1997, Yoon et al. 2007).

Final lengthening is one of the best documented processes affecting the duration of speech units. It affects preboundary segments, especially vowels, and has been evidenced for English by Klatt (1975, 1976), Lehiste (1972), Martin (1970), Oller (1973), Edwards et al. (1991), Gussenhoven and Rietveld (1992), Byrd (2000), Byrd et al. (2006); Dutch by Nooteboom and Doodeman (1980), and Cambier-Langeveld (2000); and Hebrew by Berkovits (1994). Preboundary segments display the strongest lengthening effect in the phrase-final, and especially utterance-final position (e.g. Lindblom and Rapp 1973). However, Klatt (1976) claims that lengthening only affects sonorant and fricative codas. Wightman et al. (1992), and Gussenhoven and Rietveld (1992) have observed longer duration of preboundary syllables. Generally, before lower-level unit boundaries the effect is less systematic (e.g. Harris and Umeda 1974).

Although final lengthening is a well-documented process, it is not easy to precisely establish the locus and relative magnitude of the effects. The most important observations are based on studies by Berkovits (1994) and Shattuck-Hufnagel and Turk (1998), who examined lengthening in phrase-final words and found that lengthening of segments regularly increases towards the phrase boundary and the effect starts with the stressed syllable. This means that if the last syllable is stressed, the penult is not affected. If the penultimate syllable is stressed, then the lengthening begins there but the magnitude still increases in final segments. Therefore it may be claimed that final lengthening begins with the last stressed syllable (or vowel — cf. Wightman et al 1992) and "continues to the phrase break." (White and Mattys 2007a: 514).

Initial strengthening is manifested in wider linguopalatal contact in plosives and nasals (stops) and longer contact duration. Research by Fougeron and Keating (1997), Byrd and Saltzman (1998), Fougeron (2001), Cho and Keating (2001), Keating et al. (2003) proved the existence of the process in English, French, Korean and Taiwanese. Moreover, initial strengthening is magnified at boundaries of higher-rank domains. Still, the effect is not detectable for all speakers if adjacent rank domains are compared. The span of initial strengthening is not clearly defined but, according to Kuzla (2009: 19, quoting Bombien et al. 2006), "initial strengthening effects are strongest in the segments immediately following the prosodic boundary and decay rapidly in subsequent segments." Likewise, Cho and Keating (2007) state that initial strengthening affects mainly consonants, while vowels are more sensitive to stress and accent. Oller (1973) also observed that the magnitude of initial lengthening might depend on the type of onset consonants.

Certain observed strengthening processes refer to prosodically strong positions rather than domain edges. Roughly speaking, these positions can still be defined as domain-initial, although certain domains are not counted. Inkelas and Rose (2008) mention word-initial and stressed-syllable-initial positions as strong, which automatically embraces all domain-initial positions except non-foot-initial syllables. Naturally, whenever a position is defined as stressed-syllable-initial, it is equivalent to foot-initial. Prosodically strong positions may be language-specific and some accounts may regard the word-initial position as weak (e.g. Steriade 2001).

Prosodically strong positions are used, for instance, in formulating the Uniform Strengthening Account (Cole et al. 2007), which predicts that all plosives are more fortis-like in such contexts in that both their occlusion phase and VOT are longer (Pierrehumbert and Talkin 1992, Cho and Keating 2007), but see Choi (2003) and Cole et al. (2007), whose research does not confirm this prosodic position effect.

1.3.2 Domain-specific processes

Most processes influencing the duration of utterance constituents are reported to operate at many levels of the prosodic hierarchy. Those associated with specific domains are discussed below with reference to the relevant constituents.

1.3.2.1 The foot

The foot is a unit whose definition depends on the notion of stress, or relative prominence relations between its components (see Chapter 2). The length of a foot naturally depends on the number and complexity of constituent syllables. Faure et al. (1980) claim that the duration of individual syllables does not depend on how many constitute a foot, but other studies suggest that the mean syllable duration is inversely proportional to the foot complexity expressed in the number of constituent syllables, e.g. Huggins (1972) and Fowler (1977). The same conclusion was drawn by Campbell (1992) and Eriksson (1991 — for Swedish), who observed that both stressed and unstressed syllables are shorter in longer feet.

Rakerd et al. (1987) described **foot-level shortening**, which predicts shortening of a stressed syllable, proportional to the number of unstressed syllables following it. This is in line with the compression hypothesis advocated by as early studies as Jones (1918), confirmed, for instance, by Nooteboom (1972), Fowler (1981) and Hoequist (1983). Kim and Cole (2005) also observed

shorter duration of stressed syllables in more complex feet in read American English speech but only within the intermediate phrase. A similar observation was made by Van Lancker et al. (1988), who stated that a stressed syllable tends to be shorter before an unstressed syllable than before another stressed one. Certainly, the same kind of syllable duration variability underlies Abercrombie's (1967) definition of stress-timed languages.

1.3.2.2 The word

Word boundaries may lengthen their peripheral segments both initially and finally. Oller (1973) and Cooper (1991) point out that word-initial syllable onset consonants are longer than corresponding word-medial ones. A syllable rhyme may be longer word-finally, even in non-phrase-final positions (e.g. Beckman and Edwards 1990, Wightman et al. 1992). Klatt (1975) observes that in this type of environment whole word-final syllables tend to be longer than initial and medial ones.

Apart from these domain-edge processes, **polysyllabic shortening** has been reported by Lehiste (1972), Port (1981), Turk and Shattuck-Hufnagel (2000), and White (2002), who have suggested that the primary stress syllable is shortened in longer lexical words such as /meɪs/ in "mace," "mason" and "masonry." Port (1981) investigated /i:/ and /ɪ/ in two similar syllables: "d__p" and "d__b" which allowed him to observe vowel duration before fortis and lenis codas. The two high vowels appeared in monosyllabic words and with "-er" and "-erly" attached to the tested structure, to form two- and three-syllable long words. In each case the vowel was significantly longer in monosyllables, while the difference in disyllabic and trisyllabic words was much smaller. Moreover, both vowels before /b/ were shortened by over 30ms with the suffixes added, whereas the corresponding difference before /p/ was approximately 50% smaller. This suggests that the influence of the phonetic context on vowel duration may be stronger than the effects of polysyllabic shortening.

Polysyllabic shortening is hard to distinguish from foot-level shortening because in English feet and words are often coextensive. The relative nature of duration and interaction of factors influencing unit duration makes it possible to question polysyllabic shortening and interpret the length differences purely in terms of phrase-final lengthening (cf. Nakatani et al. 1981).

Word duration also depends on the speaker's estimated recognition facility connected with its frequency of occurrence and previous mention. Differences in lexical probability influencing word duration were suggested by Lindblom (1990), Jurafsky et al. (2001), Bell et al. (2002), Aylett and Turk (2004, 2006). Bybee (2000) demonstrated that English word final stops are deleted or unreleased more often in more frequent words. Baker and Bradlow (2009)

showed that more probable words are considerably shorter in clear and plain speech styles. A shortening in the second mention of a word also occurred under controlled accent conditions suggesting independent word probability effect. This phenomenon may partially account for significantly longer duration of English words and phrases pronounced by foreign learners, and high rhythmic irregularity in FL performance especially at lower proficiency levels in read speech, where the reader is often faced with less familiar lexical items.

Although a large number of studies point to acoustic cues to prosodic boundaries, including word boundaries, perceptual segmentation is not easy for foreign learners if the most salient boundary cue — a pause — is absent. In natural speech words are not normally separated by pauses, which causes serious comprehension problems in FL learners. This may suggest that boundary cues other than pauses are either subtle or language-specific, or must be supported by other cues in order to work effectively.

1.3.2.3 The intonational phrase and the utterance

The intonational phrase, the largest utterance constituent, is often delimited by pauses and displays a complete intonation contour or, in terms of the AM framework, includes at least one nuclear tone (H* or L*) and a boundary tone (L% or H%), followed by break index 4 in the ToBI system. At this level, the domain-edge lengthening processes described in section 1.3.1 occur with relatively greater magnitude. Naturally, the strongest lengthening effect occurs at the utterance boundary, where it can be spread over several syllables (Klatt 1976, Cummins 1999).

1.3.3 The pauses

A pause is the most salient cue to speech chain segmentation. However, as mentioned before, it is typically used to delimit higher level units such as utterances and intonational phrases.

Both pause occurrence and pause duration is determined by the speaker, speech rate, discourse, prosodic structure, phrase length, and syntax (Krivokapić 2007). Pause duration increases at longer and more complex syntactic units (but varied results have been found with respect to phrases which precede or follow a pause). Krivokapić (2007) also observes the effect of discourse organisation, i.e., referring to Smith (2004), longer pauses at topic shift than other discourse boundaries. In complex sentences, pauses are more likely to appear if they follow long subjects (Grosjean et al. 1979, Cooper and Paccia-Cooper 1980, Ferreira 1991, Selkirk 1995b, Strangert 1997).

Generally, greater syntactic complexity and length increase the duration of pauses surrounding the structures in question, although the strength of this influence varies across studies (cf. Krivokapić 2007).

1.4 IP-level duration determinants

Apart from segmental duration and domain structure, there are also other factors that influence speech timing of larger domains. Besides prominence patterns, they include largely speaker-specific factors such as speech rate, and, less individually biased, speaking style and intonation.

1.4.1 Intonation

The relative independence of intonation, stressed by autosegmental phonology, is also manifested in the relations between the contours or pitch targets and utterance units which carry them. The realisation of intonation contours may be incomplete, truncated (cf. Grabe et al. 2000) or compressed (completed at faster speed — cf. Grønnum 1991).

On the other hand, the necessity to realise an intended pitch excursion may cause lengthening of the relevant units (e.g. Beckman and Edwards 1992). This phenomenon can be regarded as one aspect of accentual lengthening because pitch excursions are strongly correlated with accent. However, studies by Ohala and Ewan (1973), Sundberg (1979) and Xu and Sun (2002) showed that even the direction of pitch change alone can be a duration determinant.

1.4.2 Speech rate

The duration of speech units naturally depends on speech rate. Slower speech will render particular utterance constituents longer. However, a proportional duration change relative to overall speech rate is not obvious (e.g. Turk and White 1999). Beckman and Edwards (1990) claim that differences related to lexical and morphological structure are magnified in slower speech. At higher speed, because phones and larger constituents are not equally compressible, timing proportions between them may also be modified (e.g. Dellwo 2008).

1.4.3 Speaking style

Temporal organisation of speech may be altered not only with different speech rates but also with different speaking styles. One example of possible changing duration proportions is the study of clear speech by Smiljanić and Bradlow (2008), who demonstrated the tendency of English and Croatian speakers to enhance their respective languages' temporal contrasts in terms of absolute durations while keeping the general timing ratios. However, a cooperation of two lengthening factors, e.g. clear speech and final position may not be additive, which was observed by the researchers in weaker final-lengthening effect in clear speech, inevitably changing the timing proportions.

Reading a story, the type of language performance used in the present study, will normally be associated with clear speech at a moderate rate. It is worth noting that FL speech is generally slower and clearer than conversational performance, also in the sense of intentional completeness of articulatory gestures. Therefore, native read speech provides a reasonable target for pronunciation learning before the learners' articulatory proficiency allows them to seek more advanced, faster speech performance models.

1.5 Conclusion

According to White (2002: 2), "[m]uch of the systematic variation in English speech timing can be described either in terms of constituent structure or in terms of the distribution of prominences." This chapter has handled the former aspect of timing by summarising the main views of prosodic organisation of speech in terms of prosodic domains and main factors influencing the timing of utterances. It has been established that the duration of particular speech units is a result of intrinsic duration of their constituents depending on the articulatory complexity of segments and the effects of unit structure and position within larger domains. The other aspect of timing, prominence distribution, is discussed in Chapter 2.

CHAPTER 2

PROMINENCE AS THE MAIN SPEECH TIMING FACTOR

2.1 Introduction

At all levels of prosodic hierarchy, one of the constituents is more prominent than the others. Such a constituent is called the head. Moreover, each head is aligned to all the heads at successive lower levels of the hierarchy (Halle and Vergnaud 1987, Hayes 1995).

Among the factors motivating the "highlighting" of selected elements, two seem especially important. First and foremost, the speaker marks a word as prominent if it is pragmatically more significant than the surrounding words (e.g. Jones 1918). The other factor is the omnipresence of rhythm in human existence, where language is no exception. Auer et al. (1999: 3) emphasise the significance of rhythm in the following way:

> Rhythm is one of the most pervasive aspects of the human condition; it is in the world around us and in the world within us, in our bodies and our minds, our living and our thinking. [...] human language, quite predictably, is deeply rhythmic as well. [...]
>
> Rhythms — patterned recurrences of events in time — form a large part of our lives. The alternation of night and day, of the seasons, of the phases of the moon, determine fundamentally the world we live in, and indeed the lives of most organisms in this world: the rhythms of the body (our heartbeat, breathing, hormonal cycles) as well as its basic needs (hunger, thirst, sleep) shape our experience of ourselves. The most automatic, basic human movement patterns, walking or running — and even the acquired and more complex movements of swimming, rowing or skiing — imply rhythmicity. Rhythmic movements provide comfort to the newborn baby and self-induced stimulation to the autistic child. Our minds and our perception rely on

> rhythmic patterns to such a degree that physically identical series of stimuli are quasi automatically heard or seen as rhythmically grouped, and that information is better understood and recalled when it comes in rhythmic chunks. The "process of civilization" (N. Elias) is in essential ways a process of imposing sociocultural rhythms on our lives, from the strict timing of life under the Benedictine rule up to modern working hours. And some of the most genuine and widespread cultural achievements of mankind (music, dance, poetry) are inextricably linked to rhythm. It seems in a way that the old *variatio delectat* is even more true when it is understood as *alternatio delectat*.

Similarly, Dziubalska-Kołaczyk (2002: 82) observes, referring to Allen (1975) and Fijałkowska (2000), that "a preference for isochrony and for the rhythmic structuring of a sequence in general is rooted in universal principles of human perceptual and motor behaviour." Such a preference leads to the organising of speech units in sequences of alternating stronger and weaker elements. This pattern is displayed even at the level of segments, where the salient, sonorant vocalic intervals alternate with less salient consonants. If a theory accepts the syllable as a prosodic unit, this is the unit that bears prominence at the word level as well as higher prosodic levels. Naturally, within the syllable, it is the vowel that provides the listener with the most important information concerning the prominence status, but the articulation of nucleus and coda is also strengthened in prominent syllables. Moreover, as discussed in the following sections, certain prominence markers actually exceed syllable boundaries.

2.2 Acoustic correlates of prominence

Early studies referring to syllable prominence understood as stress realisation defined it in terms of the force with which a syllable is articulated (Sweet 1906, Bloomfield 1933, Jones 1949) corresponding with the degree of loudness perceived by the hearer (Bloch and Trager 1942). Fry (1958) points out that perceptual salience of syllables is achieved by modifying one or (normally) a combination of the following features:

— the length of syllables
— the loudness of syllables
— the pitch of syllables
— the sound qualities occurring in the syllables
— the kinaesthetic memories associated with one's own production of perceived syllables.

Fry (1958: 127) says of these features that "the listener is never concerned exclusively with one of them" but rather "his linguistic judgments are determined by their interaction." The results of the experiments described therein, devoted to the role of duration, intensity and pitch in the perception of word stress (as well as Fry 1955) suggest that duration cues are more indicative of stress than intensity, and pitch changes can override the effects of the former two. Bolinger (1958), Morton and Jassem (1965), Jassem et al. (1968) also found pitch variation to be the most important cue. Streeter's (1978) empirical studies of stress perception by adult English speakers yielded similar conclusions.

Apart from these three parameters, traditionally recognised as prominence cues (e.g. Lieberman 1960), Fry (1965) investigated the influence of vowel quality (F1-F2 structure) on stress perception. This influence was found to be smaller than that of duration and intensity, although Fry always insisted on the crucial role of interaction between individual features.

References to vowel quality as a prominence cue are also present in studies on hyperspeech (Lindblom 1990) or hyperarticulation carried out by de Jong (1995) and Erickson (2002), who found that under accent the phonetic space of phonemic contrast expands, and vowels become more peripheral. A similar conclusion, called sonority expansion, was presented by Beckman et al. (1992), who claimed that under accent articulators move to increase sonority (the vocal tract becomes more open). These observations were supported by Cho (2005), who confirmed that the two processes are combined. However, the acoustic studies of the 1950s and 1960s yielded results which could be roughly illustrated by a hierarchy of prominence cues in Table 2.1 below. The table displays both acoustic cues and corresponding perceptual cues (cf. Jensen 2004).

Table 2.1. A hierarchy of perceptual and acoustic prominence cues[1]

Acoustic cues	Perceptual cues
fundamental frequency (F0)	pitch
duration	length
intensity	loudness
formant structure	quality

The views regarding pitch as the main cue to prominence have lately been challenged, again favouring intensity as the main prominence correlate. In a recent study, Kochanski et al. (2005) argue that intensity is a better predictor

[1] Besides the traditionally acknowledged features, more recent research has also highlighted the role of sub-band intensities and spectral tilt (e.g. Sluijter and van Heuven 1996).

of prominence than duration, while F0 is a very weak one unless the speaker produces large pitch excursions, concluding that pitch movements and prominence are largely independent. They point out that these results do not stand in opposition to synthesis-based experiments by Gussenhoven et al. (1997), Rietveld and Gussenhoven (1985), and Terken (1991), where larger, 1/2 octave F0 excursions induced perception of prominence.

Generally, discrepancies between various studies have been attributed to possibly different experimental conditions, especially in cases where synthetic speech and parameter manipulation were used. For instance, Beckman (1986: 157) claims that

> [t]he mapping between the physical attributes being manipulated and the psychological attributes supposedly being tested is too complicated for direct comparisons of the results. For example, it is possible that the fundamental frequency ranges tested in some of these experiments correspond to larger pitch ranges than do the others. Such differences in pitch range could easily be produced even in experiments that use the same F0 range if there are uncontrolled differences in amplitude slope patterns or F0 movement patterns.

She also observes that if certain parameters strongly interact, treating them separately may produce misleading results:

> An especially important fact to note in this regard is that all of the experiments on stress perception treat duration and amplitude as separate attributes. None of the experiments vary the intensity integral as a single physical parameter, even though automatic stress-detecting algorithms have found this to be a better cue for stress than either duration or intensity alone. (Beckman 1986: 157)[2]

Finally, a difficulty may not only lie in establishing relations between the physical cues but also in the subjective psychological reaction of the listener. According to Cooper (1998: 24), "the perception of stress in speech is, like the perception of rhythm generally, a subjective and interpretive activity rather than simply the registering of some objective feature."

The acoustic correlates of prominence may work in different ways at different levels of phonological structure. Therefore it is useful to introduce the distinction between the notions of stress and accent, representing two different kinds of prominence.

[2] See also Turk and Sawusch (1993).

2.3 Types of prominence Stress vs. accent

Prominence may be manifested in various ways depending on the prosodic hierarchy plane. Most often a distinction is made between the word level and higher prosodic levels (e.g. phrasal level).

The lexical aspect of prominence consists in the fact that every word has one syllable referred to as the stressed syllable, which is inherently stronger than the others. This inherent property is exhibited in speech production by means of the cues mentioned in the previous section. On the lexical level, it is often connected with increased articulatory effort or energy expenditure (Abercrombie 1967, Ladefoged 1975). On higher levels (phonological phrase, intonational phrase) each stressed syllable has a potential for accent (cf. Beckman and Edwards 1994), realised in pragmatically important words.

The influence of individual acoustic correlates of stress and accent on syllable salience may differ. For instance, Wells (1990: 683) states, in line with the research discussed in section 1.4.1, that a stressed syllable is "marked by greater loudness than unstressed syllables, and often by pitch prominence, or greater duration, or more clearly defined vowel qualities." On the level of IP, prominent (accented) syllables carry a distinct pitch movement. The notion of accent with reference to pitch changes was introduced by Bolinger (1958) in order to indicate the crucial role of F0 in signalling prominence.

In the present paper, "stress" will indicate lexical prominence, and "accent" will refer to pitch accented syllables introducing the head and the nucleus in the classical framework of the British School prosodic structure (cf. section 1.1.1)[3].

2.4 Focus and prominence

As indicated in the previous section, pragmatically important words or phrases are more salient in an utterance. This kind of prominence, depending on both the syntactic structures of utterances and the pragmatic context, is called focus. Focus in English is connected with pitch accent (e.g. Gussenhoven 2008: 83), which is in turn strongly correlated with longer duration of speech units, but also overall and sub-band intensity (Mo, 2008a, 2008b) and spectral tilt (Sluijter and van Heuven, 1996a, 1996b).

The role of pitch in signalling focus is generally recognised, but other factors, including syntactically and pragmatically motivated anticipation of the hearer, e.g. syntactic category information (Cole, Mo and Baek 2008) or

[3] Cf. also Cho and Keating (2009).

word repetition and frequency (Gahl 2006, cited in Nowak 2006: 46; Cole, Mo and Hasegawa-Johnson 2008) cannot be neglected. For instance, Heldner and Strangert (1997) demonstrated in an experimental study that the perception of focal status does not change if F0 in non-focused words is gradually raised.

In the British school framework, the tone group structure indicates that the nuclear tone (pitch accent) normally appears close to the right edge of that unit and is associated with lexically stressed syllable of the last content word. The last major-class word is also assigned focus by Chomsky and Halle's Nuclear Stress Rule (1968). This 'unmarked' distribution of sentence prominence was described by Newman (1946), Chomsky and Halle (1968) and Bresnan (1971, 1972), who referred to it as 'normal' stress,[4] a property that is strongly determined by syntactic factors and thus largely predictable from utterance structure. 'Normal' stress is opposed to 'contrastive' stress, which is related to conversational, pragmatic context.

Bolinger (e.g. 1958b, 1972a, 1972b), however, argued that categorical division of focus types is not justified because both normal and contrastive accents depend in the same way on the information structure of the utterance. Moreover, the acoustic differences caused by focus should be treated as gradient rather than categorical. In his later studies, Bolinger (1985, 1987) also advocated a direct relationship between the accented word and sentence focus.

Even though there was evidence related to other languages that contrastive and neutral focus could be expressed in different ways, Schmerling (1974), Ladd (1980), Gussenhoven (1983) and Selkirk (1984) supported the idea of treating focus relations as subordinate to one mechanism. Still, Bolinger was criticised for rejecting focus projection, "the ability of an accented word to signal the focus for a higher constituent, like the phrase or clause, causing differently sized focus constituents to have the same form" (Gussenhoven 2008: 85, with references to Chomsky 1971, Jackendoff 1972 and Selkirk 1984). Most recent focus studies distinguish more than one type of focus, e.g. broad and narrow (cf. Ladd 1980) relative to the size of the informationally relevant unit, often syntactically indicated (Selkirk 1984, Gussenhoven 1985).

2.5 Prominence — a phonological category or gradient property of syllables?

One of the serious problems in studying prominence relations is the status of prominence as a phonological category. The contrastive function of stress and accent motivates a categorical approach (stressed/unstressed,

[4] 'Stress' is equivalent to 'accent' in this understanding.

accented/unaccented), but the complexity of prominence relations often requires the distinguishing of more levels of prominence.

Roach (2009) argues for three levels of stress, i.e. primary, secondary and unstressed. Although he admits that "[i]t is also possible to suggest a tertiary level of stress in some polysyllabic words," he concludes that "[w]hile this may be a phonologically correct account of some pronunciations, the introduction of tertiary stress seems to introduce an unnecessary degree of complexity" (2009: 75). This approach is the most popular in foreign language teaching practice where, traditionally, primary stress and secondary stress are marked in handbook glossaries and learner's dictionaries, while unstressed syllables remain unmarked.

Halliday (1967) and Vanderslice and Ladefoged (1972) propose two separate categories of stress and accent referring to separate though interacting phonetic correlates. This also yields three prominence conditions, i.e.:

[–stressed, –accented]<[+stressed, –accented]<[+stressed, +accented]

The fourth possible combination [-stressed, +accented] is disallowed because it is natural for stressed syllables only to receive accent (Liberman 1975, Ladd 1980, Pierrehumbert 1980, Selkirk 1984, Hayes 1995).

Sluijter and van Heuven (1996a, 1996b) also advocate the two general categories of stress and accent. Moreover, they observe that increased syllable length and spectral balance changes (increased intensity in higher spectrum regions) are strong correlates of stress, but overall intensity is rather a cue of accent. Similar results for stress and accent in English have also been obtained by Turk and Sawusch (1997) and Turk and White (1999).

In contrast to these two-dimensional models of prominence relations, Campbell and Beckman (1997) propose a one-dimensional scale which, however, like many of the two-dimensional ones, ultimately relies on three categories: stressed-accented, stressed and unstressed.

Liberman and Prince (1977) and Hayes (1989) provide the following hierarchy of phonologically relevant stress levels:

— nuclear pitch accent
— pitch accent
— lexical stress
— syllable/full vowel.

This hierarchy renders five prominence levels: two accent categories that roughly correspond to the British School's nucleus and head, and apart from the stressed and unstressed levels, an additional category of unstressed unreduced syllables is distinguished.

The categorical approach to prominence is phonologically well motivated but it poses some problems for empirical studies of phenomena which depend

on prominence levels. First of all, the categorisation of prominence is performed[5] on the basis of a number of continuous parameters, each measured in different units (e.g. Hz, dB, s), while the actual contribution of individual parameters in the interaction is still unresolved. For this reason it is impossible to establish very precisely the correlation between prominence or even its individual parameters and speech timing.

Theoretically, it is possible to use a continuous scale referring to the interaction between duration and one other selected parameter, e.g. amplitude or F0, but the measured values of these acoustic cues do not necessarily reflect the listener's auditory perception of prominence, both in terms of alignment and magnitude, especially if it is based on one parameter only. Consequently, in line with the majority of contemporary studies, we will follow the traditional methodology, which employs categorical prominence levels established by analysing typical combinations of semantic, syntactic, auditory and acoustic cues.

2.6 Domain lengthening under accent and stress

The relationship between prominence and duration has been confirmed by numerous studies. The longer duration of stressed syllables is mentioned as being among language universals (Maddieson 1997, Waniek-Klimczak 2005). This correlation exists in English both in the case of word stress (Klatt 1974, Umeda 1977, Crystal and House 1988a, Sluijter and van Heuven 1996b) and phrasal accent (Pierrehumbert and Talkin 1992, Beckman and Edwards 1994, Turk and White 1999, Choi 2003). The former phenomenon is often described as accentual lengthening.

2.6.1 The domain and locus of lengthening

Accentual lengthening affects the syllable with which a pitch accent is associated. Crystal and House (1988a) claim that although consonants lengthen in accented syllables, the magnitude of lengthening is greater in vowels. This has not been universally supported, as in Eefting's (1991) study of Dutch, where consonants both in onsets and codas were lengthened more, but since different consonants have been used in different studies, it is difficult to resolve this general issue.

[5] In both nuclear tone approach (Cruttenden 1997) and AM theory (Wightman 2002) it is based on auditory impression.

It is more important, however, that the phenomenon goes beyond syllable boundaries to lengthen the adjacent syllables within the word (Sluijter 1995, Turk and Sawusch 1995, 1997; Turk and White 1999). White (2002) found that the lengthening effect is proportionally stronger in monosyllabic words. Furthermore, Turk and Sawusch (1997) claimed that a word boundary blocks or at least attenuates lengthening. They showed, for example, that /f/ is much longer in "bee FARM" than in "BEE farm",[6] i.e. in the onset of an accented word. Another interesting observation concerns unstressed syllables in phrases "bacon force" and "bake enforce," where the length of the second, unstressed syllable differed only in the former example depending on accentual pattern. This also suggests the efficiency of word boundaries in blocking accentual lengthening. Further studies (Sluijter 1995, Turk and White 1999) provide more evidence suggesting the following conclusions:

— Accentual lengthening extends beyond the syllable boundary.
— The effects of lengthening are weaker in unstressed syllables which precede the accented syllable than in the ones that follow it.[7]
— Consonants are lengthened more in the onset of an accented syllable than in its coda.
— Accentual lengthening is attenuated by word boundaries but not always completely blocked.

It is not absolutely clear how accentual lengthening interacts with final lengthening but the studies mentioned above demonstrate that preboundary elements are significantly longer if they belong to a pitch accented lexical item. In a recent study, Turk and Schattuck-Hufnagel (2007) underline the strong relationship between the location of lexical stress and final lengthening. It is interesting to notice that the locus of accentual lengthening may differ from the locus of final lengthening in that the former strongly affects the accented syllable onset (cf. Wightman et al. 1992, Turk and White 1999, White 2002).

2.6.2 The scale of lengthening

A number of studies have been conducted to establish the actual length proportions between stressed and unstressed syllables. Dellatre (1966) found that non-final stressed open syllables in English are 50% longer than unstressed ones. Crystal and House (1988a) claimed that stressed syllables are twice as long as the same syllables in unstressed positions. Although both vowels and consonants are lengthened (cf. section 2.5.2), the ratio between consonants in

[6] Pitch-accented words capitalised.

[7] The experiments included words of up to three syllables with lexical stress on the antepenult to observe the influence of lengthening on more distant segments.

stressed and unstressed syllables is approximately 1.3:1, while stressed vowels are 1.8—2 times longer (Crystal and House 1988a). The same study showed that the duration ratio between stressed and unstressed diphthongs is 1.5:1. Moreover, Crystal and House (1988c) found that the effect of stress in English is stronger in inherently long vowels.

Setter (2006) measured syllable length in read English speech in four prominence conditions (weakened, unstressed, stressed and tonic). Native British speakers showed a significant difference especially between the former two and the latter two contexts. The stressed and tonic syllables were approximately twice as long as the unstressed ones.

Crystal and House (1990) also claimed that speech rate may influence the proportions between stressed and unstressed syllables. They obtained a ratio of 1.85:1 in the basic (CV) syllable type produced by a fluent fast speaker.[8]

Sluijter (1995) measured accentual lengthening in disyllabic words (lexical

Table 2.2. Mean accentual lengthening in English lexical and reiterant disyllables (Sluijter 1995)

Duration / Accent	Syll 1	Syll 2
No pitch accent	100%	100%
Syll 1 accented	+20.5%	+18.5%
Syll 2 accented	+8.0%	+18.5%

compact and reiterant *baba*) and found that in comparison with unaccented items all syllables in pitch accented words were lengthened as in Table 2.2 above.

Finally, the influence of sentence focus in English has been found to be smaller than that of lexical stress. Studies conducted by Turk and Shattuck-Hufnagel (2000) and Xu and Xu (2005) show 25% and 14% duration difference, respectively. On the other hand, Cho and Keating have found that "vowel duration is greater in a secondary-stressed syllable in an accented word, than in a primary-stressed syllable in an unaccented word" (Cho and Keating 2009: 481), which suggests greater effects of accent than stress.

2.6.3 Accentual lengthening in Polish and other languages

Duration and the other prominence cues are not used in the same way in different languages (e.g. Cho and McQueen 2005). Dogil (1995, 1999) investigated acoustic correlates of word stress but apart from F0 high levels and

[8] This may also influence speech rhythm (e.g. Dellwo 2008), to be discussed in Chapter 3.

sharp slopes, he found no such cues in Polish. Ortega-Llebaria and Prieto (2007) demonstrated that pitch accent causes less lengthening in Spanish than in English, but L2 speakers tend to retain their native lengthening patterns. A similar relation was observed in the case of final lengthening (Delattre 1966). A possible link between final lengthening magnitude and stress/syllable timing (see Chapter 3) is not obvious considering the fact that phrase-final lengthening exists in Italian (Frota et al. 2007). However, as White (2002) observes, the research examined Neapolitan Italian, a more stressed-timed variety of the language. Still, the part of Delattre's study devoted to non-final open syllables, which revealed 50% lengthening in English, a comparable 60% in German, but only 10% in Spanish, supports the hypothesis of a stronger influence of prominence on syllable duration in stressed-timed languages, to be discussed in Chapter 3.

Beckman (1986) argues that stress-accent languages, including English (cf. also Beckman and Pierrehumbert 1986), make greater use of factors other than pitch changes, especially amplitude and duration, but also spectral variation, to mark prominence in comparison with non-stress-accent languages.

2.7 Lack of prominence as a duration determinant

The previous sections of this chapter were devoted to prominence and accentual lengthening. In many languages, the effect of accentual lengthening is further enhanced by the reduction of unstressed speech units. Reduction may be observed both in the spectral qualities of the signal and its shorter duration.

2.7.1 The locus of reduction

Qualitative and quantitative reduction of speech units is a universal phenomenon connected with informal speech and lack of prominence. In English phonetics textbooks (e.g. Clark and Yallop 1990, Cruttenden 2001), vowel reduction is presented as an important feature of the language.

Although less research has been done on consonant reduction, there is evidence that, to a degree depending on their type, consonants are also reduced in unstressed contexts (eg. Byrd 1994, Keating et al. 1994, Schmidt and Flege 1995, Van Son and Van Pols 1996).

Consonantal reduction may consist only in shorter articulation of a segment, but it may also comprise phonological processes generally described as lenitions, such as spirantisation, vocalisation, debuccalisation and the ultimate lenition — elision. Consonant elision (apart from examples presented in section 1.2.2) is commonplace in weak forms of English function

words, affecting those beginning with /w/ and /h/, i.e. "will/would," all forms of auxiliary verb "have" and many personal and possessive pronouns. This kind of reduction normally occurs in postconsonantal positions.

Quantitative consonant reduction is a result of less articulatory space taken by consecutive gestures and more gesture overlap in prosodically weak positions (cf. Lindblom 1990, Cole et al. 2007). The scale of duration differences, apart from position and prominence level, depends on the type of consonant. Jassem et al. (1984 — cf. Table 1.3) distinguished separate phone classes with respect to prominence and position-related features, i.e. "aspirated/unaspirated," "accented/unaccented," "initial/non-initial/final," and "syllabic/non-syllabic."

The complex relations between prominence level and consonant duration, including such phenomena as Voice Onset Time or the articulation of consonant clusters, are especially difficult to grasp if native and non-native speech samples are to be compared. It is probably the multitude of factors governing the timing of consonants and the lesser efficiency of conscious control of consonant length in foreign language learners that makes textbook authors concentrate more on vowel duration and syllable duration. For the same reasons, the focus of this study, apart from higher prosodic level units, is upon vowel duration and syllable duration, with only some indirect references to consonant length, where differences between native English speakers and Polish learners appear evident.

2.7.2 Reasons for vowel reduction

The motivation for vowel reduction is explained in terms of modulating attention to consecutive parts of the speech signal. Information-light weak syllables (cf. Harris 2005) are reduced in order to enhance the prominence of strong positions (de Jong 2000). Harris (2005) treats reduction as planned behaviour. Alternatively, it is regarded as a failure to achieve hypothetical articulatory targets in shorter time (Flemming 1995, Kirchner 1998). This idea of articulatory undershoot was proposed by Lindblom (1963), who also suggested "hyperarticulation" and "hypoarticulation" (Lindblom 1990) as alternate stages of speech production, invoking the hearer's "hyperperception" and "hypoperception" (cf. Cole et al. 1978). One argument for regarding vowel reduction as planned behaviour is the fact that although it is a universal tendency (e.g. Van Son and Van Pols 1996), it is not used in the same way in all languages. It may be realised by qualitative centrifugal (Russian) and centripetal (English) reduction or the magnitude of reduction may be very small (Polish).

2.7.3 Vowel reduction in English and Polish

Vowel reduction mainly consists in shortening the segment and centralising it. The strong tendency to reduce unstressed vowels in English very often results in complete vowel elision. Because it occurs in information light syllables, the vowel schwa, the most frequent result of the process, may be regarded as having no articulatory target (Keating 1988). Even if the reduced vowel disappears altogether, it usually prevents assimilation processes between onset and coda consonants, which take place in underlying consonant clusters. Catford (1988) describes the two types of consonant interaction as open (onset-coda) and closed (cluster) transitions.

As mentioned above, the scale of reduction in English and Polish differs considerably. According to Lindblom's (1990) undershoot theory, languages which show little positive correlation between stress and vowel duration do not reduce their unstressed vowels significantly. Polish may be regarded as an example of that language type (e.g. Crosswhite 2001), if we consider studies of stressed and unstressed vowel duration. Jassem (1959) claims that Polish stressed vowels in citation forms are 17% longer. He also concludes that there is little correlation between stress and vowel reduction in Polish. Nowak (2006) finds 22% difference in a more extensive study comprising various speech styles, while Rojczyk (2010b) observes a 30% difference in reiterant speech. Dogil (1999) claims that vowel duration is not a word stress correlate in Polish, but even Rojczyk's results for reiterant speech do not approach the ratio of 2:1 for English monophthongs and 1.5:1 for diphthongs reported by Crystal and House (1988b).

In conclusion, considering the research presented above, it must be acknowledged that the durational "flexibility" of Polish vowels is rather low in comparison with such languages as English. According to Nowak (2006: 378), the lack of vowel reduction in Polish may be "constrained by the admittedly limited durational variation of Polish vowels (vis-à-vis many other languages)." This statement, supported by Dogil's (1995, 1999) observations of the effects of word stress on (among other things) vowel duration, is a basic assumption for the formulation of the research hypotheses investigated in the present study.

Chapter 3

Speech timing and the notion of rhythm

3.1 Introduction

Similar to many basic notions related to language structure, the existence of rhythm in speech is a controversial issue. If we assume that the impression of rhythm depends on the distribution of prominence to selected speech units in an utterance and the durational relations between alternating prominent and non-prominent elements, then we must acknowledge the existence of "rhythmisation" mechanisms which assign prominence to particular speech units according to certain "global" patterns. Indeed, as signalled in the introduction to Chapter 2, it is claimed that a universal human preference for rhythm in perception and motor behaviour is an important factor motivating regular patterns of alternating stronger and weaker elements. The importance of this preference is reflected, for instance, in metrical stress theory (Liberman 1975, Liberman and Prince 1977, Hayes 1995), where stress is understood as a manifestation of the rhythmic structure of language.

Such an approach is not necessarily in conflict with the statement that the distribution of prominent elements primarily depends on their informational value. The analysis of syntactic structures characteristic of particular languages yields evidence that the arrangement of important lexical items is highly regular and predictable (cf. the notions of the Nuclear Stress Rule or "normal stress" discussed in section 2.4).

The description of rhythm requires distinguishing a basic organisation unit, serving the purpose of durational regularisation of speech, naturally connected with the general utterance structure described in Chapter 1. The selection of such a basic rhythm unit depends on the type of language, as it has been generally accepted that rhythmic organisation of speech forms a basis for a classification of languages.

3.2 Rhythm Class Hypothesis

3.2.1 Stress-timing and syllable-timing

Pike (1945) claimed that language rhythm is based on isochrony of either syllables or interstress intervals. Abercrombie (1967) also used this distinction to suggest that all spoken languages display isochrony of their basic rhythmic units. If a language is spoken with regular reoccurrence of chest pulses corresponding with the production of syllables, it is called syllable-timed. The only alternative, according to Abercrombie, is to produce stressed syllables at regular intervals. Languages characterised by isochronous interstress intervals are classified as stress-timed.[1] A natural consequence of recognising stress-timed languages is the adoption of the foot (cf. sections 1.1.1—1.1.2) as the main timing unit. According to this classification, English, Russian and Arabic belong to stress-timed languages, while French, Telugu and Yoruba are syllable-timed. Alternative terminology, justified by the ambiguous relations between words and feet (within-word and cross-word feet), and the importance or word-boundaries for prosodic processes, distinguishes word-stress languages and syllable-stress languages (cf. Donegan and Stampe 1983).

3.2.2 The criticism of RCH

Although Abercrombie naturally realised that basic rhythm units are not perfectly isochronous, he ascribed this fact to performance-specific hesitations and other disruptions. The Rhythm Class Hypothesis (RCH) became very popular, including in FL teaching and learning. However, no subsequent instrumental research managed to support it convincingly.

Roach (1982) measured standard deviation of syllable duration and foot duration in the above-mentioned languages classified by Abercrombie as prototypically stress-timed and syllable-timed. He found that English displays the most variability of both syllable and foot duration, even if preheads and tails are excluded. This clearly brings into question the traditional distinction, which required that foot variability be smaller in stress-timed languages. The final conclusion of this study was that languages do not systematically differ in the variability of syllables and feet.

[1] Additionally, a third type, mora-timed languages, has been distinguished (e.g. Bloch 1950, Ladefoged 1975, Port et al. 1987), but the focus of this book on the Polish learner's acquisition of English prosody does not embrace the notion of mora and mora-timing.

The fact that instrumental research rejected the idea of isochrony in stressed-time languages (Ladefoged 1967, Roach 1982, Dauer 1983) as well as in syllable-timed ones (Wenk and Wioland 1982) has never caused linguists to abandon the intuitively valid typological distinction. To some extent, it could be defended by studies which demonstrated that English listeners perceive accent distribution as being more isochronous than it actually is (Lehiste 1977, Donovan and Darwin 1979, Darwin and Donovan 1980). Therefore it has been claimed that isochrony exists as a perceptual phenomenon and, as has been suggested in the case of stress perception, the perception of rhythm may not always be connected with any physical cues (e.g. Hay and Diehl 1999).

Furthermore, according to Cauldwell, many authors persist in using the distinction (e.g. Laver 1994, Crystal 1996, Dalton and Seidlhofer 1994, Cruttenden 1997, Ball and Rahilly 1999, Rogers 2000) because "it remains the prevailing view and still features in accounts of the rhythms of speech because no other hypothesis matches its deceptively bewitching power" (Cauldwell 2002: 1).

It is not the purpose of the present study to provide evidence supporting or rejecting any form of the Rhythm Class Hypothesis, but its important role in the development of prosodic research compels us to take into account the distinction between stress-timing and syllable-timing as a possible source of discrepancies between native English and Polish learners' oral production.

3.2.3 Alternative approaches to rhythm typology

Approaches that followed Pike's and Abercrombie's proposals have incorporated other features that possibly affect rhythm typology. Donegan and Stampe (1983) understood the distinction between word rhythm and syllable rhythm not as a result of timing alone, but rather of a network of relations among more features of the language, e.g. the function and distribution of accent, syllable structure, phonemic distinctions, or phonological and phonetic processes.

Dauer (1983) also observed that the traditional classification groups together languages that share specific features which may be associated with one or the other kind of rhythm. In order to keep interstress intervals even, stress-timed languages need to shorten their unstressed syllables, using vowel reduction (cf. Bolinger 1981). In syllable-timed languages this is undesirable, if syllable duration is to remain stable. However, an uneven length of syllables may also be caused by the varying complexity of the unit in question. A syllable by definition contains exactly one vowel, so syllable-timed languages prefer less variable syllable structures, especially the basic CV model. Dauer's (1983) and subsequent studies have revealed that Abercrombie's

fundamental assumptions concerning the binary division of languages should be modified by introducing a two-dimension scale reflecting the degree of vowel reduction and allowed consonant cluster complexity.

Although prototypical representatives of each class may conform to the outlined characteristics, others are less consistent in observing the general principles. Polish, for instance, allows very complex consonant clusters, outranking even English and German, but hardly any vowel reduction. Catalan is an opposite example, with a simple syllable structure on the one hand and vowel reduction on the other. In consequence, the two-dimension scale seems better suited as a descriptive device for more complex languages.

As mentioned above, despite the instrumental research results, linguists have been reluctant to reject RCH. Dauer only borrowed from Allen (1975) and O'Connor (1973) the notion of stress-based languages to replace the abandoned stress- and syllable-timing. Within the new approach, instead of being categorically assigned to one of the two classes, a language could be described as more stress-based or less stress-based, with reference to the role of word stress, syllable structure and vowel reduction.

Apart from the "mainstream" research that assumes the existence of rhythm in world languages, there are studies that question the existence of any type of timing patterns in human speech. Cauldwell (2002) argues against any kind of rhythm in language. He claims that language is not only anisochronic but entirely irrhythmical, and any "patches of rhythm" are purely coincidental rather than intentional. To support this view, he points out methodological problems with the consistent identification of prominences, the location of interstress interval boundaries, the need to control for tempo,[2] and irregularities of preheads and post-tonic syllables.

3.2.4 Recent rhythm/timing measures

The search for temporal patterns instead of isochrony called for new rhythm measures, gradient rather than categorical in nature. Considering the role of vowel reduction and consonant clusters suggested by previous studies, Ramus et al. (1999) argued that measuring the proportion of the vocalic content (%V) in a speech sample and the standard deviation of consonantal (ΔC) and vocalic (ΔV) intervals provide an adequate profile of speech timing in a language. All three measures reflect the segmental inventory and the phonotactic principles of the language under investigation.

[2] Dauer (1983) found different timing relations in fast and slow speakers (cf. later studies by Dellwo and Wagner 2003, Dellwo 2008).

These features of language are also decisive for results obtained by the Pairwise Variability Index (PVI) (Low et al. 2000, Grabe and Low 2002). The PVI (raw — rPVI, or normalised for changes in speaking rate — nPVI) relates the duration difference of two successive vocalic or intervocalic intervals to the mean duration of such a pair. The arithmetic mean of those ratios for all successive pairs shows in fact serial variation in the duration of vowels and consonant intervals in a spoken text. Normalised PVI is used for vocalic intervals, while rPVI has been found more suitable for consonantal periods because it is difficult to separate the effects of speaking rate from the effects of syllable complexity on the duration of intervocalic intervals (Grabe and Low 2002).

VarcoV and VarcoC (Dellwo 2006) are other measures for rhythm which calculate the standard deviation of vocalic or intervocalic interval durations. They differ, however, from the metrics proposed by Ramus et al. (1999) in that they normalise for speech rate. Because Barry et al. (2003) and Dellwo and Wagner (2003) found that ΔC negatively correlates with speech rate (cf. Dauer 1983), it was considered necessary to use relative rather than absolute measures for the ΔC variation coefficient.

The continuous measures described above capture differences between languages characterised by various syllable structures and presence or absence (and degree) of vowel reduction but there are doubts whether devices neglecting prominence (beats) and prosodic structures can actually provide information about rhythm in language. Another problem is the treatment of syllabic consonants and devoiced vowels, which, together with their environment, yield unnaturally long consonant clusters.

Both beats and hierarchic structures of speech chunks are taken into account in Cummins and Port's (1998) Speech Cycling Paradigm, which models rhythmic organisation of speech as the relations between lower level units (feet) nested in larger units (phrases). The relations are displayed in experiments with Phrase Repetition Cycles (PRC), where stressed syllables are aligned with rhythmical metronome beats. An interesting finding was that Italian and Spanish speakers found the speech cycling task more difficult than English speakers (Cummins 2002). This approach offers a more comprehensive view of dynamic speech organisation, but some reservations, as in the case of reiterant speech, may concern the problem of possible deviations from natural timing, and may also reflect Cauldwell's argument that the apparent foot isochrony in English advocated in many pronunciation textbooks (Cauldwell 2002: 3) only illustrates speech plasticity.

3.2.5 Summary

The Rhythm Class Hypothesis has survived criticism derived from the refutation of isochrony in language, and the traditional terminology is still in use with reference to the timing contrasts between language types such as Western European Germanic and Romance (e.g. White and Mattys 2007a, 2007b, White et al. 2007, Wagner 2007, Dellwo 2008). However, research on rhythm "has more recently focused on cross-linguistic variation in durational contrast between stressed and unstressed syllables" (Wiget et al. 2010: 1559).

For a foreign language learner, systematic discrepancies between L1 and FL may lead to negative transfer. The awareness of these potential difficulties helps both the teacher and the learner to concentrate on the relevant issues. Recognising a number of differences between English and Polish speech timing, we try in the next sections to establish the areas where the Polish learner of English may face problems with the temporal organisation of FL speech.

3.3 English and Polish speech timing

3.3.1 The rhythms of English and Polish

With respect to rhythm, English is classified as a typical stress-timed language. It allows up to three consonants in a syllable onset and reduces vowels in unstressed positions. Polish is known to allow hardly any degree of vowel reduction, but its phonotactic rules do not impose strong restrictions on acceptable consonant clusters, which can be composed of up to four elements in the onset, e.g. *wstrętny* /ˈfstrɛntnɨ/ ("hideous") and up to five in the coda, as in *następstw* /ˈnastɛmpstf/ ("consequences" pl., gen.). Moreover, Polish does not very strictly observe the sonority principle (e.g. *mgła* /mgwa/ "fog"), which further broadens the legitimate consonant cluster inventory. This feature of the language is reflected for instance in the highest intervocalic rPVI values among 18 languages investigated by Grabe and Low (2002).

Ramus et al. (1999) mentioned Polish and Catalan as languages which exhibit complementary levels of vocalic and intervocalic variability, with high intervocalic and low vocalic variability in the former, and the reverse relation in the latter language. Polish has also been regarded as a mixed type of language by Dauer (1987) and Nespor (1990). Wagner (2007), by visualising normalised proportions between paired consecutive syllables, found that the plotted relations resemble those characteristic of syllable-timed languages.

Indeed, it seems that the stress-timed-like scores obtained by the modern rhythm metrics for Polish would no longer differ from syllable-timing

characteristics if analysed in texts with less complex consonant clusters, while the scores for vocalic period variability would not be changed. It may even be claimed that a language which does not reduce unstressed vowels *despite* having complex syllables, has a stronger bias towards syllable timing than CV languages. Likewise, Catalan may be at least equally if not more inclined towards stress-timing rhythm than the Germanic languages, where vowel reduction may be regarded as a natural offset to the longer duration of complex consonant clusters.

More argument for a rhythm type discrepancy is provided by Dogil (1999) and Nowak (2006), who observe more influence of stress on English syllable duration, while Polish appears to show less variation of this kind. Therefore we expect this lack of temporal syllable flexibility, characteristic of syllable-timed languages, to constitute the main obstacle in the acquisition of English prosody by Polish learners.

3.3.2 Timing differences between English and Polish

The above overview of literature concerning the temporal characteristics of Polish and English speech has suggested a number of areas of potential timing problems in the English pronunciation of Polish learners. They are related to the following features of the English language, which either do not occur or are marginal in Polish:

— long/short vowel distinction (see section 1.2.1)
— unstressed syllable/vowel reduction (see section 2.7.3)
— accentual lengthening (see section 2.6.3)
— stress timing (see section 3.3.1).

A review of studies investigating the actual problems of Polish learners with English speech timing is provided in the next section.

3.4 Polish learners' timing problems

This subchapter presents research results and other reports referring to the English pronunciation of Polish learners with respect to the issues listed above. The empirical study that follows is to verify previous results and establish the scale of timing problems by showing actual differences between native English speakers and Polish learners. Finally, the longitudinal aspect of the research presents conclusions concerning the rate of development of the learners' English speech by indicating the scale of reduction of the discrepancies.

3.4.1 Long and short vowels

Szpyra-Kozłowska (2003: 200—201) remarks that the preservation of vowel length contrasts is difficult for Polish learners, "whose native tongue makes neither phonemic nor allophonic use of vowel quantity." A similar observation is made by other researchers as well (e.g. Sobkowiak 1996, Nowacka 2008, Bryła 2010).

Vocalic duration cues for fortis/lenis distinction of the following consonants are even more problematic for Polish learners (Sobkowiak 1996, Szpyra-Kozłowska 2003). Production studies based both on acoustic data (Waniek-Klimczak 2005), and auditory assessment (Nowacka 2008) have shown that learners have poor abilities to make the relevant distinctions in this respect.

Performance problems are certainly aggravated if the significant contrasts are not efficiently perceived by the learners. Slowiaczek and Szymanska (1989) found 55% of identification errors related to underlying voicing in coda obstruents. Rojczyk (2010c) also observed Polish learners' insensitivity to this cue. Furthermore, Rojczyk (2010a) demonstrated a conspicuous difference in this aspect of perception between English and Polish speakers.

An earlier study of English vowel recognition in monosyllables (Porzuczek 1998) suggested that Polish learners are quite sensitive to categorical vowel duration contrasts, but they tend to misinterpret the cues far more often if those for intrinsic vowel length and postvocalic voicing are contradictory, i.e. long vowels preceding fortis consonants or short vowels preceding lenis consonants.

3.4.2 Unstressed vowel and syllable reduction

Numerous publications (e.g. Sobkowiak 2001, Hewings 2004, Dziubalska-Kołaczyk et al. 2006, Nowacka 2008) point out Polish learners' insufficient syllable reduction in unstressed positions. Reduction is inadequate considering both vowel quality (e.g. Avery and Ehrlich 1996, Gonet et al. 2010) and quantity (e.g. Porzuczek 2007, 2010a, 2010b). Luke and Richards (1982), Sobkowiak (2001), Hewings (2004), and Gonet et al. (2010) observe the strong influence of spelling on the foreign learner's pronunciation of English vowels which are supposed to be reduced. This is unambiguously assigned to L1 interference.

Gonet et al.'s (2010) longitudinal study has estimated that only around 31% of schwas are pronounced correctly by first-year Polish students of English, who later manage to improve enough to realise 67% of their schwas in an acceptable manner. Additionally, the authors observe more correct responses in medial positions, where English unstressed vowels often disappear

altogether. Word-initial and word-final positions have proved to be more difficult.

Vowel reduction is often constrained due to another reason unrelated to the vocalic segment alone. In many cases Polish learners may fail to reduce a vowel because this yields a long consonant cluster, breaking even the lenient phonotactic principles of Polish, as in "vegetables," where radical vowel reduction would make the pronunciation of consecutive consonants /dʒ-t-b-l-z/ uncomfortable for native Polish speakers, even with open transitions. Moreover, the frequency of closed transitions is higher in English, even though Polish clusters are phonologically more complex. For instance, nasal plosion and lack of plosion are restricted to homorganic clusters in Polish, while in English the processes occur regardless of the places of articulation of the stops involved.

Finally, standard Polish does not allow syllabic consonants, although they can appear in rapid speech (cf. Rubach 1974, Biedrzycki 1978) or in the Upper Silesian dialect of Polish. This fact may constitute a serious obstacle to Polish learner's unstressed syllable reduction.

3.4.3 Accentual lengthening

Although more attention is focussed on the Polish learner's problems with vowel reduction, it is generally understood that reduction and accentual lengthening are interrelated to convey the contrast between prominent and non-prominent utterance constituents. Avery and Ehrlich (1996: 145) remark that Polish learners "should be given activities that practice reduction of unstressed syllables and lengthening of stressed ones."

3.4.4 Stress timing

Natural English stress-timing has been regarded (most often on the basis of subjective impressions) as a serious problem that Polish learners face (e.g. Śpiewak and Gołębiowska 2001). The unresolved question of Polish being either syllable- or stress-timed is irrelevant here if we assume that the main problem in stress-timing acquisition is vowel length adjustment (cf. White and Mattys 2007a, Waniek-Klimczak 2009). Then the rhythm depends on the efficiency of accentual lengthening and unstressed vowel reduction.

However, since timing does not only depend on vowel length but also on the length of consonants, their relative duration must be taken into account as well. There are two potential problems connected with consonant duration. Firstly, the duration of individual consonants may vary between native speakers

and foreign learners. Aspiration is probably the most important consonantal[3] factor affecting general timing relations (see Waniek-Klimczak 2005 and Rojczyk 2010c for production and perception of VOT by Polish learners of English). Secondly, stress-timed languages use consonant clusters, which may constitute an obstacle to speakers of languages based on the CV syllable model. This is because, unaccustomed to coarticulation and closed transitions, they may need more time to move from one consonant to another. In extreme cases, speakers of Spanish, Italian or Japanese may use epenthesis in order to facilitate such transitions. Polish learners generally have no articulatory problems with English consonant clusters, which are never more complex than those in their mother tongue.

3.5 The relevance of timing problems for EFL communication

The previous section discussed the aspects of English speech timing where Polish learners have been reported to face problems. This section provides a review of opinions concerning the significance of these aspects for the teaching/learning process.

3.5.1 Long and short vowels

The phonological functions of segmental duration are regarded as important by most authors dealing with language pedagogy. For instance, Jassem (1971: 54) writes:

> With respect to segmental quality shades (especially within vowels), the teacher can afford some tolerance in requirements given the quite wide acceptability margin. However, with respect to phonological duration, the language norms are more rigid. [translation mine][4]

In her rather permissive approach to non-native English pronunciation, Jenkins (2000) also stresses the importance of long-short vowel opposition in English. Because foreign learners with a syllable-timed L1 background usually find it difficult to attend to qualitative vocalic contrasts of English,

[3] Depending on the approach, aspiration may also be considered part of the stressed vowel following a voiceless plosive.

[4] „W zakresie odcieni głoskowych (szczególnie w obrębie samogłosek) nauczyciel może sobie pozwolić tu i ówdzie na pewną liberalność w swych wymaganiach z uwagi na wspomniany wyżej dość szeroki margines poprawności. Natomiast w zakresie iloczasu normy poprawnościowe są bardziej sztywne" (Jassem 1971: 54).

a considerable proportion of their performance may become unintelligible whenever both kinds of errors are combined. The distinction between long and short or tense and lax vowels form an important issue in practical English pronunciation courses, and most popular British English dictionaries (cf. section 1.2.1) indicate intrinsic vowel length explicitly in phonemic transcription.

Another important function of vocalic duration, also discussed in previous chapters, is to indicate the phonological voiced/voiceless consonant contrast in syllable-final position. Because it is often the only phonetic cue to the contrast (cf. section 1.2.2), it must be included in English pronunciation teaching syllabus.

3.5.2 Vowel reduction

Bogle (1996) and Kenworthy (1987) insist that vowel reduction must not be neglected in teaching English pronunciation. This is especially important in interactions with native speakers. Jenkins (2000), however, considers vowel reduction redundant in communication. The pronunciation teacher's approach to this problem should still note the difference between L1 and FL in this respect. Even if Jenkins's point of view is favoured, the learner's awareness of vowel reduction in English seems indispensable for successful comprehension of native English speech.

3.5.3 Accentual lengthening

Scholars dealing with language pedagogy are practically unanimous in recognising the crucial role of robust indication of phrasal or sentence prominence for communication efficiency (cf. Bogle 1996, Kenworthy 1987, Celce-Murcia et al. 1996, Jenkins 2000). Among the acoustic cues to prominence, duration appears to be one that is relatively easy to control by the learners. However, the dearth of unambiguous results in studies attempting to describe and explain the relations between individual acoustic cues and their perception by the listeners should remind us that the temporal relations alone may not adequately reflect the nature of encoding and decoding prominence signals.

3.5.4 Stress-timing

Timing and rhythm are regarded as high pedagogical priorities (e.g. Kenworthy 1987, Nunan 1995, Bogle 1996, Jenner 1999, Celce-Murcia et al. 1996, Roach 2002, Szpyra-Kozłowska et al. 2002, 2003). Adams (1979), Anderson-Hsieh et al. (1992), Cutler (1993) and Fear et al. (1995) claim that patterns of alternating strong and weak syllables are especially significant for speakers of stress-based languages (cf. Setter 2006). Such opinions are supported by perception studies, which have shown that timing may be more important for speech recognition than segmental details (e.g. Kozhevnikov and Chistovich 1965, Wingfield and Klein 1971, Jassem 1971, Faber 1986).

Tajima et al. (1997) manipulated the English speech sample of a Chinese speaker to match the timing of a native American speaker and in the opposite way they made native American speech syllable-timed to resemble the rhythm of Chinese. It turned out that in both cases the intelligibility of speech samples changed by up to 25% to demonstrate positive correlation with stress-timing. A similar conclusion concerning the greater intelligibility of native-like speech timing has been drawn by Smith et al. (2003) from a study of perception and production of temporal vowel contrasts by Chinese English learners. Finally, Adams (1979) even suggests that improper speech timing may lead to a communication failure.

The importance of English speech timing raises concerns about its teachability and learnability. It has been observed that rhythm differences between L1 and L2/FL may make the learning process more difficult (Anderson-Hsieh and Venkatagiri 1994). Adams (1979), Taylor (1981), Bond and Fokes (1985), and Faber (1986) consider rhythm to be one of the most difficult features of English for foreign speakers to master. However, of the rhythm correlates, syllable-duration as a significant cue to word and phrase stress can be efficiently taught (Gilbert 1984; Halliday 1989; Chela-Flores 1994, 1998; Setter 2006).

Another problem that the pronunciation teacher faces already at the stage of designing a syllabus is the order of segmental and suprasegmental topics to be introduced and practised. On the one hand, the traditional design where suprasegmental issues follow segmental ones is often challenged on account of the alleged greater significance of suprasegmentals. However, the influence of phonological structures of prosodic units is a decisive factor that determines their duration in both native and non-native speech. Without the learner's satisfactory command of foreign speech sounds and their combinations, achieving native-like speech timing patterns is a highly difficult task. According to Barry (2007: 114),

> It is important that *individually learnable* properties of language be brought into focus — informationally important (prominent) words, informationally less important words (and syllables within multi-syllabic words), long and short vowels, spectrally reduced vowels, consonant elision, etc. The contextualized introduction and practice of these properties in an optimal sequence is, of course, a non-trivial task. But their command will lead to a globally correct prosody and, in time, to a sense of prosodic "rightness" for the particular communicative intention in the same way that learning verb or noun morphology and syntactic regularities will lead to a command of the correct form and sequence of words. In neither of these areas would one think of introducing teaching points by appealing to a sense of "Morphology" or "Syntax". We suggest that the appeal to a general idea of "Rhythm" which is abstracted from the prominence pattern of the particular utterance is equally unproductive.

The general question whether the actual native timing patterns are mainly reflections of some "global" rhythmic models or rather a sum of the "local" timing determinants of their constituents will certainly not be answered in this dissertation but it will frequently be asked to make the interpretation of the results of the present research more comprehensive.

3.6 Conclusion

Although interference is not the only source of foreign-accentedness of FL performance (cf. Waniek-Klimczak 2009), it is probably the element whose influence is the most systematic and predictable. Therefore, supported by traditional contrastive analysis (Lado 1957, Brière 1968, Wardhaugh 1970), it should remain a focus of the foreign language teacher. The potential interference problems described in this chapter as being the most typical of Polish learners are investigated in the subsequent parts of the dissertation in order to evaluate the scale of the problems, interspeaker variability, and the dynamics of interlaguage change with respect to timing. They form a basis for the formulation of research hypotheses presented in Chapter 4.

Chapter 4

The diagnostic study of Polish learners' English speech timing

4.1 Introduction

Following the discussion in Chapter 3, we recognise speech timing as an important aspect of foreign language production. Timing in any language, as indicated in Chapters 1—3, depends on a range of interacting factors. Many of these factors are universal and determine speech unit length in similar ways across languages, while others perform language-specific phonological functions. As stated in the introduction, the general goal of this study is to describe the non-native timing characteristics of Polish learners' English speech. We consider the read speech of advanced learners in order to control for the content of analysed speech samples and minimise the effects of low general language proficiency on the timing of learners' oral production. We focus on the timing within an IP, a relatively independent speech unit in communication, trying to observe the contribution of constituents at various levels of the prosodic hierarchy. In order to achieve this, we analyse the timing differences between native English speakers and Polish learners, assuming that L1 interference is an important source of pronunciation errors. As has been discussed before, the timing of Polish learners' English pronunciation differs from native speakers' production systematically for a number of reasons:

1. FL speech production is generally slower and marked by more pauses and other dysfluencies. This observation can be treated as extension of H&H theory (Lindblom 1990) and other theories (e.g. Gahl 2006) which claim that reduction, including quantitative reduction, is more radical in more familiar and more frequently occurring items (cf. also Munro and Derwing 1994).
2. Even in familiar and frequent units, the learner may encounter sounds or sound clusters of greater articulatory complexity, which slow down FL

pronunciation. The difficulty level may vary individually but it may also be determined by L1-FL relations.
3. L1 sounds of longer intrinsic duration may be substituted for FL-specific ones (e.g. /x/ for /h/ — cf. Jassem's classification presented in section 1.2.1).
4. Certain durational features are used systematically to perform phonological functions in FL only, while in L1 either they are not exploited or the magnitude of length variation is different in the two languages.

The interest of this study is focused mainly on performance discrepancies resulting from the last of these points, while the other three must be taken into account as potential sources of extraneous variables.

4.2 General assumptions and hypotheses

On the basis of previous considerations, in this section we recognise three general timing differences between Polish and English. Given the assumption that these differences will be reflected in the English pronunciation of Polish learners, we also formulate a number of hypotheses to be tested in the research part of this book:

1. Polish does not use quantitative vowel contrast (section 1.2.1).

Hypothesis 1: Mean duration proportion of intrinsically "long" to intrinsically "short" vowels is smaller in Polglish production (section 3.4.1).

2. The influence of prominence level on duration is smaller in Polish (section 2.6.3).

Hypothesis 2: Unstressed vowels, syllables, and function words are longer in Polglish production in terms of both absolute and relative[1] duration (section 3.4.2).
Hypothesis 3: Vowels, syllables, and content words in nuclear accent position are relatively shorter in Polglish production (section 3.4.3).

Because in many phrases such proportions may also be obtained as a result of unstressed unit reduction alone, we also propose the following hypothesis:

[1] We expect a generally faster speech rate in English speakers, which should result in shorter absolute duration of reduced units. However, it is difficult to predict the absolute duration relations between stressed and pitch accented units.

Hypothesis 4: In an intonational phrase divided exhaustively into prehead (P), head (H), and nucleus (N), according to traditional British School parsing, P:H and H:N ratios are higher in Polglish production.

3. In English, a prototypical stress-timed language, the variability of foot duration is smaller than in Polish (section 3.3.1).

Hypothesis 5: Within a head[2] consisting of more than one foot, the standard deviation from mean foot length is larger in Polglish production (section 3.4.4).

Another goal of this study is to observe how timing relations change with general EFL proficiency development. This leads to the last hypothesis, not related to L1-FL discrepancies:

Hypothesis 6: Language proficiency development is reflected in more native-like timing of Polglish production after two semesters of study including systematic practical phonetic training.

Finally, an important objective of the research is to assess the scale of differences between native and non-native production at two stages of EFL learning. Such results can prove helpful in evaluating the importance of timing in the teaching process and chances of improvement in the course of study.

4.3 Method

4.3.1 The subjects and data collection procedure

The Polish subjects were 13 first-year students (9 females and 4 males) of the English section of Jastrzębie Zdrój Teacher Training College. They were at the age of 19—20 at the time of the recordings and had completed a primary and secondary school with 10 years of English (EFL) learning experience. A high degree of homogeneity in the sample with respect to age and learning experience allows us to ignore these extralinguistic factors in analysis and concentrate on cross-linguistic interaction in the acquisition process. All students volunteered to take part in the study and were not paid for their participation. None of them reported or otherwise indicated any speech or hearing problems.

[2] This relation cannot be measured in P, which does not contain stressed syllables. N usually comprises one foot only, otherwise final lengthening and a large difference in prominence levels make stress-timing still difficult to observe.

The task was to read a passage in English (see Appendix), which was recorded at two points in time, i.e. at the beginning (October 2006) and at the end (May 2007) of the first year of practical phonetics training. The participants were given time to rehearse the text before recording and were instructed to speak at a comfortable rate. The recordings were made in quiet conditions at the college's language laboratory, using a Panasonic RR-US360 portable digital recorder, placed at a comfortable distance of approximately 25cm from the respondent's mouth. The signal was sampled at 16kHz and transferred to a computer's hard disk in the .wav format for further analysis by means of the Praat 4.4.30 software package (Boersma 2001).

The data were compared to speech samples of 12 native British English speakers (6 females and 6 males), secondary school students in Cambridge at the age of 16, recorded for the purposes of IViE research (Grabe et al. 2001). The recordings were made in local school in a quiet room.

4.3.2 Tested units and contexts

In order to verify the hypotheses listed in section 4.2 we measured samples of read text obtained from the IViE corpus (Grabe et al. 2001) of British English speech, i.e. the Cinderella passage (see Appendix). The particular tested units are presented in the following sections. The selection includes text portions (mostly tone groups) which were performed with homogenous prosodic interpretation by the subjects, especially in terms of prominence distribution and phrasing, and allowed relatively reliable segmentation.

Following the discussion in Chapter 1, the tone groups are basically divided into three parts, following the nuclear tone tradition:

— an optional **prehead** (P), grouping unstressed syllables up to but not including the stressed syllable of the first content word (usually pitch-accented),
— a **head** (H), containing the text portion from the stressed syllable of the first content word up to but not including the nuclear accent,
— the **nucleus** (N), including the nuclear accented syllable and the rest of the foot.

The last definition embraces what is traditionally regarded as the tail, but in this study, to permit exhaustive division of the phrase into left-headed feet, we use the term **tail** (T) to signify an optional foot (headed by a stressed unaccented syllable) following the nucleus.

Unlike the other constituents, a head can include more than one foot. Each of them is assigned a separate "H" label. In this way, the tone group is divided into stress feet, and an optional prehead, treated on par with a foot.

Because the study regards prominence conditions, each foot is divided into the lexically stressed syllable (marked "S"), normally pitch-accented in the first H foot and N, and the unstressed part consisting of 0-3 syllables (marked "U"). Finally, the description of a tested unit indicates relevant foot complexity expressed as the number of constituent syllables. Thus, for instance, "UN2" stands for the traditionally defined tail of a trochaic nuclear tone. Here, a tail may itself consist of a stressed syllable (ST) and unstressed ones (UT). The example below illustrates the description convention used in the present study:

so they (P2) start(SH2)+ed(UH2) shout(SN3)+ing at(UN3) Cind(ST2)+ers(UT2)

The following part of this section introduces the units selected to verify each of the hypotheses proposed in section 4.2.

Hypothesis 1: Mean duration proportion of intrinsically "long" to intrinsically "short" vowels is smaller in Polglish production.

The text provides a number of vowels in contexts allowing relatively reliable measurements, i.e. flanked by obstruents or nasal stops (cf. section 4.3.3.2). In one case, the word *looking*, the vowel preceded by a word initial lateral consonant was also found feasible to measure. The duration of vocalic segments (below in capitals) was measured in the following contexts:

— **12 short vowels:**
(but everyone) called her CInders (N2)
CInders (lived with her mother) (H2)
in hOnour of the Queen's only sOn (H4, N1)
and he was lOOking for a bride (H4)
but their mOther said (H2)
that they had enOUgh gowns (N1)
so they started shouting at CInders (T2)
when her sIsters had gOne (H3, N1)
It was her fairy gOdmOther (N1, T2)

— **3** instances of the front open vowel classified as **intermediate** (Cruttenden 2001: 92, with reference to Wiik 1965):
and she hAd to do the cleaning (H2)
they were in a bAd mood (H1)
find my hAt (N1)

— **5 long vowels:**
and going to pARties (N2)
Lily and Rosa thOUGHt this was divine (H1)
Prince William was gORgeous (N2)

they were in a bad mOOd (N1)
so they stARted shouting at Cinders (H2)

— **6 diphthongs:**
Lily and Rosa thought this was divIne (N1)
they'd wanted to bUY some new gOWns (H2, N1)
that they had enough gOWns (T1)
so they started shOUting at Cinders (N3)
fInd my hat (H2)

All the selected vowels in their carrier words are presented in Table 4.1 below.

Table 4.1. The characteristics of tested vowels

Position	Unit type and complexity	Short vowel	"Ash"	Long vowel	Diphthong
Phrase-final accented/stressed syllable (nucleus or tail)	N1	s**o**n g**o**ne	h**a**t	m**oo**d	div**i**ne g**ow**ns
	T1	—	—	—	g**ow**ns
Phrase-final word with nuclear accent (nucleus)	N2	**Ci**nders	—	p**ar**ties g**or**geous	—
Non-final nuclear accent (nucleus)	N1	g**o**dmother en**ou**gh	—	—	—
	N3	—	—	—	sh**ou**ting
Phrase-final content word (tail)	T2	godm**o**ther **Ci**nders	—	—	—
Non-final accent (head)	H1	—	b**a**d	th**ou**ght	—
	H2	m**o**ther **Ci**nders	h**a**d to	st**ar**ted	b**uy** f**i**nd
	H3	s**i**sters	—	—	—
	H4	l**oo**king h**o**nour	—	—	—

Depending on the variable in focus, the vowels are labelled according to criteria indicating their intrinsic length (diphthong (D), long (L), "ash" (A), short (S)), position (stressed phrase-final syllable (f), stressed syllable of a phrase-final (trochaic) word (w), stressed phrase-medial position (m)), prominence condition (nuclear syllable (SN), stressed head syllable (SH), stressed tail syllable (ST)), foot complexity (1-4 syllables constituting a foot).

Hypothesis 2: Unstressed vowels, syllables, function words, preheads and unstressed foot parts are longer in Polglish production in terms of both absolute and relative duration.

The duration of unstressed units (below in capitals) was measured in the following contexts:

— **20 reduced vowels:**
There was A girl... (P3)
BUt everyone called hER CindERs. (P1, H2, N2)
CindERs lived... (H2)
...and going tO parties. (H3)
...and she had tO do the cleaning. (H2)
Prince William was gorgeOUs... (N2)
They dreamed Of wedding bells. (H2)
They were in A bad mood. (P4)
They'd wantEd tO buy sOme new gowns... (H3, H3, H2)
...so they startEd shouting At CindERs. (H2, N3, T2)
When hER sistERs hAd gone... (P2, H3, N1)
It was hER fairy godmother. (P3)

— **7 reduced syllables** in content words:
But everyone called her CindERs. (N2)
CindERs lived with her mother... (H2)
So they startED shouting at CindERs. (H2, T2)
Prince William was gorgEOUs... (N2)
They'd wantED to buy some new gowns... (H3)
When her sistERS had gone... (H3)

— **3 weak forms** of HER:
But everyone called HER Cinders (H2)
When HER sisters (P2)
It was HER fairy godmother (P3)

— **3 weak forms** of TO:
They'd wanted TO buy some new gowns... (UH3)
...and she had TO do the cleaning. (UH2)
...and going TO parties. (UH3)

— **4 weak forms** of OF, THAT, AT, and SOME (one each):
They dreamed OF wedding bells. (H2)
...but their mother said THAT they had enough gowns. (H3)
So they started shouting AT Cinders. (N3)
They'd wanted to buy SOME new gowns... (H2)

— **12 preheads:**
THEY dreamed of wedding bells. (P1)
...IN honour of the Queen's only... (P1)
...AND going to parties. (P1)

THEY'D wanted to buy... (P1)
BUT THEIR mother said... (P2)
THAT THEY had enough gowns. (P2)
SO THEY started shouting... (P2)
...THERE WAS A girl called (P3)
...AND HE WAS looking for a... (P3)
IT WAS HER fairy godmother. (P3)
THEY WERE IN A bad mood. (P4)

— **8 unstressed foot parts** in relation to the **foot-initial stressed syllables:**
called:her (H2)
ho:nour of the (H4)
loo:king for a (H4)
wan:ted to (H3)
buy:some (H2)
star:ted (H2)
shou:ting at (N3)
sis:ters had (H3)

In order to avoid the effect of final lengthening, only non-final feet have been selected for this part of analysis.

Hypothesis 3: Vowels, syllables, and content words in nuclear accent position are relatively shorter in Polglish production.

The duration of pitch-accented units in nuclear position was measured in the following contexts:

— **vowels:**
...that they had enOUgh gowns (mN1)
...started shOUting at Cinders (mN3)
fairy gOdmother (mN1)
and going to pARties (wN2)
called her CInders (wN2)
Prince William was gORgeous (wN2)

— **syllables:**
They dreamed of WEDding bells (mN2+T1)
fairy GODmother (mN1+T2)
that they had eNOUGH gowns (mN1+T1)
started SHOUTing at Cinders (mN3+T2)
called her CINDers (wN2)
and going to PARTies (wN2)
and she had to do the CLEAning (wN2)
William was GORGeous (wN2)

— **content words:**
dreamed of WEDDING bells (mN2+T1)
started SHOUTING at Cinders (mN3+T2)
that they had ENOUGH gowns (mN1+T1)
called her CINDERS (wN2) and going to PARTIES (wN2)
and she had to do the CLEANING (wN2)
William was GORGEOUS (wN2)
Queen's only SON (fN1)
When her sisters had GONE (fN1)
looking for a BRIDE (fN1)
bad MOOD (fN1)
find my HAT (fN1)
buy some new GOWNS (fN1)

Hypothesis 4: In an intonational phrase divided exhaustively into a prehead (P), head (H), nucleus (N) and tail (T), P:H and H:N ratios are higher in Polglish production.

The durations of unit constituents were measured in the following contexts:

P:H

in (P1:H4) honour of the
and (P1:H3) going to
they'd (P1:H3) wanted to
they (P1:H2) dreamed of
and she (P2:H2) had to
so they (P2:H2) started
and he was (P3:H4) looking for a
it was her (P3:H2) fairy
there was a (P3:H1) girl
they were in a (P4:H1) bad

H:N[3]

dreamed of (H2:N2) wedding (bells)
why are you (H3:N3) crying my (dear)
started (H2:N3) shouting at (Cinders)
fairy (H2:N1) god(mother)

Hypothesis 5: Within a head (H) consisting of more than one foot, the standard deviation from mean foot length is larger in Polglish production.

The duration of consecutive head feet was measured in the following contexts:

[3] Only non-phrase-final nuclei were taken into account.

girl + called (H1:H1)
had to + do the (H2:H2)
(in) honour of the + Queen's + only (H4:H1:H2)
wanted to + buy some + new (H3:H2:H1)

Hypothesis 6: General language proficiency development results in more native-like timing of Polglish production after two semesters of study including systematic practical phonetic training.

This hypothesis is to be verified on the basis of all the above study results.

4.3.3 Acoustic signal analysis and the structure of tested phrases

Measurement criteria differ across studies. For instance, Kim and Cole (2005) measure feet excluding the onset of the initial stressed syllable, but they include the onset of the stressed syllable beginning the next foot, regarding p-centres (Marcus 1981, Scott 1993), significant for rhythm perception. Because any modal length of prosodic units is irrelevant for the purpose of this study, the acoustic salience of selected landmarks is assigned primary importance and we believe that it is justifiable to exclude certain portions of prepausal acoustic signal from the analysis. As a result, the termination point of a unit is constituted by the last unambiguous acoustic landmark, such as a plosive/affricate release burst or frication onset. Continuant sonorant segments are avoided in final positions and units with final fricatives have been measured up to the onset of frication noise. The main reasons for this approach to continuants are not only the high variability of final fricative lengthening but also the fact that in the final segments the signal intensity drops below the level of perceptibility (cf. Kochanski et al. 2005) at various stages of articulation in individual speakers.

All the recordings and measurements were taken by the author. Randomly chosen units were measured again six months later in order to verify the reliability of procedures. Measurement discrepancies did not exceed 5% of the original value with respect to individual units.

4.3.3.1 General domain-related principles

The general principles concerning the division of the measured phrases introduced in section 4.3.2 are summarised below:

1. An IP consists of an optional prehead, an optional head, an obligatory nucleus and an optional tail.
2. A prehead is formed by phrase-initial unstressed syllables. It is treated as a unit of IP structure on a par with a foot.

3. The head is left-headed. It begins with the lexically stressed (usually pitch-accented) syllable of the first content word and stretches up to the last pitch-accented syllable (beginning the nucleus).
4. The nucleus is left-headed. It begins with the last pitch-accented syllable in the IP and also comprises the following unstressed syllables.
5. If a content word follows the nucleus, it is treated as a separate foot called the tail.
6. A head can be composed of one or more feet. Each foot belonging to a head is marked with an "H".
7. A foot is left-headed. It begins with a lexically stressed syllable and stretches up to the next stressed syllable (cf. Abercrombie 1967, Tatham and Morton 2001, Kim and Cole 2005). Consequently, cross-word feet are allowed.
8. A syllable is constituted by any phonologically underlying vowel. Consequently, even an entirely non-vocalic realisation of an unstressed syllable does not disqualify the unit.
9. Centring diphthongs are considered monosyllabic.
10. Syllabification is based on the maximal onset principle (Kahn 1976, Clements and Keyser 1983).
11. All higher-level domain boundaries are at the same time syllable boundaries.

4.3.3.2 Specific boundary-related principles

The above principles are supplemented by detailed criteria of acoustic signal segmentation:

1. Syllable boundaries are set at the least sonorant points of the speech signal, favouring the following:
   — plosive/affricate release burst,
   — frication onset.
2. Segmental boundaries are set according to standard criteria (cf. Peterson and Lehiste 1960, Turk et al. 2006) at the following points:
   — plosive/(semi)vowel: onset of clear vocalic formant structure (positive VOT counted as part of the consonant),
   — fricative (or affricate)/vowel: onset of clear vocalic formant structure,
   — nasal/vowel: onset of vocalic formant structure (moment of closure release),
   — vowel/plosive (or affricate): offset of vocalic formant structure (moment of occlusion),
   — vowel/fricative: onset of frication noise,
   — vowel/nasal: offset of vocalic formant structure (moment of oral occlusion).

4.3.3.3 Additional remarks

Clusters of vowels and approximants are not further divided. Because a wide range of phone combinations are involved, the measurement criteria referring to consonant constrictions are preferred to onsets and offsets of voicing (cf. Turk et al. 2006).

Aspiration and frication are treated as one type of segment. Because aspiration is a feature which may magnify speech timing differences between English and Polish speakers, plosives are divided into an occlusion phase and a burst phase including aspiration or frication noise. The landmark for segmentation is the moment of release burst. Wherever multiple bursts occur, especially in the case of velar stops, the most salient one is counted.

Since native and non-native speech is compared, the analysis inevitably includes various realisations of phonemes and phoneme combinations. However, no conclusions concerning modal segment durations will be drawn and the duration discrepancies caused by articulation differences are valuable observations for further pedagogical implications.

Segmentation has been performed by hand on the basis of waveform and spectrogram inspection with interactive playback. All segmentation and measurements have been performed by the author. No measures reported in this study have been normalized or calibrated.

4.3.4 Statistical analysis

In data analysis, we regard subject group statistics, reflected in mean values with standard deviation or medians in cases where outliers might strongly influence the means. If necessary, depending on the purpose of analysis, descriptive statistics are supported by one-way ANOVA measures where the influence of a relevant factor on the results is not obvious, and paired t-tests in order to show developmental tendencies in Polish learners.

With respect to vowel duration variability, syllable duration variability and foot-timing, individual speakers' scores are also provided. Since large individual variation is a typical feature of speech prosody (e.g. Grabe 2002), rankings of individual speakers additionally illustrate the relevance of differences in general tendencies shown by the two groups of speakers.

Both absolute values and relative timing proportions are usually needed to make the results fully informative. Otherwise controlling for speech rate and its relations to timing at various levels of the prosodic hierarchy would not be possible.

Chapter 5

Absolute and relative duration of vocalic segments — results and discussion

5.1 Introduction

A comprehensive study of pronunciation faces natural difficulties especially where small speech units are to be analysed in large, varied contexts. The duration of vowels which appear in the text depends on numerous factors described in the first three chapters. Hence the gathered data, classified according to the main length determinants, may not be enough to draw general conclusions concerning the relations within the phonological system of English, but on the other hand, they can show significant differences between the two groups of speakers, even despite large individual variation, typical of both native and non-native language performance.

The present chapter displays the results referring to the timing of vocalic segments. Apart from the figures representing absolute duration of speech units expressed in milliseconds, descriptive statistics illustrate the differences between the native speakers' and Polish learners' performance. Because of large individual variability of the values, duration ranges observed in the native speakers' pronunciation are also displayed as reference for the Polish subjects' results, and the numbers of responses out of the native range are indicated. Moreover, where applicable, one-way ANOVA or paired t-tests specify the magnitude of relationship between the factor of group (EN — English speakers, PL1 — Polish speakers, first recording, PL2 — Polish speakers, second recording) and the duration of measured units.

The total number of vowel samples to be measured was 988 (26 vowels multiplied by 12 native English respondents and twice 13 Polish respondents). Nine tokens were disqualified owing to text distortions in the subjects'

performance; therefore the data comprise 979 items. Table 5.1 displays the number of tokens collected and analysed in the study, with respect to their intrinsic properties and context. The factor of prominence is investigated in two conditions, viz. as non-nuclear phrasal accent in the heads (H), and as nuclear accent (N). In some contexts, tails (T) must serve as reference counterparts for phrase-final nuclear feet in observing the effects of phrasal accent, because all heads are phrase-medial (or initial) by definition. Foot complexity comprises

Table 5.1. The number of tokens in particular categories

Factor / Group	Total (26)	Vowel type				Position			Prominence		Foot complexity			Disqualified items
		D (6)	L (5)	A (3)	S (12)	f (7)	w (4)	m (15)	N (12)	H (11)	1 (9)	2 (12)	3-4 (4)	
PL1	334	77	64	39	154	90	52	192	154	142	115	155	52	4
PL2	336	77	64	39	156	90	52	194	155	142	116	156	52	2
EN	309	71	60	36	142	83	48	178	143	131	106	143	48	3

Abbreviations: PL1 — Polish speakers, first recording; PL2 — Polish speakers, second recording; EN — English speakers; D — diphthong; L — long vowel; A — "ash"; S — short vowel; f — prepausal stressed syllable; w — stressed syllable of a prepausal (trochaic) word; m — utterance medial position; N — nuclear syllable; H — stressed head syllable.

three classes: monosyllabic, disyllabic (trochaic) and 3-4-syllabic. The two most complex groups have been combined because of a smaller amount of data and observed similar durations of their constituent syllables. The tested vowels appear in prepausal stressed syllables (f), the stressed syllable of prepausal trochaic words (w) and in phrase-medial words (m). Vocalic categories: dipththong, long vowel, "ash" and short vowel are represented by D, L, A and S, respectively. Additionally, the following consonant can be marked as lenis (+) or fortis (–).

5.2 Stressed vowel duration

The complex interaction of multiple factors determining the length of particular speech units makes it necessary to present the mean vowel durations in individual measured items. The results are displayed in Table 5.2. Considering the diversity of contexts, the mean durations and duration ranges of vowel classes must not be treated as any kind of modal values. They only serve the purpose of across-group comparison.

Apart from the reservations made in the introduction, we must bear in mind that even apparently similar contexts may prove hardly comparable owing to a variety of factors which pertain to all levels of language structure. For instance, the diphthong and its context in ***f****ind my hat* and *b****uy*** *some new* ***gow****ns* is in both cases characterised as "mDH2", but we cannot rule out

Table 5.2. The duration of English stressed vowels

Carrier word	Vowel type and position	PL1 (N=13)	PL2 (N=13)	EN (N=12)	EN range	PL1 out of EN[a]	PL2 out of EN[a]	ANOVA[b]
divine	fDN1+	220	208	271	203—383	**4–**	**5–**	**8.6–**
gowns	fDN1+	303	262	292	223—376	0	2–	**3.8–**
gowns	fDT1+	252	228	263	205—348	1+, 3–	4–	1.6
shouting	mDN3–	134	123	137	93—174	1–	2–	1.0
buy	mDH2–	178	143	131	92—154	**8+**	**7+**	**7.7+**
find	mDH2+	112	97	131	98—176	**3–**	**9–**	**5.0–**
mood	fLN1+	153	165	206	130—287	**4–**	**2–**	**5.9–**
parties	wLN2–	155	150	166	124—211	1–	2–	1.2
gorgeous	wLN2+	173	164	169	128—207	1+, 1–	0	0.4
thought	mLH1–	142	102	91	56—123	**6+**	**2+**	**8.1+**
started	mLH2–	113	102	105	64—*186*	0	1–	0.75
hat	fAN1–	142	126	158	94—205	1–	3–	2.1
bad	mAH1+	140	143	178	116—248	**5–**	**2–**	**3.9–**
had to	mAH2+	79	75	74	44—95	1+	2+, 1–	0.2
son	fSN1+	133	108	127	94—195	**0**	**6–**	**2.6**
gone	fSN1+	166	148	154	105—206	2+, 1–	1+, 1–	1.1
Cinders	wSN2+	70	68	64	45—105	1+	1+, 1–	0.3
g**o**dmother	mSN1+	93	90	89	66—115	1+, 1–	1–	0.2
enough	mSN1–	115	108	103	74—146	1+	1–	1.0
godm**o**ther	wST2+	118	111	86	64—118	**6+**[c]	**4+**	**12.4+**
Cinders	wST2+	57	58	58	27—84	0	1+	0.01
mother	mSH2+	120	107	86	59—99	**11+**	**9+**	**8.25+**
Cinders	mSH2+	62	59	55	32—90	0	0	0.75
sisters	mSH3–	56	58	51	27—69	0	3+	0.8
looking	mSH4–	62	55	54	25—87	1+	0	1.1
honour	mSH4+	123	120	96	71—121	**8+**	**5+**	**4.1+**

[a] Number of PL responses out of EN range ("+" — longer, "–" — shorter)
[b] ANOVA F-value illustrating the effect of group on vowel duration. Statistically significant values (at p=.01) in bold. "+" — longer PL means, "–" — longer EN means
[c] One response disqualified (N=12)

the possibility of timing differences caused by the large discrepancy in the overall complexity of the two phrases.

5.2.1 Intrinsic vowel length

Despite the temporal variability of vocalic segments, attempts have been made (e.g. Wiik 1965) to establish the length relations between long and short vowels of English and other languages which use this phonological contrast (cf. section 1.2.1.2). Mean English vowel durations in accented monosyllables obtained in Wiik's study are referred to by Cruttenden (2001: 96). The values referring to the respondents' realisation of the four categories distinguished earlier are displayed in Table 5.3.

Table 5.3. Mean durations (ms) of particular vowel types

Vowel class / Group	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%)
EN	203	147	137	85	130 (SD=72=55%)

Abbreviations: D — diphthong; L — long vowel; A — "ash"; S — short vowel.

The first important observation is the comparable articulatory rate in both groups of speakers, suggested by similar mean vowel durations. The unexpected shorter value in PL2, possibly a result of the learners' belief that faster reading is tantamount to more fluent and generally more proficient speech, will be investigated in further sections devoted to higher-level prosodic domains.

All groups varied the duration of vowels, showing a general tendency to observe the expected relations between particular classes. However, the range of means for individual types shows more difference between short and long vowels in native English pronunciation.

The proportions between mean class durations obtained by dividing each value by mean short vowel duration are as follows:

Table 5.4. Mean duration proportions between vowel types

Vowel class / Group	D	L	A	S
PL1	2	1.5	1.2	1
PL2	1.9	1.5	1.3	1
EN	2.4	1.7	1.6	1

Abbreviations: as in Table 5.3.

These data and the highest standard deviation from the grand mean vowel duration (see Table 5.3) indicate more flexibility of vowel length in native English speech, which supports Hypothesis 1. It is also interesting that Polish speakers do vary vowel length, although hardly any tendency to achieve more native-like proportions has been observed in the results of PL1 and PL2. However, the tentative general observation concerning mean vowel class durations needs to be verified by investigating the impact of contextual factors.

5.2.2 Final lengthening

The final position is one of the strongest duration determinants recognised in world languages. Because both the syllable and the word have been found to be the domains of final lengthening, three conditions have been distinguished. Table 5.5 displays mean vowel durations in stressed prepausal (phrase-final) syllables (fN+fT), stressed syllables of prepausal trochaic words (wN+wT), and accented syllables followed by at least one more content word within the phrase (mN+mH[1]).

Table 5.5. Vowel duration in phrase-final and non-phrase-final syllables

Vowel position / Group	f	w	m	f:m	w:m
PL1	195	114	110	**1.77**	**1.04**
PL2	177	110	99	**1.79**	**1.11**
ENG	209	114	98	**2.13**	**1.16**

Abbreviations: f — prepausal stressed syllable; w — stressed syllable of a prepausal (trochaic) word; m — utterance medial position.

Both groups lengthened the final syllable nuclei considerably, whereas in the stressed syllables of phrase-final trochaic words they were only slightly longer in relation to medial positions. Moreover, the effect of final lengthening is stronger in native speakers (113% lengthening) than in Polish respondents (77—79% lengthening), with no significant difference between the two recording sessions.

5.2.3 Accentual lengthening

Similar results appeared in comparing accented vowels in nucleus and head positions (Table 5.6), which may reflect additive effects of accent type and

[1] Although redundant, “m” is placed before “H” for better clarity of presentation.

phrase position since, unlike many nuclear syllables, heads never "benefit" from final lengthening.

Table 5.6. Vowel duration in nuclei and heads

Group \ Accent type	N	H	N:H
PL1	155	108	**1.4**
PL2	143	96	**1.5**
EN	162	96	**1.7**

Abbreviations: N — nuclear syllable; H — stressed head syllable.

In order to observe the effect of nuclear accent alone, without the strong interference of final lengthening, we calculate the mean duration of non-phrase-final syllable nuclei in *...that they had en**ou**gh gowns*, *...so they started sh**ou**ting at Cinders* and *It was her fairy g**o**dmother* and compare them to the results for stressed vowels in phrase-final trochaic nuclei: *going to p**ar**ties*, *Prince William was g**or**geous* and *...but everyone called her Cinders* (Table 5.7). The fN values are provided for reference.

Table 5.7. Group mean vowel duration in phrase-final nuclear syllables (fN: *divine*, *gowns*, *mood*, *hat*, *son*), non-phrase-final nuclear syllables (mN: *enough*, *shouting*, *godmother*) and the accented (nuclear) syllables of phrase final trochaic words (wN: *parties*, *gorgeous*, *Cinders*) in relation to mean gone mH (11 words — cf. Tables 5.1 and 5.2) vowel duration

Group \ Position	fN	wN	mN	mH	fN:mN	wN:mN	mN:mH	fN:mH	wN:mH
PL1	186	132	114	108	**1.63**	**1.16**	**1.06**	1.72	1.22
PL2	169	127	107	96	**1.58**	**1.19**	**1.11**	1.76	1.32
EN	201	133	110	96	**1.83**	**1.21**	**1.15**	2.09	1.39

The figures display the strongest impact of final lengthening on vowel duration in both groups of speakers. This is also the factor that brings about the largest difference between Polish and English speakers' production. Needless to say, the figures show timing relations between different groups of speakers, but they do not represent actual relations within the phonological system of English since they refer to different sets of sounds in various phonological contexts. The effects of accent and position can be directly observed only in a limited number of words in the text. Table 5.8 shows the duration of stressed vowels in *Cinders* (wN, mH, wT), *gowns* (fN, fT), *mother* (mH, wT) and in words with the same vowels followed by the same consonants: *divine/find*, *parties/started*.

Table 5.8. The effect of accent and position on vowel duration. Mean durations (M) in ms, duration proportions in bold

Vowel \ Group	PL1		PL2		EN	
	M	proportion	M	proportion	M	proportion
C**i**nders (wN)	70	**wN:mH=1.13**	68	**wN:mH=1.15**	64	**wN:mH=1.16**
C**i**nders (mH)	62	**mH:wT=1.09**	59	**mH:wT=1.02**	55	**mH:wT=0.95**
C**i**nders (wT)	57	**wN:wT=1.23**	58	**wN:wT=1.15**	58	**wN:wT=1.10**
g**ow**ns (fN)	303		262		292	
g**ow**ns (fT)	252	**fN:fT=1.20**	228	**fN:fT=1.15**	263	**fN:fT=1.11**
div**i**ne (fN)	220		208		271	
f**i**nd (mH)	112	**fN:mH=1.96**	97	**fN:mH=2.14**	131	**fN:mH=2.07**
p**ar**ties (wN)	155		150		166	
st**ar**ted (mH)	113	**wN:mH=1.37**	102	**wN:mH=1.47**	105	**wN:mH=1.58**
m**o**ther (mH)	120		107		86	
(god)m**o**ther (wT)	118	**mH:wT=1.02**	111	**mH:wT=0.96**	87	**mH:wT=0.99**

The overall results presented in Tables 5.2—5.7 support the hypothesis that the effects of both accentual lengthening and (especially) final lengthening on vowel duration are stronger in native English. The data in Table 5.8 also suggest that the differences in final lengthening effects are more evident than those in the effects of nuclear accent. Moreover, final lengthening is much weaker in an accented or stressed syllable if it begins a phrase-final trochee (wN, wT) than in a phrase-final accented syllable (fN). If final lengthening affects vowel duration more than accentual lengthening, then the larger N:T ratios in Polish learners' performance (*Cinders*, *gowns*) do not contradict the claim that the effects of both processes are stronger in native English speech.

Furthermore, the effect of weaker phrasal accent (H) appears comparable to the effect of word stress in unaccented positions (wT). The durations of vowels in *Cinders* and *mother* are comparable in both groups in both contexts. This also questions the effects of final lengthening on the stressed vowel of trochaic tails.

5.2.4 Foot complexity effect

The durations of accented vowels in monosyllabic up to tetrasyllabic feet appearing in the text are presented in Table 5.9.

These results must be regarded with caution because the main difference observed between English and Polish speakers is in the length of monosyllables. Since this is the group that "benefits" the most from final lengthening, the proportions may only reflect the effect of final lengthening. In order to eliminate this effect, the same relations with respect to stressed vowels

Table 5.9. Foot complexity and vowel duration

Group \ Foot complexity	1	2	3-4	1:3-4	2:3-4
PL1	173	117	94	**1.8**	**1.2**
PL2	159	107	89	**1.8**	**1.2**
EN	187	107	84	**2.2**	**1.3**

in non-phrase-final words have been computed. Each group includes three examples:
1: b**a**d, g**o**d(mother), en**ou**gh.
2: h**a**d to, m**o**ther, C**i**nders.
3-4: s**i**sters, l**oo**king, h**o**nour.

The modified means and proportions are shown in Table 5.10.

Table 5.10. Foot complexity and the duration of stressed vowels in non-phrase-final words

Group \ Foot complexity	1	2	3-4	1:3-4	2:3-4
PL1	117	87	80	**1.46**	**1.09**
PL2	114	80	78	**1.46**	**1.03**
EN	124	71	67	**1.85**	**1.06**

Although less conspicuously than in Table 5.9, the figures still indicate a strong relationship between foot complexity and accented vowel duration, especially in the native English group. However, a revision of particular examples (Table 5.11) reveals the uneven influence of individual items on

Table 5.11. Vowel duration in stressed syllables

Vowel in	Position	PL1 (N=13)	PL2 (N=13)	EN (N=12)	PL1:EN	PL2:EN
bad	mAH1+	140	143	**178**	**0.79**	**0.8**
g**o**dmother	mSN1+	93	90	89	1.04	1.01
enough	mSN1–	115	108	103	1.12	1.05
had to	mAH2+	79	75	74	1.07	1.01
mother	mSH2+	120	107	**86**	**1.4**	**1.24**
Cinders	mSH2+	62	59	55	1.13	1.07
sisters	mSH3–	56	58	51	1.1	1.14
looking	mSH4–	62	55	54	1.15	1.02
honour	mSH4+	123	120	**96**	**1.28**	**1.25**

Abbreviations: A — "ash"; S — short vowel; m — utterance medial position; H — stressed head syllable; 1, 2, 3, 4 — number of foot constituent syllables; +/– — pre-lenis/pre-fortis vowel position.

overall proportions. Significant (>20%) differences between Polish and English subjects are shown in bold.

Interestingly, the vowels in *god(mother)* and *enough*, which bear nuclear accent, do not significantly vary across the groups. Because the evidence is too scanty to decide whether the differences in *bad*, *mother*, and *honour* depend on the intrinsic properties of the vowels or the prosodic properties of the context, the problem must be left unresolved.

5.2.5 Pre-fortis clipping

The study that Cruttenden (2001) cites ranks pre-lenis vowels as the longest, pre-fortis the shortest, and pre-nasal ones as intermediate, while the present research has found the nasal codas to coincide with the longest vowel durations. This observation pertains to vowels distributed in non-symmetrical contexts, which prevents conclusions concerning systematic relations, but the effect of coda type has been similar in both groups, again with slightly more difference between particular conditions in the native speakers' performance (cf. Table 5.12).

Table 5.12. The effect of coda on vowel duration

Vowel preceding_ / Group	_nasal	_voiced	_unvoiced	_voiced:_unvoiced	_nasal:_unvoiced
PL1	149	125	115	**1.09**	**1.30**
PL2	135	122	103	**1.18**	**1.31**
EN	150	127	108	**1.18**	**1.39**

It is noteworthy that after training the Polish learners increased their mean length difference between pre-lenis and pre-fortis vowels to a native-like proportion.

5.2.6 Within-phrase relations

All of the above figures compare mean absolute durations of vowels in different types of contexts. They give a general insight into absolute vowel durations with respect to their intrinsic properties and context. This section is devoted to the length of vowels in stressed syllables within the same phrase in order to provide a better illustration of the rhythmic differences between the groups. Table 5.13 shows mean vowel length proportions in several head-to-nucleus and nucleus-to-tail relations.

Table 5.13. Within-phrase stressed vowel quantity proportions (duration means in ms in parentheses)

Group / Tested units	PL1	PL2	EN
VH:VN			
honour:son (SH:SN)	0.9 (123:133)	1.1 (120:108)	0.8 (96:127)
thought:divine (LH:DN)	0.6 (142:220)	0.5 (102:208)	0.3 (91:271)
bad:mood (AH:LN)	0.9 (140:153)	0.9 (143:165)	0.9 (178:206)
buy:gowns (DH:DN)	0.6 (178:303)	0.55 (143:262)	0.45 (131:292)
started:shouting	0.8 (113:134)	0.8 (102:123)	0.8 (105:137)
find:hat (DH:AN)	0.8 (112:142)	0.8 (97:126)	0.8 (131:158)
sisters:gone (SH:SN)	0.3 (56:166)	0.4 (58:148)	0.3 (51:154)
VN:VT			
enough:gowns (SN:DT)	0.5 (115:252)	0.5 (108:228)	0.4 (103:263)
shouting:Cinders (DN:ST)	2.3 (134:57)	2.1 (123:58)	2.4 (137:58)
god:mother (SN:ST)	0.8 (93:118)	0.8 (90:111)	1 (89:87)

None of the EN head:nucleus vowel ratios is higher than the corresponding PL proportion and four of the seven examples show relatively longer duration of the N vowel. N:T ratios show a similar tendency, where the main phrasal accent makes the vowel relatively longer in native English pronunciation. The opposite relation in *enough:gowns* can easily be explained by the final lengthening effect overriding the power of accentual lengthening in English speakers. Final lengthening is attenuated by the unstressed final syllables in *Cinders* and *(god)mother.*

The figures displayed in Table 5.13 are still based on mean values without accounting for individual speakers' speech rhythm characteristics. Table 5.14 presents data calculated from duration proportions in individual speech samples, indicating the mean proportions, individual proportion ranges for both groups and the number of Polish learners whose vowel length proportions were out of native English ranges.

This method of presentation suggests even more strongly a tendency in native English speakers to highlight the nuclear accent by increasing the relative length of the relevant vowel.[2] It is possible that the most conspicuous differences have been enhanced by other factors, such as syntactic complexity, motivating a stronger verb-object boundary in Polish learners in longer phrases (*thought:divine*, *buy:gowns*) in comparison to simpler, shorter structures (*bad:mood*, *find:hat*), where PL ratios resembled native proportions

[2] Accented vowel length will be discussed in relation to higher level units in the subsequent sections.

despite considerable discrepancies in mean durations of individual tested vowels.

Table 5.14. Within-phrase stressed vowel proportions in individual speakers

Group / Tested units	PL1 mean (min—max)	PL1 out of EN range	PL2 mean (min—max)	PL2 out of EN range	EN mean (min—max)
VH:VN					
honour:son (SH:SN)	0.9 (0.7—1.2)	3	1.2 (0.7—2.2)	8	0.8 (0.4—1.0)
thought:divine (LH:DN)	0.7 (0.3—1.0)	10/12	0.5 (0.3—0.8)	5/12	0.3 (0.2—0.5)
bad:mood (AH:LN)	0.9 (0.6—1.2)	0	0.9 (0.4—1.8)	1	0.9 (0.4—1.2)
buy:gowns (DH:DN)	0.6 (0.4—1.0)	5	0.55 (0.4—0.7)	5	0.45 (0.3—0.6)
started:shouting (LH:DN)	0.9 (0.6—1.7)	1	0.9 (0.4—1.4)	4	0.8 (0.6—1.2)
find:hat (DH:AN)	0.8 (0.6—1.2)	0	0.8 (0.5—1.3)	1	0.9 (0.6—1.6)
sisters:gone (SH:SN)	0.4 (0.2—0.6)	1	0.4 (0.2—0.6)	2	0.3 (0.2—0.5)
VN:VT					
enough:gowns (SN:DT)	0.5 (0.3—0.7)	2	0.5 (0.3—0.7)	2	0.4 (0.3—0.6)
shouting:Cinders (DN:ST)	2.4 (1.9—3.3)	0	2.3 (1.3—4.2)	0	2.6 (1.1—5.0)
god:mother (SN:ST)	0.8 (0.5—1.2)	7/12	0.8 (0.6—1.1)	6	1 (0.8—1.2.0)

5.2.7 Summary

The results concerning native and non-native English vowel duration indicate the complexity of interaction of the duration determinants already discussed in the theoretical part of this dissertation, and the need to verify many of the suggested tendencies in more detailed studies.

The conclusions based on the gathered data are formulated as follows:

1. Mean vowel length in the read speech of native English speakers and Polish learners is comparable (130—133ms), although more advanced learners show a tendency to accelerate (122ms) especially in shorter, syntactically simpler phrases. This, however, sometimes leads to less natural duration proportions than in the first recording.
2. Standard deviations referring to the means mentioned above and differences in mean durations of vowels classified according to their intrinsic length (see section 5.2.1) indicate a more clearly specified distinction between short and long vowels, and diphthongs in native English pronunciation.
3. The scale of variation is expressed in the observation that the Polish learners make their English long vowels approximately 1.5 times longer and diphthongs twice as long as the short vowels. The corresponding proportions in native speakers amount to 2.4 and 1.7, respectively, which yields 40% more difference between particular classes.

4. Vowels in nuclear syllables are generally longer in native English pronunciation not only in terms of timing relations but even in absolute values.
5. Polish learners do not increase their English vowel duration "flexibility" after seven months of training.

The above conclusions are supported by Table 5.15, which presents the data contained in Table 5.2, rearranged in order to show a ranking of tested words according to PL2:EN mean vowel duration ratio. This ratio is preferable to PL1:EN because we believe that differences which remain after training indicate more serious problems than those signalled by some even larger initial discrepancies that disappear in the course of training.

Columns 7—8 give the number of responses out of the EN range (shorter="–", longer="+") specified in column 6.

Table 5.15. Ranked POL:ENG vowel duration proportions in individual tested words

Vowel in	Position	PL1 (N=13)	PL2 (N=13)	EN (N=12)	EN (min—max)	PL1 out	PL2 out	PL1:EN	PL2:EN
find	mDH2+	112	97	131	98—176	**3–**	**9–**	**0.9**	**0.7**
divine	fDN1+	220	208	271	203—383	**4–**	**5–**	**0.8**	**0.8**
mood	fLN1+	153	165	206	130—287	**4–**	**2–**	**0.7**	**0.8**
hat	fAN1–	142	126	158	94—205	1–	3–	**0.9**	**0.8**
bad	mAH1+	140	143	178	116—248	**5–**	**2–**	**0.8**	**0.8**
gowns	fDN1+	303	262	292	223—376	0	2–	**1.0**	**0.9**
gowns	fDT1+	252	228	263	205—348	1+, 3–	4–	**1.0**	**0.9**
shouting	mDN3–	134	123	137	93—174	1–	2–	**1.0**	**0.9**
parties	wLN2–	155	150	166	124—211	1–	2–	**0.9**	**0.9**
son	fSN1+	133	108	127	94—195	**0**	**6–**	**1.1**	**0.9**
gorgeous	wLN2+	173	164	169	128—207	1+, 1–	0	**1.0**	**1.0**
started	mLH2–	113	102	105	64—*186*	0	1–	**1.1**	**1.0**
had to	mAH2+	79	75	74	44—95	1+	2+, 1–	**1.1**	**1.0**
gone	fSN1+	166	148	154	105—206	2+, 1–	1+, 1–	**1.1**	**1.0**
g**o**dmother	mSN1+	93	90	89	66—115	1+, 1–	1–	**1.0**	**1.0**
Cinders	wST2+	57	58	58	27—84	0	1+	**1.0**	**1.0**
buy	mDH2–	178	143	131	92—154	**8+**	**7+**	**1.4**	**1.1**
thought	mLH1–	142	102	91	56—123	**6+**	**2+**	**1.6**	**1.1**
Cinders	wSN2+	70	68	64	45—105	1+	1+, 1–	**1.1**	**1.1**
enough	mSN1–	115	108	103	74—146	1+	1–	**1.1**	**1.1**
Cinders	mSH2+	62	59	55	32—90	0	0	**1.1**	**1.1**
sisters	mSH3–	56	58	51	27—69	0	3+	**1.1**	**1.1**
looking	mSH4–	62	55	54	25—87	1+	0	**1.2**	**1.1**
mother	mSH2+	120	107	86	59—99	**11+**	**9+**	**1.4**	**1.2**
honour	mSH4+	123	120	96	71—121	**8+**	**5+**	**1.3**	**1.2**
godm**o**ther	wST2+	118	111	86	64—118	**6+**	**4+**	**1.4**	**1.3**

The ranking proves that shorter mean PL durations never concern short vowels, whereas longer realisations always pertain to short vowels. The two exceptions of *buy* and *thought*, where we observe the final lengthening effect in Polish learners (especially PL1) due to phrasing reorganisation, have already been mentioned in section 5.2.6.

The influence of contextual factors, such as pre-fortis clipping, foot complexity, accentual lengthening and (especially) final lengthening has also been suggested but, owing to their interaction, it would be rather unwise to attempt to isolate any of these variables while controlling the others, given the amount of data gathered for the purpose of this study. Therefore no conclusions are formulated in this respect on account of insufficient evidence.

5.3 Unstressed vowel duration

Reduced vowels were measured in unstressed syllables of monosyllabic function words and trochaic content words that appeared in the text in phrase-medial and phrase-final positions. The total number of vowel samples to be measured was 760 (20 vowels multiplied by 12 native English respondents and twice 13 Polish respondents). Five tokens were either disqualified owing to text distortions or missing from the recordings; therefore the data comprise 755 items. Apart from word category and phrase (final/non-final) position, we consider foot type (P, H, N, T), foot complexity and the voicing of the following consonant as possible duration determinants.

5.3.1 Unstressed vowels in monosyllabic function words

Table 5.16 presents mean (M) unstressed vowel duration measured in Polish learners' and native English speakers' performance. The tested units are ranked in rising order according to mean EN durations.

Table 5.16. Mean unstressed vowel duration (ms) in function words

Schwa in	Foot type	PL1	PL2	EN
1	2	3	4	5
going t**o** parties	H3–	34	30	1
buy s**o**me	H2+	56	34	7
had t**o** do	H2+	75	37	13
wanted t**o** buy	H3+	43	40	16
b**u**t everyone	P1–	82	64	42
shouting **a**t Cinders	N3–	73	81	45

cont. tab. 5.16

1	2	3	4	5
in **a** bad mood	P4+	56	44	46
there was **a** girl	P3+	73	64	48
sisters h**a**d gone	H3+	61	60	49
dreamed **o**f	H2+	106	68	54
When h**er** sisters	H2–	82	84	63
it was h**er**	P3–	129	89	72
called h**er** Cinders	H2+	119	89	78
M		**76**	**60**	**41**
SD		41.2 (55%)	33.6 (56%)	28.4 (69%)

The grand means for PL1, PL2 and EN indicate more radical vowel reduction in English speakers, regardless of the context. However, duration means for individual items in native speakers' performance vary from 1—16ms in *to* and *some*, where many speakers do not vocalise the transition between consonants, to 63—78ms in *her*, where the longest schwa is pronounced.

The largest difference between native and non-native production has been observed in *to*. Polish learners do not tend to link /t/ to the following stop without pronouncing a vowel, which is a typical articulatory habit of native English speakers especially if /t/ is followed by another voiceless consonant across the reduced vowel (cf. Porzuczek 2010a). This is illustrated in Figure 5.1.

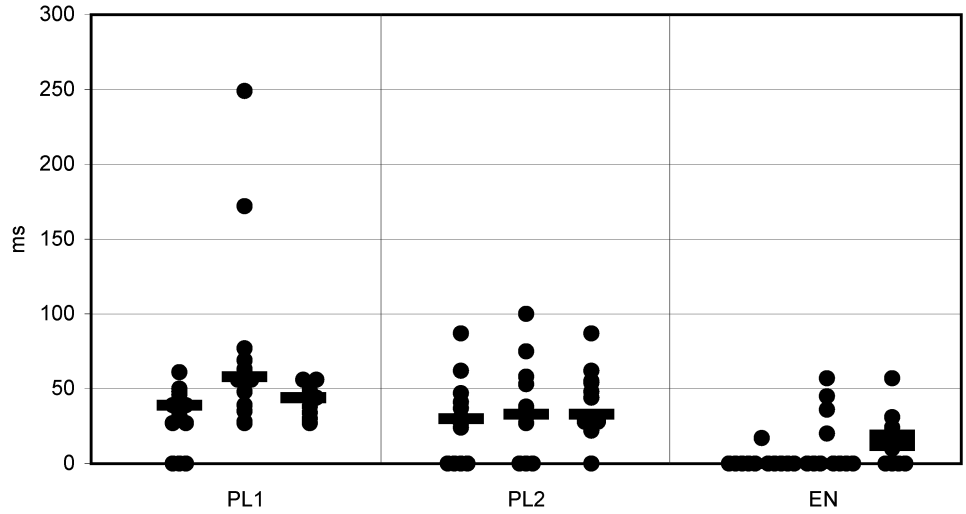

Figure 5.1. Vowel duration in ***to*** *parties* (left), ***to*** *do* (middle) and ***to*** *buy* (right). Horizontal bars indicate group medians

Voiceless transition was used in three cases (all in *to parties*) by Polish learners before the training. After the training, there were 4 such cases in *to parties*, 3 in *to do* and 1 in *to buy*. These responses and no outliers lowered the mean schwa duration in Polish learners. The vocalised realisations, however, were not significantly shorter in the second recording session.

The strong influence of the number of voiceless schwas in the data makes it pointless to use statistic measures to analyse mean values.

While the phonological structure of *to* made its vowel the shortest in the part of the study devoted to unstressed vocalic segments, the schwa in *her* turned out to be the longest among all items in this category (Figure 5.2).

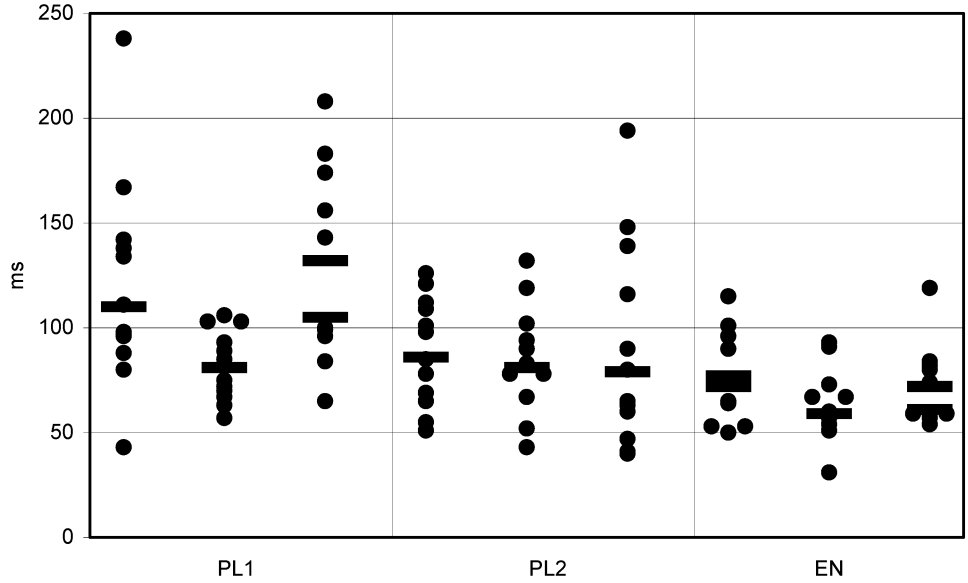

Figure 5.2. Vowel duration in *...called h**er** Cinders* (H; left), *When h**er** sisters...* (middle) and *It was h**er** fairy godmother* (P; right).[3] Horizontal bars indicate group medians

In all three positions, the English speakers reduced the vowel to 50—100ms. The context has not been found to significantly influence vowel length in PL2 and EN. A statistically significant variation (ANOVA, p=.011, F(2,35)) appeared in PL1, who consistently kept the reduced vowel in *When her* shorter or (in three cases) slightly longer than 100ms, while often failing to shorten it in the native-like manner in the other two contexts.

A comparison of vowel length in individual contexts shows a statistically significant difference between PL1 and EN in all three contexts, especially *It was her* (P).

[3] Data reported in Porzuczek (in press).

Table 5.17. The variation in ***her*** duration in relation to group and context

Schwa in	PL1:EN ANOVA, F(1, 23)	PL2:EN ANOVA, F(1, 23)	PL1:PL2 paired t-test
...called h**er**	**p=.013**	p=.259	**p=.015**
When h**er**...	**p=.011**	**p=.028**	p=.424
It was her...	**p<.001**	p=.234	**p=.019**

The statistics also indicate a considerable shortening of the vowel in two contexts, resulting in native-like durations in PL2. The duration of schwa in *When her...* remained significantly longer in PL2, and no substantial reduction occurred in comparison to PL1. However, it must be noted that PL1 mean schwa duration in this context (81.2ms) was fairly close to the grand mean of EN schwa duration in *her* (71.1ms), and the Polish learners only reduced the other two context means in their second performance. The lower mean EN vowel duration in *When her...* may still be regarded as accidental, given that no context effect has been observed in native speakers' production of the reduced vowel in question.

The unstressed vowels in the remaining function words (*a*, *a*, *had*, *but*, *at*, *of*, and *some*) were pronounced by the native speakers with vowel duration means within a rather narrow range of 42—54ms, which may be thus tentatively suggested as typical schwa duration in English read speech. It is worth noting that this narrow range has been obtained from various phrase positions, including 3 preheads, 2 heads, and one phrase-medial nucleus. As shown in Table 5.16, the duration of schwa in *her*, one function word that appeared in two different foot types, did not differ between P and H either.

An analysis of individual response dispersion (Figure 5.3) in native speakers shows only two cases of schwa longer than 80ms, and only one shorter than 20ms, except for 9 realisations of *some* without a vowel. The Polish learner's problem with vowel reduction is well illustrated by the fact that the 80ms barrier was exceeded in 30 out of 90 responses and each student made at least one schwa "too long" in their first recording. In the second session, three students kept all the seven schwas below the 80ms limit, and there were 14 longer responses in the total of 90.

Apart from the problems with reduction in contexts where the vowel is not vocalised by a significant proportion of native speakers (*to* and *some*), Polish learners do not employ native-like timing in a number of other contexts. The length of the indefinite article, 22—67ms in native speech, exceeded that range in 11/26 cases in PL1. Two more responses with no acoustic traces of *a* must be regarded as "foreign" as well because none of the native speakers dropped the article or even made it shorter than 20ms. In the PL2 session,

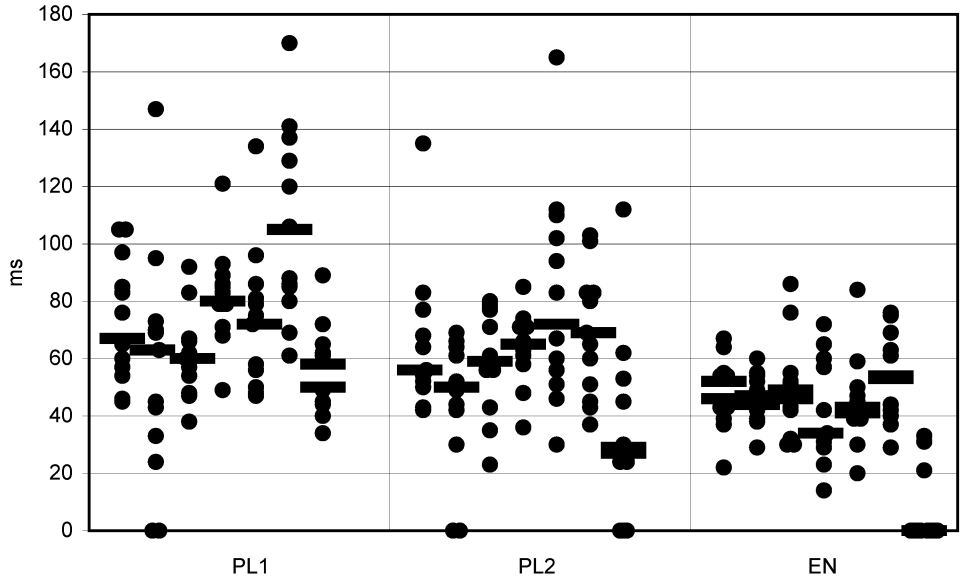

Figure 5.3. Unstressed vowel duration in (from left to right) *there was A girl*, *They were in A bad mood*, *When her sisters HAD gone*, *BUT everyone called her Cinders*, *so they started shouting AT Cinders*, *They dreamed OF wedding bells*, *They'd wanted to buy SOME new gowns*. Group medians indicated by horizontal bars

again two speakers failed to produce the word, and still five responses were longer than the longest native English schwa.

One-way ANOVA, F(1,23), confirms a significant group-related difference in the duration of *a* in *there was A girl* (PL1:EN — p=.02, PL2:EN — p=.05), but not in *They were in A bad mood* (PL1:EN — p=.41, PL2:EN — p=.826). Still, it must be noted that the very large individual variation in Polish learners' responses compared to native production reveals the instability of unstressed vowel timing. This instability may be one reason why the paired t-test results point at no statistically significant change in vowel reduction in *a* between PL1 and PL2 (p=.182 for *there was A girl*, p=.155 for *They were in A bad mood*).

A more conspicuous difference has been observed in two other weak forms with initial vowels in closed syllables, i.e. *at* and *of*. The former was much longer in both the Polish learners' recordings in comparison to native speakers' performance, as indicated by the descriptive statistics in Figure 5.3 and one-way ANOVA, F(1,23):p=.002 for PL1:EN and p=.004 for PL2:EN. The schwa in the latter word was also difficult for Polish learners, p<.001 for PL1:EN, but in this case a large improvement was observed, which diminished the difference between PL2 and EN to a statistically insignificant size (p=.06 for PL2:EN). The progress was confirmed by the results of a paired t-test (p=.003) comparing PL1 and PL2 performance of individual subjects.

The vowel in *but*, functioning as a prehead, is another example of non-native vowel reduction in Polish learners. The data in Figure 5.3 yield ANOVA results at the levels of p<.001 and p=.002 for PL1:EN and PL2:EN, respectively. A highly significant level of progress in Polish learners is indicated by a paired t-test (p<.001), reflecting the fact that 12 in 13 Polish respondents shortened the vowel in *but* in their second recording.

The importance of an individual approach to teaching makes it necessary to state that the maximum native speech duration of 86ms was exceeded in the pronunciation of reduced vowel in *at*, *of* and *but* by, respectively, 3, 11 and 10 (N=13) PL1 speakers. Then the corresponding numbers were 5, 5, and 3, respectively, for PL2.

Finally, the word *had* was pronounced with a shorter schwa by the native speakers (49ms), but the PL1 and PL2 mean values (61ms and 60ms, respectively) did not represent a statistically significant difference (p=.078, p=.153, respectively). The statistics are supported by the fact that only one Polish learner's response (92ms) exceeded the maximum native value (86ms).

Generally, the results indicate the decisive influence of segmental context on unstressed vowel duration in function words. An analysis of whole function word duration will be provided in section 6.2.2. It will shed light on the durational variability of larger units and the level of rhythmical stability of feet and phrases.

5.3.2 Unstressed vowels in trochaic content words

The effect of vowel reduction on more general timing proportions can be observed on the word level. The text provides seven trochaic content words in contexts where inter-speaker phrasing or prominence variation is practically non-existent. They are divided into two groups on the basis of phrase position: final (N, T) or non-final (H), and presented in Tables 5.18 and 5.19.

Table 5.18. Mean unstressed vowel duration in phrase-medial trochaic content words

Schwa in	H trochee	PL1	PL2	EN
want**ed**	2H3+	50	45	26
sist**er**s	2H3+	99	88	46
start**e**d	2H2+	62	46	42
Cind**er**s	2H2+	133	122	72
M		**86**	**75**	**47**
SD SD:M		42.4 (49%)	39.6 (53%)	23.9 (51%)

Table 5.19. Mean unstressed vowel duration in phrase-final trochaic content words

Schwa in	N/T trochee	PL1	PL2	EN
gorge**ou**s	2N2–	88	71	83
Cind**er**s	2N2+	152	143	123
at Cind**er**s	2T2+	145	131	109
M		**128**	**115**	**105**
SD SD:M		39.2 (31%)	46.8 (41%)	29.6 (28%)

The length difference between unstressed vowels in phrase-medial and final contexts is much larger in native English performance. The unstressed vowel in tested words is on average 55% shorter in medial position in the pronunciation of native speakers. The corresponding difference amounts to 33% and 35% in the Polish speakers.

More contextual variation in native English pronunciation is evident in the word *Cinders* measured in two phrase-final contexts (N and T) and in one medial (H) position. In each group the vowel is the longest in N. Moreover, between-group variation in this context is statistically insignificant ($p=.156$, one-way ANOVA, $F(2, 35)$), which corresponds with similar mean absolute durations (PL1 — 152ms, PL2 — 143ms, EN — 123ms). However, the discrepancies become larger if phrase position is taken into account. One-way ANOVA applied to observe the effect of context on unstressed vowel duration in *Cinders* indicates a statistically significant difference in the native speakers' group alone ($p<.001$, $F(2, 33)$). Corresponding ANOVA p-values, $F(2, 36)$, for the Polish speakers' data is $p=.242$ and $p=.411$, indicating a statistically non-significant length variation related to the context.

These results confirm the descriptive statistics shown in Tables 5.18 and 5.19, which reveal the length relation between the native English schwa in N, and 41% and 11% shorter vowels in H and T, respectively. Considering the fact that with respect to the schwa in N, the Polish learners made their vowel only 12% (H) and 5% (T) shorter in their first performance, and 15% (H) and 8% (T) shorter in the second, we acknowledge the significantly different effects of prominence level (foot type) on vowel reduction in the two groups of speakers. A one-way ANOVA test recognises the effects of context as statistically significant in native English speakers alone, and only with respect to final and non-final position (N:H and T:H, but not N:T). The results are displayed in Table 5.20.

The duration of unstressed vowels in *Cinders* and the other trochaic words in the performance of individual speakers is presented in Figure 5.4.

Table 5.20. The length variation of schwa in *Cinders* with respect to group and phrase position

Schwa in	PL1 ANOVA, F(1, 24)	PL2 ANOVA, F(1, 24)	EN ANOVA, F(1, 22)
N:T	p=.540	p=.488	p=.179
N:H	p=.124	p=.228	**p<.001**
T:H	p=.252	p=.477	**p<.001**

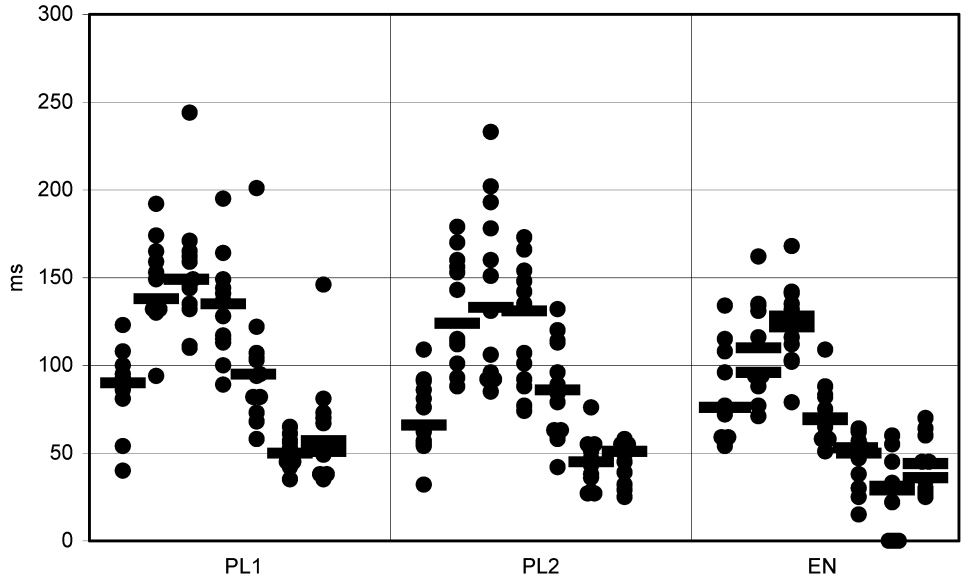

Figure 5.4. Unstressed vowel duration in (from left to right) *gorgeOUs*, *shouting at CindERs* (T), *called her CindERs* (N), *CindERs lived...* (H), *sistERs*, *wantEd*, *startEd*. Group medians indicated by horizontal bars

Native English speakers shorten the unstressed vowel the most in phrase-medial positions, where it is not affected by final lengthening. These are the contexts where statistically significant differences between Polish and English speakers appear (cf. Table 5.21).

Table 5.21. The length variation of schwa in *CindERs* (H), *sistERs*, *wantEd*, and *startEd* with respect to group

Schwa in	PL1:ENG ANOVA, F(1, 23)	PL2:EN ANOVA, F(1, 23)	PL1:PL2 paired t-test
Cind**e**rs	**p<.001**	**p<.001**	p=.118
sist**e**rs	**p<.001**	**p<.001**	p=.194
want**e**d	**p<.001**	p=.015	p=.117
start**e**d	p=.064	p=.465	p=.052

Considering the effect of underlying /r/ or trochaic vs. dactylic structure of the feet or mono- vs. bimorphemic structure of the words can lead to no conclusion because there are only four items involved in the analysis. However, if we assume the range of values obtained in the performance of native speakers as "native-like," then we must conclude that vowel reduction is more difficult for Polish learners in *Cinders* and *sisters*. The schwa in these words exceeded the native speakers' longest response in 11 and 12 cases, respectively, in the first recording of Polish learners. The number of "too long" responses was only reduced to 7 and 9, respectively, in PL2. By contrast, the schwa in the past tense suffix with no underlying /r/ in *wanted* and *started*, which appeared in a more complex, dactylic foot, was kept within the "native" limits with 2+1 exceptions in PL1 and only one longer realisation in *wanted* in PL2.

5.3.3 Relative schwa duration

In order to control for individual speech rate differences, Table 5.22 provides the ratios of tested unstressed vowels to their immediate contexts. The context for function words is constituted by the word containing the measured reduced

Table 5.22. Relative schwa duration

Schwa in	PL1	PL2	EN	EN range	PL1 out of EN range	PL2 out of EN range
	%				N	
going t**o**	8	8	**1**	0—6	**9**	**9**
buy s**o**me	11	7	**2**	0—9	**10**/12	**4**/12
had t**o**	16	12	**5**	0—20	1	2
wanted t**o**	8	9	**5**	0—17	0	0
b**u**t e(veryone)	32	28	**22**	9—36	2	2
shouting **a**t	11	13	**9**	4—15	**2**	**4**
a bad	13	11	11	7—14	**7**	**5**
a girl	28	19	**15**	10—19	**13**	**7**
sisters h**a**d	9	10	9	5—15	0	1
dreamed **o**f	16	13	**11**	7—14	**7**	**5**
h**er** sisters	16	16	**13**	6—21	2	1
h**er** fairy	19	18	17	11—28	1/12	1
called h**er**	26	27	26	17—37	1	1
want**e**d	12	13	**8**	0—19	0	2
sist**er**s	22	21	**11**	4—16	**12**	**10**
start**e**d	14	11	12	6—25	1/12	0
Cind**er**(s) (H)	37	35	**23**	15—33	**11**	**9**
Cind**er**(s) (T)	32	32	**28**	18—35	**5**	**3**
Cind**er**(s) (N)	32	32	**28**	22—35	**2**	**4**
gorge**ou**s	17	15	18	11—25	1–	1–

vowel and an adjacent content word. In line with the assumed phrase parsing approach (left-headed feet — cf. section 5.5.1), the preceding word is preferred. In preheads, however, the following content word is used. In *everyone*..., only the initial vowel is measured for reference because several respondents produced the word *everybody* instead. Unstressed vowels in trochaic content words are measured in relation to the duration of the whole word.

Figures 5.5—5.7., based on Table 5.22, graphically illustrate the relative duration of schwa in the pronunciation of Polish and English speakers (from left to right: PL1, PL2, EN for each context).

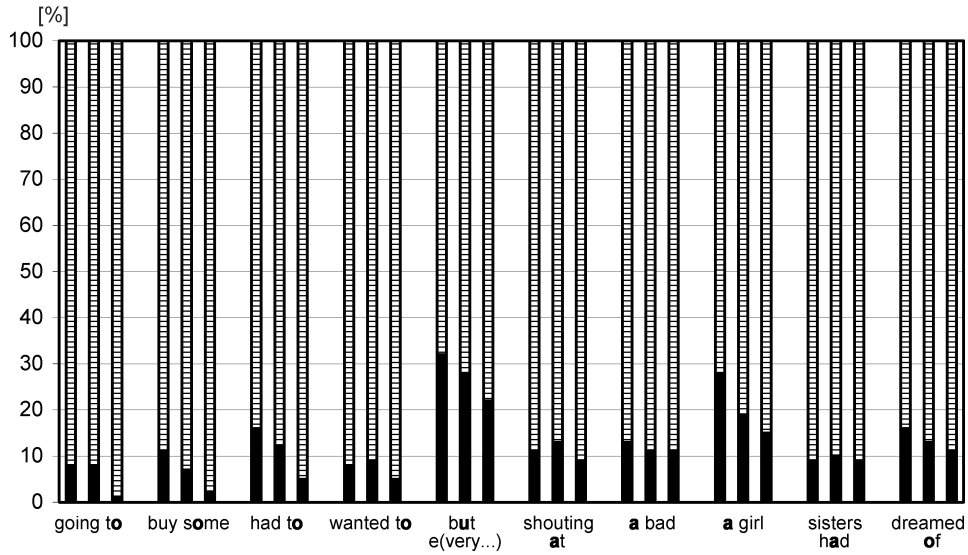

Figure 5.5. Relative schwa duration in function words

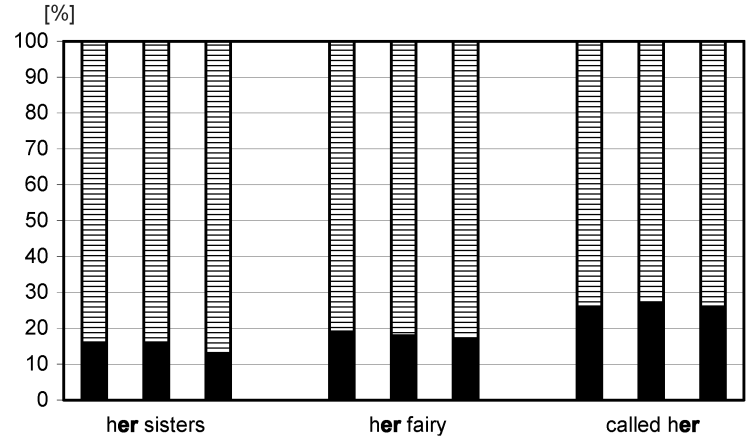

Figure 5.6. Relative schwa duration in *her*

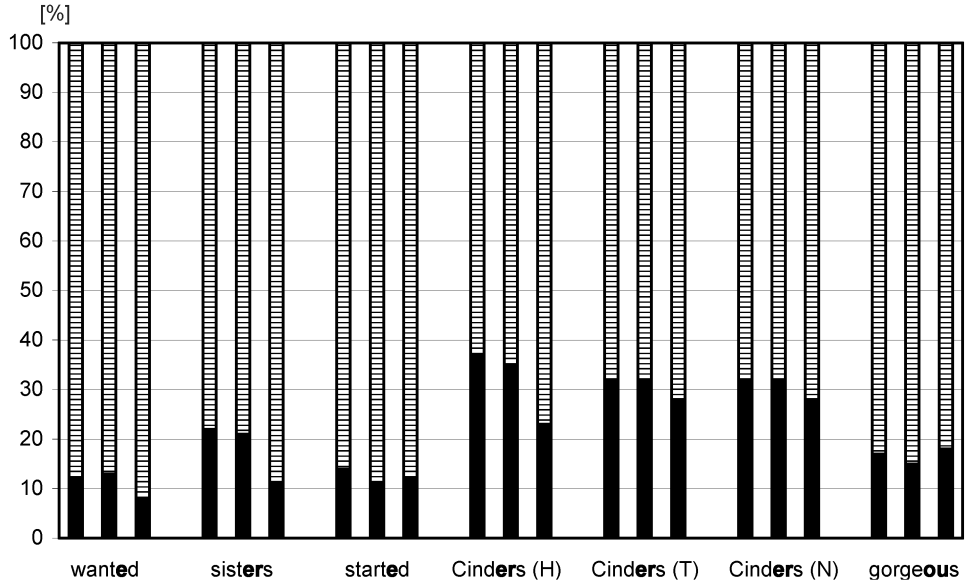

Figure 5.7. Relative schwa duration in trochaic content words

Relative schwa durations display similar patterns as those observed in absolute values. The unstressed vowel is comparable in Polish and English speakers in phrase-final positions and in past forms of verbs (*wanted*, *started*). It is considerably longer in Polish learners' realisation of phrase-medial trochaic content words (conspicuous in *Cinders*) and function words. The function words with initial /h/, which do not show obvious length differences will be discussed in section 5.5, devoted to the timing of weak forms.

5.3.4 Segmental contribution to unstressed function word duration

Chapter 5 is basically devoted to the problem of vowel length variation and its impact on temporal relations within the phrase, but consonant duration may also be varied either to suit the needs of rhythmicity of speech or to perform important phonological functions (e.g. VOT or stop closure duration). An exhaustive analysis of the duration of various consonant classes is impossible in this dissertation considering the scope of the gathered material, but this short section casts some light on the problem, the complexity of which comprises both intrinsic speech sound properties and the timing interaction within phonological structures. Table 5.23 illustrates the contribution of onset and coda to weak form duration in various function word weak forms.

Table 5.23. Segmental contribution to unstressed function word duration

Measured unit	Vowel					Consonant(s)					Word				
	PL1	PL2	EN	**PL1 /EN**	**PL2 /EN**	PL1	PL2	EN	**PL1 /EN**	**PL2 /EN**	PL1	PL2	EN	**PL1 /EN**	**PL2 /EN**
	ms					ms					ms				
shouting **at**	73	81	45	**1.6**	**1.8**	58	74	29	**2.0**	**2.6**	131	155	74	**1.8**	**2.1**
her fairy	129	89	72	**1.8**	**1.2**	80	47	17	**4.7**	**2.8**	209	136	89	**2.4**	**1.5**
her sisters	82	84	63	**1.3**	**1.3**	75	32	23	**3.3**	**1.4**	157	116	86	**1.8**	**1.4**
called **her**	119	89	78	**1.5**	**1.1**	104	39	39	**2.7**	**1.0**	223	128	117	**1.9**	**1.1**
sisters **had**	61	60	49	**1.2**	**1.2**	149	121	70	**2.1**	**1.7**	210	181	119	**1.8**	**1.5**
dreamed **of**	106	68	54	**2.0**	**1.3**	114	98	82	**1.4**	**1.2**	220	166	136	**1.6**	**1.2**
buy **some**	56	34	7	**8.0**	**4.9**	190	188	170	**1.1**	**1.1**	246	222	177	**1.4**	**1.3**
had **to**	75	37	13	**5.8**	**2.8**	42	44	55	**0.8**	**0.8**	117	81	68	**1.7**	**1.2**
going **to**	34	30	1	**34.0**	**30**	33	41	42	**0.8**	**1.0**	67	71	43	**1.6**	**1.7**
wanted **to**	43	40	16	**2.7**	**2.5**	53	36	34	**1.6**	**1.1**	96	76	50	**1.9**	**1.5**

Generally, both vowels and consonants are shorter in native English speech, although the difference varies depending on the type of segments involved. For instance, /t/ in *to* is usually longer in native performance thanks to longer VOT in comparison to Polish production. Special attention must be given to initial /h/ in weak forms of function words since it is much longer in the learners' pronunciation, which adds to timing problems connected with unstressed unit reduction. Interestingly, though, a vowel following /h/ in weak forms, especially in an open syllable (i.e. *her*), is reduced less than elsewhere, which offsets the length discrepancy between the learners and the native speakers.

In conclusion, we repeat that vowel reduction is generally smaller in Polish learner's pronunciation. The timing of consonants also contributes to the FL speech timing problems, but it depends on the type of consonant appearing in a reduced syllable. This supports the idea that segmental phonetics is a necessary prerequisite before a suprasegmental phonetics course is initiated.

5.3.5 Summary

Quantitative unstressed vowel reduction is generally more radical in native English speakers. The most significant difference in comparison to Polish learners appears in phrase medial contexts, where only the <-ed> suffix in *started* and *wanted* was not considerably longer. The apparently stronger effect of final lengthening in native speakers' production makes the schwas in phrase-final syllables comparable across the two groups.

Generally, the data confirm Hypothesis 2 in that unstressed vowels are longer in Polglish production both in absolute measures and in relation to the context. There are even contexts where very large proportions of Polish learners are unable to reduce the vowel to the level of the longest native speakers' samples (*to*, *a*, *some*, *sisters*, *Cinders* (H)). The persistence of problems with vowel reduction revealed in the second recording proves its high level of difficulty in English pronunciation acquisition by Polish learners.

5.4 Vowel duration variability in individual native speakers and Polish learners

The results presented in this chapter show large individual variation between speakers, which makes it difficult to draw general conclusions concerning native and non-native speech. This section focuses on the individual speaker's vowel length variation. If Hypotheses 1 and 2, presented in Chapter 4, are true statements, then native English speakers vary vowel duration more than Polish learners, who are believed to face interference problems with quantitative contrasts between intrinsically longer and shorter vowels, as well as contextual phenomena such as final lengthening and accentual lengthening.

5.4.1 Stressed vowel duration variability

It is obvious that among language users there are slower and faster speakers. Foreign language learners usually speak more slowly than native speakers. Our study supports this statement in general but not with respect to the articulation of vocalic segments. Table 5.24 provides a ranking of subjects according to their mean vowel duration (left), standard deviation (middle) and SD:M duration ratio (right) calculated from all the stressed monophthongs presented in Chapter 4 (20 tokens per speaker, 754 tokens obtained, 6 tokens disqualified or missing). The results of individual Polish speakers in the first recording are marked by "1" following the code. "2" indicates a second-recording result. Native speakers' codes are shown in bold type.

The results show a diversity of length means across both groups of speakers, within similar ranges: PL1 — 95—136ms, PL2 — 86—122ms, EN — 88—131ms. Because of this diversity, standard deviation from the mean does not reflect individual vocalic "flexibility" in the same way for faster and slower speakers. Therefore, a third ranking, according to SD:M ratio, is added to at least partially normalise for speech rate variability. The latter classification suggests a clear tendency in native speakers to use more vocalic

Table 5.24. Mean stressed vowel duration (ms) and its variability in English and Polish speakers

Subject	**M**	SD	SD:M	Subject	M	**SD**	SD:M	Subject	M	SD	**SD:M**
RM2	**86**	40	.47	**CJI**	131	**66**	.51	**CJE**	113	64	**.57**
CTG	**88**	37	.42	**CJE**	113	**64**	.57	**CLP**	116	62	**.53**
AS2	**88**	43	.49	**CLH**	125	**62**	.50	**CPT**	103	54	**.52**
AK2	**89**	42	.47	**CLP**	116	**62**	.53	**CJI**	131	66	**.51**
MG2	**90**	35	.39	PO1	136	**55**	.41	**CLH**	125	62	**.50**
RM1	**95**	40	.42	**CPT**	103	**54**	.52	AS1	103	51	**.50**
CSM	**98**	46	.47	**CER**	107	**51**	.48	AS2	88	43	**.49**
CMA	**98**	43	.44	AS1	103	**51**	.50	**CER**	107	51	**.48**
AJ1	**99**	36	.36	MB2	119	**51**	.43	**CHB**	101	48	**.47**
CHB	**101**	48	.47	PS2	107	**49**	.46	AK2	89	42	**.47**
CMF	**102**	43	.42	MG1	108	**49**	.45	RM2	86	40	**.47**
CPT	**103**	54	.52	AK1	114	**48**	.42	**CSM**	98	46	**.47**
LK1	**103**	37	.36	PS1	125	**48**	.39	PS2	107	49	**.46**
AS1	**103**	51	.50	**CHB**	101	**48**	.47	MG1	108	49	**.45**
LK2	**104**	41	.39	DK1	121	**47**	.39	**CMA**	98	43	**.44**
CER	**107**	51	.48	CSM	98	**46**	.47	MB2	119	51	**.43**
PS2	**107**	49	.46	MB1	133	**45**	.34	AK1	114	48	**.42**
PA2	**108**	37	.34	JK2	120	**44**	.37	**CMF**	102	43	**.42**
MG1	**108**	49	.45	PO2	119	**44**	.37	**CTG**	88	37	**.42**
AO1	**109**	33	.30	AS2	88	**43**	.49	RM1	95	40	**.42**
AJ2	**110**	38	.34	**CMF**	102	**43**	.42	PO1	136	55	**.41**
PA1	**111**	40	.36	**CMA**	98	**43**	.44	LK2	104	41	**.39**
CJE	**113**	64	.57	AK2	89	**42**	.47	DK1	121	47	**.39**
AO2	**114**	36	.32	DK2	122	**41**	.34	PS1	125	48	**.39**
AK1	**114**	48	.42	LK2	104	**41**	.39	MG2	90	35	**.39**
CLP	**116**	62	.53	RM2	86	**40**	.47	PO2	119	44	**.37**
PO2	**119**	44	.37	**CMC**	121	**40**	.33	JK2	120	44	**.37**
JK1	**119**	38	.32	PA1	111	**40**	.36	AJ1	99	36	**.36**
MB2	**119**	51	.43	RM1	95	**40**	.42	LK1	103	37	**.36**
JK2	**120**	44	.37	JK1	119	**38**	.32	PA1	111	40	**.36**
DK1	**121**	47	.39	AJ2	110	**38**	.34	AJ2	110	38	**.34**
CMC	**121**	40	.33	LK1	103	**37**	.36	PA2	108	37	**.34**
DK2	**122**	41	.34	PA2	108	**37**	.34	MB1	133	45	**.34**
PS1	**125**	48	.39	**CTG**	88	**37**	.42	DK2	122	41	**.34**
CLH	**125**	62	.50	AO2	114	**36**	.32	**CMC**	121	40	**.33**
CJI	**131**	66	.51	AJ1	99	**36**	.36	JK1	119	38	**.32**
MB1	**133**	45	.34	MG2	90	**35**	.39	AO2	114	36	**.32**
PO1	**136**	55	.41	AO1	109	**33**	.30	AO1	109	33	**.30**

variation, with average standard deviation of 47% (42—57%, except one case of 33%) of the individual mean. The Polish learners' average standard deviation was 39—40%, within the range of 30—50%.

5.4.2 Vowel duration variability in non-phrase-final syllables

Although the above results have been calculated from stressed monophthongs only, the conclusion that native English speakers attend to intrinsic vowel quantity more is still not sufficiently grounded, given the variety of phrasal contexts in which the vowels were measured. Considering only non-final syllables eliminates the effect of final lengthening, but it also reduces the number of tokens to ten per subject (377 tokens obtained, 3 tokens disqualified or missing). The resulting classification is shown in Table 5.25.

Table 5.25. The duration (ms) of non-phrase-final stressed monophthongs

Subject	**M**	SD	SD:M	Subject	M	SD	**SD:M**
1	2	3	4	5	6	7	8
RM2	**72**	24	.33	AS2	74	43	**.58**
AS2	**74**	43	.58	**CJI**	101	59	**.58**
RM1	**75**	21	.27	**CPT**	87	49	**.56**
CJE	**76**	31	.41	AS1	86	47	**.55**
AK2	**78**	33	.43	**CSM**	79	41	**.51**
MG2	**79**	32	.40	**CER**	83	40	**.48**
CSM	**79**	41	.51	**CLP**	90	43	**.48**
CTG	**80**	37	.45	**CLH**	103	48	**.47**
AJ1	**82**	30	.37	**CTG**	80	37	**.45**
CER	**83**	40	.48	PS2	90	41	**.45**
CMA	**83**	33	.39	LK2	101	45	**.45**
MG1	**83**	32	.39	**CHB**	85	38	**.44**
CHB	**85**	38	.44	PS1	109	47	**.43**
AS1	**86**	47	.55	AK1	102	44	**.43**
CPT	**87**	49	.56	AK2	78	33	**.43**
LK1	**89**	28	.32	**CJE**	76	31	**.41**
AJ2	**89**	29	.33	MG2	79	32	**.40**
CMF	**90**	35	.39	**CMA**	83	33	**.39**
CLP	**90**	43	.48	**CMF**	90	35	**.39**
PS2	**90**	41	.45	MG1	83	32	**.39**
PA1	**94**	25	.27	PA2	100	37	**.38**
MB2	**95**	29	.30	AJ1	82	30	**.37**
AO1	**98**	30	.31	DK2	111	39	**.35**

cont. tab. 5.25

1	2	3	4	5	6	7	8
PO2	**99**	33	.33	JK1	107	37	**.35**
DK1	**100**	32	.32	RM2	72	24	**.33**
PA2	**100**	37	.38	AJ2	89	29	**.33**
LK2	**101**	45	.45	PO2	99	33	**.33**
CJI	**101**	59	.58	PO1	107	35	**.33**
JK2	**102**	30	.29	DK1	100	32	**.32**
AK1	**102**	44	.43	LK1	89	28	**.32**
AO2	**102**	28	.27	MB1	120	38	**.31**
CLH	**103**	48	.47	AO1	98	30	**.31**
JK1	**107**	37	.35	**CMC**	109	33	**.30**
PO1	**107**	35	.33	MB2	95	29	**.30**
CMC	**109**	33	.30	JK2	102	30	**.29**
PS1	**109**	47	.43	RM1	75	21	**.27**
DK2	**111**	39	.35	AO2	102	28	**.27**
MB1	**120**	38	.31	PA1	94	25	**.27**

Again, at similar mean vowel length diversity in both groups, the native speakers produce vowels with more standard deviation from the mean (46%, range 39—58%, one outlier at 30%) than PL1 (36%, range 27—55%) and PL2 (38%, range 27—58%). It also appears significant that as many as 10 Polish learners produced the tested vowels with standard deviation below 39% of their mean in the first recording. In the second recording there were 8 such respondents, which still constitutes a large proportion. These results justify a more specifically focused future study of quantitative variation between English vowel classes in the pronunciation of Polish learners in comparison to native production.

Individual vowel length variation is certainly more conspicuous if reduced vowels are taken into account. Table 5.26 displays mean vowel durations with standard deviations calculated from all non-phrase final syllables measured in the study (29 tokens per speaker: 10 full vowels + 19 reduced vowels, 1,093 tokens obtained, 9 tokens disqualified or missing).

As expected, even the ranking based on mean vowel duration (left), unlike the results for stressed vowels only, groups the native speakers in the upper part of the table. The SD:M ratio (middle) leaves two native speakers below the top region of the table, but they are distinguished from the surrounding Polish speakers by shorter mean vowel duration. If, instead of the SD:M ratio, the difference between the mean and standard deviation (M – SD)[4] is used as

[4] This index "favours" more vocalic duration variability at higher speech rates.

Table 5.26. The duration (ms) of non-phrase-final stressed and unstressed monophthongs

Subject	**M**	SD	SD:M	Subject	M	SD	**SD:M**	Subject	M	SD	SD:M	**M – SD**
CSM	**51**	38	.75	**CSM**	51	38	**.75**	**CSM**	51	38	.75	**13**
CHB	**53**	39	.74	**CHB**	53	39	**.74**	**CHB**	53	39	.74	**14**
CER	**54**	38	.71	**CER**	54	38	**.71**	**CER**	54	38	.71	**16**
AS2	**54**	34	.63	**CMF**	56	39	**.69**	**CTG**	54	37	.68	**18**
CTG	**54**	37	.68	**CTG**	54	37	**.68**	**CMF**	56	39	.69	**18**
AK2	**56**	35	.62	**CPT**	60	39	**.66**	AS2	54	34	.63	**20**
CMF	**56**	39	.69	**CLH**	71	45	**.64**	**CPT**	60	39	.66	**21**
CJE	**57**	29	.51	AS2	54	34	**.63**	AK2	56	35	.62	**21**
CMA	**59**	31	.52	AK2	56	35	**.62**	**CLP**	63	38	.61	**24**
CPT	**60**	39	.66	**CLP**	63	38	**.61**	**CLH**	71	45	.64	**26**
RM2	**61**	27	.44	**CJI**	75	45	**.61**	**CMA**	59	31	.52	**28**
CLP	**63**	38	.61	**CMC**	73	43	**.59**	**CJE**	57	29	.51	**28**
PA1	**69**	32	.46	AK1	79	46	**.59**	**CJI**	75	45	.61	**30**
AS1	**70**	37	.54	PS2	75	42	**.56**	**CMC**	73	43	.59	**30**
CLH	**71**	45	.64	MG2	73	39	**.54**	AS1	70	37	.54	**32**
RM1	**72**	20	.28	AO2	73	39	**.54**	AK1	79	46	.59	**33**
MG2	**73**	39	.54	AS1	70	37	**.54**	MG2	73	39	.54	**33**
CMC	**73**	43	.59	PS1	89	47	**.53**	PS2	75	42	.56	**33**
AO2	**73**	39	.54	**CMA**	59	31	**.52**	AO2	73	39	.54	**34**
CJI	**75**	45	.61	**CJE**	57	29	**.51**	RM2	61	27	.44	**34**
PS2	**75**	42	.56	JK2	81	40	**.50**	PA1	69	32	.46	**37**
PA2	**77**	38	.49	AJ1	96	48	**.50**	PA2	77	38	.49	**39**
AK1	**79**	46	.59	PA2	77	38	**.49**	JK2	81	40	.50	**40**
PO2	**79**	32	.40	DK2	83	40	**.49**	PS1	89	47	.53	**42**
JK2	**81**	40	.50	AO1	85	40	**.47**	DK2	83	40	.49	**43**
DK2	**83**	40	.49	PA1	69	32	**.46**	AO1	85	40	.47	**45**
AJ2	**84**	34	.41	MB1	104	47	**.45**	PO2	79	32	.40	**48**
AO1	**85**	40	.47	LK2	86	39	**.45**	LK2	86	39	.45	**48**
LK2	**86**	39	.45	MG1	87	39	**.45**	AJ1	96	48	.50	**48**
DK1	**87**	37	.43	RM2	61	27	**.44**	MG1	87	39	.45	**48**
LK1	**87**	35	.40	JK1	91	40	**.44**	AJ2	84	34	.41	**49**
MG1	**87**	39	.45	DK1	87	37	**.43**	DK1	87	37	.43	**50**
PO1	**88**	35	.40	AJ2	84	34	**.41**	JK1	91	40	.44	**51**
PS1	**89**	47	.53	LK1	87	35	**.40**	RM1	72	20	.28	**51**
JK1	**91**	40	.44	PO2	79	32	**.40**	LK1	87	35	.40	**52**
MB2	**93**	29	.31	PO1	88	35	**.40**	PO1	88	35	.40	**53**
AJ1	**96**	48	.50	MB2	93	29	**.31**	MB1	104	47	.45	**57**
MB1	**104**	47	.45	RM1	72	20	**.28**	MB2	93	29	.31	**64**

a criterion (right), all native speakers' results are clustered at the top of the table. The two Polish speakers who achieved native-like scores are indeed the most successful students in the tested group.

5.4.3 General vowel duration variability

To make the picture complete, we present the results showing individual speakers' vowel duration variability, calculated from all 46 contexts analysed in this study (Table 5.27). The gathered data include 1,734 responses (PL1=591, PL2=595, EN=548). 14 responses are either disqualified or missing.

Table 5.27. Overall mean syllable nucleus duration (ms) and variability

Subject	**M**	SD	SD:M	Subject	M	**SD**	SD:M	Subject	M	SD	**SD:M**
1	2	3	4	5	6	7	8	9	10	11	12
CTG	**81**	55	.68	**CLH**	116	**89**	.77	**CHB**	87	67	**.78**
AK2	**81**	61	.75	**CJE**	103	**80**	.77	**CJE**	103	80	**.77**
AS2	**82**	56	.68	**CJI**	118	**77**	.66	**CLH**	116	89	**.77**
CSM	**82**	62	.75	PO1	126	**73**	.58	**CSM**	82	62	**.75**
RM2	**83**	47	.57	**CLP**	104	**72**	.70	AK2	81	61	**.75**
CHB	**87**	67	.78	AK1	110	**71**	.64	**CMF**	92	68	**.74**
MG2	**88**	45	.52	PS1	123	**69**	.56	**CER**	91	67	**.74**
CMA	**90**	62	.69	**CMF**	92	**68**	.74	**CLP**	104	72	**.70**
CPT	**91**	64	.70	**CER**	91	**67**	.74	**CPT**	91	64	**.70**
CER	**91**	67	.74	**CHB**	87	**67**	.78	**CMA**	90	62	**.69**
CMF	**92**	68	.74	**CMC**	105	**66**	.63	AS2	82	56	**.68**
RM1	**98**	48	.49	DK1	120	**66**	.55	**CTG**	81	55	**.68**
AO2	**99**	54	.54	PA1	105	**64**	.61	**CJI**	118	77	**.66**
AS1	**99**	61	.61	**CPT**	91	**64**	.70	AK1	110	71	**.64**
PS2	**103**	61	.59	MB2	123	**63**	.51	**CMC**	105	66	**.63**
CJE	**103**	80	.77	**CMA**	90	**62**	.69	AS1	99	61	**.61**
CLP	**104**	72	.70	**CSM**	82	**62**	.75	PA1	105	64	**.61**
PA2	**104**	59	.57	AK2	81	**61**	.75	PS2	103	61	**.59**
LK2	**104**	49	.47	PS2	103	**61**	.59	PO1	126	73	**.58**
PA1	**105**	64	.61	MB1	129	**61**	.47	PA2	104	59	**.57**
CMC	**105**	66	.63	AS1	99	**61**	.61	RM2	83	47	**.57**
PO2	**107**	53	.50	JK2	108	**60**	.56	JK2	108	60	**.56**
DK2	**108**	58	.54	PA2	104	**59**	.57	PS1	123	69	**.56**
JK2	**108**	60	.56	LK1	114	**58**	.51	DK1	120	66	**.55**
MG1	**109**	53	.48	DK2	108	**58**	.54	AO2	99	54	**.54**
AK1	**110**	71	.64	AO1	112	**57**	.51	DK2	108	58	**.54**

cont. tab. 5.27

1	2	3	4	5	6	7	8	9	10	11	12
AJ2	**111**	56	.50	AJ2	111	**56**	.50	MG2	88	45	**.52**
AJ1	**111**	55	.49	AS2	82	**56**	.68	AO1	112	57	**.51**
AO1	**112**	57	.51	JK1	115	**56**	.49	MB2	123	63	**.51**
LK1	**114**	58	.51	**CTG**	81	**55**	.68	LK1	114	58	**.51**
JK1	**115**	56	.49	AJ1	111	**55**	.49	AJ2	111	56	**.50**
CLH	**116**	89	.77	AO2	99	**54**	.54	PO2	107	53	**.50**
CJI	**118**	77	.66	PO2	107	**53**	.50	AJ1	111	55	**.49**
DK1	**120**	66	.55	MG1	109	**53**	.48	RM1	98	48	**.49**
PS1	**123**	69	.56	LK2	104	**49**	.47	JK1	115	56	**.49**
MB2	**123**	63	.51	RM1	98	**48**	.49	MG1	109	53	**.48**
PO1	**126**	73	.58	RM2	83	**47**	.57	MB1	129	61	**.47**
MB1	**129**	61	.47	MG2	88	**45**	.52	LK2	104	49	**.47**

A larger amount of data confirms the previous observations. Mean vowel duration in native English performance is generally shorter than in Polglish production if both stressed and unstressed vowels are involved. A clear difference between the groups, however, is evident in the proportion of standard deviation to mean vowel duration. With the exception of the two most proficient Polish speakers, the calculated ratio is up to 61% for Polish learners and at least 63% for all English speakers.

5.4.4 Polish learners' vowel timing diversity and development

Large individual variation in the results of individual respondents have been observed in both groups of speakers. An analysis of individual speakers' results can help verify the previous observations and separate chance differences between native and non-native respondents from regular patterns.

Table 5.28 presents the Polish learners' mean vowel durations before (M1) and after the training (M2), and vowel duration variability expressed as the proportion of standard deviation to mean vowel duration. The table focuses on stressed vowels (A — 20 items), stressed vowels and diphthongs (B — 20+6 items), and stressed vowels in non-phrase-final syllables (C — 10 items). The corresponding results obtained by native speakers are shown below the Polglish results.

In the case of native speakers, bold type indicates extreme native English personal mean durations as well as the lowest and second lowest SD:M ratios.[5]

[5] CMC's scores significantly differed from all the other results.

In the case of Polish learners, bold print indicates personal means out of the native score range as well as SD:M ratios lower than the second lowest native English score. Ratios lower than CMC's scores are additionally marked with an asterisk. The speakers are ranked according to the SD:M ratio corresponding to all stressed vowels and diphthongs (B).

Table 5.28. Polish learner's vowel timing development. Personal stressed vowel mean durations and personal durational variation (native English data provided in the lower part of the table)

Columns												
1	2	3	4	5	2	3	4	5	2	3	4	5
Subject	A. Stressed vowels (20)				B. Stressed vowels and diphthongs (26)				C. Non-final stressed vowels (10)			
PL	M1	M2	SD:M1	SD:M2	M1	M2	SD:M1	SD:M2	M1	M2	SD:M1	SD:M2
MB	**133**	119	**.34**	.43	151	140	**.39***	**.51**	**120**	95	**.31**	**.30**
JK	119	120	**.32***	**.40**	138	138	**.41***	.41*	107	102	**.35**	**.29***
AO	109	114	**.30***	**.32***	130	127	**.44**	**.39***	98	102	**.31**	**.27***
PS	125	107	**.39**	.46	149	125	**.46**	**.49***	109	90	.43	.45
MG	108	90	.45	**.40**	121	**103**	**.48**	**.43***	83	79	.39	.40
RM	95	**86**	.42	.47	112	**100**	**.48**	**.51**	**75**	**72**	**.27***	**.33**
DK	121	122	**.39**	**.30***	143	135	**.49**	**.40***	100	**111**	**.32**	**.35**
AJ	99	110	**.36**	**.30***	115	127	**.50**	**.47**	82	89	**.37**	**.33**
PA	111	108	**.36**	**.30***	132	128	**.50**	**.47**	94	100	**.27***	**.38**
PO	**136**	119	**.41**	**.40**	**160**	133	**.50**	**.41***	107	99	**.33**	**.33**
LK	103	104	**.36**	**.40**	125	119	**.51**	**.44**	89	101	**.32**	.45
AK	114	89	.42	.47	135	106	.52	.62	102	78	.43	.43
AS	103	88	.50	.49	122	106	.55	.57	86	**74**	.55	.58
EN	M		SD:M		M		SD:M		M		SD:M	
CMC	121		**.33* (min1)**		140		**.44* (min1)**		**109 (max)**		**.30* (min1)**	
CTG	**88 (min)**		**.42 (min2)**		**106 (min)**		**.52 (min2)**		80		.46	
CJI	**131 (max)**		.51		**155 (max)**		**.52 (min2)**		101		.58	
CSM	98		.47		115		.53		79		.52	
CER	107		.48		127		.53		83		.48	
CMF	102		**.42 (min2)**		126		.54		90		**.39 (min2)**	
CPT	103		.52		122		.54		87		.56	
CLP	116		.53		139		.54		90		.48	
CMA	98		.44		120		.55		83		**.39 (min2)**	
CHB	101		.47		121		.57		85		.44	
CLH	125		.50		**155 (max)**		.62		103		.47	
CJE	113		.57		139		.63		**76 (min)**		.41	

Individual Polglish vowel duration means in the first recording (column 2) were comparable to native means, with two exceptions (MB and PO) where two of the three values exceeded the maximum native mean duration. Only one such case was found in PL2 (column 3, DK, only for non-final vowels). The increased reading rate resulted in two shorter PL2 means, i.e. MG and RM, for stressed vowels and diphthongs.

A more significant difference between native and non-native timing is reflected in more durational "flexibility" of native English vowels expressed by SD:M ratios. Five Polish students obtained a lower value (column 4, asterisk) than the lowest English ratio at least in one of the three groups of tested items. Moreover, if we compare the results to the second lowest EN ratio, then the number of Polish speakers who fail to vary their English stressed vowel length in a native-like manner rises to 11 out of 13 (column 4, bold: 9 for stressed vowels, 11 for vowels and diphthongs and 9 for vowels in non-final syllables). Almost identical problems appear in the second recording (column 5, bold). Only two speakers, AK and AS, were able to obtain native-like ratios in all three groups.

Most Polish speakers shortened their vowels in the second recording, but this tendency was not very strong (cf. columns 2—3). It is hardly surprising considering the fact that only in isolated cases were the Polish mean vowel durations shorter or longer than extreme native values. Moreover, the direction of mean duration change in individual speakers was not correlated with the change of SD:M ratio.

Naturally, the most spectacular differences between Polish and English speakers appear in groups of vowels where unstressed units are taken into account. The results are displayed in Tables 5.29 (D-E) and 5.30 (F-G). All symbols are used in the same way as in Table 5.28. The speakers are ranked according to the SD:M ratio (PL2 and EN) corresponding to all monophthongs (Table 5.29, E) and monophthongs in non-phrase-final syllables (Table 5.30, G). Maximum individual English means and minimum SD:M ratios are shown in bold.

As mentioned before, the main difference between the two groups of speakers is in the reduced vowel length. This time we observe significant differences in vowel duration means, proportional to the reduced vowel share in the tested material. All 13 Polish speakers made their average schwa longer than the longest native English mean (part D, column 2). The monophthongs in non-phrase-final syllables (G), which included 10 stressed and 17 reduced vowels, were longer in 10 subjects. The influence of vowel reduction on mean vowel length was not strong enough in PL1 to reach native-like values in 7 cases for all tested monophthongs (E) and in 4 cases for all monophthongs and diphthongs (F). The second recording brought the results in parts E and F closer to native norms (2 and 1 longer means, respectively) but in parts D and G

Table 5.29. Polish learner's vowel timing development. Stressed and unstressed monophthongs (native English data provided in the lower part of the table)

Columns								
1	2	3	4	5	2	3	4	5
Subject	D. Unstressed vowels (20)				E. All monophthongs (40)			
PL	M1	M2	SD:M1	SD:M2	M1	M2	SD:M1	**SD:M2**
LK	**95**	**86**	**.43**	**.45**	99	95	**.39**	**.42**
AJ	**107**	**89**	.50	.51	**103**	**100**	**.44**	**.43**
MB	**102**	**104**	.53	**.42**	**117**	**111**	**.44**	**.43**
PO	**81**	**71**	**.36**	**.36**	**108**	95	**.48**	**.45**
PA	**70**	**70**	.63	.57	90	90	**.51**	**.47**
AO	**84**	62	.52	.58	97	88	**.42**	**.50**
MG	**95**	**70**	**.45**	.61	**102**	80	**.45**	**.50**
DK	**88**	**76**	.52	.62	**105**	98	**.46**	**.50**
RM	**76**	59	**.33**	.50	85	73	**.40**	**.52**
JK	**85**	**71**	**.48**	.59	**102**	95	**.42**	**.52**
PS	**85**	**75**	.57	.68	**105**	91	**.49**	.57
AS	**70**	48	.56	.59	86	68	**.55**	.60
AK	**76**	48	.79	.75	95	68	.60	.64
EN	M		SD:M		M		**SD:M**	
CMC	58		.67		89		**.56 (min)**	
CMA	51		**.50 (min)**		75		.57	
CTG	45		.69		66		.60	
CJI	**68 (max)**		.55		**99 (max)**		.62	
CMF	46		.81		74		.66	
CLH	64		.73		94		.66	
CPT	49		.57		76		.66	
CLP	58		.68		87		.68	
CJE	55		.64		84		.70	
CER	47		.78		76		.70	
CHB	42		.80		71		.71	
CSM	40		.81		69		.71	

Table 5.30. Polish learner's vowel timing development. All tested vocalic units and non-final monophthongs (native English data provided in the lower part of the table)

Columns								
1	2	3	4	5	2	3	4	5
Subject	**F. All vowels and diphthongs (46)**				**G. Non-final monophthongs (27)**			
PL	M1	M2	SD:M1	SD:M2	M1	M2	SD:M1	**SD:M2**
MB	**130**	**124**	**.47**	**.51**	**104**	**93**	**.45**	**.31**
PO	**125**	106	**.59**	**.50**	**88**	**79**	**.40**	**.40**
AJ	111	111	**.50**	**.51**	**96**	**84**	**.50**	**.41**
RM	96	83	**.49**	**.57**	72	61	**.28**	**.44**
LK	112	105	**.51**	**.47**	**87**	**86**	**.40**	**.45**
DK	**120**	108	**.56**	**.54**	**87**	**83**	**.43**	**.49**
PA	105	104	**.62**	**.58**	69	**77**	**.46**	**.49**
JK	115	109	**.49**	**.56**	**91**	**81**	**.44**	**.50**
MG	110	89	**.48**	**.52**	**87**	73	**.45**	.54
AO	110	99	**.51**	**.55**	85	73	**.47**	.54
PS	**121**	103	**.56**	**.60**	**89**	**75**	.53	.56
AK	109	81	.65	.76	**79**	56	.59	.62
AS	99	81	**.62**	.70	70	54	.54	.63
EN	M		SD:M		M		**SD:M**	
CJE	103		.79		57		**.51 (min)**	
CMA	90		.69		59		.52	
CMC	105		**.64 (min)**		73		.59	
CJI	**117 (max)**		.67		**75 (max)**		.61	
CLP	104		.71		63		.61	
CLH	115		.78		71		.64	
CPT	90		.71		60		.66	
CTG	80		.69		54		.68	
CMF	91		.76		56		.69	
CER	91		.75		54		.71	
CHB	87		.79		53		.74	
CSM	82		.76		51		.75	

the Polglish means remained longer than the native maximum in 9 and 8 cases, respectively.

Table 5.31 is a compact illustration of individual learners' results and development tendencies in relation to the limits set by native speakers' scores shown in detail in Tables 5.28—5.30. A-G refer to tested groups of syllable nuclei. "Y" for M1, M2 or SD:M indicates that a corresponding mean value or ratio falls between EN minimum and maximum values. If a mean is higher, or SD:M ratio lower, it is indicated by an "N". "N!" marks the latter ratios (columns 5—6) if they are even lower than the outlier mentioned above (CMC). "N–" marks two cases (RM, MG) in column 3 where the mean duration of stressed vowels and diphthongs (B) was shorter than the native minimum. Columns 4, 7, 10 and 13 (Mch and SD:Mch) show the direction of the ratio change in the second recording (+, 0, or –). As indicated in the table, columns 2—7 refer to stressed nuclei only (A-C), while columns 8—13 show results embracing the scores for stressed and unstressed vowels (D-G).

Table 5.31. Mean vowel duration, variability and developmental tendencies in Polish learners

Columns												
1	2	3	4	5	6	7	8	9	10	11	12	13
PL	M1 ABC	M2 ABC	Mch ABC	SD:M1 ABC	SD:M2 ABC	SD:Mch ABC	M1 DEFG	M2 DEFG	Mch DEFG	SD:M1 DEFG	SD:M2 DEFG	SD:Mch DEFG
MB	NYN	YYY	–––	NN!N	YNN	++–	NNNN	NNNN	+–––	YNNN	NNNN	––+–
PO	NNY	YYY	–––	NNN	NN!N	––0	NNNN	NYYN	———–	NNNN	NNNN	0––0
LK	YYY	YYY	+–+	NNN	NNY	+–+	NYYN	NYYN	———–	NNNN	NNNN	++–+
PA	YYY	YYY	––+	NNN!	NNN	––+	NYYY	NYYN	00–+	YNNN	YNNN	–––+
AJ	YYY	YYY	+++	NNN	NNN	–––	NNYN	NNYN	––0–	YNNN	YNNN	+–+–
DK	YYY	YYN	+–+	NNN	NN!Y	––+	NNNN	NYYN	———–	YNNN	YNNN	++–+
JK	YYY	YYY	+0–	N!N!N	NN!N!	+0–	NNYN	NYYN	———–	NNNN	YNNN	++++
RM	YYY	YN-Y	–––	YNN!	YNN	+++	NYYY	YYYY	———–	NNNN	YNNN	++++
AO	YYY	YYY	+–+	N!NN	N!N!N!	+––	NYYN	YYYY	———–	YNNN	YNNY	++++
MG	YYY	YN-Y	–––	YNY	NN!Y	––+	NNYN	NYYY	———–	NNNN	YNNY	++++
PS	YYY	YYY	–––	NNY	YNY	+++	NNNN	NYYN	———–	YNNY	YYNY	++++
AK	YYY	YYY	–––	YYY	YYY	++0	NYYN	YYYY	———–	YYYY	YYYY	–+++
AS	YYY	YYY	–––	YYY	YYY	–++	NYYY	YYYY	———–	YNNY	YYYY	++++

The influence of less quantitative vowel reduction resulted in a clear-cut distinction between the native and non-native SD:M ratios for E and F. Only one Polish learner (AK) varied her vowel length in a native-like manner in the first recording (column 11). That there was little progress in this aspect of timing in the second recording is illustrated by the fact that only one speaker (AS) managed to raise this ratio to a native-like level for both E and F, while another speaker (PS) only managed to place the ratio for E within the native range. None of the remaining ten Polish subjects achieved that in their second recording (column 12), even though most speakers increased their vowel variability after the training (pluses in column 13 for E and F).

SD:M ratios for vowels not affected by final lengthening (G) were within the native range in 3 PL1 speakers and apart from these, two more achieved a native-like ratio in their second recording session. The schwa length variation, where no phonological distinctions are expected, is still generally larger in English speakers, with 5 and 3 Polish speakers below the native norm in the first and second recordings, respectively. The difference is probably due to the contexts where English speakers radically reduce or even elide unstressed vowels. The leftmost Ns in columns 8 and 9 (D) again confirm the vowel reduction difficulty level, although 4 learners succeeded in producing significantly shorter schwas after the training.

It must also be noted that the variation differences between English and Polish speakers have diminished in contexts involving unstressed vowel reduction (D-G) in a vast majority of cases (pluses in column 13), especially owing to more radical quantitative reduction (minuses in column 10). These results indicate the correct direction of the learners' development despite only modest progress.

On the other hand, vowel length variability caused by factors excluding unstressed vowel reduction shows less uniform tendencies since a faster speech rate does not ensure a higher SD:M ratio. The ratio rose consistently in 6 speakers. Two of them (AS and AK) had managed to achieve native-like timing already in the first recording, one (LK) slowed down in the second recording and the remaining three (MB, PS and RM) were able to increase the ratio at a higher speed of reading. Still, besides AS and AK, only PS succeeded in reaching the native vowel length flexibility level in at least two groups of tested items (A-C) and turned out to be the only student besides AS to turn the initial "N" into a "Y" for all tested nuclei (column 12, E).

5.5 Conclusion

With respect to the hypotheses proposed in Chapter 4, we conclude that the analysis of vowel duration presented in this chapter supports the following statements in their parts referring to vowel timing:

— **Hypothesis 1** (section 4.4.1): "Mean duration proportion of intrinsically 'long' to intrinsically 'short' vowels is smaller in Polglish production" (cf. sections 5.1—5.2).
— **Hypothesis 2** (section 4.4.2): "Unstressed vowels, [...] are longer in Polglish production both in terms of absolute and relative duration" (cf. section 5.3).
— **Hypothesis 3** (section 4.4.3): "Vowels [...] in nuclear accent position are relatively shorter in Polglish production" (cf. sections 5.2.3, 5.2.6).

In section 5.2.7, summarising observations concerning the timing of stressed vowels, we conclude that articulatory rate is an individual feature and does not depend on the subject's L1 since both groups' mean durations and individual variation are comparable with respect to stressed syllables. However, Polish learners display relatively less variation of vocalic segment length with respect to vowel type, accent condition, foot complexity, phrase position (final vs. non-final syllable), and the influence of the following consonant's phonation type. The results are not uniform for all tested items since the investigated speaking style makes it impossible to prevent interaction of various duration determinants.

More categorical conclusions are drawn from the results pertaining to unstressed vowel duration (cf. section 5.3.4). These segments were regularly longer in the pronunciation of Polish learners, although the difference became practically insignificant in phrase-final syllables and in function words beginning with /h/, where the English speakers reduced the following vowel less radically than in other contexts.

Naturally, shorter unstressed vowels in the native speakers' performance further increase the difference in vowel quantity variation between the two groups (section 5.4.3). A ranking of respondents according to their individual vowel length variability expressed as the proportion of standard deviation to mean vowel length practically separates the native speakers from Polish learners (Table 5.26).

Finally, from the pedagogical standpoint, as most of the learners ranked higher in their second performance, we conclude that pronunciation training (or simply general language experience) allows them to approach native norms. However, only the most talented ones (two students who also ranked high in the first recording) managed to reach those norms after a year of study. Typically, the progress was most evident where the difference between

the learners and native speakers was initially the largest, i.e. in the duration of unstressed vowels, whereas the variability of stressed vowel duration was not significantly improved. Because the initial difference was only partially reduced in the former case, and no significant progress was made in the latter aspect, non-native timing of vowel sounds is still a characteristic feature of an advanced Polish learner of English after a year's study at a teacher training college.

Chapter 6

Timing relations within the tone group — results and discussion

6.1 Introduction

This chapter presents and analyses the timing of syllables and words, including content words and weak forms of function words (section 6.2), as well as relations between feet on the tone-group (IP) level (section 6.3). The analysis refers to these parts of Hypotheses 2 and 3 that concern syllables and words (section 6.2), and Hypotheses 4 and 5, concerning the relations between tone group constituents (section 6.3).

6.2 Word/foot-level timing relations

The timing relations within the foot depend on the relative duration of the stressed syllable and the unstressed syllables following it. As shown in Chapter 5, Polish learners display non-native timing of vowels in their English speech, which we ascribe to L1 interference and fluency problems (mainly in PL1 performance). These problems pertain to intrinsic temporal properties of vowels, prominence level and phrase position. In this section we investigate whether similar discrepancies between native and non-native English production occur at the level of word and foot.

6.2.1 Syllable duration in content words

Constituent syllable duration has been measured in seven trochaic words presented in Table 6.1. The analysis is thus based on 2×7=14 syllables, produced by 13 Polish subjects twice and 12 English subjects (2×182+168=532

Table 6.1. Group mean syllable duration in trochaic content words

Word	Stressed syllable duration (ms)					Unstressed syllable duration (ms)					Unstressed syllable share (%)		
	PL1	PL2	EN	**PL1:EN**	**PL2:EN**	PL1	PL2	EN	**PL1:EN**	**PL2:EN**	PL1	PL2	EN
wanted (H3)	256	184	179	**1.43**	**1.03**	183	154	124	**1.48**	**1.24**	42	46	41
sisters (H3)	296	261	279	**1.06**	**0.94**	128	113	66	**1.94**	**1.71**	30	30	19
started (H2)	327	299	293	**1.12**	**1.02**	128	117	97	**1.32**	**1.21**	28	28	25
Cinders (H2)	230	227	244	**0.94**	**0.93**	133	122	72	**1.85**	**1.69**	37	35	23
Cinders (N2)	318	296	313	**1.02**	**0.95**	152	142	123	**1.24**	**1.15**	32	32	28
gorgeous (N2)	328	298	318	**1.03**	**0.94**	132	114	135	**0.98**	**0.84**	29	28	30
shouting (N3)	360	321	312	**1.15**	**1.03**	222	157	123	**1.80**	**1.28**	37	32	28

syllables measured altogether). The words are grouped according to their prominence status (H or N). Following the criteria presented in Chapter 4, a syllable boundary was in each case set at the moment of release of the prevocalic stop consonant, while the word-final fricatives in *sisters*, *Cinders* and *gorgeous* were excluded from analysis.

The length of stressed syllables is comparable in the production of English and Polish speakers, while the unstressed syllables are considerably longer in Polish learners' speech with the exception of phrase-final positions (*Cinders* and *gorgeous*). Consequently, the longer articulation of unstressed syllables results in different (stressed:unstressed) length proportions within individual phrases, which is also illustrated in Table 6.2, including individual variation ranges. The table also shows the number of Polish respondents who articulated tested syllables of non-native length (out of EN min—max range obtained in the study). As in the previous tables showing the number of non-native-like responses (cf. section 6.2), a "+" following a number stands for longer duration or larger proportion; a "–" indicates values smaller than the corresponding native speakers' minimum. No mark following a number is equivalent to a "+".

Table 6.2. Stressed:unstressed syllable proportions in trochaic content words

Word	Unstressed syllable share (%)			Unstressed syllable share range (%)			PL out of EN range (N)	
	PL1	PL2	EN	PL1	PL2	EN	PL1	PL2
wanted (H3)	42	46	41	28—49	38—58	36—46	2–	6
sisters (H3)	30	30	19	25—39	21—40	5—27	9	8
started (H2)	28	28	25	17—37	22—36	16—32	3	2
Cinders (H2)	37	35	23	27—41	24—42	15—33	11	10
Cinders (N2)	32	32	28	23—44	22—44	22—35	2	5
gorgeous (N2)	29	28	30	23—33	20—37	22—35	0	1
shouting (N3)	37	32	28	28—58	20—42	24—33	8	4+, 1–

The data show similar tendencies to the results for unstressed vowel duration alone (cf. section 5.3.3). The unstressed syllables are significantly more reduced in native English speech if they are phrase-medial and/or contain an underlying /r/ in the syllable coda. Thus the unstressed syllable is similar in both groups in *gorgeous*, moderately shorter in native speakers' *Cinders* (N), and much shorter in *Cinders* (H). Likewise, the same kind of discrepancy between English and Polish speakers is more pronounced in *sisters* than in *wanted* or *started*. Apart from general timing tendencies, the unstressed *-ing* ending, as in *shouting*, is often lengthened in Polish learners' speech by a plosive consonant released after the velar nasal. The data presented in Tables 6.1 and 6.2 are graphically illustrated in Figure 6.1.

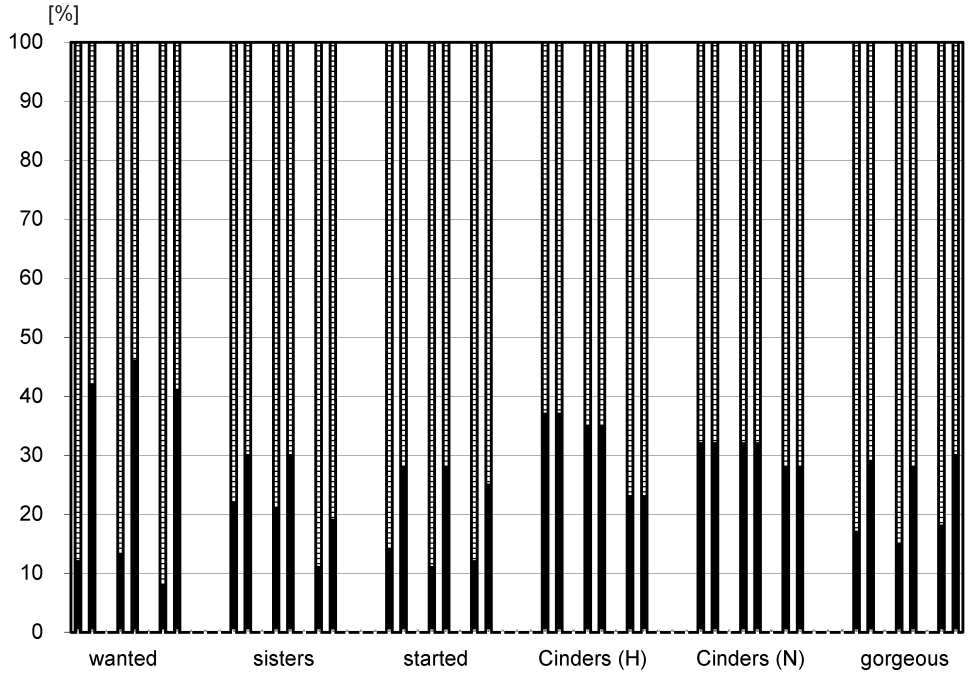

Figure 6.1. Relative duration of unstressed vowels and unstressed syllables in trochaic words

Within each pair, the left bar stands for relative unstressed vowel duration, while the right one signifies the unstressed syllable length, excluding the word-final fricative in *Cinders*, *sisters* and *gorgeous* (cf. section 4.5). Each context is represented by three pairs of bars, indicating, from left to right: PL1, PL2 and EN. The unstressed syllable in *Cinders* only comprises the vocalic signal, which keeps both vowel and syllable proportions identical.

6.2.2 The duration of unstressed function words

Absolute and relative unstressed vowel duration in monosyllabic function words was presented in sections 5.3.1 and 5.3.3. This section handles the duration of whole tested words in temporal relations to the context, which gives some insight into the role of consonant reduction in building the rhythmic patterns of phrases. Table 6.3 presents mean durations of unstressed function words and their proportions to the whole measured fragment (cf. section 5.3.3). All tested units are phrase-medial in order to eliminate the effect of final lengthening. For the purpose of this analysis, a total of 377 tokens were analysed (2×10×13=260 PL samples and 10×12=120 EN samples; 3 PL items disqualified owing to text distortions).

Table 6.3. Absolute and relative duration of unstressed function words

Measured unit	Weak form mean duration and PL:EN mean ratios					Weak form share in measured unit and PL:EN ratios				
	PL1	PL2	EN	**PL1:EN**	**PL2:EN**	PL1	PL2	EN	**PL1:EN**	**PL2:EN**
	ms					%				
shouting **at**	131	155	74	**1.77**	**2.09**	18.5	23.9	14.4	**1.28**	**1.66**
her fairy	209	136	89	**2.35**	**1.53**	31.2	27.2	20.7	**1.51**	**1.31**
her sisters	157	116	86	**1.83**	**1.35**	25.9	21.4	16.9	**1.53**	**1.27**
called **her**	223	128	117	**1.91**	**1.09**	49.6	38.5	38.9	**1.28**	**0.99**
sisters **had**	210	181	119	**1.76**	**1.52**	32.1	30.6	22.1	**1.45**	**1.38**
dreamed **of**	220	166	136	**1.62**	**1.22**	32.7	32.2	27.9	**1.17**	**1.15**
buy **some**	246	222	177	**1.39**	**1.25**	49.8	50.2	46.5	**1.10**	**1.10**
had **to**	117	81	68	**1.72**	**1.19**	28.9	27.8	30.5	**0.95**	**0.91**
going **to**	67	71	43	**1.56**	**1.65**	17.2	20.1	16.8	**1.02**	**1.20**
wanted **to**	96	76	50	**1.92**	**1.52**	18	18	14	**1.29**	**1.29**

All measured function words are longer in the pronunciation of Polish learners. Table 6.4 shows the number of Polish learners who did not reduce their weak forms enough to fall within the native speakers' ranges.

Table 6.4. Absolute weak form duration ranges

Function word in context	Mean duration (ms)			Group duration range (ms)			Responses out of EN range	
	PL1	PL2	EN	PL1	PL2	EN	PL1	PL2
shouting **at**	131	155	74	85—203	99—324	47—104	11	10
her fairy	209	136	89	83—296	47—238	57—119	11	7+, 1–
her sisters	157	116	86	103—279	43—210	57—136	9	4+, 1–
called **her**	223	128	117	139—363	84—186	88—171	11	2+, 2–
sisters **had**	210	181	119	169—255	116—241	82—152	13	10
dreamed **of**	220	166	136	124—377	131—205	106—158	10	8
buy **some**	246	222	177	182—308	172—295	137—210	10	7
had **to**	117	81	68	75—277	62—128	35—108	3	2
going **to**	67	71	43	34—92	48—111	25—69	4	4
wanted **to**	96	76	50	66—130	48—122	24—96	5	3

Table 6.4 demonstrates that a large proportion of Polish learners do not reduce unstressed function words. The scale of this discrepancy between English and Polish speakers makes the use of more advanced statistical measures redundant.

The longer pronunciation of /t/ in *to* by the native speakers results in less difference between native speakers and the learners in terms of absolute duration. The other weak forms, difficult in the first recording, remain difficult despite phonetic training, with the exception of *her*, where the instruction (including /h/-reduction) has led to considerable, in isolated cases even excessive, length reduction.

The Polish learners' weak forms are also longer in terms of timing proportions unless the reference content word is prolonged to a similar degree. This is the case in *buy some*, *had to*, *going to* (PL1) and *called her* (PL2). Table 6.5 presents individual timing variation in terms of the function word relative duration ranges.

Table 6.5. Relative function word durational variation

Measured unit	Mean weak form share (%)			Weak form share range (%)			PL responses out of EN range (N)	
	PL1	PL2	EN	PL1	PL2	EN	PL1	PL2
shouting **at**	18.5	23.9	14.4	14—30	15—41	10—19	5	10
her fairy	31.2	27.2	20.7	18—44	11—41	14—28	8	5+, 1–
her sisters	25.9	21.4	16.9	14—40	10—37	12—23	10	5+, 2–
called **her**	49.6	38.5	38.9	39—57	26—52	31—49	7	3+, 5–
sisters **had**	32.1	30.6	22.1	19—38	25—37	18—28	12	8
dreamed **of**	32.7	32.2	27.9	20—51	23—44	22—32	6+, 1–	6
buy **some**	45.9	46.3	46.5	45—56	46—58	42—54	3	2
had **to**	28.9	27.8	30.5	21—45	21—40	18—45	0	0
going **to**	17.2	20.1	16.8	9—24	14—29	11—23	1+, 1–	3
wanted **to**	18	18	14	13—22	11—25	9—23	1	1

The role of particular content words in phrase timing is investigated in the next section.

6.2.3 The duration of content words

The content words used as reference for the analysis of weak form duration are 8—45% longer in Polish learners' first recording. The group mean durations are displayed in Table 6.6. The Polish subjects' production is also compared to native English speech.

The difference is larger in fixed phrases, especially constructions functioning as alternative to modal verbs, i.e. *going to* — 54%, *had to* — 78%, which are prone for reduction in native speech. These words are also longer in PL2 recordings, where the other content words become comparable in duration

Table 6.6. Content word duration in comparison to unstressed function word duration

Measured unit	Content word duration (ms)					Weak form (WF) duration (ms)					WF share (%)		
	PL1	PL2	EN	**PL1:EN**	**PL2:EN**	PL1	PL2	EN	**PL1:EN**	**PL2:EN**	PL1	PL2	EN
shouting at	582	477	435	**1.34**	**1.1**	131	155	74	**1.77**	**2.09**	18.5	23.9	14.4
sisters had	453	412	418	**1.08**	**0.99**	210	181	119	**1.76**	**1.52**	32.1	30.6	22.1
her **sisters**	453	412	418	**1.08**	**0.99**	157	116	86	**1.83**	**1.35**	25.9	21.4	16.9
her **fairy**	453	347	341	**1.33**	**1.02**	209	136	89	**2.35**	**1.53**	31.2	27.2	20.7
wanted to	439	343	303	**1.45**	**1.13**	96	76	50	**1.92**	**1.52**	18.0	18.0	14.0
going to	325	280	211	**1.54**	**1.33**	67	71	43	**1.56**	**1.65**	17.2	20.1	16.8
dreamed of	449	359	357	**1.26**	**1.01**	220	166	136	**1.62**	**1.22**	32.7	32.2	27.9
buy some	249	219	203	**1.23**	**1.08**	246	222	177	**1.39**	**1.25**	45.9	46.3	46.5
called her	221	202	183	**1.21**	**1.10**	223	128	117	**1.91**	**1.09**	49.6	38.5	38.9
had to	276	212	155	**1.78**	**1.37**	117	81	68	**1.72**	**1.19**	28.9	27.8	30.5

to native performance. The difference of 33% and 37%, respectively, makes them similar to function words as these are 22—109% longer in PL2 production, except for *called her* (see section 5.3 for an interepretation of a longer native duration of the pronoun in this context), with 9% difference. The results displayed in Table 6.6 are graphically illustrated in Figure 6.2.

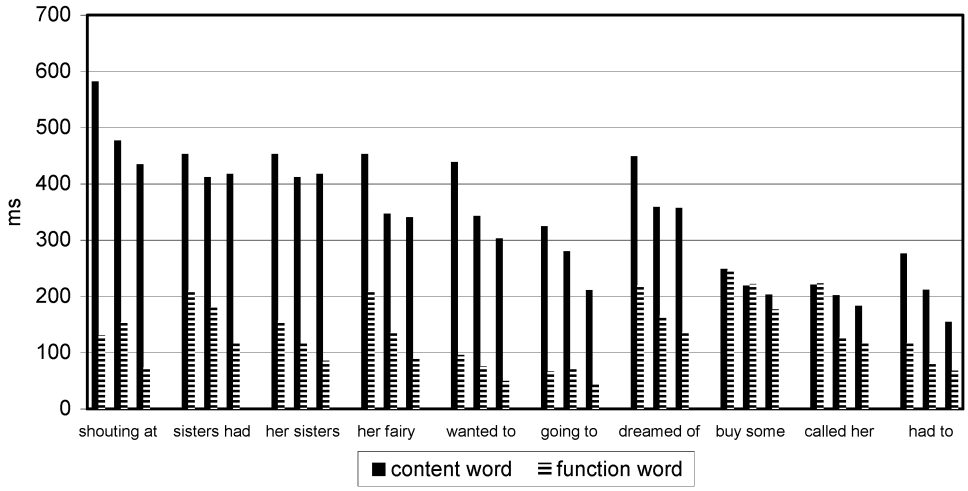

Figure 6.2. Mean length proportions between neighbouring content and function words in PL1, PL2 and EN, respectively from left to right

Finally, we present the scale of individual length variation of content words in terms of absolute values (Table 6.7) and timing proportions (Table 6.8), with the number of Polish subjects out of the native duration ranges.

Table 6.7. Absolute content word duration ranges

Word	Mean duration (ms)			Duration range (ms)			Responses out of EN range	
	PL1	PL2	EN	PL1	PL2	EN	PL1	PL2
shouting at	582	477	435	430—751	350—688	341—494	11	4
sisters[*] had	453	412	418	338—701	309—521	339—490	3+, 1–	2+, 2–
her **fairy**	453	347	341	367—685	250—433	277—425	4	1
wanted to	439	343	303	306—575	252—416	222—399	10	3
going to	325	280	211	258—476	206—364	166—239	13	10
dreamed of	449	359	357	305—571	259—495	295—493	5	2+, 1–
buy some	249	219	203	194—331	169—276	161—238	5	2
called her	221	202	183	153—288	128—265	151—256	3	1+, 1–
had to	276	212	155	175—343	137—316	100—204	12	7

[*] Since no context reference is required here, *sisters* appears only once in Table 6.7.

The absolute duration of content words was generally longer in Polish learners, and most of them pronounced a much longer *going* and *had* also in the second recording. Most subjects managed to shorten *shouting* and *wanted*, which in the former case resulted in less native-like timing of the tested word in combination with *at*. Otherwise, the number of PL responses within the native timing limits generally increased (cf. Table 6.8).

Table 6.8. Relative content word duration ranges

Measured unit	Mean content word share (%)			Content word share range (%)			PL responses out of EN range (N)	
	PL1	PL2	EN	PL1	PL2	EN	PL1	PL2
shouting at	81.5	76.1	85.6	70—86	59—85	81—90	5	10
sisters had	67.9	69.4	77.9	62—81	63—75	72—82	12	8
her **sisters**	74.1	78.6	83.1	60—86	63—90	77—88	10	5+, 2–
her **fairy**	68.8	72.8	79.3	56—82	59—89	72—86	8	5+, 1–
wanted to	82.0	82.0	86.0	78—87	75—89	77—91	1	1
going to	82.8	79.9	83.2	76—91	71—86	77—89	1+, 1–	3
dreamed of	67.3	67.8	72.1	49—80	56—77	68—78	6+, 1–	6
buy some	54.1	53.7	53.5	44—55	42—54	46—58	3	2
called her	50.4	61.5	61.1	43—61	48—74	51—69	7	3+, 5–
had to	71.1	72.2	69.5	55—79	60—79	55—82	0	0

The final part of this section refers to dactylic and trochaic cross-word feet. It reveals the level of timing difficulty concerning syllables in three positions: content word stressed, content word unstressed, and function word unstressed, with reference to Jassem's (1952) definition of narrow rhythm units and anacruses. Three dactylic feet used in the previous analysis have been segmented into individual syllables. For the four cross-word trochaic feet the data from the previous tables are used. The presentation is provided in two tables that show the number of Polish respondents who articulated tested syllables of non-native length (out of EN min—max range obtained in the study). Table 6.9 displays the trochaic content words in combination with a function word (dactylic feet), while Table 6.10 groups word combinations of a content word followed by a function word, both monosyllabic (trochaic feet). The remaining examples, i.e. *her sisters* and *her fairy*, which do not form feet, have been excluded from this part of the analysis, as has been dactylic *going to*, where a sonorant string of segments in *going* hinders reliable syllabification.

Apart from the Polish learners' problems with English speech timing, Table 6.9 again shows the complex interaction between absolute and relative

Table 6.9. Polish learners' problems with absolute and relative duration of syllables in cross-word dactylic feet

Foot	Stressed syllable				Unstressed syllable				Weak form			
	absolute		relative		absolute		relative		absolute		relative	
	PL1	PL2	PL1	PL2	PL1	PL2	PL1	PL2	PL1	PL2	PL1	PL2
shouting at	7	2	11–	11–	12	7+, 1–	5	2+, 3–	11	10	5	10
sisters had	2	1+, 1–	9–	9–	2	2	9	8	13	10	12	8
wanted to	6	2–	9–	7–, 1+	10	6	2–	2+, 1–	5	3	1	2

duration of speech units. Too long unstressed syllables result in non-native timing in most students, although the stressed syllable may also be longer. The timing of *shouting at* is an interesting example. The content word was significantly shortened by the learners in the second recording (5 more subjects within the native norms) but the apparent incompressibility of *at* resulted in non-native-like foot timing in 5 subjects more than in the first recording.

Analogical data concerning non-native durations and timing in trochaic feet, already presented for other purposes in Tables 6.6—6.8, are gathered in Table 6.10.

Table 6.10. Polish learners' problems with absolute and relative duration of syllables in cross-word trochaic feet (cf. Tables 6.6—6.8)

Foot	Absolute				Relative function[a]	
	content		function			
	PL1	PL2	PL1	PL2	PL1	PL2
dreamed of	5	2+, 1–	10	8	6+, 1–	6
buy some	5	2	10	7	3	2
called her	3	1+, 1–	11	2+, **2–**	7	3+, **5–**
had to	12	7	3	2	0	0

[a] Naturally, the number of non-native content word relative durations would be expressed by reverse values due to a dichotomic foot division.

The phrase *had to* does not conform to the general tendencies, which is caused by short native durations of fixed phrases performing grammatical functions similar to modal verbs (cf. *going to*, Tables 6.7—6.8).

Conclusions concerning the timing difference between unstressed syllables of content words and unstressed function words (Jassem's non-initial anacruses) cannot be drawn because the results are varied within each of the two classes. This indicates the impact of other factors, such as the phonological structure of unstressed syllables, a variable difficult to control in the present research material.

6.2.4 The effect of nuclear accent and phrase position

To observe the influence of nuclear accent on the duration of words[1] we need to separately analyse phrase-final and non-phrase-final accented words. Table 6.11 displays absolute durations of non-phrase-final nuclear syllables and words, and their share in the whole measured phrase. The phrases were measured without preheads in order to eliminate the influence of strong native reduction of that unit on the general timing relations. Although most of the measured phrases are not complete tone groups, they will be referred to as IPs for convenience.

Table 6.11. Mean absolute and relative duration of non-phrase-final nuclear words and their accented syllables

Unit / Group	Nuclear syllable				Nuclear word		
	god(mother)	wed(ding)	(e)nough	shout(ing)	wedding	enough	shouting
Mean absolute duration in (ms)							
PL1	267	200	275	360	380	329	582
PL2	270	156	275	320	321	311	477
EN	249	138	265	311	307	317	435
Unit duration relative to the whole measured phrase (%)							
PL1	16.5	11.8	20.0	18.3	22.4	23.8	29.4
PL2	19.4	11.2	23.5	18.1	23.2	26.8	26.6
EN	21.2	10.4	26.5	21.4	23.3	31.5	29.8

Even though each nuclear syllable and word (except *enough*) is shorter in native production, they constitute a larger proportion of the measured phrase. The differences between the two groups are quite consistent though not very large; therefore we have applied a one-way ANOVA test for group factor significance. Apart from these data, the two bottom rows of Table 6.12 display

Table 6.12. The duration of non-phrase-final nuclear words and syllables, and whole tested tone group units in cross-group relations

Relation	IP				Nuclear syllable				Nuclear word		
	FG	DoW	HEG	SSaC	god	wed	nough	shout	wedding	enough	shouting
PL1:EN	**<.001**	**<.001**	**<.001**	**<.001**	.227	**<.001**	.618	**.007**	**.009**	.534	**<.001**
PL1:PL2	**.008**	**<.001**	**.002**	.108	.864	**.004**	.968	**.043**	**.014**	.177	**.008**
PL1:PL2 t	**<.001**	**<.001**	**<.001**	**.027**	.299	**.003**	.482	**.007**	**.004**	.082	**.005**
PL2:EN	**.002**	.299	**.028**	.070	.195	.254	.620	.583	.572	.726	.144

Abbreviations: FG=fairy god(mother), DoW=dreamed of wedding (bells), HEG=had enough (gowns), SSaC=started shouting at Cinders.

[1] The durations of nuclear feet are presented in the next section, devoted to IP-level timing relations.

the significance of unit duration change in Polish speakers as a group (P1:P2 — ANOVA), and individual tendencies (P1:P2 t — paired t-test).

Statistical tests applied to longer prosodic units show a significant difference in IP duration between English speakers and untrained Polish learners (PL1:EN). There was, however, no regularity in cross-group differences in mean word and syllable duration. The second recording revealed (especially in the paired t-test) a significant shortening of those nuclear syllables and words which had been longer in PL1 (PL2:PL1, PL2:PL1 t). In effect, the duration of all syllables and words became comparable in Polish and English respondents (PL2:EN), and only two of the four measured tone groups still differed significantly.

Considering the effects of final lengthening described in the previous chapters, we expected a larger difference in phrase-final nuclear trochaic words, and a much larger difference in phrase-final nuclear monosyllables. These relations are demonstrated in Tables 6.13 and 6.14.

In line with prior expectations, the proportion of nuclear syllables and words in the whole phrases is consistently higher in native English speech, while all the measured IPs are 13—26% longer in PL1. In PL2, mean IP durations are only 5—9% longer than the native scores, but the relative duration of nuclear syllables and words, though consistently longer than in the first recording, remains considerably shorter than the native proportions. In terms of absolute values, all PL2 measured units approach native standards, but the timing relations are only slightly changed. Table 6.14 shows statistical significance of between-group mean absolute duration differences in IPs with final trochaic nuclei, the accented words and nuclear syllables. "PL1:PL2 t" refers to a paired t-test; the others are one-way ANOVA results.

The nuclear syllables and whole words, with the exception of *parties*, are similar in duration in PL1 and EN, while PL1 IPs are again considerably longer. In the learners' second recording, all measured units are similar in length to native speakers' performance. The word *parties*, shorter in PL1 as a result of weaker final lengthening effect, remains shorter in PL2, owing to the Polish learners' general tendency to accelerate.

The last group of nuclei, phrase-final monosyllabic words, show the same relations as trochaic nuclei. Slightly longer in PL1, in PL2 they often become shorter than EN responses when the learners accelerate. A similar trend is observed in the duration of whole measured portions of text. The lack of preheads means that the relative durations of monosyllabic nuclei depend on individual cases rather than showing consistent cross-group relations (Table 6.15). In most cases, however, the nuclear syllable constitutes a larger part of the phrase in native speech than in the Polish learners' performance, and the training does not significantly change these proportions. Smaller nucleus shares appear in the native speakers' *bad mood* and *find my hat*, where

Table 6.13. Absolute and relative durations of phrase-final trochaic nuclei (column 2), their accented syllables (1), and whole carrier phrases (3). Percentages in columns 3 (in bold) relate the IP mean durations to corresponding mean EN duration

Unit / Group	Columns											
	1	2	3	1	2	3	1	2	3	1	2	3
	Cin	Cinders	ChC	par	parties	GtP	clea	cleaning	HtDtC	gor	gorgeous	PWwG
Absolute duration (ms)												
PL1	318	470	914	326	440	846	248	489	1138	328	460	1370
PL2	296	439	770	314	430	809	244	464	1014	298	412	1158
EN	313	436	736	324	498	751	255	465	929	318	452	1084
Relative duration (%)												
PL1	35.1	51.9	**124**	38.7	52.2	**113**	22.1	43.3	**122**	24.0	33.8	**126**
PL2	38.8	57.0	**105**	39.3	54.0	**108**	24.2	45.9	**109**	25.8	35.7	**107**
EN	42.7	59.3	**100**	43.1	66.2	**100**	27.5	50.1	**100**	29.2	41.6	**100**

Abbreviations: ChC=called her Cinders, GtP=going to parties, HtDtC=had to do the cleaning, PWwG=Prince William was gorgeous.

Table 6.14. The duration of IPs with final trochaic nuclei (column 3), the trochaic nuclear words (2) and nuclear syllables (1) in cross-group relations (one-way ANOVA/paired t-test)

Unit / Group	Columns											
	1	2	3	1	2	3	1	2	3	1	2	3
	Cin	Cinders	ChC	par	parties	GtP	clea	cleaning	HtDtC	gor	gorgeous	PWwG
One-way ANOVA												
PL1:EN	.650	.500	**<.001**	.877	**.015**	**.010**	.608	.333	**<.001**	.625	.771	**<.001**
PL1:PL2	**.020**	.156	**.001**	.439	.583	.421	.720	.216	**.025**	.088	.032	**.001**
Paired t-test												
PL1:PL2	**.020**	.085	**<.001**	.131	.233	.214	.330	.038	**.001**	.039	.008	**<.001**
PL2:EN	.142	.906	.346	.514	**.003**	.202	.398	.973	.077	.250	.098	.148

Table 6.15. Group mean absolute and relative duration of phrase-final accented (nuclear) monosyllabic words (column 1) and their carrier IP's Head+Nucleus (2)

Unit / Group	Columns											
	1	2	1	2	1	2	1	2	1	2	1	2
	son	QOS	gone	WhShG	bride	LfaB	mood	BM	hat	FmH	gown(s)	BsNG
Mean absolute duration (ms)												
PL1	380	1199	340	1309	412	1020	315	668	356	804	397	1625
PL2	314	956	325	1194	395	905	307	651	320	699	351	1347
EN	368	1021	319	1065	401	879	305	662	355	819	371	1269
Relative duration (%)												
PL1	32.0	**117**	26.2	**123**	40.7	**116**	47.2	**101**	44.7	**115**	24.5	**121**
PL2	32.8	**94**	27.7	**112**	43.6	**103**	47.4	**98**	46.0	**85**	26.1	**106**
EN	36.1	**100**	29.9	**100**	45.9	**100**	45.7	**100**	43.4	**100**	29.2	**100**

Abbreviations: QOS=Queen's only son, WhShG=When her sisters had gone, LfaB=looking for a bride, BM=bad mood, FmH=find my hat, BsNG=buy some new gowns.

Table 6.16. The duration of phrase-final monosyllabic nuclear words and syllables (column 1) and whole tested H+N combinations (2) in cross-group relations

Unit / Ratio	Columns											
	1	2	1	2	1	2	1	2	1	2	1	2
	son	QOS	gone	WhShG	bride	LfaB	mood	BM	hat	FmH	gown(s)	BsNG
PL1:EN	.528	**.011**	.201	**<.001**	.582	**.009**	.597	.838	.971	.766	.184	**<.001**
PL1:PL2	**.001**	**<.001**	.411	.122	.489	**.028**	.597	.532	.043	**.009**	**.001**	**<.001**
PL1:PL2 t	**<.001**	**<.001**	.170	.040	.206	**.006**	.245	.192	.021	**<.001**	**<.001**	**<.001**
PL2:EN	**.008**	.178	.788	.086	.812	.569	.911	.659	.145	**.011**	.296	**.018**

the Polish learners make the head syllables less stressed and, consequently, much shorter.

The statistical analysis of durational differences shown in Table 6.16 confirms the observations that the untrained learners take more time to read an English phrase than native speakers unless it is a simple phrase with a high proportion of content words. Only *bad mood* and *find my hat* were not longer in PL1. The nuclear words, on the other hand, were all comparable in length.

Generally, the nuclear syllables are similar in absolute length in Polish learners and native English speakers, which makes them relatively longer, and therefore more prominent, in the latter group, considering the longer phrase durations in the learners' performance. This effect is additive to final lengthening effect.

6.2.5 Individual learners' results and progress

From the pedagogical point of view, we are also interested in the individual learner's performance, which is not sufficiently reflected in general results, comparisons of group means or even in data showing the general scale of observed problems. Even the figures indicating the number of subjects who meet native English standards may to some extent be misleading especially if the numbers referring to PL1 and PL2 refer to different students. Therefore this section is devoted to individual students' scores and progress.

The presentation of results refers to Hypotheses 2—3, focusing on syllable and word duration discrepancies between Polish and English speakers in unstressed positions and under nuclear accent. Table 6.17 ranks all speakers according to mean absolute duration of unstressed syllables in non-nuclear, non-phrase final trochaic content words.

Table 6.17 shows much stronger syllable reduction in native speakers. The raw durations, however, do not allow for individual speech rate differences and rank faster speakers higher. The analysis performed in Chapter 5 suggests that average speech rate does not depend on the speaker's L1. Still, the calculations are involved in stressed vowel durations alone and do not reflect longer and more complex phonological structures. Indeed, the average time a native speaker needs to pronounce a phrase is shorter, not only because of more reduced unstressed units, but also because of consonantal lenitions, less often employed by non-native speakers.

Since we aim at finding clearly defined differences between native and non-native speech in a situation where large individual variation makes it a difficult task, we will try to test whether the fast non-native speakers who rank among native ones for mean unstressed syllable duration make a temporal distinction between prominent and non-prominent units. To achieve this, we

Table 6.17. Individual speakers' mean unstressed syllable durations (MUSL) in ms. Native speakers' codes in bold

Subject	All unstressed syllables (19)	Subject	Weak forms (14)	Subject	Content words (5)
CSM	80	**CSM**	82	**CSM**	75
CTG	82	**CTG**	84	**CTG**	77
CLP	89	**CHB**	90	**CLP**	78
CHB	94	**CLP**	93	**CJE**	101
CMA	96	**CMA**	93	**CHB**	104
RM2	100	**CER**	93	**CMC**	104
CER	101	RM2	97	**CMA**	105
CPT	102	AS2	98	RM2	106
CJE	102	**CPT**	100	**CPT**	106
CMC	102	**CMC**	102	AK2	107
AS2	104	**CJE**	102	**CJI**	111
CJI	108	**CMF**	106	MG2	115
AK2	109	**CJI**	107	AS2	122
CMF	112	AK2	110	**CER**	123
CLH	115	**CLH**	111	RM1	125
AO2	123	AO2	121	MB2	125
AS1	129	DK2	121	PA1	126
RM1	131	AS1	122	PO2	126
DK2	134	PS2	132	**CLH**	127
PA2	137	PA2	133	**CMF**	127
PA1	142	RM1	134	AO2	130
PO1	142	PO1	142	PO1	141
PS2	142	JK2	145	AS1	147
MG2	143	AK1	147	AO1	147
MB2	146	PA1	149	PA2	149
PO2	149	LK2	150	AJ2	162
LK2	155	MG2	153	DK2	168
JK2	155	MB2	153	LK2	169
AO1	166	PO2	158	PS1	170
AK1	168	AJ2	171	DK1	171
AJ2	169	AO1	173	PS2	171
PS1	176	PS1	178	AJ1	175
MG1	182	MG1	180	MB1	178
JK1	182	JK1	183	JK1	180
DK1	190	LK1	192	JK2	185
AJ1	192	DK1	196	MG1	188
MB1	192	MB1	197	LK1	213
LK1	198	AJ1	199	AK1	226

calculate a speaker's personal mean syllable duration (MSL), which is a figure obtained by simply dividing the speaker's total measured duration of a reference set of six phrases[2] by the total number of syllables (37). The selected reference phrases are characterised by the lowest proportion of pauses, hesitations, text distortions and, in effect, relatively lower standard deviation in duration measures within particular groups of speakers.

This method is based on the same idea as the procedure described in Chapter 5, where we calculated the proportion of standard deviation to mean vowel length in order to demonstrate greater vowel duration "flexibility" in native speakers. The difference is that standard deviation cannot be used in this part of analysis because we do not (need to) know the actual duration of all syllables in the reference phrases. Still, given an MSL, which indicates the speaker's reading rate relative to all the other subjects, we can divide the figure by corresponding mean unstressed syllable duration (MUSL), calculated for 19 items (cf. Table 6.17), 14 weak forms of function words and 5 unstressed syllables in content words.[3] In this way, we obtain the syllable variation index (SVI), which shows the scale of the individual speaker's syllable length variation. The results and ranking of speakers are displayed in Table 6.18.

Table 6.18. Syllable variation index (SVI=MSL:MUSL) and individual speaker ranking by SVI (column 2) and MSL (column 6)

Columns					
1	2	3	4	5	6
Subject	SVI	MSL	MUSL	Subject	MSL
CSM	**1.89**	152	80	**CTG**	135
CLP	**1.82**	162	89	RM2	145
CMC	**1.73**	177	102	**CSM**	152
CHB	**1.72**	161	94	**CMA**	154
CJE	**1.70**	173	102	MG2	157
CJI	**1.66**	179	108	AS2	160
CTG	**1.66**	135	82	**CHB**	161
CMF	**1.63**	182	112	**CLP**	162
CER	**1.62**	165	101	**CER**	165
CPT	**1.62**	165	102	**CPT**	165
CMA	**1.60**	154	96	AK2	170
CLH	**1.59**	183	115	**CJE**	173
AK2	1.56	170	109	PA2	175
AS2	1.53	160	104	**CMC**	177
AS1	1.52	196	129	**CJI**	179

[2] See Appendix B.
[3] See Appendix C.

cont. tab. 6.18

Columns					
1	2	3	4	5	6
Subject	SVI	MSL	MUSL	Subject	MSL
PO1	1.51	214	142	AO2	182
DK2	1.49	200	134	**CMF**	182
AO2	1.48	182	123	**CLH**	183
PA1	1.45	206	142	RM1	191
RM1	1.45	191	131	PS2	191
RM2	1.45	145	100	AS1	196
MB2	1.36	198	146	PO2	198
PS1	1.35	238	176	MB2	198
PS2	1.34	191	142	MG1	199
JK2	1.33	206	155	DK2	200
PO2	1.33	198	149	LK2	203
AO1	1.32	220	166	JK2	206
LK2	1.31	203	155	PA1	206
AK1	1.29	216	168	PO1	214
AJ2	1.28	216	169	AK1	216
PA2	1.28	175	137	AJ2	216
DK1	1.28	242	190	AJ1	218
JK1	1.27	232	182	AO1	220
MB1	1.23	236	192	LK1	229
LK1	1.16	229	198	JK1	232
AJ1	1.13	218	192	MB1	236
MG2	1.10	157	143	PS1	238
MG1	1.10	199	182	DK1	242

As Table 6.18 (columns 1—4) illustrates, SVI cleanly separates native speakers from non-native speakers. Neither MSL nor MUSL alone are decisive for the speaker's ranking, although shorter values tend to occupy higher positions. Larger figures generally indicate a higher speech rate, usually connected with more confident, fluent use of language, which is also more likely to include more unstressed unit reduction.

The ranking of speakers according to mean syllable duration reveals other regularities. Almost half of the Polish speakers (6 in 13) in their second recording achieve a speech rate comparable to native speakers, but none of them reaches a native-like SVI value.

In terms of the learners' development, we observe a faster reading rate in all 13 speakers after the practice. Most of them (8) have improved their timing as well. Two subjects (PA and PO) have increased their distance from native timing standards, and the remaining three (PS, RM, MG) have not changed their syllable timing despite considerable acceleration.

6.2.6 Summary

In comparison to native English speakers' performance, untrained Polish learners' unstressed syllables are 32—94% longer in non-phrase-final trochaic content words, depending on syllable structure. After 7 months of training, the corresponding values are reduced to a 21—71% difference. Phrase-final unstressed syllables are comparable (cf. section 6.2.1). Tested function words are also longer, 39—135% in PL1 and 9—109% in PL2 (section 6.2.2). The Polish learners' function word relative durations are longer too, although certain fixed phrases, e.g. *had to* or *going to* show native-like proportions because the verbs are also considerably shorter in native performance. Both unstressed syllables of content words and unstressed function words are in many cases longer than the longest native durations in more than half of the Polish learners (cf. Tables 6.2, 6.4 and 6.5), which indicates a serious pedagogical problem.

The problems with weak forms and unstressed syllables are also visible in temporal relations to content word durations which, often longer in PL1, become similar to EN in PL2. The accompanying unstressed function words, despite the learners' general progress, remain significantly longer. These discrepancies are neutralised in phrase-final positions, where the native speakers show a tendency to lengthen their speech units more than do Polish learners.

This subchapter has shown generally more syllable durational variation in native English speech, as demonstrated in Table 6.18 by individual speakers' relations between their personal mean syllable duration and mean unstressed syllable duration (SVI).

6.3 Tone-group level relations

The last section of this chapter presents an analysis of timing relations between the units described in Chapter 4. The results are to verify Hypotheses 4 and 5, which concern higher prosodic levels.

6.3.1 Preheads

In this section we analyse the duration of preheads, where the shortest durations are expected in native English speech. Apart from the traditional notion of prehead, it may also be worth our investigating the timing of anacruses as defined by Jassem (1952; cf. section 1.1.1), which are claimed to be more radically reduced than the unstressed syllables in within-word feet

(cf. also Bouzon and Hirst 2004). However, the limited material of the present research contains rather few contexts where medial anacruses could be compared, and this would render potential results unreliable. The data presented in section 6.2 for the foot level only show a general difference between native English speakers and Polish learners in unstressed syllable reduction, but the actual discrepancies seem to depend on the structure of a tested unit more than on its position.

The present analysis compares the durations of anacruses in phrase-initial positions, i.e. preheads, in selected phrases.[4] Because of large individual variation, we concentrate on durational ranges in each group and the number of Polish learners who produced longer preheads than any of the native speakers.

Table 6.19. Absolute duration ranges of preheads and the number of Polish learners (N=13) exceeding the longest native durations (PL out)

Prehead in context	Syllables	PL1	PL2	EN	PL1 out	PL2 out
They dreamed of wedding bells.	1	120—344	85—195	82—147	9	3
...**in** honour of the queen's only...	1	88—413	69—145	63—168	2 (7)[a]	0 (4)[b]
...**and** going to parties.	1	169—441	118—216	53—148	13	11
...**They'd** wanted to buy...	1	115—817	95—311	90—257	8	2
...**but their** mother said...	2	226—386	199—329	146—256	12	7
...**that they** had enough gowns.	2	133—524	140—458	49—301	11	8
...**so they** started shouting...	2	246—459	199—518	124—336	9	7
...**there was a** girl called...	3	311—552	261—482	206—308	13	8
...**and he was** looking for a...	3	334—886	250—599	231—506	9 (12)[b]	2 (9)[c]
It was her fairy godmother.	3	344—819	281—598	227—357	10	11
They were in a bad mood.	4	307—754	297—837	251—407	12	8

[a] Responses longer than the second longest EN time (119ms)
[b] Responses longer than the third longest EN time (383ms)
[c] Responses longer than the second longest EN time (295ms)

The durations of preheads were varied in each of the groups, but almost all Polish learners in the first recording and still more than a half in the second recording exceeded the longest native values. These data are even more convincing than mean prehead syllable durations with respect to prehead complexity (Table 6.20).

[4] The phrase *And she had to do the cleaning* was excluded from analysis as it turned out that many subjects, especially native speakers, stressed the initial conjunction.

Table 6.20. Group mean prehead syllable duration and whole prehead mean duration. SD=standard deviation. PL:EN ratio in parentheses

Prehead	PL1	PL2	EN
P1	228 (2.04)	146 (1.30)	112
P2	179 (1.86)	152 (1.58)	96
P3	170 (1.73)	135 (1.38)	98
P4	144 (1.73)	121 (1.46)	83
Mean P duration	**369 (1.75)** SD=174 (47%)	**289 (1.37)** SD=143 (49%)	**211** SD=100 (47%)

No comparison is possible between values referring to preheads of different complexity because of various syllable structures and large individual variation, also illustrated by Table 6.19. It may only be stated that there is a tendency in Polish learners to make preheads approximately 37% longer than native speakers even after reading practice. Before the training, the difference amounts to approximately 75%. Mean prehead syllable duration tends to decrease in more complex preheads in both groups of speakers but no further conclusions are possible on the basis of the present results.

6.3.2 Heads

The absolute durations of head feet with respect to the number of constituent syllables is shown in Table 6.21.

Table 6.21. Head foot duration ranges

Unit	Syllables	PL1	PL2	EN	PL1 out	PL2 out
1	2	3	4	5	6	7
girl	1	231—402	213—463	171—394	1 (7)[a]	1 (3)[a]
called	1	201—461	128—337	170—259	9	6+, 2–
bad	1	284—478	251—446	316—407	2+, 3–	2+, 4–
queen's	1	391—600	285—517	298—491	5	1+, 1–
only	2	251—527	218—361	216—386	3	0
Cinders	2	249—480	254—438	258—360	6+, 2–	5+ ,1–
fairy	2	367—685	250—433	277—425	4	1+, 2–
started	2	345—576	294—536	256—750	0 (7)[b]	0 (4)[c]
called her	2	339—650	267—435	244—371	10	2
find my	2	321—646	281—512	302—637	1	1–
dreamed of	2	518—863	413—723	416—632	6	2+, 1–

cont. tab. 6.21

1	2	3	4	5	6	7
going to	3	312—540	263—426	216—308	13	9
wanted to	3	372—705	318—531	260—438	12	6
sisters had	3	515—870	426—786	460—600	9	6+, 1–
Why are you	3	295—551	223—491	224—349	11	7+, 1–
honour of the	4	491—1952	389—851	297—462	13	10
looking for a	4	436—812	359—666	347—577	8	3

[a] Responses longer than the second longest EN time (295ms)
[b] Responses longer than the second longest EN time (456ms)
[c] Ratios higher than EN second highest (0.71)

Head feet containing more than one unstressed syllable (3-4 syllables altogether) are significantly longer in Polish learners' pronunciation in comparison to less complex feet, which is also illustrated by group mean durations in Table 6.22.

Table 6.22. Group mean absolute and relative duration of head feet

Group / Unit	PL1	PL2	EN	PL1:EN	PL2:EN
H1	362.4	309.4	295.75	**1.23**	**1.05**
H2 (**H2:H1**)	455.07 (**1.26**)	374.78 (**1.21**)	369.93 (**1.25**)	**1.23**	**1.01**
H3 (**H3:H1**)	501.44 (**1.38**)	427.12 (**1.38**)	357.33 (**1.21**)	**1.40**	**1.20**
H4 (**H4:H1**)	751.62 (**2.07**)	567.27 (**1.83**)	427.42 (**1.45**)	**1.76**	**1.33**

The absolute length of non-nuclear accented syllables in PL2 and EN (H1) is comparable. Also the mean duration of trochaic head feet (H2) is similar in the two groups. Before the training, the learners make mono- and disyllabic feet 23% longer than the native speakers. Moreover, more complex feet (H3, H4) are 40—76% longer in their pronunciation, and this difference is reduced to 20—30% after the training. These proportions are a natural consequence of the relatively smaller foot length variation in native English production, suggesting a preference for more syllable-timing characteristics of Polish-accented English speech.

6.3.3 Prehead:Head relations

Both preheads and heads have been found longer in the pronunciation of Polish learners. The difference is reduced in the post-training performance, but it remains significant in most cases in preheads and those heads that include more than one unstressed reduced syllable.

This section shows prehead:head relations calculated in individual speakers' production. Therefore, mean values are replaced with median personal P:H ratios. They are displayed in Table 6.23 for ten tested phrases (380 tokens considered). Ratio ranges (in parentheses) show the scale of individual variation. The two rightmost columns (PL out) show the number of Polish subjects (N=13) with ratios larger than EN maximum (positive figures) or smaller than EN minimum (negative figures).

Table 6.23. Median Prehead:Head duration ratios (and ratio ranges) in native English and Polish-accented English performance

Phrase \ Group	PL1	PL2	EN	PL1 out	PL2 out
in : honour of the (P1:H4)	0.14 (0.09—0.42)	0.17 (0.13—0.3)	0.26 (0.15—0.42)	7–	5–
and : going to (P1:H3)	0.53 (0.31—1.08)	0.48 (0.42—0.73)	0.43 (0.2—0.56)	4	4
they'd : wanted to (P1:H3)	0.54 (0.25—1.44)	0.36 (0.3—0.7)	0.32 (0.22—0.64)	6	1
they : dreamed of (P1:H2)	0.28 (0.17—0.46)	0.22 (0.17—0.43)	0.22 (0.16—0.28)	6	3
and she : had to (P2:H2)	0.84 (0.59—1.22)	0.89 (0.8—1.35)	1.42 (0.87—1.73)	7–	6–
so they : started (P2:H2)	0.76 (0.61—0.98)	0.86 (0.65—0.98)	0.49 (0.34—0.88)	4 (10)[a]	4 (9)
and he was : looking for a (P3:H4)	0.96 (0.54—1.25)	0.83 (0.54—1.06)	0.63 (0.4—1.39)	0 (9)[b]	0 (3)
it was her : fairy (P3:H2)	1.16 (0.74—2.14)	1.31 (0.76—1.68)	0.89 (0.62—1.17)	6	10
there was a : girl (P3:H1)	1.33 (0.83—2.09)	0.91 (0.57—1.32)	1.04 (0.61—1.61)	3	1–
they were in a : bad (P4:H1)	1.62 (0.86—2.33)	1.43 (0.73—2.91)	0.9 (0.72—1.19)	11	8

[a] Ratios higher than EN second highest (0.71)
[b] Ratios higher than EN second highest (0.89)

Although the absolute durations of preheads are regularly longer in Polish learners' pronunciation, the timing differences between Polish and English speakers are not always significant. If a head is also longer in PL (cf. Tables 6.21—6.22), the P:H ratios are similar in native English and Polish-accented speech. The large individual variation in both groups, and its relevance for conclusions concerning the issue of rhythm, will be discussed in the next section, devoted to H:H ratios.

6.3.4 Head:Head relations

In Chapter 4 we hypothesise that if English is a more stress-timed language (which is also suggested by the results in section 6.3.2, especially Table 6.22), then the influence of the mother tongue may cause greater duration differences between successive head feet in the read speech of Polish learners. Table 6.24 shows median ratios and ratio ranges illustrating the timing of such feet. The two rightmost columns (PL out) show the number of Polish subjects (N=13) with ratios larger than EN maximum (positive figures) or smaller than EN minimum (negative figures).

Table 6.24. Within-head foot timing in native English and Polish-accented read speech. Individual H:H ratio ranges in parentheses

Phrase \ Group	PL1	PL2	EN	PL1 out	PL2 out
girl : called (H1:H1)	0.93 (0.66—1.57)	1.11 (0.68—2.97)	1.06 (0.88—1.74)	5–	1+, 3–
had to : do the (H2:H2)	1.39 (1.09—3.67)	1.15 (0.85—1.38)	0.91 (0.68—1.44)	4	0
honour of the : queen's (H4:H1)	1.85 (1.26—3.25)	1.74 (1.16—2.39)	1.02 (0.76—1.55)	10	7
queen's : only (H1:H2)	1.27 (1.11—1.67)	1.3 (0.96—1.77)	1.33 (0.95—1.62)	1	3
wanted to : buy some (H3:H2)	1.11 (0.72—2.03)	0.92 (0.75—2.48)	0.91 (0.73—1.28)	3+, 1–	1+, 1–
buy some : new (H2:H1)	2.49 (1.21—3.26)	2.62 (1.12—4.2)	2.39 (1.64—3.06)	1+, 1–	5+, 1–

Individual timing relations show a lot of variation also in native speakers, which opens the floor for discussion about the importance of rhythm for language production. Apparently, the tendency for rhythmic performance is not strong enough to override the influence of other factors, such as prosodic unit complexity or prominence distribution, on read speech timing. Significant timing differences are visible only in one example (H4:H1) with the largest foot complexity disproportion. Possible smaller discrepancies in foot pairs consisting of similar syllable numbers seem not to have been levelled off by individual variation considering similar median ratios in both groups.

Another kind of evidence for rhythmic differences between Polish and English speakers is provided by the analysis of H:H:H timing relations between the three feet in *(in) honour of the — queen's — only (son)*, and *(they'd) wanted to — buy some — new (gowns)*. The relative duration of each foot was

measured in individual speakers as a proportion (%) of the sum of all three feet. Then standard deviation from individual speakers' mean foot duration share ((F1+F2+F3)/3) was calculated. The group median personal SD values and SD ranges (in parentheses) obtained in this way are shown in Table 6.25. The lower lines (italics) for both phrases display standard deviation (in ms and %) from group mean foot duration calculated from all responses (12×3 for each group). In order to eliminate the possible influence of the uneven number of subjects in the groups on standard deviation value, we have not counted one Polish subject who had the highest standard deviation score in both recordings.

Table 6.25. The timing of three-foot heads. Upper line: group personal median SD (%). Lower line: SD from group mean foot duration

Group / Phrase	PL1 (N=12)	PL2 (N=12)	EN (N=12)
honour of the : queen's : only	15.8 (8.3—23.4)	14.7 (5.3—22.7)	6.61 (0.17—11.2)
Group mean SD in ms (and %)	*221.3 (41%)*	*177.9 (42%)*	*67.7 (20%)*
wanted to : buy some : new	15.4 (10.1—18.7)	15.6 (11.9—19.4)	13.5 (8.78—16.3)
Group mean SD in ms (and %)	*159.6 (39%)*	*142.3 (42%)*	*105.6 (35%)*

Overall standard deviation from group mean foot length is considerably lower in native speakers' production even after the most non-native-like performance for each phrase has been excluded from analysis. The figures in parentheses, representing relative foot duration flexibility ranges for individuals, indicate far more uniform foot length in native speakers. What is even more interesting is that less variation has been observed in native speakers' recordings of *(in) honour of the queen's only (son)*, despite the more varied number of syllables in individual feet.

The typical foot timing in the two phrases is shown graphically in Figures 6.3 and 6.4, based on group mean absolute durations of particular feet (provided in the graphs).

Finally, we focus on timing relations in individual subjects, since mean or median values show general timing tendencies but do not reflect actual performance timing. The data are displayed in Table 6.26. Individual speakers are ranked according to standard deviation from mean absolute foot duration and foot duration proportion (%) of the measured fragment, in increasing order.

The ranking differences between standard deviations calculated from absolute and relative durations are due to the fact that smaller figures, obtained in faster speech, yield smaller SD, which places native speakers slightly higher in the table. Generally, the native speakers are clustered in top ranks of the table in both classifications.

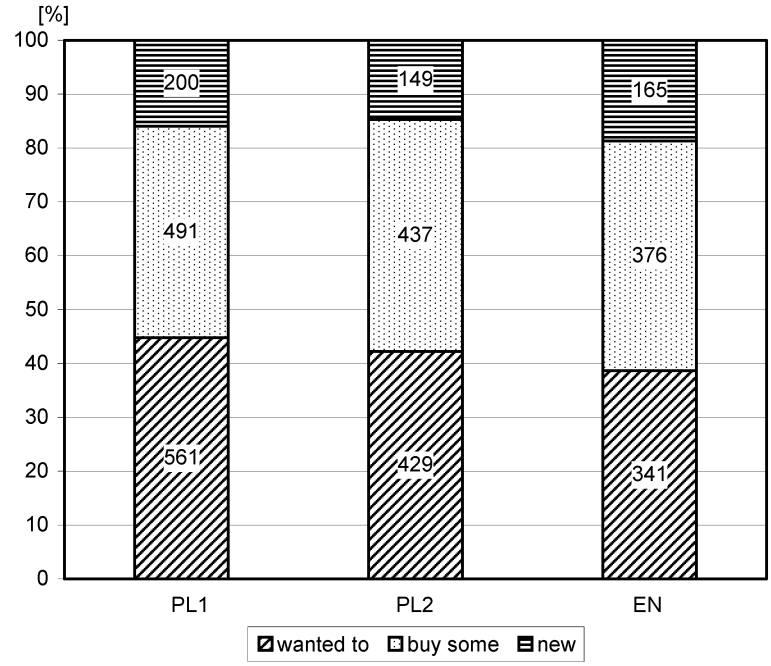

Figure 6.3. The timing of *wanted to* + *buy some* + *new*

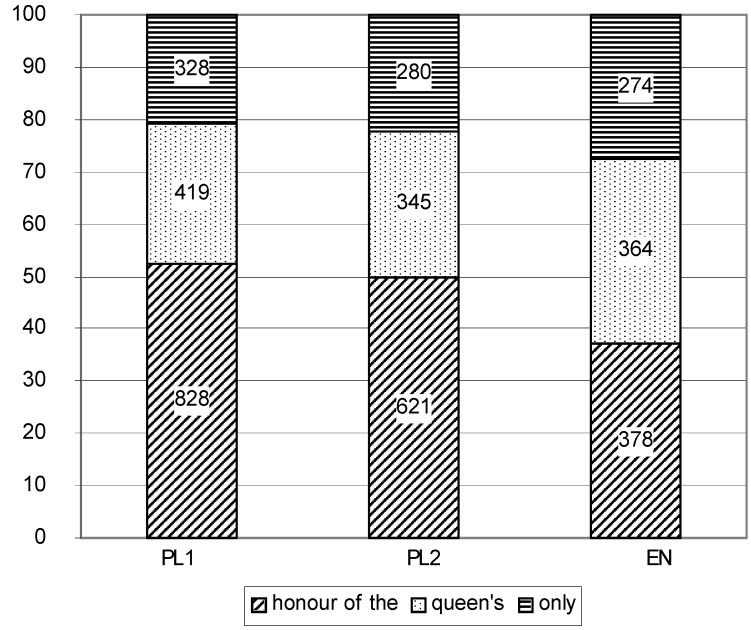

Figure 6.4. The timing of *honour of the* + *queen's* + *only*

Table 6.26. Individual speakers' head foot durational variation (SD)

...honour of the — Queen's — only...				...wanted to — buy some — new...			
SD (%)		SD (ms)		SD (%)		SD (ms)	
CLP	0.8	**CLP**	1.5	**CTG**	8.8	**CTG**	65.7
CJI	2.0	**CJI**	22.1	PS1	10.1	**CMC**	103.2
CMF	3.4	**CMF**	41.8	**CMC**	10.6	AS2	109.2
CJE	5.0	**CJE**	50.1	**CLH**	11.6	**CLH**	110.5
MB2	5.3	**CSM**	55.7	AS2	11.9	**CHB**	111.0
CSM	5.7	MB2	56.1	**CJI**	12.1	RM2	111.2
CER	6.4	**CMA**	59.8	AO1	12.4	**CJI**	112.9
AS2	6.7	**CER**	64.4	RM1	12.9	**CSM**	121.8
CMA	6.8	AS2	65.6	**CPT**	12.9	**CLP**	129.7
CMC	7.4	**CTG**	68.2	**CHB**	13.0	AK2	130.2
CTG	7.7	**CMC**	76.2	AO2	13.2	**CJE**	130.3
CLH	7.8	**CPT**	91.8	PO2	13.7	AO2	134.8
CPT	8.2	**CLH**	94.2	RM2	13.9	**CPT**	134.8
MB1	8.3	RM2	101.3	AK1	14.0	PS1	138.9
PA1	9.2	PA1	107.1	**CJE**	14.0	**CER**	140.6
AK2	9.4	**CHB**	112.7	AK2	14.5	**CMA**	140.7
RM2	11.0	AK2	135.7	AS1	14.5	RM1	144.0
CHB	11.2	MB1	161.6	AJ2	14.9	**CMF**	144.6
JK1	11.2	DK2	174.4	**CER**	15.0	PO2	157.6
LK1	12.1	JK2	175.2	**CMF**	15.1	AS1	157.8
JK2	14.0	JK1	179.6	PO1	15.4	MG2	160.7
DK2	14.5	AO2	190.5	DK1	15.4	AK1	166.6
PO2	14.6	LK2	191	**CMA**	15.8	DK2	167.1
AO2	14.8	PO2	207.4	**CSM**	16.2	AO1	171.7
LK2	14.8	PA2	213.1	**CLP**	16.3	AJ2	173.3
AO1	15.3	AK1	229.8	DK2	16.3	PA2	178.2
AK1	15.3	AO1	236.8	JK1	16.3	PS2	180.7
PA2	16.1	LK1	247	PA1	16.5	PA1	181.1
PS1	16.2	MG1	256	LK1	16.9	MG1	187.8
PO1	16.2	PO1	256.7	MB1	16.9	PO1	193.8
MG1	17.3	PS2	267.9	MG1	17.3	DK1	195.1
PS2	18.6	PS1	270.9	LK2	18.2	JK2	197.1
RM1	18.7	RM1	283.9	PS2	18.5	LK2	203.1
AS1	19.2	AJ2	284.4	MG2	18.7	JK1	207.8
AJ2	20.2	AS1	299.6	AJ1	18.7	MB2	217.2
MG2	22.7	MG2	328.5	MB2	19.4	MB1	221.2
AJ1	23.4	AJ1	388.5	JK2	19.4	AJ1	222.6
DK1	27.0	DK1	817.4	PA2	20.3	LK1	225.8

Table 6.27. Individual speaker's head foot durational variation (SD) in *...honour of the Queen's only... and ...wanted to buy some new...* Scores within native norms in bold

Score / Speaker	Phrase							
	honour of the Queen's only				wanted to buy some new			
	SD (ms)		SD (%)		SD (ms)		SD (%)	
	PL1	PL2	PL1	PL2	PL1	PL2	PL1	PL2
AJ	388.5	284.4	23.4	20.2	222.6	173.3	18.7	**14.9**
AK	229.8	135.7	15.3	**9.36**	166.6	**130.2**	**14.0**	**14.5**
AO	236.8	190.5	15.3	14.8	171.7	**134.8**	**12.4**	**13.2**
AS	299.6	**65.58**	19.2	**6.72**	157.8	**109.2**	**14.5**	**11.9**
DK	817.4	174.4	27.0	14.5	195.1	167.1	**15.4**	**16.3**
JK	179.6	175.2	**11.2**	14.0	207.8	197.1	**16.3**	19.4
LK	247.0	191.0	12.1	14.8	225.8	203.1	16.9	18.2
MB	161.6	**56.13**	**8.3**	**5.3**	221.2	217.2	16.9	19.4
MG	256.0	328.5	17.3	22.7	187.8	160.7	17.3	18.7
PS	270.9	267.9	16.2	18.6	**138.9**	180.7	**10.1**	18.5
PA	**107.1**	213.1	**9.24**	16.1	181.1	178.2	16.5	20.3
PO	256.7	207.4	16.2	14.6	193.8	157.6	**15.4**	**13.7**
RM	283.9	**101.3**	18.7	**11.0**	**144.0**	**111.2**	**12.9**	**13.9**

Table 6.27 displays individual Polish speakers' foot durational variation in the two phrases as standard deviation from mean foot length. Apart from proportion variation, standard deviation in milliseconds is also presented, which shows the combined effect of timing and speech rate.

The results (Table 6.26 rearranged) suggest more foot length variation in Polish learners' performance. There is also a tendency in most speakers (11 and 12 in respective phrases) to reduce the variation in the second recording. In some cases, however, it is not reflected in % of SD scores, which indicates a higher speech rate in the second recording, without arriving at more balanced foot relations. Three speakers out of the eleven (*honour of the Queen's only*) and nine out of the twelve (*They'd wanted to buy some new*) who reduced the foot variation in ms displayed higher variation indices calculated with reference to foot length proportions.

Only one subject (PA) in the first recording session obtained a native-like score in both measures, which was not repeated in the second, post-training attempt. AS and RM were the only Polish learners who managed to arrive at both SD in ms and SD in % within the limits of native speaker performance in both phrases.

6.3.5 Head:Nucleus (non-phrase-final) relations

The timing relations between a non-final nucleus and the preceding head foot are similar in both groups (Table 6.28). The results are presented as group median personal timing ratios.

Again, the two rightmost columns (PL out) show the number of Polish subjects (N=13) with ratios larger than EN maximum (positive figures) or smaller than EN minimum (negative figures).

Table 6.28. Head:Non-Phrase-Final-Nucleus timing in native English and Polish-accented read speech. Personal H:N ratio ranges in parentheses

Phrase \ Group	PL1	PL2	EN	PL1 out	PL2 out
dreamed of : wedding (H2:N2)	1.62 (1.34—2.66)	1.59 (1.11—2.1)	1.63 (1.16—2.41)	1	0
started : shouting at (H2:N3)	0.63 (0.47—0.86)	0.63 (0.53—0.94)	0.7 (0.48—1.4)	1–	0
why are you : crying my (H3:N3)	0.56 (0.43—0.76)	0.52 (0.34—0.77)	0.59 (0.32—0.7)	1	1
fairy : god(mother) (H2:N1)	1.62 (1.19—2.94)	1.31 (0.98—1.72)	1.41 (1.04—1.64)	5	1+, 2–

The results suggest little difference in accentual lengthening between the two groups of speakers. The H:N ratios are not higher in the production of Polish learners, a finding which does not support Hypothesis 4.

6.3.6 Prehead:Head:Nucleus relations

This last section of the chapter is an illustration of whole tone group timing characteristic of the two groups of subjects. It shows in a series of graphs the timing relations between tone group constituents based on group median durations of individual units. The relations are shown as proportions, but the actual median durations are also provided.

The phrase in Figure 6.5 illustrates several typical features of Polish-accented English production, i.e. a longer prehead, a longer head consisting of a fixed verbal construction, and a shorter phrase-final nucleus containing an intrinsically long vowel.

The conjunction *and* in *And she had to do the cleaning* (see Figure 6.6) is often stressed by the native speakers in this position, which makes it relatively longer. The difference in timing mainly depends on the shorter realisation of *had to*, another verbal fixed phrase which is reduced in length possibly as a result of lexical frequency effect.

The short prehead *in* (see Figure 6.7) is hardly compressible, which is reflected in only slightly shorter native durations. The decisive timing difference is connected with the inability of Polish learners to reduce the unstressed syllable sequence in *honour of the*. We can also observe relatively weaker phrase-final nucleus lengthening in Polish learners, who reduce the duration of the phrase in the second recording with no significant timing change.

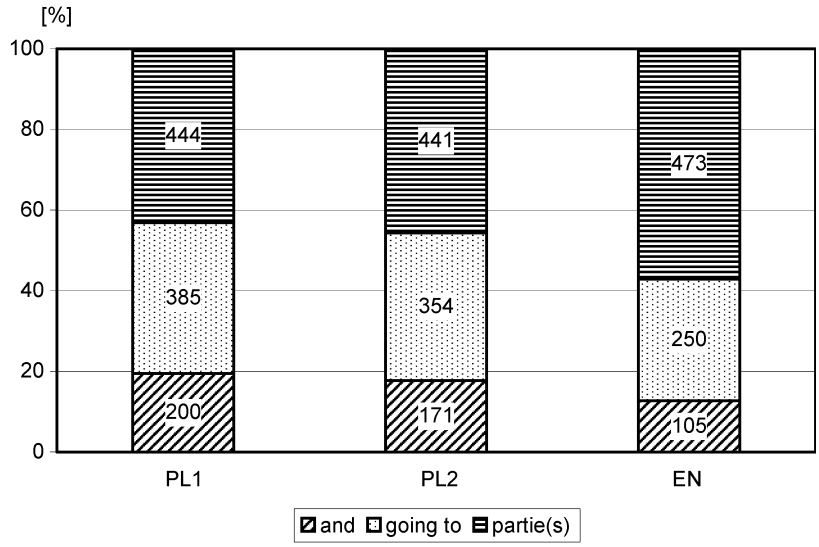

Figure 6.5. The timing of *and going to partie(s)*

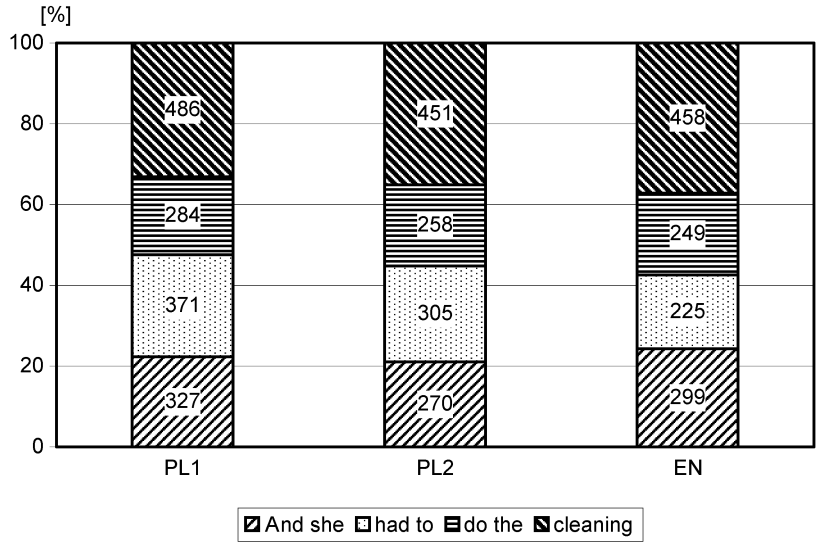

Figure 6.6. The timing of *And she had to do the cleaning*

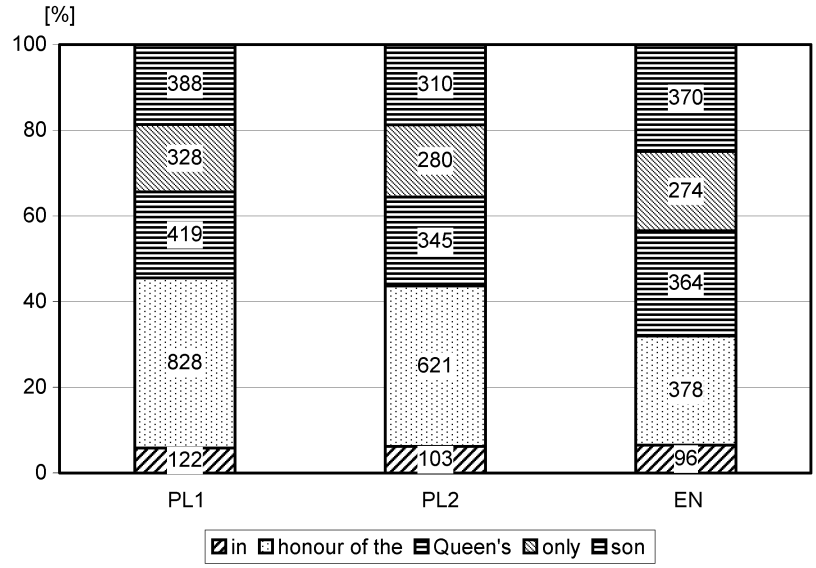

Figure 6.7. The timing of *in honour of the Queen's only son*

The phrase shown in Figure 6.8 is another example of less prehead reduction by Polish learners. In line with the previous example, the nucleus is relatively longer in relation to the head in native English performance, although in this case the learners' timing in the second recording is closer to native standards owing to significant shortening of the prehead and the complex, 4-syllable head.

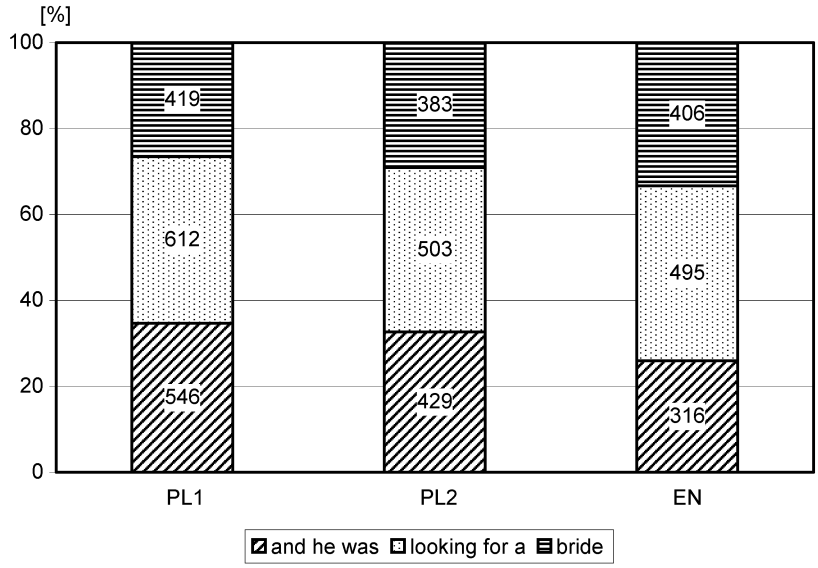

Figure 6.8. The timing of *and he was looking for a bride*

They dreamed of wedding bells (Figure 6.9) provides more evidence for stronger prehead reduction and shorter duration of heads in relation to nuclei and tails.

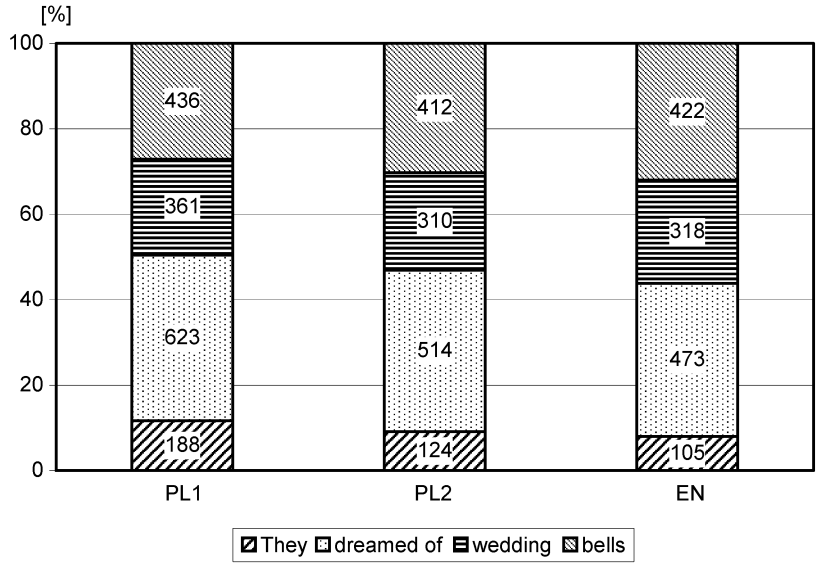

Figure 6.9. The timing of *They dreamed of wedding bells*

The complexity of prehead (see Figure 6.10) again turns out to be a major obstacle in acquiring native-like timing even for more advanced learners:

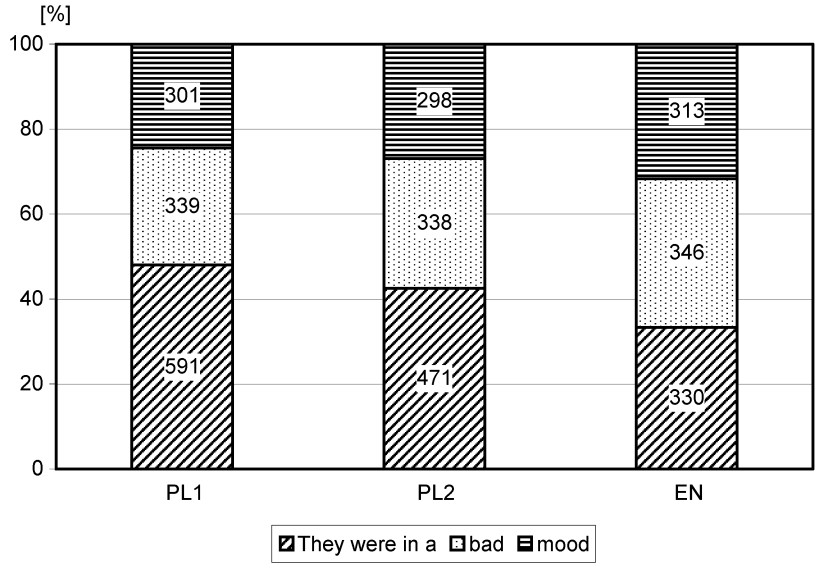

Figure 6.10. The timing of *They were in a bad mood*

They'd wanted to buy some new gowns (Figure 6.11) also illustrates Polish learners' tendency to make the early parts of an utterance (prehead, head) relatively longer, thus making the nucleus less prominent.

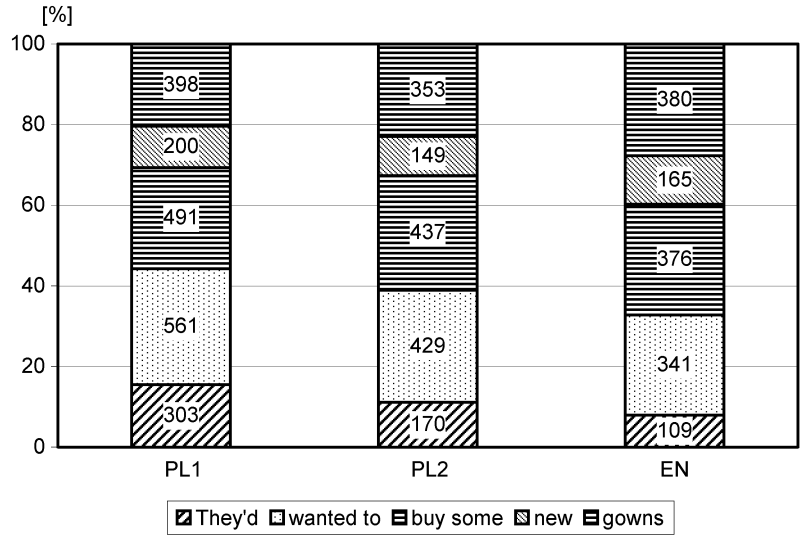

Figure 6.11. The timing of *They'd wanted to buy some new gowns*

Apart from the prehead problems, the following example suggests more discrepancy between Polish and English speakers in final lengthening than accentual lengthening:

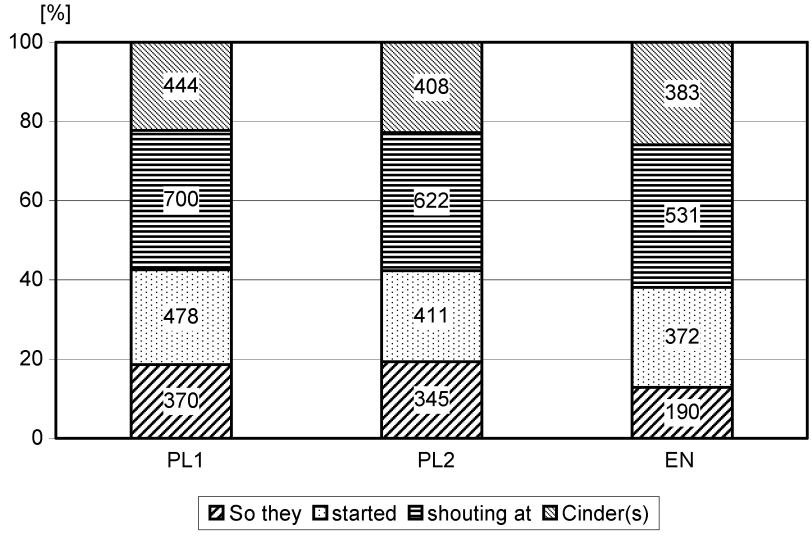

Figure 6.12. The timing of *So they started shouting at Cinders*

The last phrase, one with a monosyllabic nuclear foot, shows that this time it is the Polish learners who make the tail longer in relation to the nucleus:

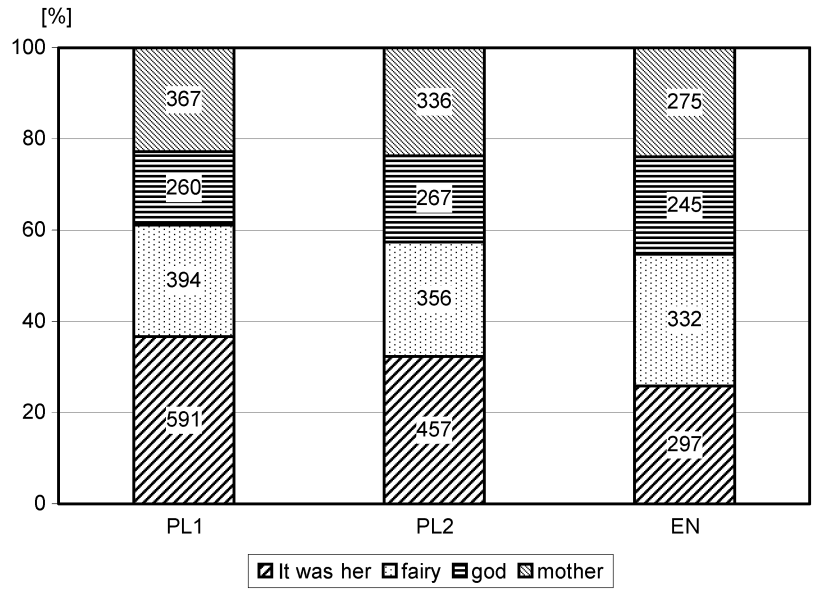

Figure 6.13. The timing of *It was her fairy godmother*

6.4 Conclusion

The analysis presented in Chapter 6 leads to the following conclusions concerning Hypotheses 2—5:

Hypothesis 2: Unstressed syllables, function words and preheads are longer in Polglish production in terms of both absolute and relative duration.

The hypothesis is supported by the data presented in section 6.2.1 (unstressed syllables in content words), section 6.2.2 (unstressed function words), although there is no significant difference in phrase-final syllables. Preheads are also much shorter in native speech (cf. section 6.3.1).

Hypothesis 3: Syllables and content words in nuclear accent positions are relatively shorter in Polglish production.

The hypothesis is not supported by the data presented in section 6.2.4. Significant differences occur in phrase-final position alone, which again suggests greater magnitude of final-lengthening in native speech in comparison to Polish-accented production.

Hypothesis 4: P:H and H:N ratios are higher in Polglish production.

The first part of the hypothesis is supported by the data presented in section 6.3.3. The ratios may not be significantly different, however, if the H contains a number of unstressed syllables or a fixed phrase, eg. *had to*, *going to*, etc. The hypothesis is not supported in its second part. H:N ratios are higher in Polglish production only if the nucleus is phrase-final. Otherwise, no significant difference has been observed between English and Polish speakers (cf. section 6.3.5).

Hypothesis 5: Within a head consisting of more than one foot, the standard deviation from mean foot length is larger in Polglish production.

The hypothesis is supported by the data presented in section 6.3.4. Despite large individual variation which may question the importance of rhythm in speech, there is a tendency in native English speakers to keep foot timing more regular than in the case of Polish learners. Conclusions referring to rhythmic differences, however, require further study.

Chapter 7

Conclusions and pedagogical implications

7.1 Limitations of the present study

Read speech analysis appears to be a reasonable compromise between the studies of natural, spontaneous speech and well-controlled experimental laboratory research. On the one hand, it allows direct inter-subject comparison of samples of the same text produced in the same conditions, which is difficult in natural conversational performance studies. On the other, it deals with more realistic language performance than is possible in experimental laboratory research.

Moreover, in the case of foreign language learners, confident spontaneous speech may be the ultimate goal of the teaching/learning process, but otherwise it generally consists in monitored speech practice, and reading texts is a typical production activity that reflects the learners' current FL pronunciation quality. At the same time, it displays their ability to consciously incorporate the native features introduced as part of the pronunciation training, which are often lost in conversational performance.

The choice of research procedures always involves the shortcomings inevitably connected with particular methods. In comparison to spontaneous speech samples, read speech often lacks natural communicative and emotional impact on the speaker's prosody. Moreover, compared to experimental studies, read speech research involves far weaker control of the interaction of individual independent variables, including extraneous ones. Duration is an acoustic parameter that can be measured at a relatively high level of reliability, but in a read speech study it is also practically impossible to isolate individual independent variables determining the length of measured units. In addition, separate continuous qualities such as intensity or pitch need to be categorised as one feature (e.g. stress or accent), often with some arbitrariness.

Furthermore, a read text still provides rather limited material for extensive investigation of selected units in various contexts, which would allow for better control of particular independent variables.

Finally, apart from the above limitations, speech timing displays a wide scale of individual variation, which makes regular differences between two groups of speakers difficult to observe.

All these problems lead us to interpret the findings very cautiously. We tend to draw conclusions only where differences between the two groups clearly appear to override the ranges of within-group individual variation. We do not rely solely on conventional statistical measures such as ANOVA, which are based on mean values and as such are rather sensitive to differences that in this specific study may still not be large enough to be considered significant. Therefore, we are particularly interested in the numbers of Polglish responses located outside of the limits set by the native speakers' results.

Such restrictions on possible conclusions have made us abandon certain lines of investigation, such as the impact of post-vocalic voicing on segmental duration, detailed comparison of phrase-medial anacruses and unstressed syllables of within-word feet, and, especially, more conclusive observations referring to accentual lengthening.

7.2 General conclusions: The Polish learner's English speech timing characteristics

Despite the limitations presented above, including large individual variation, certain timing characteristics of Polish-accented English speech have turned out to differ significantly from native production. The conclusions of Chapters 5 and 6 to various degrees confirm research hypotheses proposed before the study. Generally, we recognise the complexity of factors that influence the timing of native English read speech and the great individual variation which makes it difficult to establish certain timing norms that the Polish learners' results could refer to. However, as already mentioned, we assume that those learners' results which are not confined within the limits set by minimum and maximum native values can be perceived as "non-native-like". Considering this criterion, we gather conclusions referring to the six hypotheses proposed in Chapter 4. The conclusions are illustrated by spectrographic and waveform pictures showing typical differences[1] between native English and Polish learners' performances. Unless otherwise indicated, the pictures are scaled to show relative duration proportions between particular units in comparison to the juxtaposed examples.

[1] Actual examples are selected to reflect the general timing tendencies described in statistical terms in Chapters 5 and 6.

7.2.1 Hypothesis 1: Stressed vowel duration

Hypothesis 1: Mean duration proportion of intrinsically "long" to intrinsically "short" vowels is smaller in Polglish production.

The hypothesis is supported by the results and global statistics reflecting group mean durations of individual vowel types (short, "ash", long and

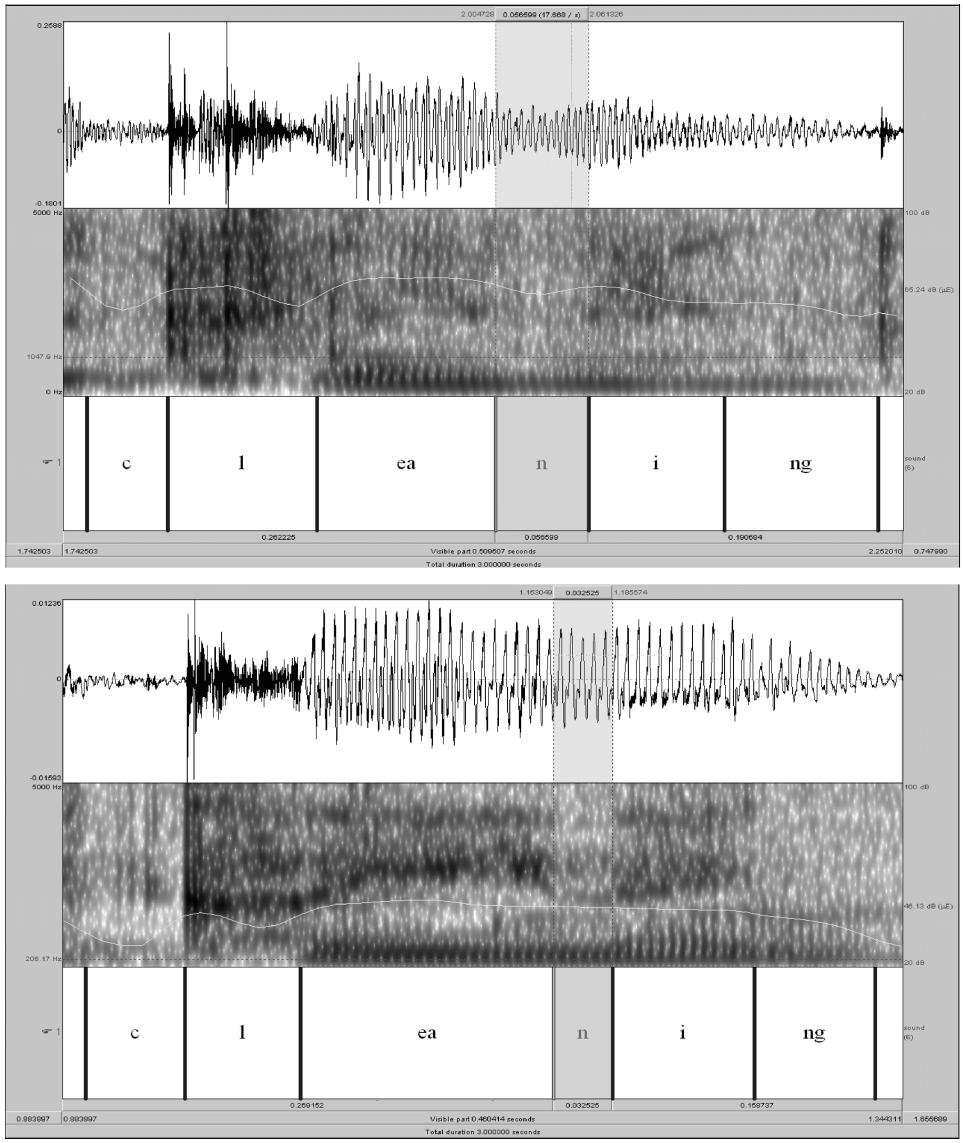

Figure 7.1. The timing of /i:/ and /ɪ/ in Polglish and native English (bottom) pronunciation of *cleaning*

diphthong) and more vocalic length variation calculated for individual speakers. The interaction of various other duration determinants does not show a clear difference in individual cases. Figure 7.1 shows an example of typical native English and Polglish timing of vowels in the word *cleaning* in an accented position.

For better measurement reliability, /l/ includes the signal from the release of /k/ to the onset of voicing only, thus adding the final voiced phase of lateral articulation to the vowel.

7.2.2 Hypothesis *2:* Unstressed unit duration

Hypothesis 2: Unstressed vowels, syllables, and function words are longer in Polglish production in terms of both absolute and relative duration.

Hypothesis 2 is supported with reference to all mentioned units and the unstressed IP constituent — the prehead (illustrated in section 7.2.4). Unstressed unit reduction is significantly stronger in the pronunciation of native English speakers. The typical differences between the two groups are illustrated in Figures 7.2—7.3.

Figure 7.2 presents two unstressed function words (*her*, *had*) and one unstressed syllable (*sisters*) in a trochaic content word in non-phrase-final position.

Each unstressed vowel (*h**e**r*, *sist**e**rs*, *h**a**d*) is shorter in native English speech. The unstressed syllable in *sisters* is relatively shorter in native pronunciation even though the plural ending is often omitted by the learners. The Polish learner also used a strong fricative /h/ in both function words, but she elided it in *her* in the second recording. Native English /h/ is hardly discernible in *her*, with clearly visible vocalic formants, while *had* is pronounced without the initial consonant. Another factor that prolongs Polglish *had* is the more salient plosive burst of the final /d/ before *gone*. In effect, in the native production, the stressed syllable of *sisters* is significantly longer than the following two unstressed syllables. In PL1 and PL2 these durations are similar.

Figure 7.3 shows typical English voiceless realisation of the weak form of *to*. Polish learners typically vocalise the syllable peak. We also observe relatively longer vowel duration in both syllables of *parties* pronounced by native speakers. The longer realisation of the unstressed syllable illustrates stronger final lengthening in native English speech, while the duration discrepancy between native and Polglish articulation of the accented vowel may be attributed to its intrinsic length combined with the different effects of nuclear accent and the phrase-final position of the word.

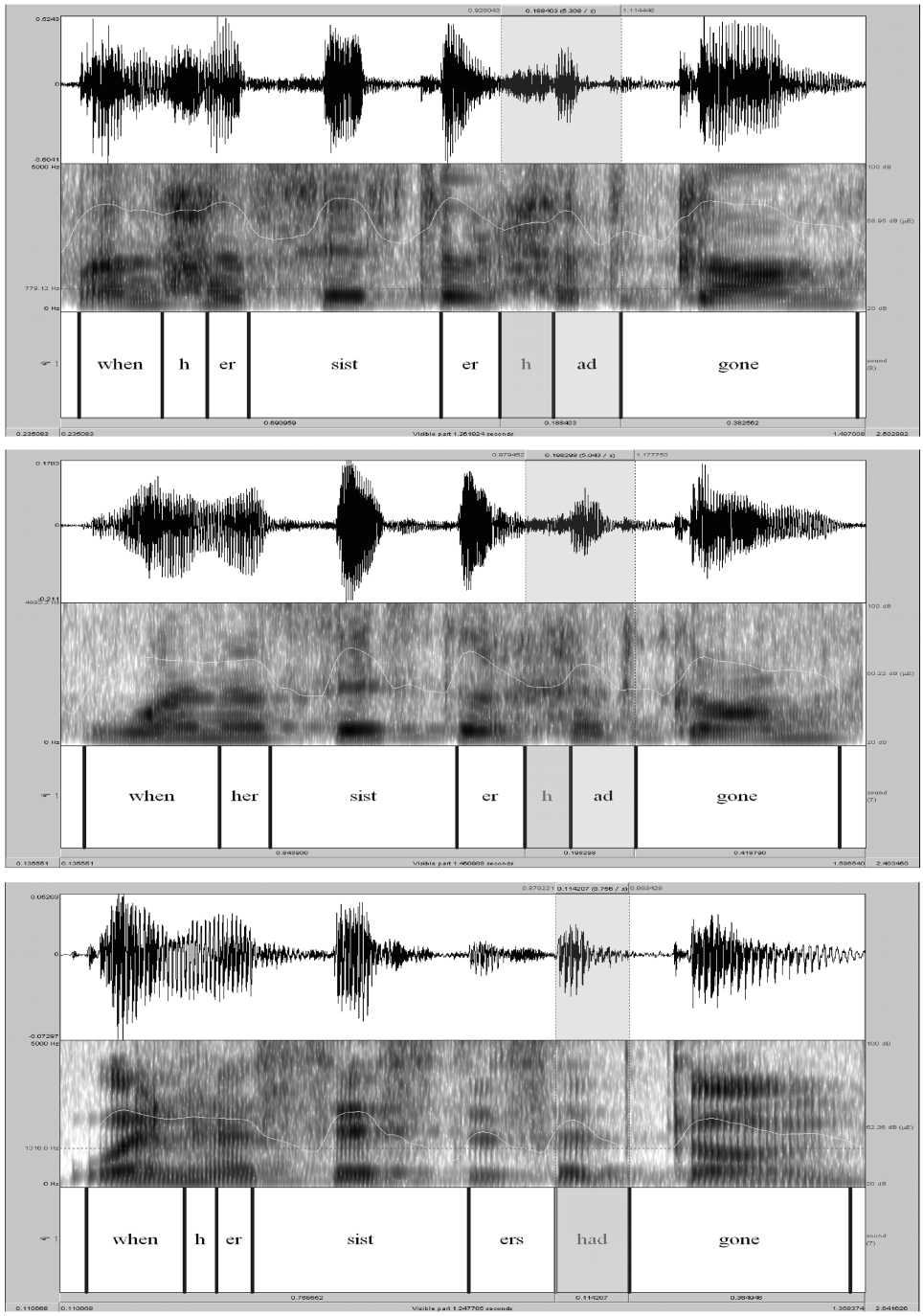

Figure 7.2. Typical PL1, PL2 (the same person) and EN (bottom) realisations of *When her sisters had gone*

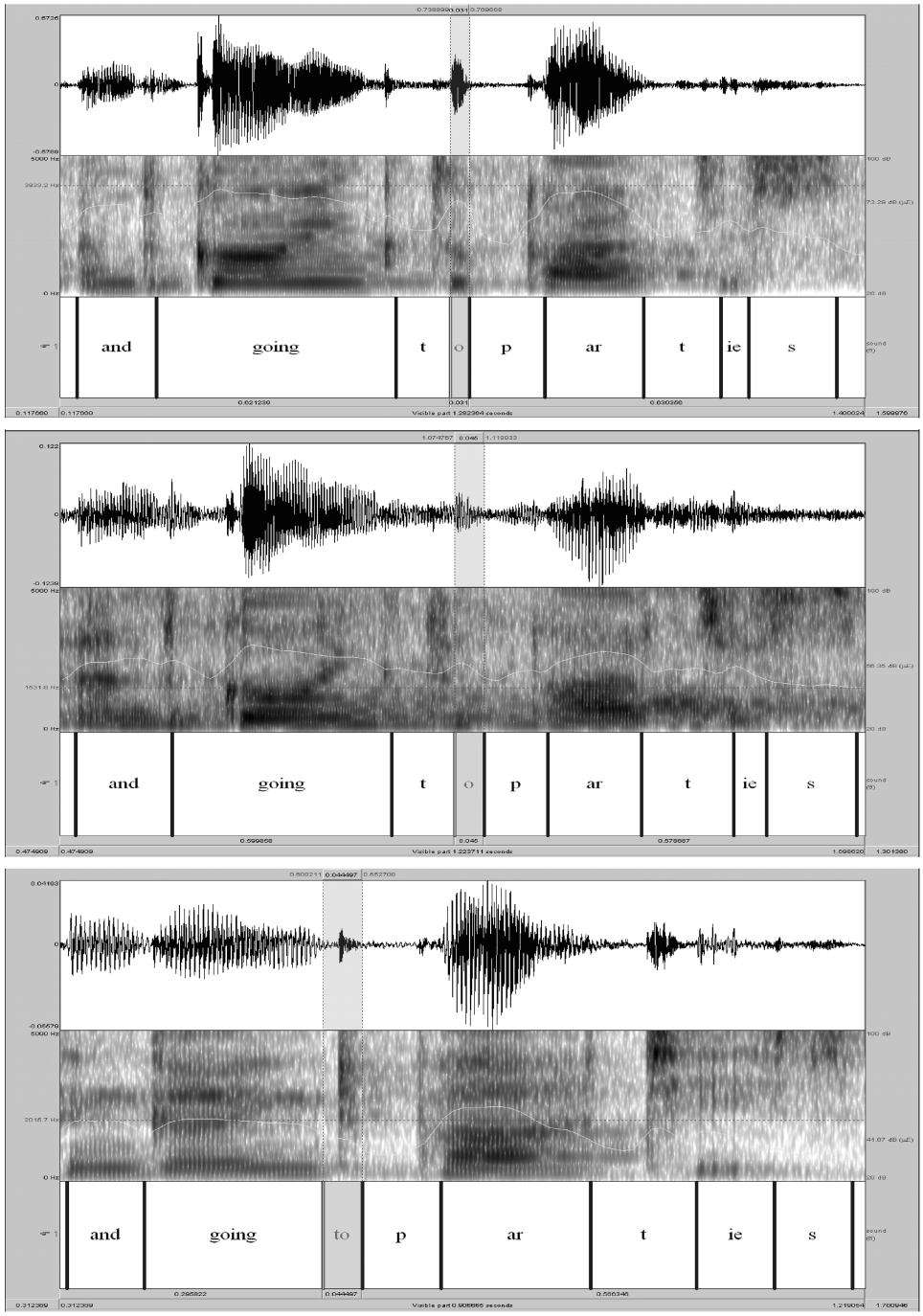

Figure 7.3. Typical PL1, PL2 (the same person) and EN (bottom) realisations of *and going to parties*

The examples of *her*, *had* and *to* illustrate the importance of weak form reduction, which drastically change the timing patterns. They also point out the significance of segmental issues in the acquisition of foreign language prosody. The different articulatory mechanisms in the production of /t/ and /h/ (both intrinsic and context-dependent) lead to various length proportions at successive levels of prosodic hierarchy.

7.2.3 Hypothesis 3: Nuclear accent effect on timing

Hypothesis 3: Vowels, syllables, and content words in nuclear accent position are relatively shorter in Polglish production.

As mentioned above, lack of prominence does not make the learners shorten the units in question to a native-like degree. Strong, nuclear units, on the other hand, are not lengthened as much as in native production. Yet this conclusion mainly applies to phrase-final position, the most frequent locus of focal prominence. Phrase-medial nuclei do not significantly differ in length in the performance of the two groups of subjects. Without the additive effect of final lengthening, the differences are not statistically significant. However, as illustrated in Figure 7.4, which shows the timing of a phrase-medial nuclear foot, the accented syllable is relatively longer in native pronunciation both in the onset /ʃ/ and nuclear syllable peak /aʊ/. The following two unstressed syllables in this foot combined are shorter than the accented one in the English subject's pronunciation. The opposite length relation in the Polish speaker makes us reluctant to reject Hypothesis 3 even with respect to non-final positions.

Similar relations are shown in Figure 7.5, where the 3-syllable nuclear foot is actually shorter in native speech, but the vocalic part constitutes a larger portion of the phrase.[2] Moreover, we must note again the influence of additional articulatory gestures such as plosive bursts following the velar nasal in PL performance, which adds to phrase duration.

Generally, no significantly stronger accentual lengthening has been found in native English performance unless where reinforced by final position. Phrase-final lengthening affects Polish learners' production less also if unstressed syllables occupy this position (cf. Figure 7.3). As a result, the otherwise large differences in unstressed syllable length disappear phrase-finally (cf. the comparison of *sisters*, *gorgeous*, and *Cinders* in final and medial positions). Because most of the tested nuclear units occupied

[2] This nucleus is probably more prominent than in the PL versions owing to greater relative intensity. However, two different sources of recordings used in this study and quality differences would render comparative analysis of intensity or F0 excursions unreliable.

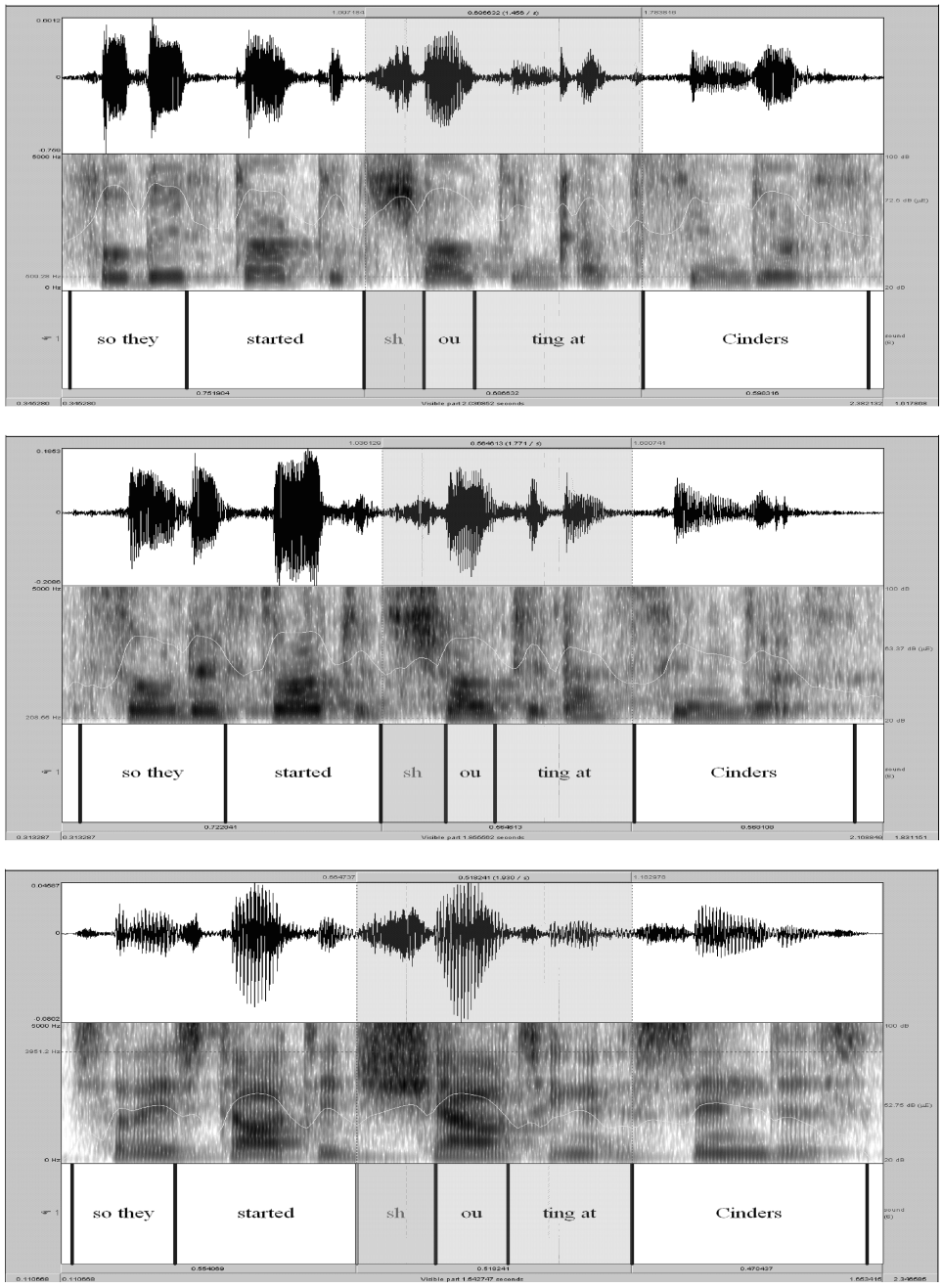

Figure 7.4. The timing of a non-phrase-final nuclear foot in *so they started shouting at Cinders*. Typical PL1, PL2 (the same person) and EN (bottom) realisations

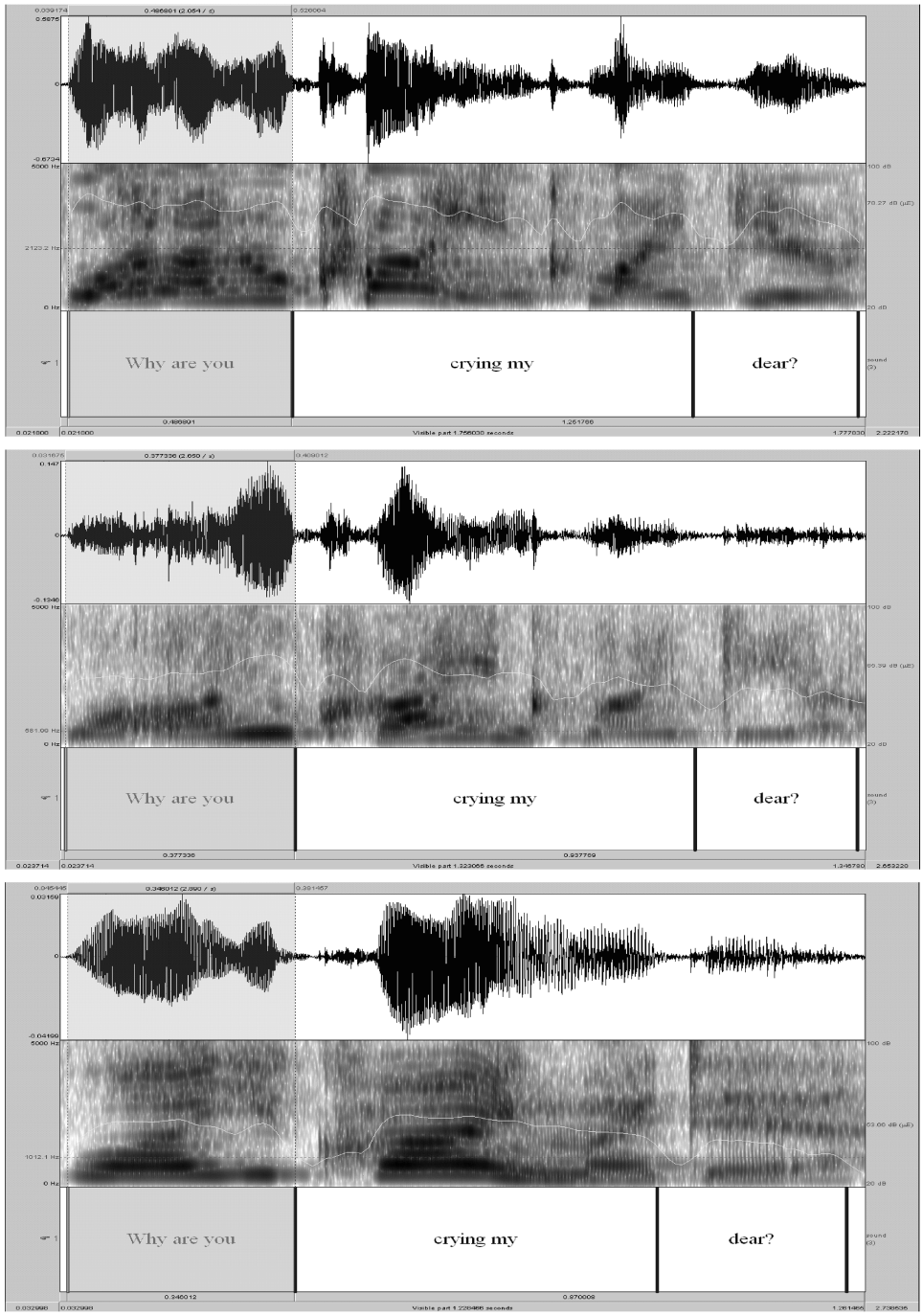

Figure 7.5. The timing of *Why are you crying, my dear?* Typical PL1, PL2 (the same person) and EN (bottom) realisations

phrase-final positions, we have been unable to gather enough evidence to claim significant differences in accentual lengthening between Polish and English speakers.

7.2.4 Hypothesis 4: IP constituent timing

Hypothesis 4: In an intonational phrase divided exhaustively into prehead (P), head (H), and nucleus (N), according to traditional British School parsing, P:H and H:N ratios are higher in Polglish production.

Hypothesis 4 has been confirmed by the results described in the previous two chapters. An example of typical timing of tone group constituents in native and Polish-accented English speech is provided in Figure 7.6.

Both ratios mentioned in Hypothesis 4 are typically larger in Polish learners' speech. Most of the H:N ratios in this study are affected by the final lengthening, but the most spectacular differences between the two groups of speakers generally concern unstressed elements: vowels, syllables (except in prepausal positions) and preheads. Apart from the phrase in Figure 7.6, the prehead duration discrepancy is illustrated in Figures 7.7, 7.8 and 7.9, which picture complex, 3-4-syllable phrase-initial units.

Relative length discrepancies between Polish and English speakers are often magnified in longer strings of unstressed units, where they can be aggravated by articulatory fluency problems. Figure 7.10 shows the absolute duration of *They were in a bad mood* produced by another pair of speakers.

The learner's timing becomes far more native-like when the prehead is articulated more fluently. The duration of monosyllabic content words is not considerably changed.

7.2.5 Hypothesis 5: Foot timing

Hypothesis 5: Within a head consisting of more than one foot, standard deviation from mean foot length is larger in Polglish production.

The limited research material gathered for this study supports Hypothesis 5. The general difference between English and Polish speakers is illustrated in Figure 7.11.

More head foot durational variation combined with less vowel and syllable durational variation constitute important evidence confirming a persistent tendency for syllable timing in the English speech of Polish learners.

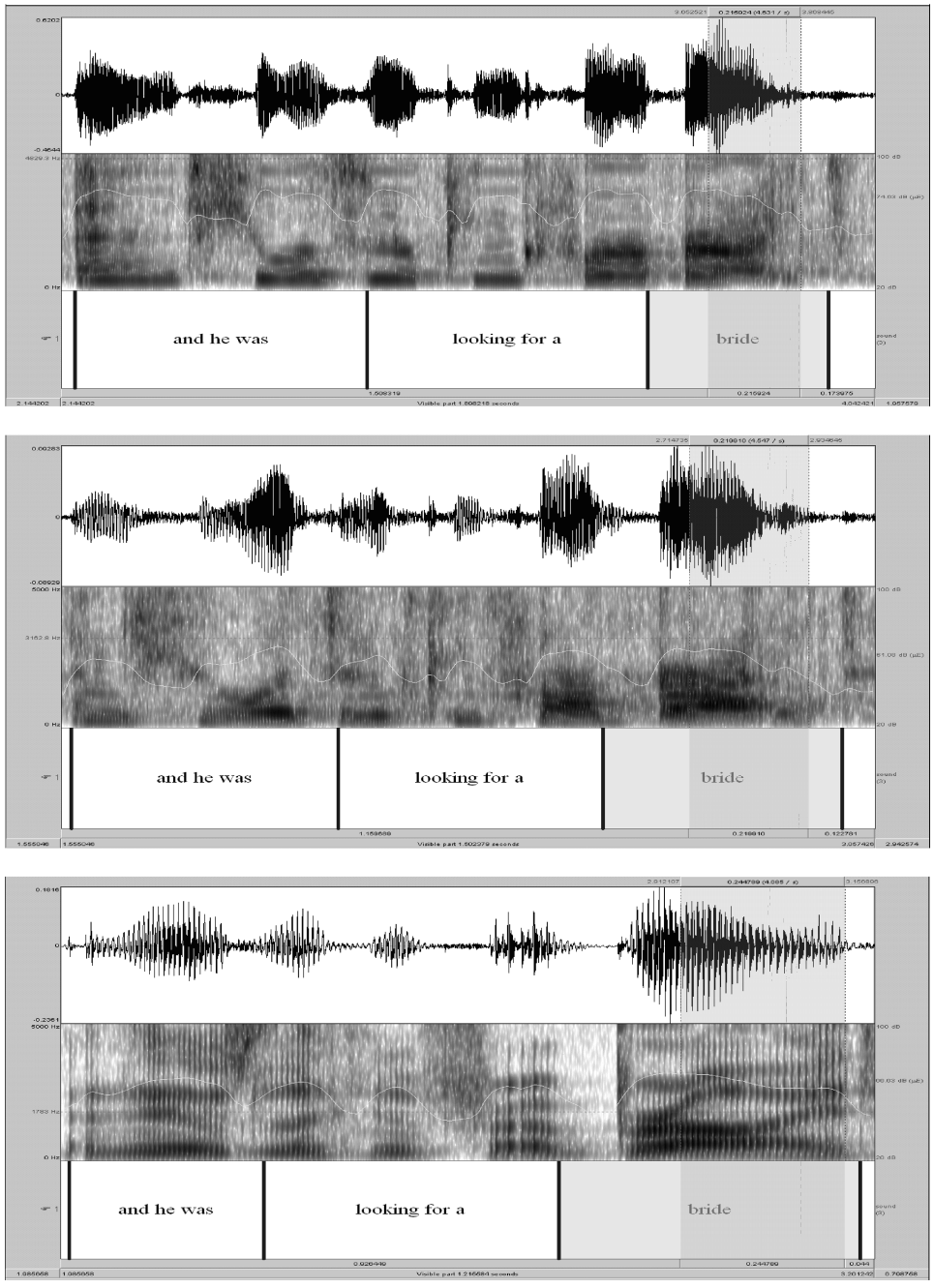

Figure 7.6. Typical PL1, PL2 (the same person) and EN timing of IP constituents in *and he was looking for a bride*

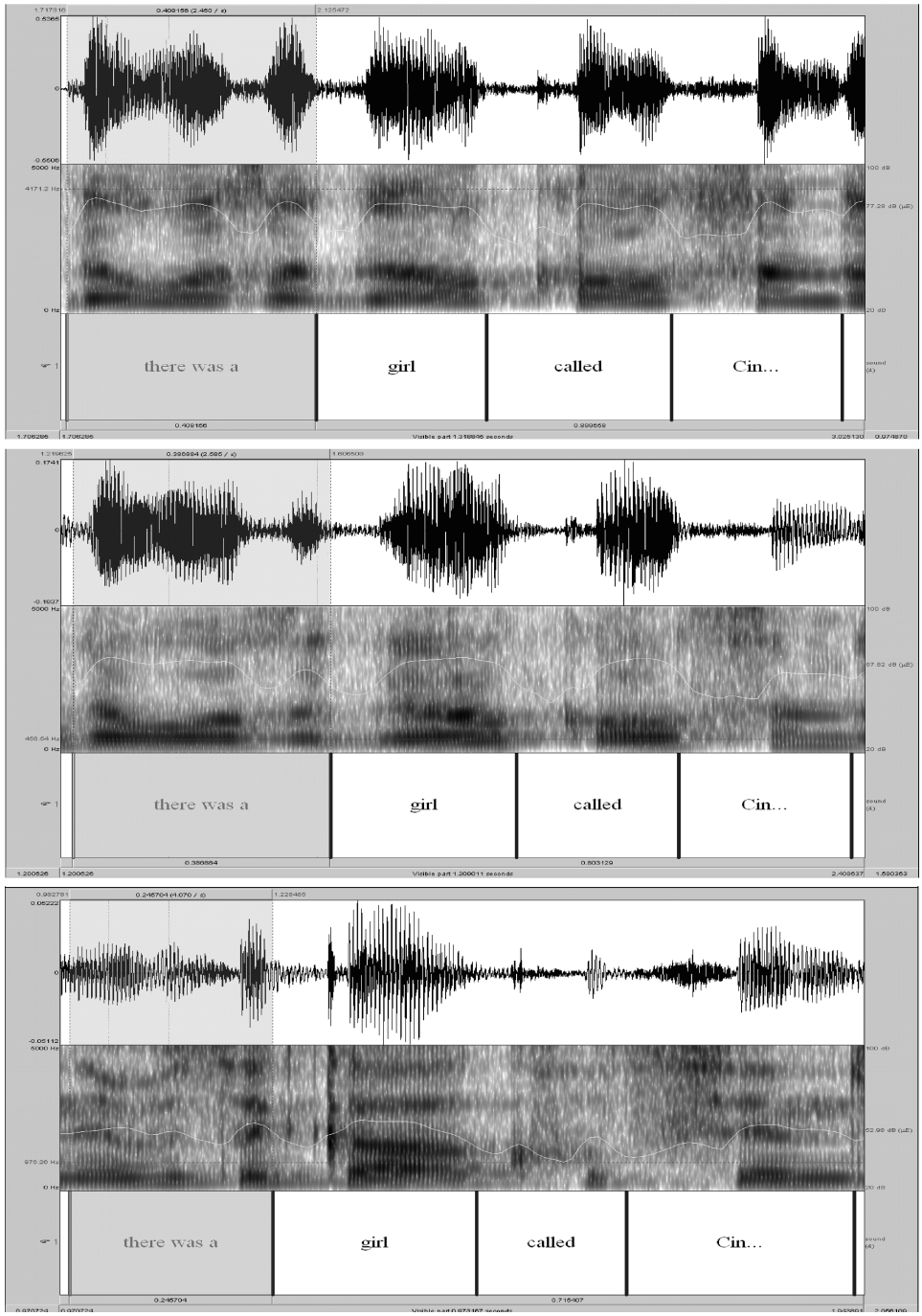

Figure 7.7. Typical PL1, PL2 (the same person) and EN (bottom) timing of *There was a girl called Cin(derella)*

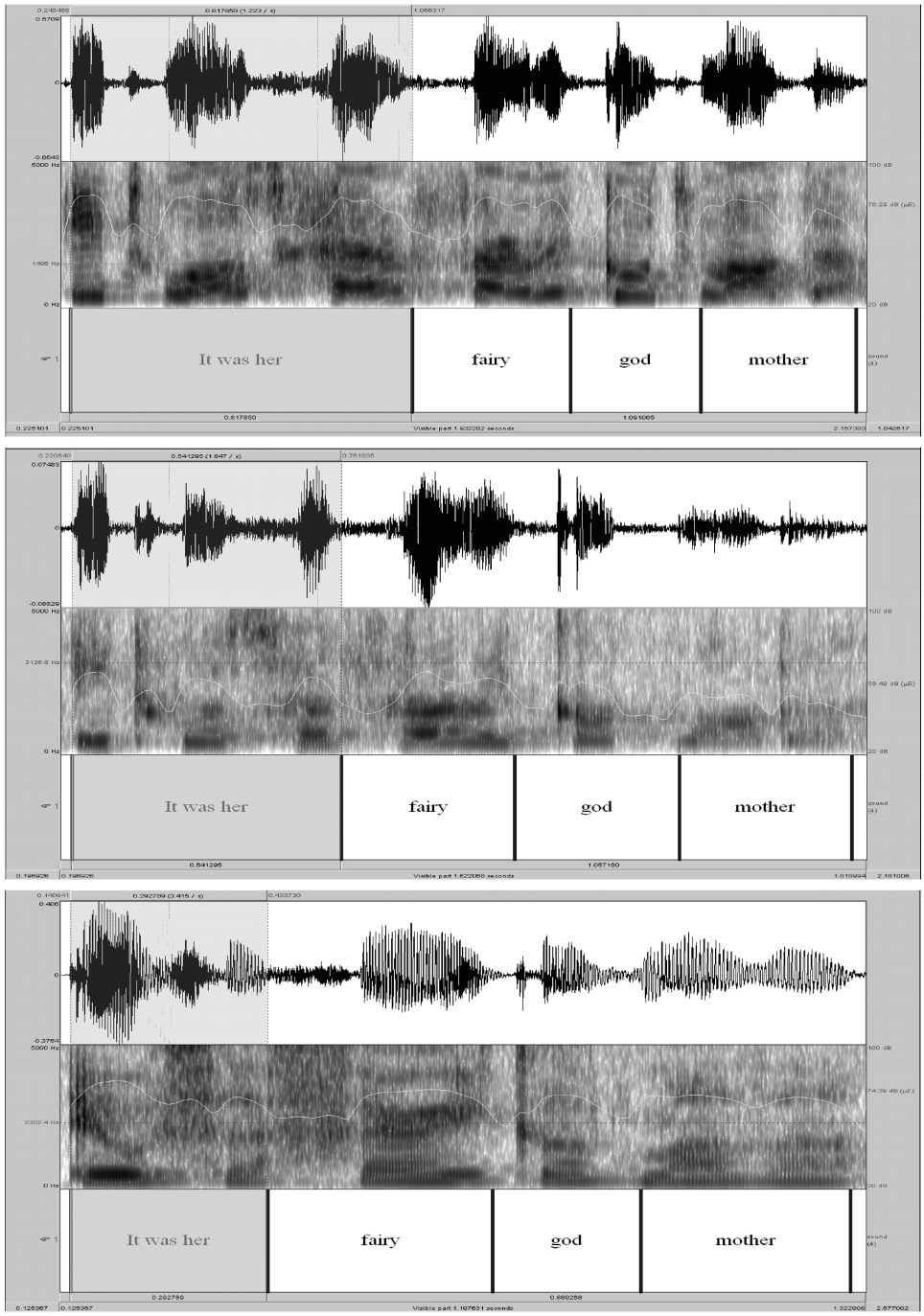

Figure 7.8. Typical PL1, PL2 (the same person) and EN (bottom) timing of *It was her fairy godmother*

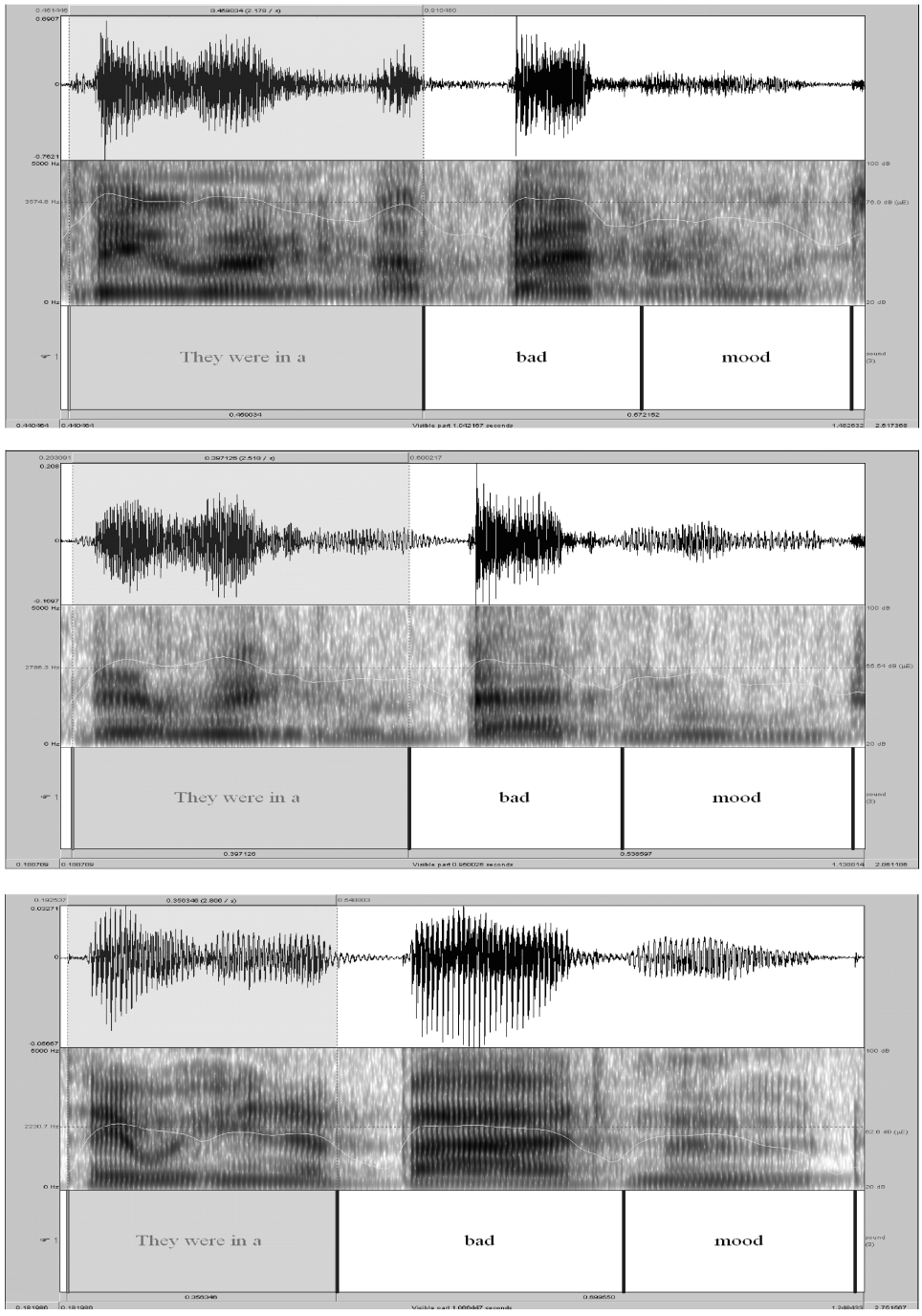

Figure 7.9. Typical PL1, PL2 (the same person) and EN (bottom) timing of *They were in a bad mood*

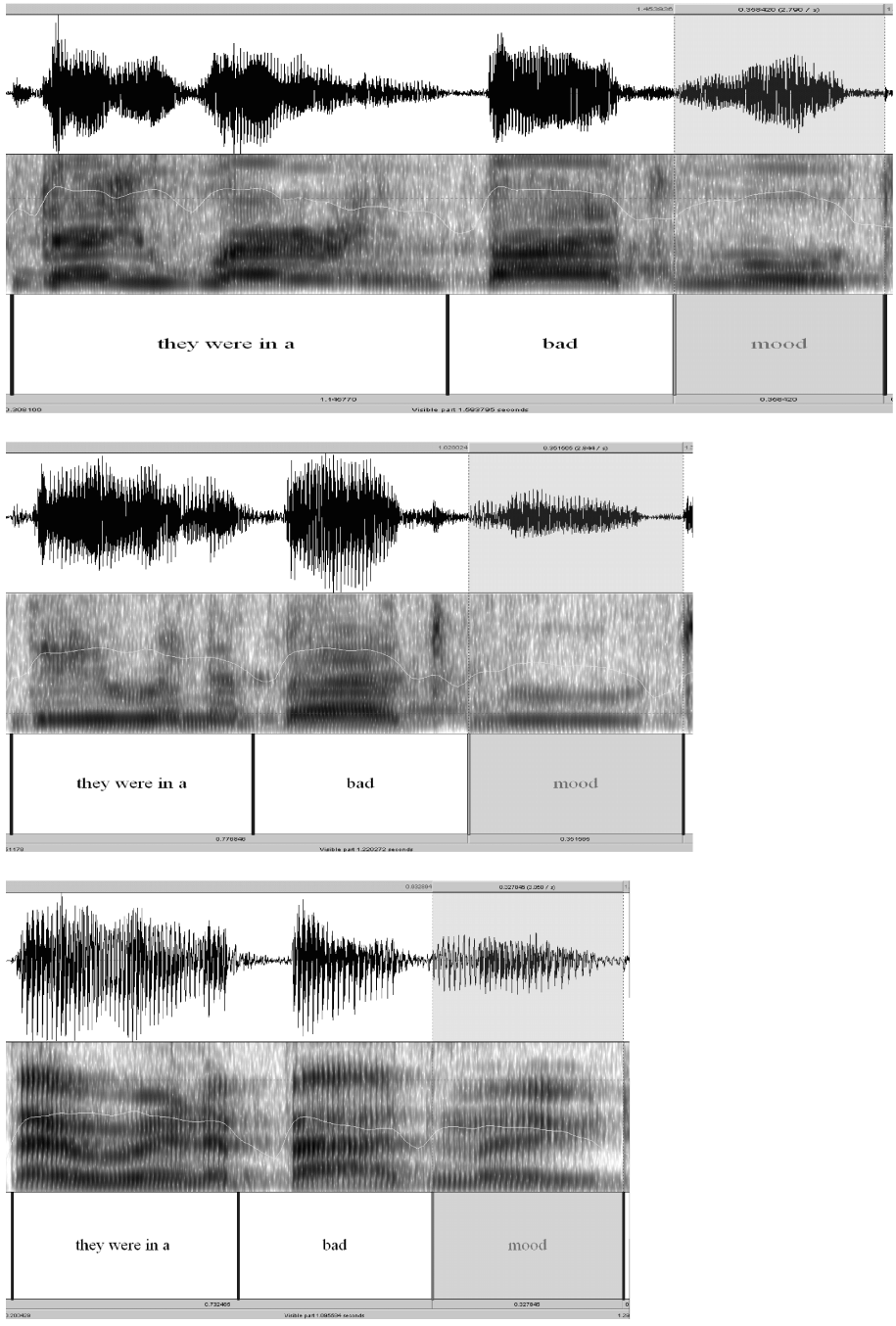

Figure 7.10. Absolute duration of IP constituents in typical PL1, PL2 (the same person) and EN (bottom) realisations of *They were in a bad mood*

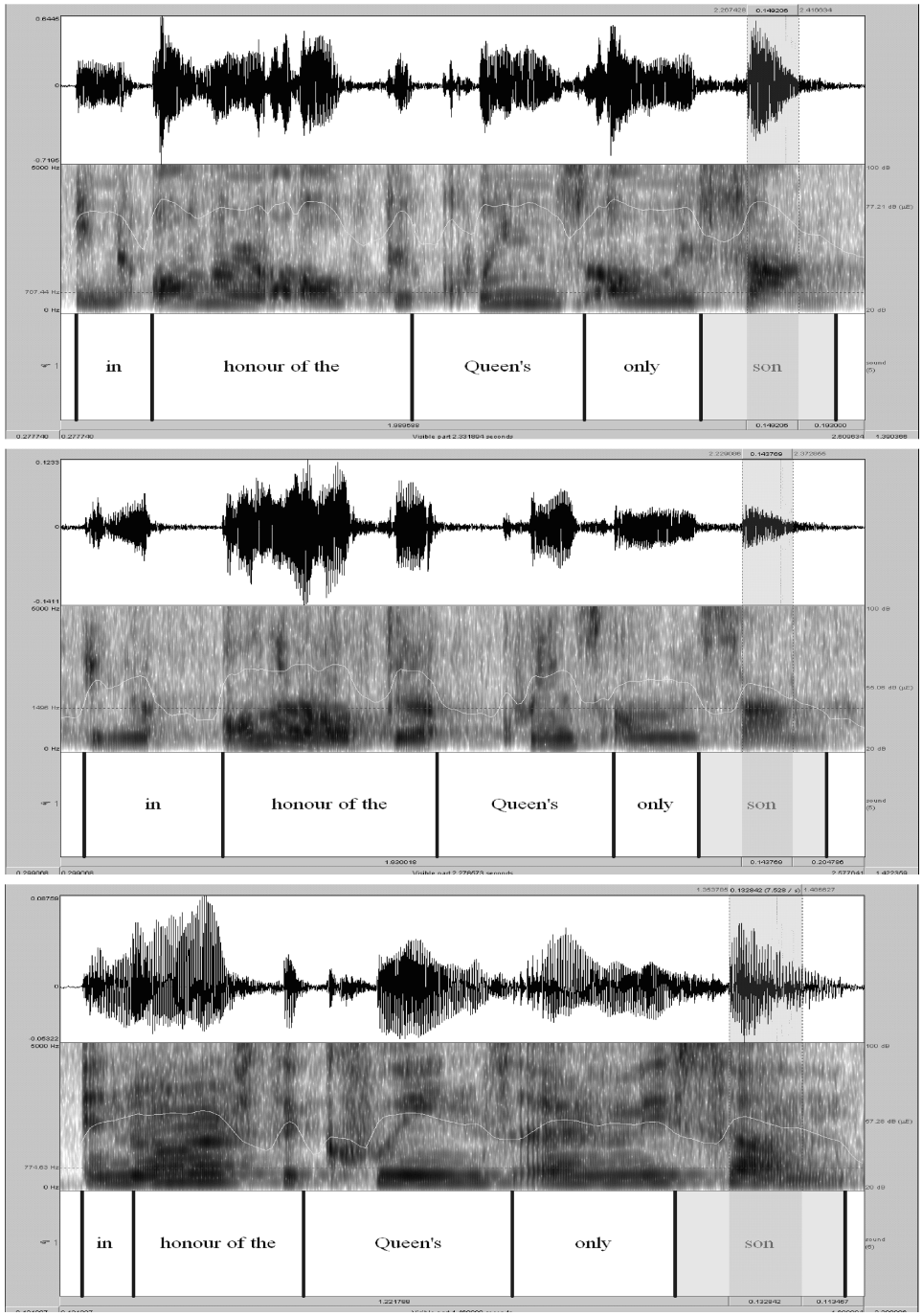

Figure 7.11. Typical PL1, PL2 (the same person) and EN (bottom) timing of head feet in *...in HONOUR OF THE (H4): QUEEN'S (H1): ONLY (H2) son*

7.2.6 Hypothesis 6: The Polish learner's development

Hypothesis 6: Language proficiency development is reflected in more native-like timing of Polglish production after two semesters of study including systematic practical phonetic training.

The global figures obtained in this research show significant reduction of the discrepancy between Polish and English speakers after increased language learning practice including practical phonetics training. Much of this progress may only be connected with better fluency in reading a familiar text for the second time, but tendencies to approach native standards also appear in contexts where no initial fluency problems were found.

The main conclusions concerning the observed developmental tendencies in Polish learners' speech timing are summarised below.

A. Stressed vowels

The learners' timing of stressed vowels did not change significantly. General vowel type mean length proportions remained non-native-like (section 5.2.1, Table 5.4). Despite a slight increase in N:H vowel duration ratio, it did not reach typical native English values (section 5.2.3, Table 5.6). The smaller durational difference between vowels in monosyllabic and 3-4-syllable feet in Polish learners' speech remained unchanged (section 5.2.4, Tables 5.9—5.10). The only index that matched typical native values was vowel length proportions in pre-fortis and pre-lenis positions (section 5.2.5, Table 5.12). Individual learners' general vowel length variation did not significantly increase, and remained at a non-native-like level in 8 subjects in comparison to 10 (out of 13) subjects before the training (cf. sections 5.4.1—5.4.2, Tables 5.23—5.24). Only two Polish subjects ranked among native speakers for stressed vowel length variation in three measured unit categories (section 5.4.4, Table 5.27). Since both these learners obtained similar results in both recordings, and there was no consistency in personal variation ratio changes, we assume that this feature of Polish learners' pronunciation did not approach native English pronunciation standards.

B. Unstressed vowels

The large initial discrepancy between Polish and English subjects in the duration of unstressed vowels was significantly reduced in the second recording (approximately 50% reduction) but the learners' durations still remained considerably longer (sections 5.3.1—5.3.2, Tables 5.16—5.18). The personal mean duration of unstressed vowels in all 13 cases exceeded the highest native English personal mean. Four learners reduced their means to native English levels in the second recording. The impact of insufficient vowel reduction was also reflected in individual speakers' overall vowel duration

variability ratios, which showed a native-like level in one PL1 speaker and two more speakers in PL2 recording (section 5.4.4, Table 5.29).

C. Syllables, weak forms and content words

The unstressed syllables were longer in Polglish production both in trochaic content words and unstressed function words. In content words, the unstressed syllables also constituted a larger proportion of the word than in native English recordings. The length of Polglish unstressed syllables was reduced in the second recording, but the stressed:unstressed syllable ratio did not substantially change owing to the shortening of stressed syllables in PL2, often below average native durations (cf. section 6.2.1, Tables 6.1—6.2).

Considering individual speakers, only 1 Polish learner did not shorten unstressed syllables in the second performance. Still, only 3 PL2 scores ranked within the native limits (none in PL1).

The weak forms of function words in the pronunciation of the learners were relatively even longer than the unstressed syllables of content words, although the actual difference mainly depended on individual syllable structure and phrase position. Despite a general tendency to shorten the weak forms in PL2, the discrepancy between the two groups of speakers remained conspicuous (section 6.2.2, Table 6.3) and in most cases a large majority of Polish subjects (typically more than 9 in 13) produced responses exceeding the native English maximum values. After the training, there were fewer such responses but the problem still remained noteworthy (cf. Table 6.4). The absolute duration difference was reduced in the second recording, but owing to the generally increased speech rate, the average unstressed unit constituted approximately the same proportion of the measured speech portion (Tables 6.5—6.6).

Finally, the personal syllable duration variability (SVI — cf. section 6.2.5, Table 6.18) in the second recording was higher in 11 learners and lower in 2, but none of them reached the level of the lowest native speaker's syllable variability.

D. Preheads

The Polish learners' preheads were significantly longer than the native norms. If a phrase-initial unstressed cluster contained more than one syllable, very few Polish subjects made it shorter than the longest corresponding native response. Even the post-training performance in most contexts left 7—10 Polish speakers outside of native English norms (section 6.3.1, Table 6.19).

E. Stress-timing

A weaker tendency for stress-timing in Polish learners' English speech is manifested in less vowel and syllable length variation, which can also be

attributed to problems with unstressed syllable reduction. This regularity is visible in the timing of head feet. The learners managed to reduce the duration of 1-2-syllable-long head feet (H1, H2) to native English norms, but more complex feet (H3, H4) were still significantly longer although the discrepancy between Polglish and native mean scores was reduced by 50% in the second recording (section 6.3.2, Table 6.22). Such feet were "too long" in 8—13 Polish learners before the training and approximately 6—10 in PL2.

The learners also did not manage to reduce the phrase-medial foot durational variation, again showing more tendency for syllable-timing (cf. section 6.3.4, Table 6.25 and Figures 6.3—6.4). Accordingly, no regularity was observed in this respect in individual speakers' PL1:PL2 relations. Therefore we can repeat the conclusion that, considering less vowel and syllable duration flexibility and more foot length variability, Polish learners have problems developing a more stressed-timed rhythm in their English speech.

F. Speech rate

The general results showed slightly longer duration of stressed units in the first recording of Polish learners' performance. The durations became comparable in the second recording, with a few examples of shorter Polglish durations. Still, the whole measured phrases were generally longer in the learners' pronunciation (cf. section 6.2.4, Tables 6.12—6.16) as a result of longer unstressed units. These units, though reduced in the second recording, remained longer, which allowed the learners only to reduce the discrepancy but not eliminate it.

7.3 Directions for further research

This picture of the Polish learner's read speech is far from complete. Duration adds to the perception of prominence and rhythm, but it cooperates with other auditory cues related to intensity (or spectral tilt), pitch, and vowel quality, which are not included in the present study. Despite the rapid development of laboratory phonetics in the past decades and a number of recent findings, the correlation of the features mentioned above is still largely unclear.

The general differences in native and non-native English speech timing demonstrated in the present study suggest possible directions for research aimed at explaining their underlying reasons. Such research, related in the first place to areas where the largest difference has been found, might concentrate on articulatory and acoustic details corresponding to duration differences. This would require isolation of individual prominence cues and investigation of their mutual relations, preferably using scalar measures.

Undoubtedly, the differences found in this research are also related to intrinsic differences in the articulation of particular segments in various positions of the prosodic structure, including the effects of domain boundaries and prominence. Hence the acoustic data should also be correlated with articulatory research including

— EPG studies of occlusion in plosive (voiced and voiceless) and nasal stops
— correlation between linguopalatal contact area, seal duration and VOT duration
— energy distribution in relation to distribution of prominent syllables.

These, however, would require precisely controlled contexts and laboratory conditions, which, as already mentioned, often results in less natural performance samples.

7.4 Pedagogical implications

Despite the need for further research to verify the findings of this study, in some points they appear convincing enough to be taken into account in English pronunciation teaching. Obviously, most of the results reflect aspects of pronunciation regularly addressed in handbooks of English phonetics. This study, however, draws the teacher's attention to certain details about the design of the course. First of all, though recognising the importance of prosody and not rejecting its alleged priority over segmental issues, we must bear in mind that practising foreign language speech rhythm must be preceded by segmental training because the learners' problems with the execution of sequences of individual articulatory gestures, aggravated by nonnative-for-native sound substitution (e.g. /x/ for /h/) or limited use of coarticulation and elision (e.g. initial /h/, preconsonantal /r/, unnecessary plosions following /ŋ/ or preceding stops and nasals) will inevitably adversely affect their foreign speech timing.

Exercises at higher proficiency levels, after the quality of principal allophone articulatory targets has been mastered by the learners, should concentrate on vowel reduction, coarticulation, and consonantal elision processes, especially in preheads and more complex head feet. This suggests a gradual shift of exercise scope to higher and higher prosodic domains. Consequently, learners should also be warned against fast speaking before they have internalised assimilation processes allowing legitimate unstressed unit reduction.

The results pertaining to various examples of content: function word relations suggest considering pronunciation exercises dealing with the timing of fixed constructions, collocations, and especially syntactic structures which are often strictly connected with regular prosodic (rhythmic and melodic) patterns (e.g. dummy subjects IT/THERE with BE, perfective constructions with weak

forms of HAVE or verbal constructions such as GOING TO or HAVE TO). Such exercises can easily be combined with regular grammatical practice, which means there need not be an increase in the amount of time the teacher devotes to pronunciation training alone.

As the general didactic conclusion from the research, even though we acknowledge the existence and importance of general rhythmic tendencies in spoken production, we would like to advocate the idea of constructing EFL pronunciation courses with an upward shift of focus through consecutive levels of the prosodic hierarchy.

APPENDIX A

The test passage. Phrases used in the analysis are highlighted in bold:

Once upon a time ***there was a girl called Cinderella***. ***But everyone called her Cinders***. ***Cinders*** *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 to parties***. *Poor Cinders had to wear all their old hand-me-downs!* ***And she had to do the cleaning!***

One day, a royal messenger came to announce a ball. The ball would be held at the Royal Palace, ***in honour of the Queen's only son****, Prince William.* ***Lily and Rosa thought this was divine***. ***Prince William was gorgeous, and he was looking for a bride! They dreamed of wedding bells!***

When the evening of the ball arrived, Cinders had to help her sisters get ready. ***They were in a bad mood***. ***They'd wanted to buy some new gowns, but their mother said that they had enough gowns***. ***So they started shouting at Cinders***. *'Find my jewels!' yelled one.* ***'Find my hat!'*** *howled the other. They wanted hairbrushes, hairpins and hair spray.*

When her sisters had gone*, Cinders felt very down, and she cried. Suddenly, a voice said:* ***'Why are you crying, my dear?'***. ***It was her fairy godmother!***

The girl poured her heart out: 'Lily and Rosa have it all!' she cried, 'even though they're awful, and fat, and they're dull! And I want to go to the ball, and meet Prince William!'

'You will, won't you?' laughed her fairy godmother. 'Go into the garden and find me a pumpkin'. Cinders went, and found a splendid pumpkin which the fairy changed into a dazzling carriage.

Appendix B

Reference phrases used for calculating personal mean syllable durations MSL=(1+2+3+4+5+6)/37

1. ...there was a girl called... (5)
1. ...and going to partie(s). (6)
2. Prince William was gorgeous... (6)
3. ...and he was looking for a bride. (8)
4. They dreamed of wedding bells. (6)
5. Why are you crying, my (dear?) (6)

APPENDIX C

Reference words used for calculating personal mean unstressed syllable durations (MUSL).

Function words:

1. ...and she had to do THE cleaning.
2. ...in honour of THE Queen's only son...
3. ...but everyone called HER Cinders.
4. When HER sisters had gone...
5. It was HER fairy godmother.
6. So they started shouting AT Cinders.
7. Lily and Rosa thought this WAS divine.
8. Prince William WAS gorgeous...
9. It WAS her fairy godmother.
10. ...and going TO parties.
11. ...and she had TO do the cleaning.
12. They'd wanted TO buy some new gowns...
13. They'd wanted to buy SOME new gowns...
14. They dreamed OF wedding bells.

Non-phrase final unstressed syllables of content words:

1. CinDERs lived...
2. When her sisTERs had gone...
3. So they started shouTING at Cinders.
4. So they starTED shouting at Cinders.
5. They'd wanTED to buy some new gowns...

References

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

ADAMS, C. 1979. *English Speech Rhythm and the Foreign Learner*. The Hague: Mouton.

ALLEN, G.D. 1975. "Speech rhythm: its relation to performance universals and articulatory timing". *Journal of Phonetics* 13 (2): 75—87.

ALLEN, J.S., and J.L. MILLER 1999. "Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words". *Journal of the Acoustical Society of America* 106: 2031—2039.

ANDERSON-HSIEH, J., R. JOHNSON and K. KOEHLER 1992. "The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure". *Language Learning* 42 (4): 529—555.

ANDERSON-HSIEH, J., and H. VENKATAGIRI 1994. "Syllable duration and pausing in the speech of Chinese ESL speakers". *TESOL Quarterly* 28: 807—812.

ANISFELD, M., N. BOGO and W.E. LAMBERT 1962. "Evaluational reactions to accented English speech". *Journal of Abnormal Social Psychology* 65: 223—231.

ARTHUR, B., D. FARRAR and G. BRADFORD 1974. "Evaluation reactions of college students to dialect differences in the English of Mexican-Americans". *Language Speech* 17 (3): 255—270.

AUER, P., E. COUPER-KUHLEN and F. MÜLLER 1999. *Language in Time. The Rhythm and Tempo of Spoken Interaction*. New York, Oxford: Oxford University Press.

AVERY, P., and S. EHRLICH 1996. *Teaching American English Pronunciation*. Oxford: Oxford University Press.

AYLETT, M., and A. TURK 2004. "The Smooth Signal Redundancy Hypothesis: A functional explanation for relationships between redundancy, prosodic prominence duration in spontaneous speech". *Language and Speech* 47: 31—56.

AYLETT, M., and A. TURK 2006. "Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei". *Journal of the Acoustical Society of America* 119: 3048—3058.

BAKER, R.E., and A.R. BRADLOW 2009. "Variability in word duration as a function of probability, speech style, and prosody". *Language and Speech* 52 (4): 391—413.

BALL, M.J., and J. RAHILLY 1999. Phonetics: *The Science of Speech*. London: Edward Arnold.

BARRY, W.J., B. ANDREEVA, M. RUSSO, S. DIMITROVA and T. KOSTADINOVA 2003. "Do rhythm measures tell us anything about language type?" *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, pp. 2693—2696.

BARRY, W.J. 2007. "Rhythm as an L2 problem: How prosodic is it?" In J. TROUVAIN and U. GUT (eds.), *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin, New York: Mouton de Gruyter, pp. 97—120.

BECKMAN, M.E. 1986. *Stress and Non-Stress Accent*. Dordrecht/Riverton: Foris Publications.

BECKMAN, M.E., and J. EDWARDS 1990. "Lengthenings and shortenings and the nature of prosodic constituency". In J. KINGSTON and M.E. BECKMAN (eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, pp. 152—178.

BECKMAN, M., and J. EDWARDS 1992. "Intonational categories and the articulatory control of duration". In Y. TOHKURA, E. VATIKIOTIS-BATESON and Y. SAPISAKA (eds.), *Speech Perception, Production and Linguistic Structure*. Tokyo: Ohmsha, pp. 356—375.

BECKMAN, M., and J. EDWARDS 1994. "Articulatory evidence for differentiating stress categories". In P.A. KEATING (ed.), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*. Cambridge: Cambridge University Press, pp. 7—33.

BECKMAN, M.E., and J.B. PIERREHUMBERT 1986. "Intonational structure in Japanese and English". *Phonology Yearbook* 3: 255—309.

BELL, A., M.L. GREGORY, D. JURAFSKY, C. GIRARD, J. BRENIER and A. IKENO 2002. "Which predictability measures affect content word durations?" In W. BYRNE, E. FOSLER-LUSSIER and D. JURAFSKY, D. (eds.), *Proceedings of the Workshop on Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology (PMLA)*. Estes Park, CO, pp. 65—70.

BERKOVITS, R. 1994. "Durational effects in final lengthening, gapping, and contrastive stress". *Language and Speech* 37: 237—250.

BIEDRZYCKI, L. 1978. *Fonologia angielskich i polskich rezonantów*. Warszawa: Państwowe Wydawnictwo Naukowe.

BIRDSONG, D. 1992. "Ultimate attainment in second language acquisition". *Language* 68: 706—755.

BLEVINS, J. 1995. "The syllable in phonological theory". In J.A. GOLDSMITH (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 206—244.

BLOCH, B. 1950. "Studies in colloquial Japanese IV: Phonemics". *Language* 26: 86—125.

BLOCH, B., and G.L. TRAGER 1942. *Outline of Linguistic Analysis*. Baltimore: Waverley Press.

BLOOMFIELD, L. 1933. *Language*. New York: Henry Holt.

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

BOGLE, D. 1996. Practical Phonology. Edinburgh: Moray House Publications.

BOLINGER, D.L. 1958a. "A theory of pitch accent in English". *Word* 14: 109—149.

BOLINGER, D.L. 1958b. "Stress and information". *American Speech* 33: 5—20.

BOLINGER, D.L. 1972a. "Accent is predictable (if you're a mind reader)". *Language* 48: 633—644.

BOLINGER, D.L. 1972b. *Intonation: Selected Readings*. Harmondsworth: Penguin.
BOLINGER, D.L. 1981. *Two Kinds of Vowels, Two Kinds of Rhythm*. Bloomington, In Indiana University Linguistics Club.
BOLINGER, D. 1985. "Two views of accent". *Journal of Linguistics* 21: 79—123.
BOLINGER, D. 1987. "More views on 'Two Views of Accent'". In IDEM, *On Accent*. Bloomington, IN. Reproduced by the Indiana University Linguistics Club, pp. 124—146.
BOMBIEN, L., C. MOOSHAMMER, P. HOOLE, B. KUHNERT and J. SCHNEEBERG 2006. "An EPG study of initial /kl/ clusters in varying prosodic conditions in German". *Proceedings of the 7th International Seminar on Speech Production*. Ubatuba, Brazil.
BOND, Z., and J. FOKES 1985. "Non-native patterns of English syllable timing". *Journal of Phonetics* 13: 407—420.
BONGAERTS, T. 1999. "Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners". In D. BIRDSONG (ed.), *Second Language Learning and the Critical Period Hypothesis*. Mahwah: Lawrence Erlbaum Associates, pp. 133—159.
BOUZON, C., and D. HIRST 2004. "Isochrony and Prosodic Structure in British English". In B. BEL and I. MARLIEU (eds.), *Proceedings of the Second International Conference on Speech Prosody*. Nara, Japan, pp. 223—226.
BRESNAN, J. 1971. "Sentence stress and syntactic transformations". *Language* 47: 257—281.
BRESNAN, J. 1972. "Stress and syntax: A reply". *Language* 48: 326—342.
BRIÈRE, E.J. 1968. *A Psycholinguistic Study of Phonological Interference*. The Hague: Mouton.
BROWMAN, C.P., and L. GOLDSTEIN 1986. "Towards an articulatory phonology". *Phonology Yearbook* 3: 219—252.
BROWMAN, C.P., and L. GOLDSTEIN 1989. "Articulatory gestures as phonological units". *Phonology* 6: 201—251.
BROWMAN, C.P., and L. GOLDSTEIN 1992. "Articulatory phonology: An overview". *Phonetica* 49: 155—180.
BRYŁA, A. 2010. "Phonetic properties of Euro-English — empirical evidence". In E. WANIEK-KLIMCZAK (ed.), *Issues in Accents of English 2: Variability and Norm*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, pp. 37—60.
BURNS, T.C., K.A. YOSHIDA, K. HILL and J.F. WERKER 2007. "The development of phonetic representation in bilingual and monolingual infants". *Applied Psycholinguistics* 28: 455—474.
BYBEE, J. 2000. "The phonology of the lexicon: evidence from lexical diffusion". In M. BARLOW and S. KEMNER (eds.), *Usage-based Models of Language*. Stanford, CA: CSLI, pp. 65—85.
BYRD, D. 1994. "Relations of sex and dialect to reduction". *Speech Communication* 15: 39—54.
BYRD, D. 2000. "Articulatory vowel lengthening and coordination at phrasal junctures". *Phonetica* 57: 3—16.
BYRD, D., A. KAUN, S. NARAYANAN and E. SALTZMAN 2000. "Phrasal signatures in articulation". In M.B. BROE and J.B. PIERREHUMBERT (eds.), *Acquisition and*

the Lexicon (Papers in Laboratory Phonology V). Cambridge: Cambridge University Press, pp. 70—87.

BYRD, D., J. KRIVOKAPIĆ and S. LEE 2006. "How far, how long: On the temporal scope of prosodic boundary effects". *Journal of the Acoustical Society of America* 120: 1589—1599.

BYRD, D., and E. SALTZMAN 1998. "Intragestural dynamics of multiple phrasal boundaries". *Journal of Phonetics* 26: 173—199.

BYRD, D., and E. SALTZMAN 2003. "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening". *Journal of Phonetics* 31: 149—180.

CAMBIER-LANGEVELD, T. 2000. *Temporal Marking of Accents and Boundaries*. Leiden: Holland Institute of Generative Linguistics.

CAMPBELL, N. 1992. *Multi-level Timing in Speech*. PhD Thesis, University of Sussex.

CAMPBELL, N., and BECKMAN, M. 1997. "Stress, prominence and spectral tilt". In A. BOTINIS, G. KOUROUPETROGLOU and G. CARAYANNIS (eds.), *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Implication*. Athens: ESCA and University of Athens, pp. 67—70.

CATFORD, J.C. 1988. *A Practical Introduction to Phonetics*. New York: Oxford University Press.

CAULDWELL, R. 2002. "The functional irrhythmicality of spontaneous speech: A discourse view of speech rhythms". *Applied Language Studies* 2 (1): 1—24.

CELCE-MURCIA, M., D. BRINTON and J. GOODWIN 1996. *Teaching Pronunciation: A Reference for Teachers of English to speakers of other languages*. New York, NY: Cambridge University Press.

CHELA-FLORES, B. 1994. "On the acquisition o English rhythm: Theoretical and practical issues". *International Review of Applied Linguistics in Language Teaching* 32 (3): 232—242.

CHELA-FLORES, B. 1998. *Teaching English Rhythm: From Theory to Practice*. Caracas: Fondos Editorial Tropykos.

CHEN, M. 1970. "Vowel length variation as a function of the voicing of the consonant environment". *Phonetica* 22: 125—159.

CHO, T. 2005. "Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English". *Journal of the Acoustical Society of America* 117(6): 3867—3878.

CHO, T., and P. KEATING 2001. "Articulatory and acoustic studies on domain-initial strengthening in Korean". *Journal of Phonetics* 29: 155—190.

CHO, T., and P. KEATING 2009. "Effects of initial position versus prominence in English". *Journal of Phonetics* 37: 466—485.

CHO, T., and P. LADEFOGED 1999. "Variation and universals in VOT: evidence from 18 languages". *Journal of Phonetics* 27: 207—229.

CHO, T., and J. MCQUEEN 2005. "Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress". *Journal of Phonetics* 33: 121—157.

CHOI, H. 2003. "Prosody-induced acoustic variation in English stop consonants". *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, pp. 2662—2664.

CHOMSKY, N. 1971. "Deep structure, surface structure and semantic interpretation". In D.D. STEINBERG and L.A. JAKOBOVITS (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge: Cambridge University Press, pp. 183—216.

CHOMSKY, N. 1986. *Barriers. (Linguistic Inquiry Monograph 13)*. Cambridge, MA, London: MIT Press.

CHOMSKY, N., and M. HALLE 1968. *The Sound Pattern of English*. New York: Harper and Row.

CLARK, J., and C. YALLOP 1990. *An Introduction to Phonetics and Phonology*. Oxford: Blackwell Publishing.

CLEMENTS, G.N., and S.J. KEYSER 1983. *CV Phonology*. Cambridge, MA: MIT Press.

COLE, J., H. KIM, H. CHOI and M. HASEGAWA-JOHNSON 2007. "Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech". *Journal of Phonetics* 35: 180—209.

COLE, J., Y. MO and S. BAEK 2008. "The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech". Presentation at the Experimental & Theoretical Approaches to Prosody, Cornell University, Ithaca, NY. Retrieved 25 June 2010. http://prosody.beckman.illinois.edu/pubs/40_Cole_etal_2008_ETAP_slides.pdf.

COLE, J., Y. MO and M. HASEGAWA-JOHNSON 2008. "Signal-based and expectation-based factors in the perception of prosodic prominence". Presentation at Laboratory Phonology 11, Wellington, New Zealand. Retrieved 25 June 2010. http://prosody.beckman.illinois.edu/pubs/41_Cole-Mo-HJ_2008_LabPhon11.pdf.

COLE, R.A., J. JAKIMIK and W.E. COOPER 1978. "Perceptibility of phonetic features in fluent speech". *Journal of the Acoustical Society of America* 64: 44—56.

COOPER, A.M. 1991. "Laryngeal and oral gestures in English /p, t, k/". *Proceedings of the 12th International Congress of Phonetic Sciences*, vol. 2. Aix-en-Provence, pp. 50—53.

COOPER, G.B. 1998. *Mysterious Music. Rhythm and Free Verse*. Stanford, CA: Stanford University Press.

COOPER, W.E., and J. PACCIA-COOPER 1980. *Syntax and Speech*. Cambridge, MA: Harvard University Press.

CROSSWHITE, K. 2001. *Vowel Reduction in Optimality Theory*. New York: Routledge.

CRUTTENDEN, A. 1997 [1986]. *Intonation*. Cambridge: Cambridge University Press.

CRUTTENDEN, A. 2001. *Gimson's Pronunciation of English*, 6th edition. London: Edward Arnold.

CRYSTAL, D., 1969. *Prosodic Systems and Intonation in English*. London: Cambridge University Press.

CRYSTAL, D. 1996. "The past, present and future of English rhythm". *Speak Out. Newsletter of the IATEFL Pronunciation Special Interest Group* 18: 8—13.

CRYSTAL, T.H., and A.S. HOUSE 1988a. "Segmental duration in connected speech signals: syllabic stress". *Journal of the Acoustical Society of America* 83 (4): 1574—1585.

CRYSTAL, T.H., and A.S. HOUSE 1988b. "The duration of American English consonants: an overview". *Journal of Phonetics* 16: 285—294.

Crystal, T.H., and A.S. House 1988c. "The duration of American English vowels: an overview". *Journal of Phonetics* 16: 263—284.

Crystal, T.H., and A.S. House 1990. "Articulation rate and the duration of syllables and stress groups in connected speech". *Journal of the Acoustical Society of America* 88: 101—112.

Cummins, F. 1999. "Some lengthening factors in English speech combine additively at most rates". *Journal of the Acoustical Society of America* 105: 476—480.

Cummins, F. 2002. "Speech rhythm and rhythmic taxonomy". In B. Bel and I. Marlin (eds.), *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence: Laboratoire Parole et Langage, pp. 121—126.

Cummins, F., and R.F. Port 1998. "Rhythmic constraints on stress timing in English". *Journal of Phonetics* 26 (2): 145—171.

Cunningham, U. 1986. *A Linguistic Theory of Timing*. PhD Thesis, Department of Linguistics, University of Nottingham.

Cunningham, U. 2008. "Acoustic variability in the production of English vowels by native and non-native speakers". In E. Waniek-Klimczak (ed.), *Issues in Accents of English*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, pp. 1—15.

Cutler, A. 1993. "Segmenting speech in different languages". *The Psychologist* 6 (10): 839—845.

Dalton, C., and B. Seidlhofer 1994. *Pronunciation*. Oxford: Oxford University Press.

Darwin, C.J., and A. Donovan 1980. "Perceptual studies of speech rhythm: Isochrony and intonation". In J.C. Simon (ed.), *Spoken Language Generation and Understanding*. Dordrecht: Reidel, pp. 77—85.

Dauer, R. 1983. "Stress timing and syllable timing reanalyzed". *Journal of Phonetics* 11: 51—62.

Dauer, R. 1987. "Phonetic and phonological components of language rhythm". *Proceedings of the 11th International Congress of Phonetic Sciences*, vol. 5. Tallinn, pp. 447—450.

de Jong, K.J. 1991. "An articulatory study of vowel duration changes in English". *Phonetica* 48: 1—18.

de Jong, K.J. 1995. "The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation". *Journal of the Acoustical Society of America* 97 (1): 491—504.

de Jong, K.J. 2000. "Attention modulation and the formal properties of stress systems". In J. Boyle, J-H. Lee and A. Okrent (eds.), *Chicago Linguistuic Society* 36 (1): 71—91.

Delattre, P. 1962. "Some factors of vowel duration and their cross-linguistic validity". *Journal of the Acoustical Society of America* 34: 1141—1142.

Delattre, P. 1966. "A comparison of syllable length conditioning among languages". *International Review of Applied Linguistics in Language Teaching* 4 (3): 183—198.

Dellwo, V. 2006. "Rhythm and speech rate: A variation coefficient for ΔC". In P. Karnowski and I. Szigeti (eds.), *Language and Language-processing*. Frankfurt am Main: Peter Lang, pp. 231—241.

Dellwo, V. 2008. "The role of speech rate in perceiving speech rhythm". Speech Prosody 2008, *Proceedings of Speech Prosody 2008 series*. Campinas, pp. 375—378.

Dellwo, V., and P. Wagner 2003. "Relationships between speech rhythm and rate". *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, pp. 471—474.

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

Dogil, G. 1995. "Phonetic correlates of word stress". *Phonetic AIMS*. Stuttgart: University of Stuttgart, 1—60.

Dogil, G. 1999. The phonetic manifestation of word stress in Lithuanian, Polish, German and Spanish. In H. van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 273—311.

Donegan, P., and D. Stampe 1983. "Rhythm and the holistic organization of language structure". In J. F. Richardson, M. Marks and A. Chukerman (eds.), *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*. Chicago: Chicago Linguistic Society, pp. 337—353.

Donovan, A., and C.J. Darwin 1979. "The perceived rhythm of speech". *Proceedings of the 9th International Congress of Phonetic Sciences*, vol. 2. Copenhagen: Institute of Phonetics, pp. 268—274.

Duanmu, S. 2008. *Syllable Structure: The Limits of Variation*. Oxford University Press.

Dziubalska-Kołaczyk, K. 2002. *Beats-and-Binding Phonology*. Frankfurt: 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.

Edwards, J.E., M.E. Beckman and J. Fletcher 1991. "The articulatory kinematics of final lengthening". *Journal of the Acoustical Society of America* 89: 369—382.

Eefting, W. 1991. "The effect of information value and accentuation on the duration of Dutch words, syllables, and segments". *Journal of the Acoustical Society of America* 89: 412—424.

Elert, C.-C. 1964. *Phonological Studies of Quantity in Swedish*. Stockholm: Almqvist and Wiksell.

Erickson, D. 2002. "Articulation of extreme formant patterns for emphasized vowels". *Phonetica* 59 (2—3): 134—149.

Eriksson, A. 1991. *Aspects of Swedish Speech Rhythm*. PhD Thesis, University of Göteborg, Sweden.

Faber, D. 1986. "Teaching the rhythms of English: A new theoretical base". *International Review of Applied Linguisticsin Language Teaching* 24: 205—216.

Fach, M.L. 1999. "A comparison between syntactic and prosodic phrasing". *Proceedings of Eurospeech* 1: 527—530.

Faure, G., D.J. Hirst and M. Chafcouloff 1980. "Rhythm in English: Isochronism, Pitch and Perceived Stress". In L.R. Waugh and C.H. van Schooneveld (eds.), *The Melody of Language. Intonation and Prosody*. Baltimore, MD: University Park Press, pp. 71—79.

Fear, B.D., A. Cutler and S. Butterfield 1995. "The strong/weak syllable distinction in English". *Journal of the Acoustical Society of America* 97 (3): 1893—1904.

Felps, D., H. Bortfeld and R. Gutierrez-Osuna 2009. "Foreign accent conversion in computer assisted pronunciation training". *Speech Communication* 51: 920—932.

FERREIRA, F. 1991. "Effects of length and syntactic complexity on initiation times for prepared utterances". *Journal of Memory and Language* 30: 210—233.

FIJAŁKOWSKA, A. 2000. *Rhythm in Speech and Music*. MA Thesis. Adam Mickiewicz University, Poznań.

FLEGE, J.E. 1984. "The detection of French accent by American listeners". *Journal of the Acoustical Society of America* 76 (3): 692—707.

FLEGE, J.E., E.M. FRIEDA, A.C. WALLEY and L.A. RANDAZZA 1998. "Lexical factors and segmental accuracy in second language speech production". *Studies in Second Language Acquisition* 20: 155—187.

FLEMMING, E. 1995. *Auditory representations in phonology*. PhD Thesis, University of California, Los Angeles.

FOUGERON, C. 2001. "Articulatory properties of initial segments in several prosodic constituents in French". *Journal of Phonetics* 29: 109—136.

FOUGERON, C., and P.A. KEATING 1997. "Articulatory strengthening at edges of prosodic domains". *Journal of the Acoustical Society of America* 101 (6): 3728—3740.

FOWLER, C. 1977. *Timing Control in Speech Production*. Bloomington, IN: Indiana University Linguistics Club.

FROTA, S., M. D'IMPERIO, G. ELORDIETA, P. PRIETO and M. VIGÁRIO 2007. "The phonetics and phonology of intonational phrasing in Romance". In P. PRIETO, J. MASCARÓ and M.-J. SOLÉ (eds.), *Segmental and Prosodic Issues in Romance Phonology (Current Issues in Linguistic Theory)*. Amsterdam, Philadelphia: John Benjamins, pp. 131—153.

FRY, D.B. 1955. "Duration and intensity as physical correlates of linguistic stress". *Journal of the Acoustical Society of America* 27: 765—768.

FRY, D.B. 1958. "Experiments in the perception of stress". *Language and Speech* 1: 126—152.

FRY, D.B. 1965. "The dependence of stress judgments on vowel formant structure". In E. ZWIRNER and W. BETHGE (eds.), *Proceedings of the 6th International Congress of Phonetic Sciences*. Karger, pp. 306—311.

FUDGE, E.C. 1969. "Syllables". *Journal of Linguistics* 5: 253—287.

GAHL, S. 2006. "Why do frequent words shorten? Homophones in spontaneous speech". *Talk Presented at the Annual Meeting of the Linguistic Society of America*. Albuquerqe, NM.

GEE, J.P., and F. GROSJEAN 1983. "Performance structures: a psycholinguistic and linguistic appraisal". *Cognitive Psychology* 15: 411—458.

GILBERT, J.B. 1984. *Clear Speech: Pronunciation and Listening Comprehension in American English*. Cambridge: Cambridge University Press.

GIMSON, A.C. 1980 [1974, 1962]. *An Introduction to the Pronunciation of English*. London: Edward Arnold.

GIMSON, A.C. 1978. "Towards an international pronunciation of English". In P. STREVENS (ed.), *In Honour of A.S. Hornby*. Oxford: Oxford University Press, pp. 45—53.

GOLDSTEIN, L., M. POUPLIER, L. CHEN, E. SALTZMAN and D. BYRD 2007. "Dynamic action units slip in speech production errors". *Cognition* 103 (3): 386—412.

GONET, W. 1997. "Duration of R.P. monophthongs in monosyllabic words". In W. JASSEM and C. BASZTURA (eds.), *Speech and Language Technology*, vol. 1.

Poznań: Polish Phonetics Association/Wrocław: Wydawnictwo Prac Naukowych FORMAT, pp. 195—228.

GONET, W., J. SZPYRA-KOZŁOWSKA and R. ŚWIĘCIŃSKI 2010. "The acquisition of Vowel Reduction by Polish students of English". In E. WANIEK-KLIMCZAK (ed.), *Issues in Accents of English 2: Variability and Norm*. Newcastle-upon-Tyne: Cambridge Scholars Publishers, pp. 291—308.

GRABE, E. 2002. "Variation adds to prosodic typology". In B. BEL and I. MARLIN (eds.), *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence: Laboratoire Parole et Langage, pp. 127—132.

GRABE E. and E.L. LOW 2002. "Durational variability in speech and the rhythm class hypothesis". In C. GUSSENHOVEN and N. WARNER (eds.), *Laboratory Phonology 7*. Berlin, New York: Mouton de Gruyter, pp. 515—546.

GRABE, E., B. POST and F. NOLAN 2001. *The IViE Corpus*. Department of Linguistics, University of Cambridge. Retrieved 7 September 2006. http://www.phon.ox.ac.uk/IViE.

GRABE, E., B. POST, F. NOLAN and K. FARRAR 2000. "Pitch accent realization in four varieties of British English". *Journal of Phonetics* 28: 161—185.

GRØNNUM, N. 1991. "Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German". *Phonetica* 47: 188—214.

GROSJEAN, F., L. GROSJEAN and H. LANE 1979. "The patterns of silence: Performance structures in sentence production". *Cognitive Psychology* 11: 58—51.

GUSSENHOVEN, C. 1983. "Focus, mode and the nucleus". *Journal of Linguistics* 19: 377—417.

GUSSENHOVEN, C. 1985. "Two views of accent: A reply". *Journal of Linguistics* 21: 125—138.

GUSSENHOVEN, C. 2008. "Types of focus in English". In C. LEE, M. GORDON and D. BÜRING (eds.), *Topic and Focus. Cross-linguistic Perspectives on Meaning and Intonation*. Dordrecht: Springer, pp. 83—100.

GUSSENHOVEN, C., B.H. REPP, A. RIETVELD, H.H. RUMP and J. TERKEN 1997. "The perceptual prominence of fundamental frequency peaks". *Journal of the Acoustical Society of America* 102 (5): 3009—3022.

GUSSENHOVEN, C., and A. RIETVELD 1992. "Intonation contours, prosodic structure and preboundary lengthening". *Journal of Phonetics* 20: 283—303.

HALLE, M., and J.-R. VERGNAUD 1987. *An Essay on Stress*. Cambridge: MIT Press.

HALLIDAY, M.A.K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.

HALLIDAY, M.A.K. 1970. *A Course in Spoken English: Intonation*. London: Oxford University Press.

HALLIDAY, M.A.K. 1989. *Spoken and Written Language*. 2nd edition. Oxford: Oxford University Press.

HARRIS, J. 2005. "Vowel reduction as information loss". In P. CARR, J. DURAND and C.J. EWEN (eds.), *Headhood, Elements, Specification and Contrastivity*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 119—132.

HARRIS, M.S., and N. UMEDA 1974. "Effect of speaking mode on temporal factors in speech: vowel duration". *Journal of the Acoustical Society of America* 56: 1016—1018.

HAY, J., and R. DIEHL 1999. "Effect of duration, intensity and f0 alternations on rhythmic grouping". *Proceedings of the 14th International Congress of the Phonetic Sciences*. San Francisco, pp. 245—248.

HAYES, B. 1989. "The prosodic hierarchy in meter". In P. KIPARSKY and G. YOUMANS (eds.), *Phonetics and Phonology*. Rhythm and Meter. New York: Academic Press, pp. 201—260.

HAYES, B. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

HELDNER, M., and E. STRANGERT 1997. "To what extent is perceived focus determined by F0-cues?" *Proceedings of EUROSPEECH 97* 2: 875—878.

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

HIRST, D., and C. BOUZON 2005. "The effect of stress and boundaries on segmental duration in a corpus of authentic speech (British English)". *Proceedings of Interspeech 2005*. Lisbon, pp. 29—32.

HOEQUIST, C. 1983. "Durational correlates of linguistic rhythm categories". *Phonetica* 40: 19—23.

HOGAN, J.T., and A.J. ROZSYPAL 1980. "Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant". *Journal of the Acoustical Society of America* 67: 1764—1771.

HOONHORST, I., C. COLIN, E. MARKESSIS, M. RADEAU, P. DELTENRE and W. SERNICLAES 2009. "French native speakers in the making: From language-general to language-specific voicing boundaries". *Journal of Experimental Child Psychology* 104: 353—366.

HORNBY, A.S. (ed.) 2000. *Oxford Advanced Learner's Dictionary of Current English*. 6th edition. Oxford: Oxford University Press.

HOUSE, A.A. 1961. "On vowel duration in English". *Journal of the Acoustical Society of America* 33: 1174—1178.

HOUSE, A.S., and G. FAIRBANKS 1953. "The influence of consonantal environment upon the secondary acoustical characteristics of vowels". *Journal of the Acoustical Society of America* 25: 105—113.

HUGGINS, A.W.F. 1972. "On the perception of temporal phenomena in speech". *Journal of the Acoustical Society of America* 51: 1279—1290.

INKELAS, S., and Y. ROSE 2008. "Positional neutralization: A case study from child language". *Language* 83 (4): 707—736.

JACKENDOFF, R.S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

JACKENDOFF, R. 1977. *X-bar Syntax*. Cambridge, MA: MIT Press.

JASSEM, W. 1952. *Intonation of Conversational English (Educated Southern British)*. Wrocław: Wrocławskie Towarzystwo Naukowe.

JASSEM, W. 1959. "The Phonology of Polish Stress". *Word* 15: 252—269.

JASSEM, W. 1962. "Akcent języka polskiego". *Prace Językoznawcze* 31: 1—116.

JASSEM, W. 1971. *Podręcznik wymowy angielskiej*. Warszawa: Państwowe Wydawnictwo Naukowe.

JASSEM, W., D.R. HILL and I.H. WITTEN 1984. "Isochrony in English speech: its statistical validity and linguistic relevance". In D. GIBBON and H. RICHTER (eds.),

Intonation, Accent and Rhythm. Studies in Discourse Phonology. Berlin, New York: Walter de Gruyter, pp. 203—225.

JASSEM, W., J. MORTON and M. STEFFEN-BATÓG 1968. "The perception of stress in synthetic speech-like stimuli by Polish listeners". In W. JASSEM (ed.), *Speech Analysis and Synthesis*, vol. 1. Warszawa: PWN, pp. 289—308.

JASSEM, W., and L. RICHTER 1989. "Neutralization of voicing in Polish obstruents". *Journal of Phonetics* 17: 205—212.

JENKINS, J. 1998. "Which pronunciation norms and models for English as an International Language?" *ELT Journal* 52 (2): 119—126.

JENKINS, J. 2000. *The Phonology of English as an International Language*. Oxford: Oxford University Press.

JENKINS, J. 2007. *English as a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press.

JENNER, B. 1999. "Teaching pronunciation: the common core". *Newsletter of the IATEFL Pronunciation Special Interest Group* 1.

JENSEN, C. 2004. *Stress and Accent. Prominence Relations in Southern Standard British English*. PhD Thesis, University of Copenhagen.

JONES, D. 1964 [1918, 1949]. *An Outline of English Phonetics*. 9th edition. Cambridge: Heffer.

JUN, S-A. 1993. *The Phonetics and Phonology of Korean*. PhD Thesis. Ohio State University.

JUN, S-A. 1998. "The Accentual Phrase in the Korean prosodic hierarchy". *Phonology* 15 (2): 189—226.

JUN, S-A. 2005. "Prosodic typology". In IDEM (ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, pp. 430—458.

JURAFSKY, D., A. BELL, M. GREGORY and W.D. RAYMOND 2001. "Probabilistic relations between words: Evidence from reduction in lexical production". In J. BYBEE and P. HOPPER (eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins Publishing Company, 229—254.

KAHN, D., 1976. *Syllable-based Generalizations in English Phonology*. Bloomington, IN: Indiana University Linguistics Club.

KEATING, P.A. 1985. "Universal phonetics and the organization of grammars". In V.A. FROMKIN (ed.), *Phonetic Linguistics: Essays in Honour of Peter Ladefoged*. Orlando, FL: Academic Press, pp. 115—132.

KEATING, P.A. 1988. "Underspecification in phonetics". *Phonology* 5: 275—292.

KEATING, P.A., T. CHO, C. FOUGERON and C.-S. HSU 2003. "Domain-initial articulatory strengthening in four languages". In J. LOCAL, R. OGDEN and R. TEMPLE (eds.), *Papers in Laboratory Phonology*, vol. 6. Cambridge: Cambridge University Press, pp. 143—161.

KEATING, P.A., B. LINDBLOM, J. LUBKER and J. KREIMAN 1994. "Variability in jaw height for segments in English and Swedish VCVs". *Journal of Phonetics* 22: 407—422.

KENWORTHY, J. 1987. *Teaching English Pronunciation*. Harlow, Essex: Longman.

KIM, H., and J. COLE 2005. "The stress foot as a unit of planned timing evidence from shortening in the prosodic phrase". *Interspeech 2005*. Lisbon, Portugal, pp. 2365—2368.

KINGDON, R. 1958. *The Groundwork of English Intonation*. London: Longman.
KIRCHNER, R.M. 1998. *An Effort-based Approach to Consonant Lenition*. PhD Thesis, University of California, Los Angeles.
KLATT, D.H. 1973. "Interaction between two factors that influence vowel duration". *Journal of the Acoustical Society of America* 54: 1102—1104.
KLATT, D.H. 1974. "The duration of [s] in English words". *Journal of Speech and Hearing Research* 17: 51—63.
KLATT, D.H. 1975. "Vowel lengthening is syntactically determined in a connected discourse". *Journal of Phonetics* 3: 129—140.
KLATT, D.H. 1976. "Linguistic uses of segmental duration in English: acoustic and perceptual evidence". *Journal of the Acoustical Society of America* 59: 1208—1221.
KLESSA, K., and D. ŚLEDZIŃSKI 2007. "A study of chosen temporal relations within syllable structure in Polish". *Speech and Language Technology* 9/10: 87—92.
KO, E. 2007. "Acquisition of vowel duration in children speaking American English". *Proceedings of Interspeech* 2007: 1881—1884.
KOCHANSKI, G., E. GRABE, J. COLEMAN and B. ROSNER 2005. "Loudness predicts prominence: fundamental frequency lends little". *Journal of Acoustical Society of America* 118 (2): 1038—1054.
KOZHEVNIKOV, V.S., and L.A. CHISTOVICH 1965. *Speech: Articulation and Perception*. Washington, D.C.: Joint Publication Research Service No. 30,543 (translated from *Rech: Artikulyatsiya i Vospriyatiye*. Moscow, Leningrad 1965).
KRAUSE, S.E. 1982. "Developmental use of vowel duration as a cue to postvocalic stop consonant voicing". *Journal of Speech and Hearing Research* 25: 388—393.
KRIVOKAPIĆ, J. 2007. "Prosodic planning: effects of phrasal length and complexity on pause duration". *Journal of Phonetics* 35 (2): 162—179.
KUZLA, C. 2009. *Prosodic Structure in Speech Production and Perception*. PhD Thesis, Radboud University Nijmegen.
LADD, D.R. 1980. *The Structure of Intonational Meaning*. Bloomington, IN: Indiana University Press.
LADD, D.R. 1986. "Intonational phrasing: The case of recursive prosodic structure". *Phonology* 3: 311—340.
LADD, D.R. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
LADEFOGED, P. 1967. *Three Areas of Experimental Phonetics*. London: Oxford University Press.
LADEFOGED, P. 1975. *A Course in Phonetics*. New York: Harcourt Brace Jovanovich, Inc.
LADEFOGED, P. 2004. "Phonetics and phonology in the last 50 years". *UCLA Working Papers in Phonetics* 103: 1—11.
LADO, R. 1957. *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press.
LAEUFER, C. 1992. "Patterns of voicing-conditioned vowel duration in French and English". *Journal of Phonetics* 20: 411—440.
LAVER, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.
LEHISTE, I. 1970. *Suprasegmentals*. Cambridge, MA: The MIT Press.

LEHISTE, I. 1972. "The timing of utterances and linguistic boundaries". *Journal of the Acoustical Society of America* 51: 2018—2024.

LEHISTE, I. 1977. "Isochrony reconsidered". *Journal of Phonetics* 5 (3): 253—263.

LENNEBERG, E.H. 1967. *Biological Foundations of Language*. New York: Wiley.

LEV-ARI, S., and B. KEYSAR 2010. "Why don't we believe non-native speakers? The influence of accent on credibility". *Journal of Experimental Social Psychology* 46: 1093—1096.

LIBERMAN, M., 1975. *The Intonational System of English*. PhD Thesis, Massachusetts Institute of Technology.

LIBERMAN, M., and A. Prince 1977. "On stress and linguistic rhythm". *Linguistic Inquiry* 8: 249—336.

LIEBERMAN, P. 1960. "Some acoustic correlates of word stress in American English". *Journal of the Acoustical Society of America* 32: 451—454.

LINDBLOM, B. 1963. "Spectrographic study of vowel reduction". *Journal of the Acoustical Society of America* 35: 1773—1781.

LINDBLOM, B., and K. RAPP 1973. *Some Temporal Regularities of Spoken Swedish*. Stockholm: University of Stockholm, Institute of Linguistics.

LINDBLOM, B. 1990. "Explaining phonetic variation: a sketch of the H&H theory". In H.J. HARDCASTLE and A. MARCHAL (eds.), *Speech Production and Speech Modeling*. Dordrecht: Kluwer Academic Publishers, pp. 403—439.

LIPPI-GREEN, R. 1997. *English With an Accent: Language, Ideology, and Discrimination in the United States*. London: Routledge.

LISKER, L. 1974. "On 'explaining' vowel duration variation". *Glossa* 8: 223—236.

LISKER, L. 1986. "Voicing in English: a catalogue of acoustic features signalling /b/ versus /p/ in trochees". *Language and Speech* 29: 3—11.

LISKER, L., and A.S. ABRAMSON 1964. "A cross-language study of voicing in initial stops: acoustic measurements". *Word* 20: 384—422.

LISKER, L., and A.S. ABRAMSON 1967. "Some effects of context on voice onset time in English stops". *Language and Speech* 19: 1—28.

LOW E.L., E. GRABE and F. NOLAN 2000. "Quantitative characterisations of speech rhythm: syllable-timing in Singapore English". *Language and Speech* 43: 377—401.

LUCE, P.A., and J. CHARLES-LUCE 1985. "Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production". *Journal of the Acoustical Society of America* 78: 1949—1957

LUKE, K.K., and J.C. RICHARDS 1982. "English in Hong-Kong: Functions and status". *English World-Wide* 3: 147—164.

MACK, M. 1982. "Voicing-dependent vowel duration in English and French: Monolingual and bilingual production". *Journal of the Acoustical Society of America* 71: 173—178.

MADDIESON, I. 1985. "Phonetic cues to syllabification". In V.A. FROMKIN (ed.), *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. New York: Academic Press, pp. 203—221.

MADDIESON, I. 1997. "Phonetic Universals". In W.J. HARDCASTLE and J. LAVER (eds), *The Handbook of Phonetic Sciences*. Oxford: Blackwell, pp. 619—639.

MAJOR, R.C. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

MARCUS, S.M. 1981. "Acoustic determinants of perceptual center (P-center) location". *Perception and Psychophysics* 30: 247—256.

MARKHAM, D. 1997. *Phonetic Imitation, Accent, and the Learner*. PhD Thesis (Travaux de l'Institut de linguistique de Lund, vol. 33). Lund: Lund University Press.

MARTIN, J.G. 1970. "On judging pauses in spontaneous speech". *Journal of Verbal Learning and Verbal Behavior* 9: 75—78.

MO, Y. 2008a. "Duration and intensity as perceptual cues for naïve listeners' prominence and boundary perception". *Speech Prosody 2008*. Campinas, Brazil, pp. 739—742.

MO, Y. 2008b. "Acoustic correlates of prosodic prominence for naïve listeners of American English". *Proceedings of the 34th Annual Meeting of the Berkeley Linguistic Society*. Retrieved 30 September 2011. http://prosody.beckman.illinois.edu/pubs/43_Mo_2008_bls34.pdf.

MORTON, J., and W. JASSEM 1965. "Acoustic correlates of stress". *Language and Speech* 8: 150—181.

MOYER, A., 1999. "Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction". *Studies in Second Language Acquisition* 21: 81—108.

MUNRO, M., and T. DERWING 1995. "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners". *Language Learning Technology* 45 (1): 73—97.

MUNRO, M.J., and T.M. DERWING 1994. "Evaluations of foreign accent in extemporaneous and read material". *Language Testing* 11: 253—266.

NAKATANI, L.H., K.D. O'CONNOR and C.H. ASTON 1981. "Prosodic aspects of American English speech rhythm". *Phonetica* 38: 84—106.

NEMSER, W. 1971. "Approximative systems of foreign language learners". *IRAL* 9 (2): 115—123.

NERI, A., C. CUCCHIARINI, H. STRIK, and L. BOVES 2002. "The pedagogy-technology interface in computer assisted pronunciation training". *Computer Assisted Language Learning* 15 (5): 441—467.

NESPOR, M. 1990. "On the rhythm parameter in phonology". In I.M. ROCA (ed.), *Logical issues in language acquisition*. Dordrecht: Foris Publications, pp. 157—175.

NESPOR, M, and I. VOGEL 1986. *Prosodic Phonology*. Dordrecht, Riverton: Foris Publications.

NEWMAN, S. 1946. "On the stress system of English". *Word* 2: 171—187.

NOOTEBOOM, S.G. 1972. *Production and Perception of Vowel Duration: A Study of Properties of Vowel in Dutch*. Eindhoven: Philips Research Lab.

NOOTEBOOM, S.G., and G.J.N. DOODEMAN 1980. "Production and perception of vowel length in spoken sentences". *Journal of the Acoustical Society of America* 67: 276—287.

NOWACKA, M. 2003. "'Hot bits' or 'hobbits'? — English speech in Polish 'ears'". In W. SOBKOWIAK and E. WANIEK-KLIMCZAK (eds.), *Dydaktyka fonetyki języka obcego*. Płock: Wydawnictwo Naukowe PWSZ w Płocku, pp. 96—113.

NOWACKA, M. 2008. *The Phonetic Attainment in Polish University and College Students of English. A Study in the Productive and Receptive Pronunciation Skills*. PhD Thesis, Uniwersytet Marii Curie-Skłodowskiej.

NOWAK, P. 2006. *Vowel Reduction in Polish*. PhD Thesis, University of California, Berkeley.

NUNAN, D. 1995. *Language Teaching Methodology*. Phoenix: ELT.

O'CONNOR, J.D., and G.F. ARNOLD 1961. *Intonation of Colloquial English*. London: Longman.

O'CONNOR, J.D., and G.F. ARNOLD 1973. *Intonation of Colloquial English*. 2nd edition. London: Longman.

OHALA, J.J., and W.G. EWAN 1973. "Speed of pitch change (A)". *Journal of Acoustical Society of America* 53 (1): 345.

OLLER, D.K. 1973. "The effect of position in utterance on speech segment duration in English". *Journal of the Acoustical Society of America* 54: 1235—1247.

ORTEGA-LLEBARIA, M., and P. PRIETO 2007. "Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables". In P. PRIETO, J. MASCARÓ and M.-J. SOLÉ (eds.), *Segmental and Prosodic Issues in Romance Phonology* (Current Issues in Linguistic Theory). Amsterdam, Philadelphia: John Benjamins, pp. 155—176.

PALMER, H.E. 1922. *English Intonation, with Systematic Exercises*. Cambridge: Heffer.

PENNINGTON, M.C. 1999. "Computer-aided pronunciation pedagogy: promise, limitations, directions". *Computer Assisted Language Learning* 12: 427—440.

PETERSON, G.E., and I. LEHISTE 1960. "Duration of syllable nuclei in English". *Journal of the Acoustical Society of America* 32: 693—703.

PIERREHUMBERT, J. 1980. *The Phonology and Phonetics of English Intonation*. PhD Thesis, Massachusetts Institute of Technology.

PIERREHUMBERT, J., and D. TALKIN 1992. "Lenition of /h/ and glottal stop". In G.J. DOCHERTY and D.R. LADD (eds.), *Papers in Laboratory Phonology II*. Cambridge: Cambridge University Press, pp. 90—117.

PIKE, K. 1945. *The Intonation of American English*. Ann Arbor: University of Michigan Press.

PORT, R.F. 1979. "The influence of tempo on stop closure durations a cue for voicing and place". *Journal of Phonetics* 7: 45—56.

PORT, R.F. 1981. "Linguistic timing factors in combination". *Journal of the Acoustical Society of America* 69: 262—274.

PORT, R.F. 2007. "The graphical basis of phones and phonemes". In O.-S. BOHN and M.J. MUNRO (eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam, New York: John Benjamins Publishing Company, pp. 349—365.

PORT, R.F., J. DALBY and M. O'DELL 1987. "Evidence for mora timing in Japanese". *Journal of the Acoustical Society of America* 81: 1574—1585.

PORZUCZEK, A. 1998. *The Acquisition of the Vocalic Component of English by Advanced Polish Learners*. PhD Thesis, University of Silesia, Katowice.

PORZUCZEK, A. 2007. "English vowel quantity in Polish learner's speech perception and production". In J. ARABSKI, D. GABRYŚ-BARKER, A. ŁYDA (eds.), *PASE Papers*

2007. Studies in Language and Methodology of Teaching Foreign Languages. Katowice: PARA, pp. 96—105.

Porzuczek, A. 2010a. "The weak forms of TO in the pronunciation of Polish learners of English". In E. Waniek-Klimczak (ed.), *Issues in Accents of English 2. Variability and Norm*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, pp. 309—324.

Porzuczek, A. 2010b. "The effect of IP constituent complexity on unstressed syllable duration in Polish learner's English pronunciation". In E. Waniek-Klimczak (ed.), *Research in Language* 8. Warszawa: Versita, pp. 149—162.

Porzuczek, A. (in press). "Segmental and prosodic reasons for non-native timing of her in Polish learner's English speech". In E. Waniek-Klimczak and L.R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*. Berlin, Heidelberg: Springer-Verlag.

Pulgram, E. 1970. *Syllable, Word, Nexus, Cursus*. The Hague: Mouton.

Rakerd, B., W. Sennett and C.A. Fowler 1987. "Domain-final lengthening and foot-level shortening in spoken English". *Phonetica* 44: 147—155.

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

Raphael, L.J. 1972. "Preceding vowel duration as a cue to the voicing characteristics of word-final consonants in English". *Journal of the Acoustical Society of America* 51: 1296—1303.

Rietveld, A.C.M., and C. Gussenhoven 1985. "On the relation between pitch excursions and prominence". *Journal of Phonetics* 13: 299—308.

Rivera-Gaxiola, M., J. Silva-Peyrera and P.K. Kuhl 2005. "Brain potentials in native and non-native speech contrasts in 7- and 11-month-old American infants". *Developmental Science* 8: 162—172.

Roach, P. 1982. "On the distinction between 'stressed-timed' and 'syllable-timed' languages". In D. Crystal (ed.), *Linguistic Controversies*. London: Edward Arnold, pp. 73—79.

Roach, P. 2002. "Studying rhythm and timing in English speech: Scientific curiosity or a classroom necessity?" In E. Waniek-Klimczak and J. Melia, (eds.), *Accents and Speech in Teaching English Phonetics and Phonology*. Frankfurt: Peter Lang, pp. 199—206.

Roach, P. 2009. *English Phonetics and Phonology. A Practical Course*. 4th edition. Cambridge: Cambridge University Press.

Rogers, H. 2000. *The Sounds of Language: An Introduction to Phonetics*. Harlow: Longman.

Rojczyk, A. 2010a. "Preceding vowel duration as a cue to the consonant voicing contrast: Perception experiments with Polish-English bilinguals". In E. Waniek-Klimczak (ed.), *Issues in Accents of English 2: Variability and Norm*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, pp. 341—360.

Rojczyk, A. 2010b. "Vowel quality and duration in stressed and unstressed positions in Polish". Poster presented at GLOW Colloquium, Wrocław, Poland, 13 April 2010.

Rojczyk, A. 2010c. *Temporal and Spectral Parameters in Perception of the Voicing Contrast in English and Polish*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

RUBACH, J. 1974. "Syllabic consonants in Polish". *Journal of Phonetics* 22: 109—116.
RYAN, E.B., and M.A. CARRANZA 1975. "Evaluative reactions of adolescents toward speakers of standard English and Mexican American accented English". *Journal of Personality and Social Psychology* 31 (5): 855—863.
SALTZMAN, E., and D. BYRD 1999. "Dynamical simulations of a phase window model of relative timing". *Proceedings of the 14th International Congress of the Phonetic Sciences*. San Francisco, pp. 2275—2278.
SALTZMAN, E., and K.G. MUNHALL 1989. "A dynamical approach to gestural patterning in speech production". *Ecological Psychology* 1: 333—382.
SCHAIRER, K.E. 1992. "Native speaker reaction to non-native speech". *Modern Language Journal* 76 (3): 309—319.
SCHMERLING, S.F. 1974. *Aspects of English Sentence Stress*. Austin: Texas University Press.
SCHMIDT, A.M., and J.E. FLEGE 1995. "Effects of speaking rate changes on native and non-native speech production". *Phonetica* 52: 41—54.
SCOTT, S.K. 1993. *P-centers in Speech: An Acoustic Analysis*. PhD Thesis, University College, London.
SCOVEL, T. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Cambridge, New York: Newbury House.
SELINKER, L. 1972. "Interlanguage". *IRAL* 10 (3): 209—231.
SELKIRK, E.O. 1972. *The Phrasal Phonology of English and French*. PhD Thesis, Massachusetts Institute of Technology.
SELKIRK. E.O. 1980. "Prosodic domains in phonology: Sanskrit revisited". In M. ARONOFF and M.L. KEAN (eds.), *Juncture*. Saratoga, CA: Anma Libri, pp. 107—129.
SELKIRK, E.O. 1982. "The syllable". In H. VAN DER HULST and N. SMITH (eds.), *The Structure of Phonological Representations, Part II*. Dordrecht: Foris, pp. 337—383.
SELKIRK, E.O. 1984. *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge, MA: MIT Press.
SELKIRK, E.O. 1986. "On derived domains in sentence phonology". *Phonology* 3: 371—405.
SELKIRK, E. 1995a. "The prosodic structure of function words". In J. BECKMAN, L.W. DICKEY and S. URBANCZYK (eds.), *Papers in Optimality Theory*. Amherst, MA: GLSA Publications, pp. 439—469.
SELKIRK, E. 1995b. "Sentence prosody: Intonation, stress and phrasing". In J. GOLDSMITH (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 550—569.
SELKIRK, E. 2005. "Comments on Intonational Phrasing in English". In S. FROTA, M. VIGARIO and M.-J. FREITAS (eds.), *Prosodies (Selected Papers from the Phonetics and Phonology in Iberia Conference, 2003). Phonetics and Phonology*. Berlin: Mouton de Gruyter, pp. 11—58.
SELKIRK, E., and K. TATEISHI 1988. "Constraints on minor phrase formation in Japanese". In L. MACLEOD, G. LARSON and D. BRENTARI (eds.), *Proceedings of the 24th Annual Meeting of the Chicago Linguistics Society*. Chicago: Chicago Linguistics Society, pp. 316—336.

SETTER, J. 2006. "Speech rhythm in world Englishes. The case of Hong Kong". *TESOL Quarterly* 40 (4): 763—782.

SHATTUCK-HUFNAGEL, S., and A. TURK 1998. "The domain of phrase-final lengthening in English". *The Sound of the Future: A Global View of Acoustics in the 21st Century (Proceedings 16th International Congress on Acoustics and 135th Meeting Acoustical Society of America)*. New York, pp. 1235—1236.

SILVERMAN, K., M. BECKMAN, J. PITRELLI, M. OSTENDORF, C. WIGHTMAN, P. PRICE, J. PIERREHUMBERT and J. HIRSCHBERG 1992. "ToBI: a standard for labeling English prosody". In J.J. OHALA, T.M. NEAREY, B.L. DERWING, M.M. HODGE and G.E. WIEBE (eds.), *Proceedings of the International Conference on Spoken Language Processing*. Banff: University of Alberta, pp. 867—870.

SLOWIACZEK, L., and D. DINNSEN 1985. "On the neutralizing status of Polish word-final devoicing". *Journal of Phonetics* 13: 325—341.

SLOWIACZEK, L.M., and H.J. SZYMANSKA 1989. "Perception of word-final devoicing in Polish". *Journal of Phonetics* 17: 205—212.

SLUIJTER, A.M.C. 1995. *Phonetic Correlates of Stress and Accent*. PhD Thesis, Rijksuniversiteit te Leiden.

SLUIJTER, A.M.C., and V.J. VAN HEUVEN 1996a. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100: 2471—2485.

SLUIJTER, A.M.C., and V.J. VAN HEUVEN 1996b. "Acoustic correlates of linguistic stress and accent in Dutch and American English". *Proceedings of ICSL, 96*. Philadelphia: Applied Science and Engineering Laboratories, Alfred I du Pont Institute, pp. 630—633.

SMILJANIĆ, R., and A.R. BRADLOW 2008. "Stability of temporal contrasts across speaking styles in English and Croatian". *Journal of Phonetics* 36: 91—113.

SMITH, B., A.R. BRADLOW and T. BENT 2003. "Production and perception of temporal contrasts in foreign accented English". In M.J. SOLE, D. RECASENS and J. ROMERO J. (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, pp. 519—522.

SMITH, C. 2004. "Topic transitions and durational prosody in reading aloud: Production and modeling". *Speech Communication* 42: 247—270.

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

SOBKOWIAK, W., 2001 [1996]. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

STERIADE, D. 2001. "Directional asymmetries in place assimilation". In E. HUME and K. JOHNSON (eds.), *The Role of Speech Perception in Phonology*. San Diego: Academic Press, pp. 219—50.

STRANGERT, E. 1997. "Relating prosody to syntax: Boundary signaling in Swedish". In G. KOKKINAKIS, N. FAKOTAKIS and E. DERMATAS (eds.), *Proceedings of the 5th European Conference on Speech Communication and Technology*. Rhodes, Greece pp. 239—242.

STREETER, L.A. 1978. "Acoustic determinants of phrase boundary perception". *Journal of the Acoustical Society of America* 64: 1582—1592.

SUNDBERG, J. 1979. "Maximum speed of pitch changes in singers and untrained subjects". *Journal of Phonetics* 7: 71—79.

SWEET, H. 1891. *A Handbook of Phonetics*. Oxford: Henry Frowde.
SWEET, H. 1906. *A Primer of Phonetics*. Oxford: Clarendon.
SZPYRA-KOZŁOWSKA, J. 2003. "The lingua franca core and the Polish learner". In W. SOBKOWIAK and E. WANIEK-KLIMCZAK (eds.), *Dydaktyka fonetyki języka obcego*. Płock: Wydawnictwo Naukowe PWSZ, pp. 193—210.
SZPYRA-KOZŁOWSKA, J., J. FRANKIEWICZ and W. GONET 2002. "Aspekty fonetyki angielskiej nauczane w szkołach średnich". In W. SOBKOWIAK and E. WANIEK-KLIMCZAK (eds.), *Dydaktyka fonetyki języka obcego*. Płock: Wydawnictwo Naukowe PWSZ, pp. 9—27.
SZPYRA-KOZŁOWSKA, J., M. NOWACKA, L. BĄK, I. CHABER and R. ŚWIĘCIŃSKI 2003. "Priorytety w nauczaniu fonetyki języka angielskiego". In W. SOBKOWIAK and E. WANIEK-KLIMCZAK (eds.), *Dydaktyka fonetyki języka obcego*. Płock: Wydawnictwo Naukowe PWSZ, pp. 211—224.
ŚPIEWAK, G., and L. GOŁĘBIOWSKA 2001. "Polish speakers". In M. SWAN and B. SMITH (eds.), *Learner English. A Teacher's Guide to Interference and Other Problems*. 2nd edition. Cambridge: Cambridge University Press, pp. 162—178.
TAJIMA, K., R. PORT and J. DALBY 1997. "Effects of temporal correction on intelligibility of foreign-accented English". *Journal of Phonetics* 25: 10—24.
'T HART, J., R. COLLIER and A. COHEN 1990. *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.
TATHAM, M., and K. MORTON 2001. "Intrinsic and adjusted unit length in English rhythm". *Proceedings of the Institute of Acoustics — WISP 2001*. St Albans: Institute of Acoustics, pp. 189—200.
TAYLOR, D.S. 1981. "Non-native speakers and the rhythm of English". International *Review of Applied Linguistics in Teaching* 19 (3): 219—226.
TERKEN, J. 1991. "Fundamental frequency and perceived prominence of accented syllables". *Journal of the Acoustical Society of America* 89 (4): 1768—1776.
TURK, A., S. NAKAI and M. SUGAHARA 2006. "Acoustic segment durations in prosodic research: A practical guide". In S. SUDHOFF, D. LENERTOVA, R. MEYER, S. PAPPERT, P. AUGURZKY, I. MLEINEK, N. RICHTER and J. SCHLIESER (eds.), *Methods in Empirical Prosody Research*. Berlin, New York: De Gruyter, pp. 1—28.
TURK, A.E., and J.R. SAWUSCH 1993. "On the perceptual integrality of duration and amplitude cues to stress (A)". *Journal of the Acoustical Society of America* 93 (4): 2371.
TURK, A.E., and J.R. SAWUSCH 1995. "The domain of the durational effects of accent". *Speech Communication Group Working Papers* 10: 42—71.
TURK, A.E., and J.R. SAWUSCH 1997. "The domain of accentual lengthening in American English". *Journal of Phonetics* 25: 25—41.
TURK, A.E., and S. SHATTUCK-HUFNAGEL 2000. "Word-boundary-related duration patterns in English". *Journal of Phonetics* 28: 397—440.
TURK, A.E., and S. SHATTUCK-HUFNAGEL 2007. "Phrase-final lengthening in American English". *Journal of Phonetics* 35 (4), 445—472.
TURK, A.E., and L. WHITE 1999. "Structural influences on accentual lengthening in English". *Journal of Phonetics* 27: 171—206.

TWADDELL, W.F. 1935. "On defining the phoneme". *Language Monograph* 16. Reprinted in M. JOOS (ed.) 1957. *Readings in Linguistics I*. Chicago: University of Chicago Press, pp. 55—80.

UMEDA, N. 1975. "Vowel duration in American English". *Journal of the Acoustical Society of America* 58: 434—445.

UMEDA, N. 1977. "Consonant duration in American English". *Journal of the Acoustical Society of America* 61: 846—858.

VAN LANCKER, D., J. KREIMAN, and D. BOLINGER, D. 1988. "Anticipatory lengthening". *Journal of Phonetics* 16: 339—347.

VAN SANTEN, J.P.H. 1992. "Contextual effects on vowel duration". *Speech Communication* 11: 513—546.

VAN SON, R.J.J.H., and L.C.W. VAN POLS 1996. "An acoustic profile of consonant reduction". *Proceedings of the 4th International Conference on Spoken Language Processing*. Philadelphia, pp. 1529—1532.

VANDERSLICE, R., and P. LADEFOGED 1972. "Binary suprasegmental features and transformational word-accentuation rules". *Language* 48: 819—838.

WAGNER, P. 2007. "Visualising levels of rhythmic organization". *Proceedings of 16th International Congress of Phonetic Sciences*. Saarbrücken, pp. 1113—1116.

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

WANIEK-KLIMCZAK, E. 2009. "Sources of difficulty in the acquisition of English rhythm by Polish learners: A developmental perspective". In M. WYSOCKA (ed.), *On Language Structure, Acquisition and Teaching. Studies in Honour of Janusz Arabski on the Occasion of His 70th Birthday*. Katowice: Wydawnictwo Uniwersytetu Śląskiego, pp. 369—378.

WARDHAUGH, R. 1970. "The Contrastive Analysis Hypothesis". *TESOL Quaterly* 4 (2): 123—130.

WEINREICH, U. 1953. *Languages in Contact. Findings and Problems*. The Hague: Mouton Publishers.

WELLS, J.C., 2008 [1990]. *Longman Pronunciation Dictionary*. 3rd edition with CD-ROM. Harlow: Pearson Education Limited.

WENK, B.J., and F. WIOLAND 1982. "Is French really syllable-timed?" *Journal of Phonetics* 10: 193—216.

WHITE, L. 2002. *English Speech Timing: A Domain and Locus Approach*. PhD Thesis, University of Edinburgh.

WHITE, L., and S.L. MATTYS 2007a. "Calibrating rhythm: First language and second language studies". *Journal of Phonetics* 35: 501—522.

WHITE, L., and S.L. MATTYS 2007b. "Rhythmic typology and variation in first and second languages". In P. PRIETO, J. MASCARÓ and M.-J. SOLÉ (eds.), *Segmental and Prosodic Issues in Romance Phonology*. Amsterdam: John Benjamins Publishing Company, pp. 237—257.

WHITE, L., S.L. MATTYS, L. SERIES and S. GAGE 2007. "Rhythm metrics predict rhythmic discrimination". *Proceedings of 16th International Congress of Phonetic Sciences*. Saarbrücken, pp. 1009—1012.

WIGET, L., L. WHITE, B. SCHUPPLER, I. GRENON, O. RAUCH and S.L. MATTYS 2010. "How stable are acoustic metrics of contrastive speech rhythm?" *Journal of the Acoustical Society of America* 127 (3): 1559—1569.

WIGHTMAN, C.W. 2002. "ToBI or not ToBI". In B. BEL, I. MARLIN (eds.), *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence: Laboratoire Parole et Langage, pp. 25—29.

WIGHTMAN, C.W., S. SHATTUCK-HUFNAGEL, M. OSTENDORF and P.J. PRICE 1992. "Segmental durations in the vicinity of prosodic phrase boundaries". *Journal of the Acoustical Society of America* 91: 1707—1717.

WIIK, K. 1965. "Finnish and English vowels: a comparison with special reference to the learning problems met by native speakers of Finnish learning English". *Annales Universitatis Turkensis, Series B, No 94*. Turku: Turun Yliopisto.

WINGFIELD, A., and J.F. KLEIN 1971. "Syntactic structure and acoustic pattern in speech perception". *Perception and Psychophysics* 9/1A: 23—35.

WREMBEL, M. 2002. "Miejsce fonetyki języka angielskiego w szkole — implikacje dla kształcenia nauczycieli". In W. SOBKOWIAK and E. WANIEK-KLIMCZAK (eds.), *Dydaktyka fonetyki języka obcego*. Płock: Wydawnictwo Naukowe PWSZ, pp. 29—40.

XU, Y., and X. SUN 2002. "Maximum speed of pitch change and how it may relate to speech". *Journal of the Acoustical Society of America* 111 (3): 1399—1413.

XU, Y., and C.X. XU 2005. "Phonetic realization of focus in English declarative intonation". *Journal of Phonetics* 33: 159—197.

YOON, T-J., J. COLE and M. HASEGAWA-JOHNSON 2007. "On the edge: Acoustic cues to layered prosodic domains". *Proceedings of 16th International Congress of Phonetic Sciences*. Saarbrücken, pp. 1017—1020.

INDEX OF NAMES

SUBJECT INDEX

Index of languages

Andrzej Porzuczek

Relacje czasowe pomiędzy elementami frazy intonacyjnej w wymowie angielskiej zaawansowanego ucznia polskiego

S t r e s z c z e n i e

Niniejsza praca poświęcona jest analizie relacji czasowych pomiędzy elementami składowymi frazy intonacyjnej w wymowie angielskiej zaawansowanego ucznia polskiego. Celem pracy jest wykazanie i opisanie różnic w tym zakresie między polskim uczniem a rodzimym użytkownikiem języka angielskiego oraz ich interpretacja w kontekście glottodydaktycznym.

W części teoretycznej omówiono historię i stan badań nad prozodią języka mówionego oraz metodologię akustycznych badań mowy. Rozdział pierwszy przedstawia modele struktury prozodycznej wypowiedzi w celu ustalenia jednostek istotnych dla analizy relacji czasowych, czyli tych elementów frazy, które mogą stanowić odrębną domenę procesów wpływających na czas trwania artykulacji. Rozdział ten charakteryzuje również owe procesy, opisując domenę i zasięg ich oddziaływania. Drugi rozdział poświęcony jest pojęciu akcentu, który jest kluczowym zjawiskiem decydującym o ogólnym kształcie prozodycznym wypowiedzi, a więc rytmie, intonacji i tytułowych relacjach czasowych między poszczególnymi elementami. Trzeci rozdział przedstawia historię badań nad rytmem języka od momentu przedstawienia przez Kennetha Lee Pike'a idei podziału języków świata na dwie klasy według ogólnych tendencji rytmicznych mowy, do współczesnych metod określania rytmu w języku na podstawie parametrów, takich jak zróżnicowanie długości samogłosek czy stopień złożoności zbitek spółgłoskowych. W rozdziale trzecim przedstawione są również problemy ucznia polskiego z opanowaniem angielskiej prozodii wynikające z różnic pomiędzy językami.

Rozdział czwarty rozpoczyna badawczą część książki. Opisuje empiryczne badanie porównawcze relacji czasowych w tekście czytanym przez polskich słuchaczy pierwszego roku kolegium nauczycielskiego w odniesieniu do analogicznych relacji w mowie czytanej rodzimych użytkowników standardowej angielszczyzny brytyjskiej. Dodatkowo, nagrania słuchaczy kolegium powtórzono po siedmiu miesiącach w celu uzyskania danych na temat kierunku i tempa rozwoju ich wymowy angielskiej w warunkach nauczania obejmującego standardowy akademicki kurs praktycznej fonetyki angielskiej. Poszczególne sekcje przedstawiają oparte na dyskusji z części teoretycznej założenia metodologiczne, hipotezy badawcze, materiał językowy wybrany do analizy, strukturalno-akustyczne kryteria podziału analizowanych fraz intonacyjnych

na mniejsze jednostki (stopy, sylaby, segmenty) oraz techniczne procedury badawcze. Piąty rozdział koncentruje się na przedstawieniu wyników odnoszących się do czasu trwania segmentów wokalicznych w wymowie obu grup respondentów. Analizie poddano zarówno bezwzględną długość samogłosek, jak również ich relatywną długość w odniesieniu do kontekstu. Rozdział szósty przedstawia wyniki odnoszące się do wyższych poziomów hierarchii prozodycznej: relacje czasowe pomiędzy sylabami w obrębie stopy, jak również proporcje czasu trwania stóp w różnych pozycjach frazy intonacyjnej. W rozdziale siódmym dokonano podsumowania wyników i przedstawiono propozycje odnośnie do kierunków przyszłych badań i wnioski dydaktyczne mogące poprawić skuteczność przyswajania wymowy angielskiej przez Polaków.

Na podstawie przeprowadzonych badań stwierdzono wyraźnie dłuższy czas trwania elementów nieakcentowanych (samogłosek, sylab, wyrazów funkcyjnych, anakruzy) w wymowie Polaków, z wyjątkiem końcowej sylaby frazy intonacyjnej. Istotne różnice wystąpiły zarówno w wartościach absolutnych, jak i w proporcjach czasowych. Nie zaobserwowano natomiast wyraźnych różnic w bezwzględnej długości samogłosek i sylab akcentowanych pomiędzy obiema grupami respondentów, z wyjątkiem sylab akcentowanych na końcu frazy, gdzie są one znacznie dłuższe w wymowie rodowitych Anglików. Większy niż u Polaków kontrast między elementami akcentowanymi a nieakcentowanymi wynika prawdopodobnie z bardziej radykalnej redukcji elementów nieakcentowanych w angielskiej wymowie rodzimej.

Relacje czasowe w obrębie stopy oraz w jednostkach wyższych poziomów struktury prozodycznej, mogące wskazywać na tendencje rytmiczne w mowie, również sugerują rozbieżności między grupami respondentów w miejscach, gdzie decydujący wpływ na czas trwania jednostek ma redukcja elementów nieakcentowanych. Istotne różnice znaleziono także w przypadku jednostek leksykalnych, stanowiących stały element często używanych konstrukcji gramatycznych, np. *have to* czy *going to*. Zaobserwowano ponadto większą u rodzimych użytkowników języka angielskiego tendencję do wyrównywania czasu trwania stopy rytmicznej obejmującej ciąg sylab nieakcentowanych oraz poprzedzającą je sylabę akcentowaną. Największe rozbieżności dotyczyły czasu trwania anakruzy, która w wymowie respondentów angielskich jest wyraźnie krótsza.

W odniesieniu do tendencji rozwojowych polskich uczniów, stwierdzono znaczące zbliżenie się wyników do norm wymowy rodzimej po siedmiu miesiącach od pierwszego badania. Wzrosło ogólne tempo mowy, które jednak nie zawsze szło w parze z uzyskaniem bardziej "angielskich" proporcji czasu trwania składowych elementów wypowiedzi. O około połowę zmniejszyła się różnica między Polakami i Anglikami w bezwzględnych wartościach czasu trwania jednostek nieakcentowanych, choć w niektórych kontekstach (np. w anakruzie) większości uczniów nie udało się uzyskać wyników zbliżonych do wymowy rodzimych użytkowników języka angielskiego. Nie zmieniły się również istotnie wskaźniki określające zróżnicowanie długości samogłosek akcentowanych, co wskazuje na trudność w wykorzystaniu różnic czasowych do kontrastowania samogłosek napiętych i nienapiętych oraz sygnalizowania dźwięczności wygłosu sylaby i granic domen prozodycznych.

Wyniki badań oraz jakościowa analiza pojedynczych kontekstów sugerują duży wpływ artykulacji segmentów na relacje czasowe na poziomie frazy i zdania.

W związku z tym zalecane jest utrzymanie tradycyjnej kolejności wprowadzanych ćwiczeń fonetycznych, polegającej na treningu wymowy segmentów w stopniowo rozszerzanym kontekście, a następnie koncentracji na kolejnych, wyższych poziomach struktury prozodycznej wypowiedzi.

Przedstawione w niniejszej pracy rezultaty badań oraz wykorzystanie zastosowanych w nich metod mogą posłużyć do identyfikacji konkretnych problemów w przyswajaniu obcej wymowy, jak również wprowadzić element obiektywizmu do zazwyczaj impresjonistycznej oceny warstwy prozodycznej wymowy języka obcego.

Andrzej Porzuczek

Temporale Relationen zwischen den einzelnen Elementen der Intonationsphrase in der englischen Aussprache eines fortgeschrittenen polnischen Schülers

Z u s a m m e n f a s s u n g

Die vorliegende Monografie ist der Untersuchung von temporalen Relationen zwischen den Bestandteilen der Intonationsphrase in der englischen Aussprache eines fortgeschrittenen polnischen Schülers gewidmet. Der Verfasser hat sich zum Ziel gesetzt, die Unterschiede zwischen dem polnischen Schüler und dem Muttersprachler aufzuweisen und zu beschreiben und sie glottodidaktisch zu interpretieren.

Im theoretischen Teil werden die Geschichte und der Stand von den Forschungen über die Prosodie der gesprochenen Sprache, und die Methodologie der akustischen Sprachuntersuchungen erörtert. Das erste Kapitel präsentiert Strukturmodelle der prosodischen Aussage, welche die für die Analyse der temporalen Relationen wichtigen Phrasenelemente unterscheiden lassen, also Elemente, die eine separate Domäne der die Artikulationszeit beeinflussenden Prozesse bilden können. An der Stelle charakterisiert der Verfasser diese Prozesse, indem er ihre Domäne und ihren Einwirkungsbereich darstellt. Das zweite Kapitel ist dem Akzent gewidmet, der für die allgemeine prosodische Form der Aussage, d.h.: Rhythmus, Intonation und temporale Relationen zwischen den einzelnen Elementen, entscheidenden Erscheinung. Im dritten Kapitel wird die Geschichte der Forschungen über den Sprachrhythmus geschildert, angefangen von der von Kenneth Lee Pike vorgestellten Idee der Einteilung der Weltsprachen in zwei Klassen, den allgemeinen rhythmischen Sprachtendenzen entsprechend, bis zu gegenwärtigen Methoden der Sprachrhythmusbestimmung aufgrund solcher Parameter, wie: differenzierte Vokallänge oder Komplexitätsgrad der Konsonantenhäufung. Hier erörtert der Verfasser auch die aus den Unterschieden zwischen den beiden Sprachen entstehenden Probleme der polnischen Schüler mit der Beherrschung der englischen Prosodie.

Mit dem vierten Kapitel beginnt der Forschungsteil des Buches. Der Verfasser beschreibt empirische vergleichende Forschung von temporalen Relationen in einem von polnischen Hörern des ersten Jahres des Lehrerkollegs und von den Muttersprachlern des britischen Standardenglischen gelesenen Text. Die Tonaufnahmen

wurden noch nach sieben Monaten wiederholt, damit die Richtung und das Tempo der Entwicklung von der englischen Aussprache bei den Hörern im Rahmen des Universitätskurses der praktischen englischen Phonetik überprüft werden konnten. Die einzelnen Forschungssektionen enthalten methodologische Richtlinien, Forschungshypothesen, den zur Forschung ausgewählte Sprachstoff, strukturell-akustische Kriterien der Einteilung von den untersuchten Intonationsphrasen in kleinere Einheiten (Versfüße, Silben, Segmente) und technische Forschungsverfahren. Im fünften Kapitel werden Ergebnisse der Forschungen über die Laufzeit der Vokalsegmenten in der Aussprache der beiden Gruppen der Befragten präsentiert. Analysiert wurden sowohl die absolute Länge der Vokale, wie auch deren kontextbezogene relative Länge. Das sechste Kapitel beinhaltet die, die höheren Stufen der prosodischen Hierarchie betreffenden Ergebnisse: temporale Relationen zwischen den Silben innerhalb eines Versfußes und das Verhältnis von der Laufzeit der Versfüße in verschiedenen Positionen der Intonationsphrase. Im siebten Kapitel werden alle Ergebnisse zusammengefasst, die Richtungen der künftigen Forschungen bestimmt und die der besseren Aneignung der englischen Aussprache von den Polen dienenden didaktischen Schlüsse gezogen.

Aufgrund der durchgeführten Forschungen wurde eine deutlich längere Laufzeit von unbetonten Elementen (Vokalen, Silben, Funktionswörtern, Auftakten), ausgenommen die Endsilbe der Intonationsphrase, in der Aussprache der Polen festgestellt. Wesentliche Unterschiede beobachtete man sowohl in absoluten Werten wie auch im Zeitverhältnis. Im Bereich der absoluten Länge von Vokalen und betonten Silben dagegen wurden keine deutlichen Unterschiede zwischen den beiden Gruppen der Befragten festgestellt, ausgenommen die betonten Endsilben der Phrase, wo sie bei gebürtigen Engländern viel länger waren. Der bei ihnen beobachtete größere als bei den Polen Kontrast zwischen den betonten und unbetonten Elementen folgt wahrscheinlich daraus, dass unbetonte Elemente in der englischen Aussprache von den Muttersprachlern viel radikaler reduziert werden.

Temporale Relationen innerhalb des Versfußes und auf höheren Stufen der prosodischen Struktur, die rhythmische Tendenzen in der Sprache anzeigen können, lassen auch bestimmte Diskrepanzen zwischen den einzelnen Gruppen der Befragten an den Stellen vermuten, wo die Laufzeit der Sprecheinheiten vor allem von der Reduktion der unbetonten Elemente abhängt. Wesentliche Unterschiede wurden auch bei den lexikalen Einheiten festgestellt, die ein festes Element von solchen häufig gebrauchten grammatischen Konstruktionen, wie: have to oder going to sind. Es wurde darüber hinaus eine größere als bei Muttersprachlern Tendenz zur Kompensation der Laufzeit von dem rhythmischen Versfuß beobachtet, welcher die Folge von unbetonten Silben und die vorhergehende betonte Silbe umfasst. Die größten Diskrepanzen betrafen die Laufzeit des Auftaktes, der von englischen Befragten viel kürzer ausgesprochen wurde.

Wenn es um Entwicklungstendenzen der polnischen Schüler geht, wurde festgestellt, dass diese Ergebnisse den Normen der muttersprachlichen Aussprache nach sieben Monaten von der ersten Untersuchung ähnlich waren. Das allgemeine Sprechtempo wurde gestiegen, doch es ging nicht immer damit einher, dass die Laufzeit von den Bestandteilen der Aussage der in der Aussprache der Muttersprachler beobachteten Laufzeit ähnlich ist. In absoluten Werten der Laufzeit von unbetonten Sprecheinheiten war der Unterschied zwischen Polen und Engländern etwa um eine

Hälfte kleiner, obwohl in manchen Kontexten (z.B.: in dem Auftakt) ist es den meisten Schülern nicht gelungen, sich an englische Muttersprachler heranzureichen. Die die Differenzierung von der Länge der betonten Vokale bezeichnenden Kennzahlen haben sich auch nicht erheblich geändert, was davon zeugt, dass es schwierig ist, temporale Unterschiede zur Kontrastierung von gespannten und ungespannten Vokalen, und zur Signalisierung der Stimmhaftigkeit des Silbenauslautes und der Grenzen von prosodischen Domänen anzuwenden.

Die Forschungsergebnisse und die qualitative Analyse von den einzelnen Kontexten lassen einen großen Einfluss von der Artikulation der Segmente auf temporale Relationen im Phrasen- u. Satzbereich vermuten. Es wird deshalb empfohlen, die gebräuchliche Reihenfolge von phonetischen Übungen aufrechtzuerhalten und die Aussprache von den einzelnen Segmenten in einem schrittweise erweiterten Kontext zu trainieren und sich dann auf nächste, höhere Stufen der prosodischen Aussage zu konzentrieren.

Die in vorliegender Monografie dargestellten Forschungsergebnisse und die in den Forschungen angewandten Methoden können der Identifizierung von konkreten Problemen mit der Aneignung der fremden Aussprache dienen und die gewöhnlich impressionistische Beurteilung der prosodischen Ebene der fremden Aussprache objektivieren.

Executive Editor: Krystian Wojcieszuk

Cover Designer: Paulina Dubiel

Technical Editor: Małgorzata Pleśniar

Proofreader: Danuta Stencel

Computer-generated forms: Edward Wilk

ISSN 0208-6336
ISBN 978-83-226-2090-8

Published by
Wydawnictwo Uniwersytetu Śląskiego
ul. Bankowa 12B, 40-007 Katowice
www.wydawnictwo.us.edu.pl
e-mail: wydawus@us.edu.pl

First impression. Printed sheets: 13.0. Publishing sheets: 13.0.
Offset paper, grade III, 90 g Price: 16 zł (+ VAT)

Printing and binding: PPHU TOTEM s.c.
M. Rejnowski, J. Zamiara
ul. Jacewska 89, 88-100 Inowrocław

ISSN 0208-6336 • ISBN 978-83-226-2090-8 • Cena 16 zł (+ VAT)