Państwowa Wyższa Szkoła Zawodowa w Nowym Sączu

The acquisition of English intonation by Polish adult learners

Monika Zięba-Plebankiewicz

Nowy Sącz 2013

Table of Contents

INTRODUCTION

The present dissertation has been driven by the assumption that the melody of speech is one of the first aspects of language to be acquired by infants but the last to be mastered by second language learners. What is more, as a bulk of research reveals, while the growing proficiency in the L2 has an increasing influence on the mother tongue and may cause a loss of nativeness in L1 speech, L1 intonation tends to remain unaffected. However, most of the existing studies into phonological acquisition concentrated on the segmental aspect of language. Intonation has received greater attention only recently, mostly because of technological advance which allows for a detailed and precise analysis of speech. The thesis has aimed at providing an explanation to the phenomenon, answering the need of devising a theoretical model of acquisition of intonation.

Many studies into intonation concentrated on its language-specific character being the main cause of errors committed by L2 learners (e.g. Backman 1977; Willems 1982; Mennen 1998). Acquisition, however, is affected by universal influences to a much greater extent than by the mother tongue patterns; therefore a model of the process should incorporate both types of forces.

The dissertation is composed of four chapters and an appendix. Chapter 1 presents definitions of prosody, suprasegmentals and intonation, followed by a historical overview of approaches to this aspect of language. The main components of intonation, i.e. duration, loudness and pitch, are discussed, along other prosodic elements: stress, accent, prominence and rhythm. The chapter proceeds to the description of intonational meanings and functions, including those universal and paralinguistic, encompassed by the three biological codes. Finally, the intonation of Polish and English is briefly described.

Chapter 2 offers an account of the studies into the acquisition of intonation. Maturational, social, psychological and individual factors affecting the process are addressed, as well as various views on the notion of nativeness. An overview of perceptual studies is followed by the description of the influence of the universals and the L1.

Chapter 3 provides the reasons for selecting Natural Phonology as the theoretical foundation for the model of the acquisition of intonation. The basic tenets of Natural Phonology and its concepts related to first and second language acquisition are presented. Intonational typology based on the holistic approach of Donegan and Stampe (1983) is established for English and Polish. Finally, a model of the acquisition of intonation is proposed with the emphasis on the phonological processes influencing learning: the universal processes and typological preferences.

Chapter 4 presents an empirical verification of the assumptions formulated in chapter 3. The experiment investigated the influence of universal processes and language-specific preferences on the development of L2 intonational system in the L2 learners' interlanguage. An evaluation of the experiment follows and topics for future research are provided.

The dissertation concludes with the recapitulation of the results of the experiment and the evaluation of the proposed model of the acquisition of intonational.

List of abbreviations

CAH Contrastive Analysis Hypothesis

CP Critical Period

DIAL dialogue

f0 fundamental frequency

IP intonational phrase

L1 first language

L2 second language

MDH Markedness Differential Hypothesis

MON monologue

OM Ontogeny Model

OPM Ontogeny-Phylogeny Model

SCH Structural Conformity Hypothesis

SDRH Similarity Differential Rate Hypothesis

SLA Second Language Acquisition

CHAPTER 1: Prosody and prosodic features

1.1. Defining prosody and intonation

Pitch variations in natural languages have been the part of many phonological and phonetic theories since Joshua Steele (1775). However, they were usually treated as a marginal aspect of language. The turn of the 20th century witnessed a revival in the study of intonation, which constitutes a vital part of the prosodic level of language (e.g. Ladd 1996; Steffen-Batogowa 1996; Hammond 1997; Hirst – Di Cristo 1998; Botinis 2000; Carmichael 2000; Gussenhoven 2000, 2001, 2002; Chun 2002; Grabe – Karpiński 2003; Grabe 2004, Karpiński 2006).

The term *prosody* is often confusing. It is derived from the Greek word *προσωδία*, 'prosodia', denoting "song sung to a musical instrument" (*Dictionary of Language and Linguistics* 1996), so prosody is "the musical accompaniment to the words" (Fox, 2000:1). It is concerned with units greater than the individual segments, namely length, accent, stress, tone and intonation, hence another name: suprasegmentals. However, the precise description of the scope and current linguistic meaning of prosody is far from being easy.

Many linguists, e.g. Cruttenden (1986), use the terms "prosody" and "suprasegmentals" interchangeably, since, as he admits, they coincide to a large extent. Other linguists claim, though, that the phonological structure above the segment is too complex to consider prosodic features as the ones simply superimposed on sounds. Thus, one can draw a distinction between "suprasegmental" and "prosodic" mode of description (Fox 2000). "Suprasegmental", a term introduced by American structuralists who believed that the speech continuum consists of minimal units, is applied to "distinctive features that, unlike a phoneme, cannot be segmented individually from linguistic utterances" (*Dictionary of Language and Linguistics* 1996). The processes at the suprasegmental level vary from language to language: in English they include stress, rhythm and intonation, as well as vowel length, sound reduction, elision, coarticulation, and assimilation, while other languages, such as Turkish or Finnish, involve vowel harmony (Hyman 1975). Therefore, this term can be used to "refer to a particular formalisation in which a phonological feature or process is conceived of in non-segmental terms; in theory, any phonological feature can be analysed in this way, whether prosodic or not" (Fox 2000: 2).

According to Laver (1994), the analysis of the suprasegmental level of language concerns the description of the domains larger than a segment, such as, among others, settings[1], the syllable or the utterance. The suprasegmental patterns include pitch and loudness, the perceptual correlates of fundamental frequency, rhythm, "the complex perceptual pattern produced by the interaction in time of the relative prominence of

---

[1] "Any tendency for the vocal apparatus to maintain a given configuration of featural state over two or more segments in close proximity in the stream of speech" (Laver 1994: 153).

stressed and unstressed syllable" (Laver 1994: 152), rate, that is the tempo of speech, and continuity, depending on the occurrence of pauses in the utterance.

On the other hand, the term "prosodic" can be seen as evolving from the British tradition which typically analyses spoken language in situational context, that is, the situation in which an utterance has occurred. The term thus refers to this part of an utterance which is left once the single segments of speech are set aside, i.e. length, accent, stress, tone and intonation. What should also be excluded from prosody are the paralinguistic and non-linguistic auditory aspects of speech since they are not systematic and not used conventionally or intentionally in communication. The former refer to temporary modifications of voice, resulting in e.g. whisper, falsetto, etc. (voice qualifiers) and a giggle, cry, etc. (voice qualifications) (Crystal 1969: 133, 138), while the latter – to voice quality, sneezes or coughs. On the whole, prosody can be defined as "those auditory components of an utterance which remain once segmental as well as non-linguistic and paralinguistic effects have been removed" (Couper-Kuhlen 1985: 2-3) or as "sets of mutually defining phonological features which have an essentially variable relationship to the words selected" (Crystal 1969: 5). That means that the denotative meaning of a word will be dependent on its segmental structure but not on loudness, duration or pitch[2].

The place of stress is, at least in English, problematic. Stress does differentiate words, e.g. *'project* (noun) from *pro'ject* (verb), therefore it does not belong to the scope of prosody, unlike the type of stress which decides about the rhythmic beat and/or about the placement of the nucleus within an utterance. The latter does not influence word meaning, therefore it constitutes a part of prosody (Couper-Kuhlen 1985: 4). The components of speech are presented in Figure 1.1.
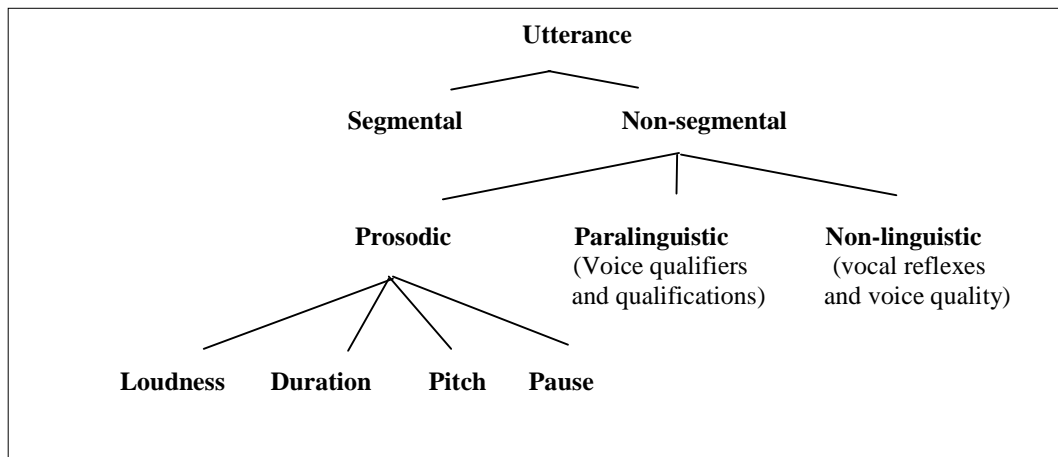
Figure 1.1 Auditory components of speech in communication (Couper-Kuhlen 1985: 4)

---

[2] Only in intonational languages.

There is no single accepted definition of intonation. Bolinger (1989: 1) remarked that "intonation is a symptom of how we feel about what we say and how you feel when you say it." Other linguists are more precise; for them the term refers to "the distinctive use of patterns of pitch, or melody" (Crystal 2003). Phonetically, intonation can be defined as "continuous, and continuously varying, pitch pattern" which is perceived as uninterrupted, since voiceless breaks are not salient enough for the listener to register (Fox 2000: 274). Ladd (1996: 6) offers his own definition of intonation which includes certain key points: suprasegmental phonetic features that carry sentence-level or postlexical pragmatic meanings in a linguistically structured way. Suprasegmental elements are restricted to fundamental frequency (f_0), intensity and duration. Sentence-level pragmatic meanings refer to whole phrases and utterances, namely various sentence types, speech acts, focus and information structure, thus excluding word stress, accent and tone. Finally, intonation is linguistically structured, i.e. it is a system of categorically distinct entities, e.g. low vs. high tones, and relations, e.g. strong vs. weak.

In many sources, though, the term "intonation" is used interchangeably with "prosody". The differences between definitions result from the ambiguity of the word, since intonation can be defined in its broad sense as comprising such lexical characteristics as word stress, tone and quality, or in its narrow sense, without these elements (Hirst – Di Cristo 1998: 4). Another source of ambiguity concerns the levels of analysis and description. On the phonetic level, intonation relates to the variations of such acoustic parameters as fundamental frequency (f_0), which is the primary parameter, intensity and duration (see Figure 1.1). Some authors include even rhythm which is reflected in variations of spectral characteristics (Crystal 1969). On the other hand, the linguistic level of representation, at least in English, often included the lexical (word stress) and non-lexical prosodic features (intonation) corresponding directly to the physical parameters of intensity and fundamental frequency respectively. It soon occurred, however, that the correspondence is more complicated, as fundamental frequency is also an important perceptual cue for stress (e.g. Jassem 1952; Bolinger 1958; Fry 1958; Lehiste 1970; after Hirst – Di Cristo 1998).

Hirst and Di Cristo (1998: 1-2) remark that the description of intonation poses a challenge for a researcher, since it is one of the most universal and, at the same time, one of the most language specific features of natural languages. Intonation is universal because it is present in all languages and some of its functions are shared by many unrelated linguistic systems, e.g. raised pitch marking non-finality, in contrast to lowered pitch. On the other hand, intonation is language-specific, as there exists no universal intonational meanings and the choice of a particular pattern depends not only on the language but also on the dialect, style, mood or attitude of the speaker.

Another source of difficulty in the description of intonation is the fact that the phenomenon indeed has meaning, though frequently elusive, contrary to other phonological features, both prosodic and segmental. For example, a falling intonation usually denotes 'completeness' or 'statement', while a rising intonation tends to be

interpreted as 'incompleteness' or 'question'. The meaning of intonational patterns is often associated with speaker's attitudes and emotions rather than with grammatical categories, which contributes to problems in describing intonation systems (Fox 2000). Speakers vary considerably in vocal expression of their emotions, therefore it is difficult to establish any direct relationship between emotions and particular intonation patterns. Liberman et al. (1967) claim that there is no single acoustic correlate of the emotional modes and prosody: a given emotion can be expressed by a number of contours and vice versa, some intonational patterns can depict a number of attitudes. Another problematic aspect of intonation is the fact that the distinctions it creates are not discrete but gradient, e.g. the falling contour can be of various heights paralleled by relevant gradience of meaning. All these problems led Bolinger (1978) to describe intonation as "a half-tamed savage", since the phenomenon cannot be accounted for in terms of the structures employed in the analysis of segments. Other linguists claimed that intonation belongs "around the edges of language", on the verge of paralanguage (Fox 2000: 270).

In the present work, the terms "suprasegmental" and "prosodic" will be used interchangeably, as the minute distinctions between them are of no great importance for this project and will be reserved for the broad sense of the above-the-segment level of language. Both will refer to the abstract cognitive systems and the physical parameters they are mapped onto. "Intonation" will refer to the non-lexical features that include both the phonetic features of utterances and the phonological level of pitch patterns, declination, boundary phenomena, etc., which are described later in this chapter.

1.2. The phonetics of prosody

A phonetic basis for the description of prosodic features includes certain components of the speech process. On the whole, three components of the physiology of speaking have been specified:
   a) the subglottal component: the lungs and trachea which produce the air stream;
   b) the larynx: responsible for voicing, aspiration, glottalisation and regulation of the pitch;
   c) the supralaryngeal component: the pharynx, mouth and nose, most of the segments.

For the segmental level of language the supralaryngeal component is relevant, as the place and manner of articulation are defined in terms of the positions of speech organs. The feature of voice, aspiration and glottalization are included in the laryngeal component. What differentiates prosodic features from segmental features, though, is the fact that they mostly result from laryngeal or subglottal activity: pitch, intensity and other components of suprasegmentals depend on the activity of the laryngeal and respiratory muscles (Fox 2000: 3-4).

A phonetic description of intonation in terms of such speech components is not adequate, though, due to the overlapping described above. Since certain segmental and

suprasegmental features are localised in the laryngeal component, the description does not provide a clear distinction between what is and what is not prosody. Therefore, other features must be taken into consideration.

1.2.1. Duration

Prosodic features differ acoustically from one another quite considerably. Both tone and intonation are realised in terms of the same parameter of pitch, fundamental frequency (f_0) measuring the rate of vibration of the vocal folds, unlike accent and stress. When it comes to duration[3], however, its accurate phonetic description appears to be almost impossible. First of all, in order to measure the relative duration of sounds, one should be able to draw clear syllable boundaries which are not always easy to determine[4], especially in the framework that treats intervocalic consonants as ambisyllabic. Secondly, the duration of speech sounds are conditioned by global and local factors: the former include the accent of the speaker, the overall tempo and continuity of the utterance, as well as the paralinguistic state of the language user at the moment of speaking. Local factors are exemplified by the influence of the neighbouring sounds, the segmental structure of the syllable, stress placement, and the position of the sound within the utterance: utterance-initially, medially or finally. (Laver 1994: 432). Therefore, no absolute duration for a particular sound can be established, for which reason duration is usually described as relative. For a detailed discussion of duration and length see Laver (1994).

1.2.2. Loudness

Acoustically, loudness is identified with intensity[5] (or power) of a sound or a sequence of sounds depending on the varying pressure of the air stream. This feature is related to *amplitude*, another component of sound, along frequency. As in the case of duration, absolute intensity of a syllable often depends on different influences which makes it difficult to assess. As for linguistic purposes for which loudness can be used, it adds emphasis to an utterance or expresses speaker's emotions (Cruttenden 1986). Moreover, the average loudness of speech varies not only between individual speakers, but also between accent-communities. Abercrombie (1967: 95) tentatively assumes that the degree of loudness might be language-specific, e.g. Egyptian Arabic can be perceived as relatively louder than Scots Gaelic.

---

[3] Here *duration* and *length* are used interchangeably, but the two terms are often treated separately, with *duration* being a phonetic feature and *length* – a phonological one (Laver 1994: 436).
[4] For a detailed discussion of the syllable see, e.g. Dziubalska-Kołaczyk (2002).
[5] According to Laver (1994: 501), loudness is "the perceptual feature relating to the physical concept of intensity."

1.2.3. Pitch

Pitch is a perceptual concept, i.e. relating to listener's judgements, involved in tone and intonation[6]. Its acoustic correlate, as mentioned in 1.2.1., is fundamental frequency. It is estimated that in European languages f_0 values in conversational speech range from 120 Hz for men to 220 Hz for women and 330 Hz for children below ten years of age (Laver 1994: 451). That is why pitch is relative: perceptual judgements will be made "in terms of a hypothesized placement within the general range of pitch over which the speaker's voice is believed to move" (Laver 1994: 457), therefore the lowest pitch produced by a child will be perceived as higher than the highest pitches of a man. Moreover, the perception of the pitch of a particular syllable in an utterance (in intonational languages) depends on the pitch of the neighbouring syllables (Laver 1994).

The composition of syllables influences an overall pitch pattern. First of all, only voiced sounds possess a repetitive waveform, however, in the flow of speech gaps in voicing are not perceived. Voicing, though, influences the f_0 of neighbouring voiced sounds, namely vowels, will have a higher f_0 after voiceless consonants. What is more, the f_0 peak will fall on the beginning of the vowel placed after a voiceless consonant but if the vowel follows a voiced consonant, the peak is in the middle of it. As far as pitch changes are concerned, a given tone (pitch pattern), e.g. a fall (HL%), is realised differently if it falls on a word consisting of a single syllable voiced throughout, e.g. *John*, or *Betty*, composed of two syllables with two short vowels and a voiceless consonant in between. In the former case, the fall is rather continuous, in the latter – it is a step between a high and a low tone (Cruttenden 1986).

The melody of an utterance depends on the sequence of the relative pitch values within the speaker's pitch-range as perceived by the listener. According to Laver (1994), the listener assesses the speaker's pitch-range, and all the meanings associated with the phenomenon, on the basis of a number of assumptions.

a) Organic range: the maximum range which the speaker is able to produce, depending on their individual anatomy and physiology.
b) Paralinguistic range: the adjustments of pitch-range within the organic range in order to convey attitudinal information (which is culture-specific).
c) Linguistic range: the habitual adjustments of the speaker's range of pitch in paralinguistically unmarked, attitudinally neutral utterances.
d) The phonological pitch span: the local range utilised for prosodic purposes within an utterance or part of an utterance.

Pitch span involves two concepts: the baseline, "which is the series of pitch-values that is perceived as forming the floor of the current pitch span" and the plateau (topline),

---

[6] *Tone* refers to "speech melody when it is a property of the word" and *intonation* – "when it is the property of the sentence" (Abercrombie 1967: 105).

"which is the series of pitch-values that is perceived in forming the ceiling of the current pitch-span" (Laver 1994: 459), both being trends over several syllables. The width of the pitch span is defined as the interval between the baseline and the plateau.

In order to decipher the information conveyed by the speaker's pitch-range, the listener must possess at least basic knowledge of the speaker's organic range of pitch, their current attitudinal state, the usual linguistic range, as well as the value of the pitch span currently used.

Another feature of pitch is declination, a phenomenon present in many languages, though not in all. It refers to "the downward slope of both the baseline and the topline, with the progressive narrowing of the pitch span within the linguistic range" (Laver 1994: 459).

A description of intonation must also take into account the notions of pitch height and pitch contour. The former relates to the relative height of the pitch of a given syllable within the values of pitch span (e.g. "high", "mid", "low", etc.), and the latter to "the shape and direction of the trajectory shown by any perceptible change in pitch-value though the duration of the syllable", e.g. "level", "rise", "fall", etc. (Laver 1994: 461).

Finally, pitch contours are aligned with the segmental element of an utterance: a change in pitch may occur in the early, middle or late part of the syllable, or it may extend over more than a single syllable ('t Hart – Collier – Cohen 1990: 153). Figure 1.2 presents the dependencies between various aspects of prosody.

Articulatory dimension	Acoustic dimension	Auditory dimension
Vibration of ----------------------- vocal folds	Fundamental -------------------- frequency (f_0)	Pitch
Physical effort -------------------- (intensity)	Amplitude -----------------------	Loudness

Figure 1.2   Relationship between the articulatory, acoustic and auditory dimensions of prosody (after Couper-Kuhlen 1985: 7)

1.2.4. Stress, accent and prominence

Intonation, like music, is part of a larger metrical structure involving the grouping, accentuation and rhythm of its components (Laver 1994). However, since the present work concentrates on intonation, these components of metrical structure will be treated briefly.

Stress refers to the degree of force used in producing a syllable in a word or sentence (Crystal 2003) so that the syllable is more prominent for communicative purposes (Chun 2002). There are several phonetic correlates of stress, i.e. intensity (stressed syllables have higher intensities than unstressed ones caused by the quicker closure of the glottis during vocal folds vibrations), fundamental frequency, although f_0 contours depend more on the intonational grammar than on a word's stress pattern, quality (vowels in unstressed syllables tend to be centralised and less rounded), finally duration, as consonants and vowels are tend to be longer in stressed syllables (Gussenhoven 2004: 15). For many authors (e.g. Cruttenden 1986; Chun 2002), stress refers to prominence in its general sense excluding pitch change as a primary factor, while prominence involving pitch is called accent. However, in various approaches to intonation those terms are frequently ambiguous.

1.2.5. Rhythm

The perceived regularity of prominent syllables in an utterance is referred to as rhythm. It can be discussed in terms of a contrast between long and short syllables, or high and low pitch, yet its description in terms of stressed vs. unstressed syllables is the most frequent (Crystal 2003).

The traditional view of rhythm (Pike 1946; Abercrombie 1965, 1967) holds that the rhythmic organisation varies from language to language, with some systems, e.g. English, German, Russian or Arabic, revealing a tendency to stress-timing, i.e. pronouncing stressed syllables at roughly equal intervals of time, no matter how many unstressed syllables between the stressed ones exist. In other languages, e.g. French, Spanish, Japanese, Italian, Turkish or Yoruba, syllables, not stresses, tend to be pronounced at regular intervals. Such languages are referred to as syllable-timed (cf. Donegan – Stampe's (1983) "isosyllabic" and "isoaccentual" dichotomy). Therefore, the length of words depends on the number of syllables, not the number of stresses (Chun 2002). The third group of languages reveals a tendency for mora-timing (e.g. Japanese), in which rhythm depends on morae, or sub-units of syllables which consist of one short vowel and any preceding onset consonants. The duration of successive morae is said to be near-equal, therefore, mora-timed languages bear more resemblance to syllable-timed languages than to stress-timed languages.

Empirical studies, however, proved that strict isochrony is not absolute, but speakers "merely perceive intervals between beats as isochronous" (Couper-Kuchlen 1985: 52). A more recent rhythmic classification (e.g. Dauer 1983, 1987; after Grabe 2004) places languages on a rhythmic continuum and states that languages can be more or less stress-based. For instance, prominent syllables occur at regular intervals in English, a stress-timed language, yet a similar phenomenon can be observed also in Spanish, a syllable-timed language. However, in English prominent syllables are perceptually more salient than in Spanish. Therefore, what influence the rhythmic structure of a language are

the phonological, phonetic, lexical and syntactic factors, such as the structure of the syllable, vowel reduction and word stress. The fact that syllable structures of stress-timed languages are more varied than in syllable-timed languages, and that in syllable-timed languages vowel reduction is rare causes the languages of these types to differ.

Nespor (1990) rejects traditional rhythmic categories since they fail to account for "intermediate languages", such as Polish or Catalan, which are rhythmically mixed, i.e. they possess the properties of both stress-timing and syllable-timing (cf. Ramus et al. 1999).

For a detailed discussion of rhythm see, e.g. Donegan – Stampe (1983), the model of Beats-and-Binding by Dziubalska-Kołaczyk 2002a, and Zborowska 2001.

1.3. The phonology of prosody

Prosodic features must also be described in terms of their linguistic role (Fry 1968). As Crystal claims (1969: 5),

> we may define prosodic systems as sets of mutually defining phonological features which have an essentially variable relationship to the words selected [7], as opposed to those features … which have a direct and identifying relationship to such words.

What distinguishes prosodic from segmental features is that, since they relate to phenomena larger than a segment (Lehiste 1970, Laver 1994), the linguistic distinctions they evoke differ from those of segments, namely they are both paradigmatic and syntagmatic: unlike segments, they can be "established by a comparison of items in sequence" (Lehiste 1970: 2), in other words, they must be defined "in relation to other items in the same utterance" (Ladefoged 1975).

It should be noted that the relationship between prosodic and segmental features from the phonological point of view is the reverse of the phonetic perspective. Phonetically, prosodic features constitute the more basic part of speech production; phonologically, they have been considered as secondary to segments, since they are seen as modifications of the latter (cf. Bloomfield's (1935) "secondary phonemes"). The consequence of this stance is that prosodic features are devoid of inherent meaning and that they are meaningful only if they occur in larger forms. Bloomfield's judgement is unsatisfactory, though, since it was based only on English; in other languages prosodic features can be primary phonemes, e.g. pitch in Chinese or duration in German (Fox 2000).

Various approaches to the phonology of intonation will be discussed in the following subchapter, presenting a brief overview of intonational models.

---

[7] Except tone languages.

1.4. Models of intonation

A number of approaches to intonation have been suggested so far, some adopting whole tonal contours, others only parts, i.e. tonal sequences, as the fundamental components of intonation. For one of the most exhaustively described prosodic systems, English, two main approaches exist: the contour analysis devised in Europe, particularly in Britain, and the levels analysis that originated in the United States. The former approach describes pitch patterns according to their overall shape, grouping intonational phenomena into tone units, tone groups or sense groups which are further analysed into such components as pre-head, head, nucleus and tail. Such approach depends to a large extent on auditory judgements made by listening to recorded data.

On the contrary, the latter approach developed in the USA is based on acoustic analysis. American researchers decompose intonational contours into sequences of tones or pitch levels, except Bolinger (1951) whose theory of pitch accent is similar to the British school. The most recent theories include generative models based on metrical representations of utterances and on sequences of two tones, H (high) and L (low).

1.4.1. The British tradition

Although the British approach can be said to have begun with the works of a German scholar Hermann Klinghardt (1927), the names commonly associated with the origin of the tradition are Jones (1909) and Armstrong and Ward (1926) whose means of description focusing on the overall shape of the contour is referred to as the tune analysis. They established a set of contrasting tunes, or tones, enclosed within a domain of a sense group which is said to express a single thought. Each group has to contain a nucleus, i.e. the syllable with the greatest prominence. A tune pattern begins on the stressed syllable of the last stressed word in a group and includes the following syllables. Pauses mark boundaries between sense groups. Pitch range and pitch height are vital for this description. Five tones are proposed (Sweet 1890), level (–), rising (′), falling (ˋ), falling-rising or compound rise (ˇ) and rising-falling or compound fall (ˆ). In transcription, the tone marks are put before the word they occur in or at the end of a sentence they modify.

The tonetic approach (Palmer 1922) continues the above conventions, yet it develops the analysis of tunes by dividing them into smaller segments. The largest intonational unit is the tone group which consists of a head and a nucleus, each showing pitch contours. The head includes all the syllables before the nuclear syllable and bears a pitch pattern that differs from that of the nucleus. Tone diacritics are written before each stressed syllable which indicates where a given pitch movement begins. The tonetic approach is further developed by O'Connor and Arnold (1961), whose work is still used in foreign language classrooms. They represent tunes graphically by means of dots, large for

the nucleus and smaller for unstressed syllables, between two horizontal lines representing the range of the speaker voice. They grouped tones according to the attitude they convey:

a) *Low Fall*: the voice falls during the word from a medium to a very low pitch.
b) *High Fall*: the voice falls during the word from a high to a very low pitch.
c) *Rise-Fall*: the voice first rises from a fairly low to a high pitch, and then quickly falls to a very low pitch.
d) *Low rise*: the voice rises during the word from a low to a medium pitch or a little above.
e) *High rise*: the voice rises during the word from a medium to a high pitch.
f) *Fall-rise*: the voice first falls from a fairly high to a rather low pitch, and then, still within the word, rises to a medium pitch.

Another follower of the tonetic approach, Halliday (1967), based his phonological categories on meaningful grammatical contrasts, therefore different intonational choices in such utterances as *I'm GOing* and *I'M going* should be treated as different grammatical distinctions. According to him, even though a single sentence can be produced with a number of diverse pitch contours, not all pitch movements are of significance. He establishes five significant patterns along with their general meanings.

Tone	*Symbol*	*Tonic movement*	*Terminal pitch tendency*
1	`	falling	low
2	´	rising	high
	ˇ	falling-rising	high
3	¯´	rising	mid
4	^´	(rising)-falling-rising	mid
5	ˇ`	(falling)-rising-falling	low

Figure 1.3 Halliday's (1967) intonational patterns

According to Halliday, falling tone 1 indicates that the information communicated is known and stated, while rising tone 2 implies that unknown polarity and requesting information. Low rising tone 3 carries uncertainty. Falling-rising tone 4 is used to communicate information known but with the air of doubt or reservation. On the contrary, doubt is indicated by rising-falling tone 5 in order to be dismissed.

Apart from the general meanings ascribed to the above tones, an intonational pattern can carry an additional specific, even attitudinal, meaning, resembling the attitudinal significance described by O'Connor and Arnold, namely the degree of involvement of the speaker. For instance, falling tone 1 applied to a sentence *I saw him yesterday* tends to be perceived as neutral, rising tone 2 in the same sentence sounds contradictory or challenging and low rising tone 3 is non-commital and disengaged. Falling-rising tone 4 makes the sentence reserved on the part of the speaker, while fising-falling tone 5 indicates the speaker's involvement.

What makes this type of analysis problematic for theoretical accounts is the lack of a clearly stated distinction between the phonetic and phonological segmentation of the pattern, or the inability to establish a definite number of significant contours in a language, e.g. the exact number of falling contours. Moreover, it often does not explain how contours are mapped onto words and syllables, although O'Connor and Arnold's model does. These drawbacks are largely caused by the pedagogical orientation of the approach and the lack of sound theoretical foundation, although the analysis has some phonological basis. In spite of this, the British tradition, especially Halliday's explicitly phonological framework, has remained influential (Fox 2000; Chun 2002).

1.4.2. American structuralist approach

American linguists represent the phonemic approach to intonation, influenced by the phonemic treatments of segments. They intended to arrive at an intonational model that could be described by a small number of discrete pitch accents.

For Bloomfield (1935), intonation consists of "secondary phonemes" of pitch. He distinguished five of them: fall [.], rise [?], lesser rise [¿], which occur sentence-finally, and exclamatory pitch [!] and suspension [,] which can either combine with the former ones or are placed in a non-final position in an utterance. He did not refer to the melody of speech as "intonation" but rather "differences of pitch" (1935: 114).

The linguists who created the basis for the phonemic analysis of intonation, though, are Pike (1945) and Wells (1945) who established four distinctive pitch levels, with 1 meaning extra-high, and 4 implying low. The pitch levels indicate contour points and are grouped into patterns or contours. The most significant patterns, referred to as primary contours, are those at the end of the utterance. The first stressed syllable in a contour carries the first contour points and is called the beginning point. The end point is located at the end of the contour. Additionally, in between some contours there occurs a direction-change point. If an unstressed precontour precedes the primary contour, the two form a total contour. The pitch levels transcribed to a sentence are depicted in example 1 below.

(1)

The doctor   bought a car.

3–   °2–4–3   4–   °2–4

Example 1 presents two total contours. Hyphens join contour points in each contour and the symbol of ° indicates the beginning point of the primary contour.

Pitch levels alone carry no meaning; they merely provide "end points, beginning points or direction points to the contours" (Pike 1945: 26), thus contributing to the meaning. Therefore it is the contour that is meaningful. Pike attempts to describe the

general meanings of primary contours. The falling contours constitute contrastive pointing, i.e. the syllable which is the beginning point of this contour receives selective attention of the speaker and hearer; on the other hand, the contour falling to pitch four located before the final type pause carries the meaning of finality (Pike 1945: 44). What is worth noting, the contour falling to pitch three indicates nonfinality.

The model created by Wells (1945) resembles the one of Pike's, though he uses the reverse numbering of pitch levels. His approach is based on the theoretical framework to a greater extent than Pike's. In the analysis of pitch he applies the principles of the analysis of segments. As he groups pitch phonemes into morphemes (sequences of pitch phonemes), Wells deprives the contour of its phonological status. His approach influenced Trager – Smith (1951) who decompose the contour into four pitch phonemes, each with possible allophones. The modifications of the final part of the contour are due to one of three terminal junctures: 'single-bar' /, 'double-bar' // and 'double cross #.

The main problem with the level approach, not only for linguistic theory but also for language teachers, is the relativeness of the pitch levels. In the absence of any absolute pitch within an intonation group, one may assume four, five or more levels equally possible, which makes the analysis wholly arbitrary. Moreover, it was shown that very small pitch movements can convey meaning, while larger pitch variations may not reveal any significant differences in meaning, which the four levels frequently fails to represent. Similarly, the model does not represent clearly the phonetic details of the pitch movement from one level to another, as in the sentence ^{4}*George couldn't do it*[1] # (the numbers represent Wells's levels: 4 'extra high' and 1 'very low'), it is not obvious on which syllable exactly the transition took place. As for the terminal junctures, their usage is justifiable when they mark the reversal of the contour, as in 41// (rising terminal). On the other hand, the terminal is not significant in 41# (falling terminal) (Cruttenden 1986).

1.4.3. Bolinger's theory of pitch accents

Bolinger (1957, 1958) was the first to conduct quantitative acoustic experiments on the intonation of American English. In 1951 he proposed a theory in which intonation contours consist of a sequence of pitch accents, each pitch accent[8] being a marker of prominence. According to him, neither pitch nor stress are phonemically independent, since it is pitch that is the main cue to stress. Consequently, intonational morphemes are affected by both pitch and stress: pitch accents. Apart of these elements, intonation includes non-accentual aspects of pitch, i.e. gradient patterns, such as steep or gradual falls.

Bolinger (1958) distinguishes three types of pitch accents for American English, the most frequent Accent A, Accent B, and the least frequent Accent C, later (1989) called

---

[8]Similarly to the assumptions of the British approach, Bolinger (1951) claimed that the configurations of pitch are more meaningful than pitch levels.

profiles of the accent or "shapes determined by how the pitch jump cueing the accent is realised" (1986: 139). Accent, or Profile A, represents a configuration with "an abrupt fall in or from the syllable that is made to stand out by the fall" (1989: 3). Profile B can be related to the rise in the British tradition, as it is "marked by a jump to the syllable that is made to stand out by the jump, with any following unaccented syllables usually continuing with a gradual rise but often staying level or even falling slightly" (1989: 3). Profile C can be regarded a reverse of Profile A, since it is "marked by down to rather than down from" (1989: 4). Combinations of the profiles are possible, namely CA, AC and CAC (1986). In order to represent intonation, Bolinger illustrates pitch movements with the very utterance.

1.4.4. Generative approaches

Generative phonology constitutes a contrast to the structuralist approaches. The classical framework does not focus much on intonation. Stockwell (1960) attempted to employ Chomsky's model (1957) by adapting his rules to show that each sentence had an intonational contour (Figure 1.4).

1.	S	→	Nuc + IP (Intonation Pattern)
2.	IP	→	C (Contour) + JP (Juncture Point)
3.	C	→	Disc(ontinuity) / Cont(inuity)
4.	Disc →		001 ↓ (= any fall to 1 + terminal fade; 0 = any pitch phoneme)
5.	Cont →		002 ↓ / 003 ↓ / 004 ↓ / 021 ↑ / 032– ↑ / 043– ↑
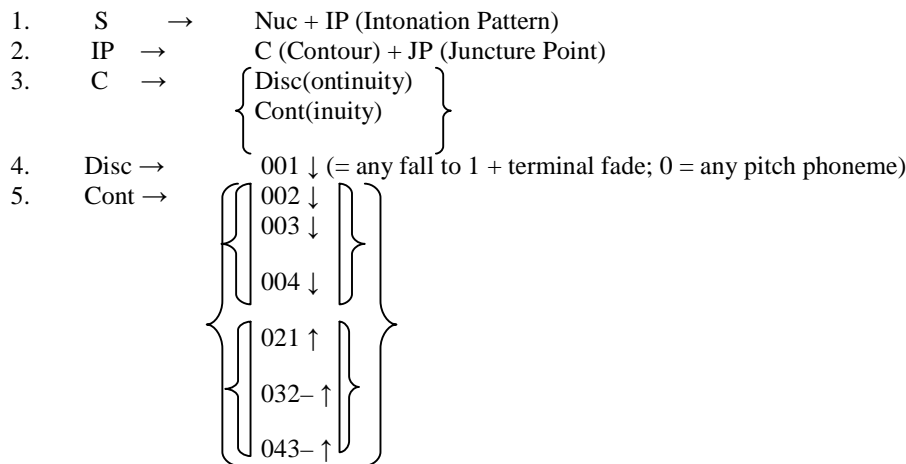
Figure 1.4. Stockwell's (1960) rules generating intonation patterns for a sentence

Nuc represents the syntactic structure of the sentence. Rules 2 to 5 generate the Intonation Pattern (IP) which includes a Contour (C) and a Juncture Point (JP). The Contour showing Discontinuity implies a falling pitch with a terminal fade, while the Contour representing Continuity results in a non-low pitch with a terminal fade or a terminal rise (Fox 2000).

The model, clear as it may be, has some flaws. It is, in fact, an adaptation of Trager and Smith's (1951) categories; therefore it cannot be considered a significant contribution to the phonology of intonation. Moreover, the model assumes that there exists only one

intonation pattern for a sentence. Another problem is the fact that the rules separate syntax from intonation, which Stockwell (1972) addresses in his later work, where he tries to incorporate the rules for intonation with the ones of syntactic structure. Despite the flaws, the idea that an appropriate algorithm can generate intonation contours has been influential (e.g. Fujisaki – Nagashima 1969) (Fox 2000).

In the 1970s, the issue of whether the placement of the sentence stress can be derived from syntactic rules, as described in Chomsky and Halle's *The Sound Pattern of English* (1968), was under debate, though no agreement was achieved.

1.4.5. The Lund School

Gårding and her colleagues at Lund (Bruce 1977; Bruce – Gårding 1978; Gårding 1981, 1983) are among those influenced by the generative approach to intonation. Since their analysis was created for Swedish, it describes both intonation and tonal accents of this language. The model treats lexical prosody separately from phrase- and sentence prosody. The basis of the analysis is the sentence, marked for "lexical accents or tones, accents (tones) at phrase and sentence level, morphological and phrase boundaries, and the mode of sentence intonation" (Gårding 1983). The final pattern results from the combination of the previously mentioned factors. The model is summarised in the algorithm in Fig.6. In Rule 1 a tonal grid, described as "the global frame for the sentence intonation" is created. Rules 2,3 and 4 allow to insert the pitch peaks on the grid.

Rule 1. *Sentence and phrase intonation*.

Draw the tonal grid using sentence and major grid boundaries.

Rule 2. *Sentence and phrase boundaries*.

Insert highs and lows on the grid according to language and dialect.

Rule 3. *Sentence and phrase accent*.

Insert highs and lows on the grid according to language and dialect.

Rule 4. *Word accent*.

Insert highs and lows on the grid according to language and dialect.

Rule 5. *Contrastive word accent*.

Adjust highs and lows according to language and dialect.

Rule 6. C*ontext rules*.

Insert highs and lows on the grid according to context.

Rule 7. *Concatenation*.

Connect neighbouring generated highs and lows.

Fig.1.5 Gårding's algorithm of pitch assignment (Gårding 1983)

The weakness of the model is that it fails to identify those characteristics of the intonation pattern that are phonologically relevant, therefore its applicability is limited. What is more, the generation of an intonational contour is based on accentual and tonal features, while the phonological features of intonation are not in focus. As a consequence, distinctions between significant and insignificant pitch features are not drawn consistently.

1.4.6. The Dutch school

The approach that originated in the Institute for Perception Research at Eidhoven (IPO) differed from the other schools in that it is based not on phonologically distinct categories but on perceptually relevant features of the acoustic signal (Cohen – 't Hart 1967; 't Hart – Cohen 1973; 't Hart – Collier – Cohen 1990). The features were determined by the resynthesis of intonation patterns: in the analysis stylized patterns were created in such a way that they were simpler than, but perceptually equivalent to the original patterns. Thus "microintonation phenomena", that is fluctuations of pitch that are imperceptible and probably involuntary, were eliminated on the assumption that "the f_0 curves that do contribute essentially to the perception of the speech melody are just those changes that are programmed and voluntary executed by the speaker" ('t Hart – Collier – Cohen 1990: 40).

The smallest unit of perceptual analysis, in contrast to 'pitch-levels' models, is pitch movement, which can be divided into perceptual features with discrete values and precise phonetic definitions, along the dimensions of direction, timing, rate of change and size. In their analysis of Dutch, 't Hart, Collier and Cohen (1990) established five rising movements, marked with Arabic numerals, and five falling movements, which were labelled with capital letters. For example, movement 1 indicates an early, fast, full rise, whereas movement E an early, fast, half fall. The symbols can be combined to represent more than one movement occurring on a single syllable, e.g. A&2, 5&A.

't Hart, Collier and Cohen (1990: 78) formulated a set of restrictions which limit the possible configurations of pitch movements, e.g. "rise 1 can be followed by fall A or B but never by fall C." The configurations are grouped into the following classes: optional prefix or suffix, or obligatory root, together forming a contour. The prefix can be recursive. Figure 1.6 presents the contour of the sentence *The meeting has lasted three hours*.

De vergadering heeft drie uur geduurd

The meeting has lasted three hours
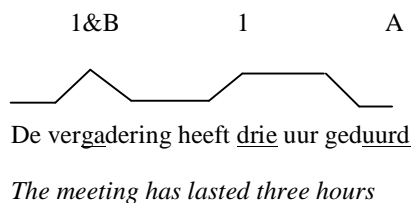
Figure 1.6 The IPO analysis of the intonational pattern (after Fox 2000: 286)

22

There are certain theoretical problems connected with the IPO approach due to which the model is not much different from other theories trying to find phonologically relevant categories for intonation. First of all, the model is said to be perceptually based, yet perception is determined by the phonological categories. Moreover, some pitch movements, even though perceptible, may not be phonologically significant (Fox 2000).

1.4.7. Non-linear approaches

More recent approaches interpret intonation as a non-linear phenomenon. The earliest works (Liberman 1975; Goldsmith 1976, 1978) introduced "tone melodies" similar to tone patterns found in tone languages. Their framework together with a new approach to stress (Liberman – Prince 1977), which contrasts stressed and unstressed syllables in an utterance, became the basis of the most influential models to date.

The standard generative model of intonation created by Pierrehumbert (Pierrehumbert 1980) assumes that the underlying components of tunes are pitch levels, namely high (H) and low (L) (Pierrehumbert, 1980; Pierrehumbert – Beckman 1986; Ladd 1996), since it is believed that the contours are "a phonetic by-product of the simpler tonal primitives, which are the real targets for pitch" (Roca 2000: 388). These primitive elements, referred to as tones, are assigned to an autosegmental tier, called the tone tier, by means of association[9] lines. The tone tier is distinct from the tiers housing segmental features, therefore association lines integrate tones to syllable head segments: nuclear vowels. Thus, an utterance traditionally represented as *undóubtedly* and *kangároo* can be transcribed as follows:

(2)   H   L   L H

           | | \/

      un.doub.ted.ly   kan.ga.roo   (Roca 2000: 389)
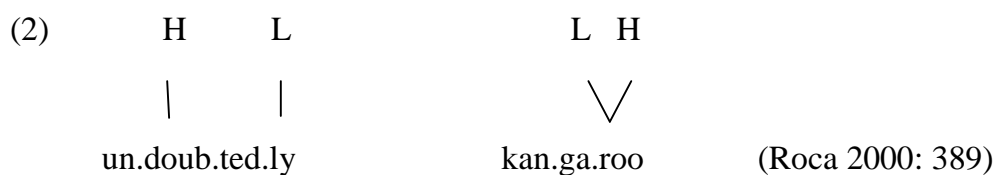
The first tone associates to the main stress and the second one – to the last syllable, while in the last word the tones converge on the last syllable. This does not mean, though, that only these syllables are intonated: those unmarked for tone will receive an intermediate pitch, mid-level (M). Moreover, the syllables between toned ones will reveal a gradual change of pitch from the tone on the left to the tone on the right, which is referred to as interpolation.

---

[9]Association – the abstract structural property of "belonging together" (Ladd 1996: 55). Association lines – lines connecting units on different levels (Crystal 2003). Cf. "tune-text association" (Liberman 1975).

The tones in the above examples are simplified; it should be noted that each tune contains three kinds of tone: the word tone[10] (Pierrehumbert's pitch accent) marked with an asterisk[11] (e.g. H*), the phrase tone (phrase accent) marked with '¯' (e.g. L¯), and the boundary tone marked with '%' (e.g. L%). Therefore, the examples in (2) can be rewritten in the following way:

(3)   H* L¯L%   L H¯ H%

             │   │ │   \\/

       un.doub.ted.ly   kan.ga.roo   (Roca 2000: 393)
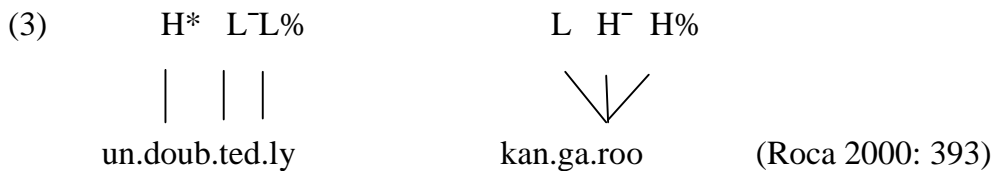
Pitch accents are the most important elements of intonational contours. In English (and Polish) they can comprise two tones, e.g. H*+L or H+L*. Pitch accents associate to stressed syllables of words, the boundary tone – to the edge of the domain[12] of intonational association and functions as a marker of intonational phrase boundaries, while the phrase accent links to the syllable shortly after the pitch accent, if such syllables are present. Otherwise, the phrase accent links to the same syllable as the pitch accent (Roca 2000: 387-392). Contrary to pitch accents, the edge tones, phrase accents and boundary tones, are single tones.

As far as utterances longer than one word are concerned, intonational representation can be shown as follows:

(4)   H*   H*   H*   H*   H*   H* L¯ L%

          │   │   │   │   │   \\/

      the dark clouds in the sky threaten imminent rain   (Roca 2000: 393)
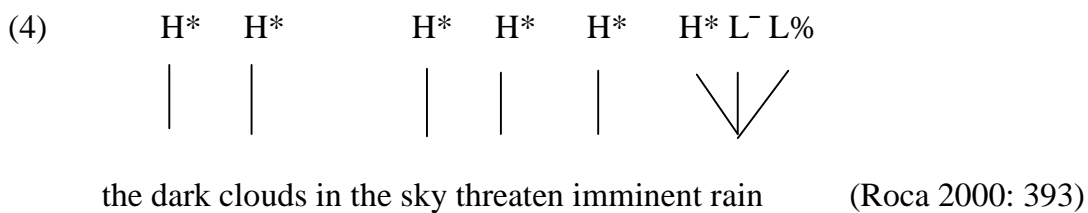
The nuclear tone is of the statement type (H*L¯L%) on the right, associated to a monosyllable, and H*s mark prenuclear word tones.

The structure of tunes can be summarised in the following rule (Figure 1.7) which could generate all possible English tunes.

---

[10]The term "word tone" will be replaced with the "pitch accent" in the present work.
[11](*) indicates the dominant component of the word tone, associating to the stressed syllable.
[12]For the explanation of the term "domain" see 1.5.1.
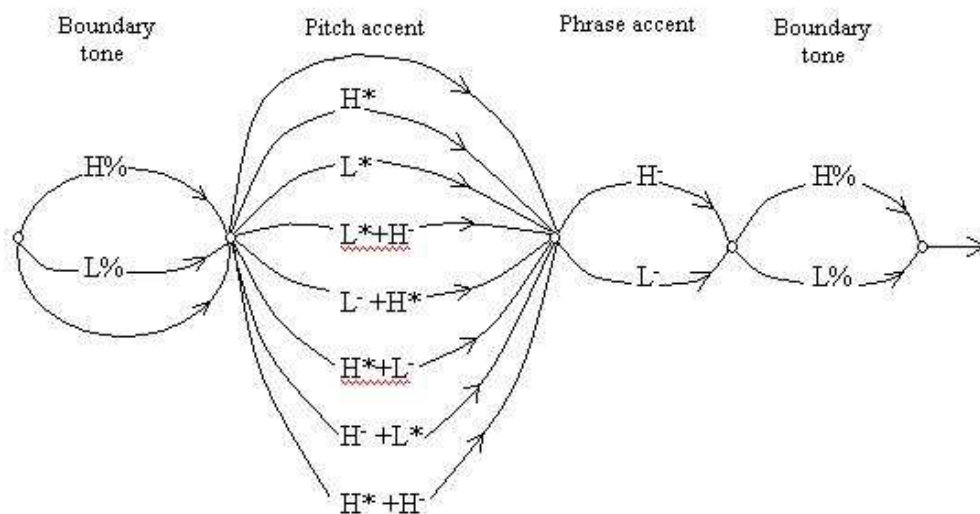
Figure 1.7   Rule generating legal tunes in English (Pierrehumbert 1980, after Ladd 1996: 81).

The diagram reveals that an intonational phrase consists of a sequence of tones, H and L. The intonational phrase consists of four components: a boundary tone at the beginning of the phrase, a pitch accent, a phrase accent and a final boundary tone. The tones marked with asterisks are central to the accent, the tones marked with '‾' represent 'leading' or 'trailing' tones. Phrase accents are located at the end of the word carrying the last pitch accent and indicate the movement in pitch from the last pitch accent. The final boundary tones (labelled %) account for the pitch movement on the last syllable of the intonational phrase.

It is possible to trace certain parallels between Pierrehumbert's nuclear accents and the structure of intonation in the British tradition (Figure 1.8).

However, as Ladd (1996: 82) noted, an attempt to find a complete correspondence between the two notations does not make sense; the two systems view certain phenomena differently, e.g. the notions of "heads" and "pre-heads" (discussed below). Moreover, inventories of tone-types are too divergent to be compared which is shown by the way the tones are grouped: in Pierrehumbert's system there are five sets of four types plus two more. As already mentioned, she rejected the notion of the nucleus[13] and any correspondences mentioned in figure 4 are ones between the sequences of the last accent, phrase tone and boundary tone, and British nuclear tones.

---

[13] Yet it cannot be denied that at least one of the pitch accents in an IP is more prominent than other pitch accents, so the rejection of the notion of nucleus cannot be justified.

Pierrehumbert	British-style
H* L L%	fall
H* L H%	fall-rise
H* L H%	stylised fall-rise
H* H H%	high rise
L* L L%	low fall
L* L H%	low rise (narrow pitch range)
L* H L%	stylised low rise
L* H H%	low rise
L+H* L L%	rise-fall
L+H* L H%	rise-fall-rise
L+H* H L%	stylised high rise (with low head)
L+H* H H%	high rise (with low head)
L*+H L L%	rise-fall (emphatic)
L*+H L H%	rise-fall-rise (emphatic)
L*+H H L%	stylised low rise
L*+H H H%	low rise
H+L* L L%	low fall (with high head)
H+L* L H%	low fall-rise (with high head)
H+L* H L%	stylised high rise (low rise?) with high head
H+L* H H%	low rise (high range)
H*+L H L%	stylised fall-rise ('calling contour')
H*+L H H%	fall-rise (high range)

Figure 1.8 Correspondences between Pierrehumbert's (1980) and British-style nuclear tones
(after Ladd 1996: 82)

Another fundamental difference between Pierrehumbert's and the British approaches is the way they treat intonational contours. The British school chunks the contour into parts: pre-head, head, nucleus and tail. In Pierrehumbert's framework, the division does not exist, since contours are decomposed into strings of accents. They comprise one obligatory accent and any preceding accents are identical:

(5)   L*   L*   L*   H*L L%

I walked home the whole night.

These preceding accents cannot be equated to heads, since they are not a global shape or even constituents; they are only a substring of the contour. What makes this analysis of intonation more relevant than the British one is that there seems to be no phonetic difference between nuclear and pre-nuclear accents in Western European languages. Ladd (1996: 211) claims, therefore, that there is no reason for continuing the traditional division. On the other hand, one may claim that the lack of nuclear and prenuclear accents in the model is simply the result of the terminology which Pierrehumbert introduces; moreover, it can be argued that the distinction should be maintained, as it proves to be helpful in marking certain pragmatic/discoursal meanings of intonation and in describing information structure.

Pierrehumbert's model bears more resemblance to the approaches in the American tradition. The term "pitch accent" is used to mark the interrelation between pitch and stress, like in Bolinger's analysis. It is also claimed that her model with only two levels allows to avoid the problems which Pike's four-level approach has to face, for instance the issue of relative and absolute pitch. As in the metrical theory of Lieberman and Prince where the levels of stress were not absolute but relative, the height of pitch in Pierrehumbert's system is relative to the baseline, the degree of prominence the speaker chooses to apply and to the preceding tones. Thus the problem of accounting for the fact that small pitch movements can convey significant differences in meaning is solved.

The weakness of the model is that it focuses on the description of a system that can generate typical English contours but disregards intonational meanings (Cruttenden 1997: 64-66). Pierrehumbert – Hirschberg (1990) deal with this aspect of intonation in their work on the interpretation of discourse (see section 1.4.8).

Pierrehumbert's analysis has been adopted to the description of German intonation (Féry 1993) and its conventions for transcription have been known as the ToBI (Tone and Break Index) system (cf. Beckman – Hirschberg 1994). ToBI gave rise to other transcriptional systems: ToDI for the description of Dutch and IviE (*Intonational Variation in English*), which was devised for the transcription of intonational variation in English dialects (e.g. Grabe – Post – Nolan 2001; Grabe 2004).

Modern phonological frameworks, such as metrical and prosodic phonology, provide insight into syntagmatic relationships within utterances: they no longer represent the phonological structure of a word or utterance in a linear way[14]; they assume that prosodic features are interrelated and interdependent, e.g. rhythm is related to prominence, prominence to accent which, like intonation, is based on pitch. Metrical phonology assumes that prosodic features share a common, multi-dimensional structure which is an abstraction in mental representations of utterances, but can be graphically represented by means of metrical trees. The frameworks utilise the notions of the syllable, the foot, the Intonational Phrase and the Utterance (Selkirk 1980, Nespor – Vogel 1986), as well as the phenomena of extrametricality. In autosegmental theory, phonological representations form a set of parallel and interactive tiers. The prosodic structure differs from language to language; along its universal features, such as the syllable, some characteristics are language-specific, e.g. various timing or the presence of feet[15]. It must be noted, though, that the concept of "structure" does not refer to the the Saussurean "system" in structuralist frameworks, but rather to structure as "organization".

Figure 1.9 represents one of the possible pronunciations of the phrase *Too many cooks spoil the broth*. This mental construct contains the most common set of prosodic constituents, although other elements, for instance the Intermediate Phrase, the Accentual

---

[14]Hence the name "non-linear" models. The approach began with Firthian prosodic phonology. Other non-linear models include autosegmental phonology or dependency phonology.
[15]Languages without Level 1 accentuation, i.e. without a string of accented and unaccented syllables forming the foot. Level 1 accentuation is often referred to as "word stress".

Phrase or a mora tier for such languages as Japanese, can also appear. The speaker, in the process of language-specific phonetic alignment, translates the tones into f_0 values (Gussenhoven 2002a). This model is referred to as Autosegmental-Metrical (Ladd 1996): autosegmental because of the separate tiers for segments and tones, and metrical because of the hierarchical organisation of the constituents.
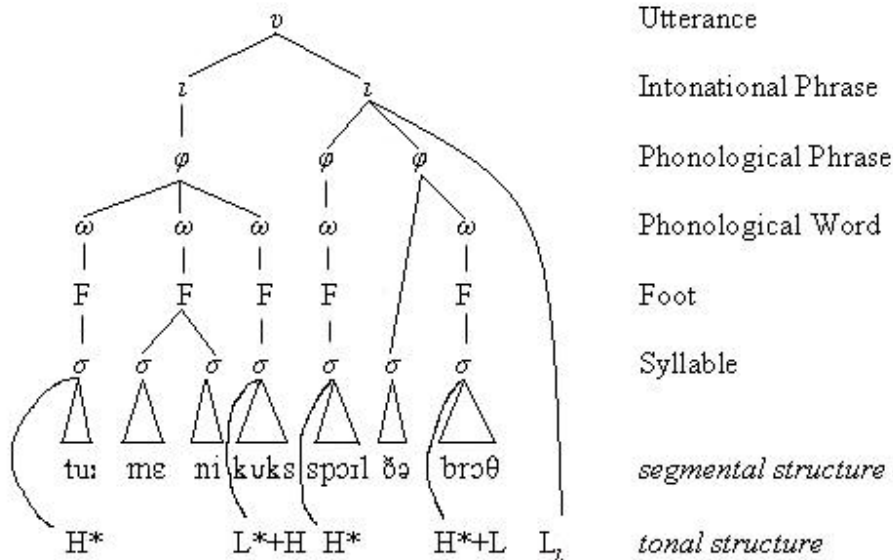
Figure 1.9 Surface phonological structure of a sentence *Too many cooks spoil the broth* (based on

Gussenhoven 2002a: 271)

The non-linear models are generative, with universal rules constrained by an independently motivated, language-specific structure. Thus, the task of non-linear models is to discover the nature of the structure of prosodic features and to describe their role. According to generative frameworks, prosodic structure can be understood either as resulting from "an elaborate conspiracy between phonological rules (Kisseberth 1970) in which formal processes lead to the same result;" or "as a complex set of output conditions or surface phonetic constraints on the rules of grammar constraining the rules to produce the desired outcome (Shibatani 1973);" or the structure could be interpreted "as a set of filters, allowing well-formed structures to pass but blocking ill-formed ones" (Fox 2000: 332). Optimality Theory (Prince – Smolensky 1993; McCarthy – Prince 1993; Archangeli – Langendoen (eds.) 1997) develops these principles. Hammond (1997) specifies the following constraints which take the place of rules: Syllable Licensing (grouping segments into syllables), the Peak Constraint (allowing for the existence of a peak in each syllable), the Complex Constraint (there should be only one consonant in the syllable margin)[16], the

---

[16]The Complex Constraint is by no means universal – it is violated by such languages as English and Polish.

Rooting Constraint (words must be stressed), the Trochaic Constraint (feet are trochaic) and the Parse-Syllable Constraint (two unfooted syllables cannot be adjacent).

The constraints are universal but not absolute, since they are regarded violable. They are ranked, so that some can take precedence over others, yet the ranking is language-specific. Only the forms which obey the constraints can be licensed by the grammar (Fox 2000: 332).

Since the description of a prosodic structure involves larger stretches of speech than the segment, a concept of a domain has been introduced. According to Crystal (2003: 148) the term indicates those parts of the tree diagram which derive from any one node; in other words, "the stretch of speech to which we assign a particular feature" (Fox 2000: 334) or "an array of elements under a common scope" (Roca 2000). As a consequence of syntagmatic relationships between features, the domain is dual in nature: one can distinguish the domain of application and domain of relevance to which features can be assigned, e.g. the domain of application of the feature accent is the syllable (it applies to the syllable) since accent is the syllable's phonetic property. At the same time, the domain of relevance is a larger constituent, the foot, since the contrasts relevant for the accent can be established there (Fox 2000: 334).

The phonological word (ω), which immediately dominates the foot, "represents the mapping between the phonological and the morphological components of the grammar" (Nespor – Vogel 1986: 109). Although morphological elements used to define it are language-specific, the following general definition can be given (Nespor – Vogel 1986: 141-142): "the domain of ω is terminal syntactic node (Q), or:

a) "the domain of ω consists of a stem, any element identified by specific phonological and/or morphological criteria, and any element marked with a diacritic [+W]";
b) "any unattached elements within Q form part of the adjacent ω closest to the stem; if no such ω exists, they form a ω on their own. In other words, the phonological word groups affixes with stems."

The definition indicates that in no language can there be a phonological word that is larger than the terminal element of the syntactic tree; moreover, there can be no more than one ω in a single stem.

The smallest phrasal phonological domain is the Phonological Phrase (φ)**,** identified with syntactic constituents, e.g. *fond of princess Diana*. Phonological phrase obligatorily comprises the syntactic head and non-recursive elements which are not themselves syntactic heads, e.g. *very fond of princess Diana, not very fond of princess Diana* (Roca 2000: 474). In English, it is the domain of application of two phonological rules: Iambic Reversal (e.g. *thirtéen mén → thírteen mén*) and the Monosyllabic Rule, reducing

monosyllabic (weak) words that do not belong to such categories as noun, verb or adjective (Nespor – Vogel 1986: 177-178).

A larger phonological domain is the Intonational Phrase (IP), the one where tunes associate to the segmental tiers. The IP is most often grouped on the basis of syntactic factors. Other criteria that specify the boundaries of IPs include semantic information related to prominence, and performance factors, i.e. speech rate and style. Nespor and Vogel (1986: 189) specify the scope of IP domain as the one consisting of "all the φs in a string that are not structurally attached to the sentence tree on the level of s-structure." The domain larger than IP is the Phonological Utterance (U), the domain where such phonological rules as *r*-insertion or flapping occur.

The phenomenon of stress belongs to the foot, a constituent which in unmarked cases comprises a strong (stressed) and weak (unstressed) syllable. Goldsmith (1990: 171) starts the organisation of the metrical tree from syllable rhymes which are grouped into feet, daughter nodes of the phonological word (see example 6 below).

(6)

```
                        word
                       /      \
                   w            s
                   F            F
                  / \          / \
                 s   w        s   w
                 |   |        |   |
                 σ   σ        σ   σ
                 |   |         \   \
                 A th a   b   as k an
```
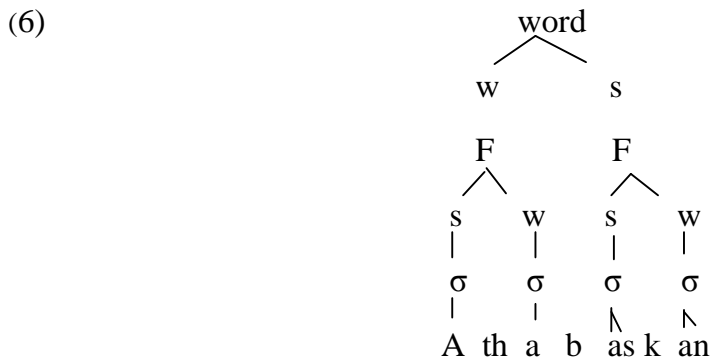
The feet are normally bisyllabic and can be either left- or right-headed (also called left- and right-dominant, respectively), i.e. either its leftmost or the rightmost rhyme is stressed and becomes the head. Moreover, feet can contain no more than two syllables (bounded feet). When a foot contain only one syllable, it is called degenerate (examples 7-9).

(7) Left-headed foot (trochee):

```
              F
            /   \
          s       w
          σ       σ
```
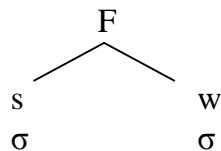
(8) Right-headed foot (iamb):

```
              F
             / \
            w   s
            σ   σ
```
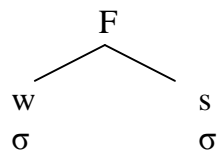
(9) Degenerate foot (marked):

```
            F
            |
            σ        (Archibald 1993)
```

The left-headed feet are also called trochaic or trochees, while right-headed ones are referred to as iambic or iambs (Goldsmith 1990).

The inventory of feet is different for each language. In quantity-sensitive languages (see section 1.6.1) long vowels or long vowels and closed syllables will not appear in weak positions. In the case of languages which possess extrametricality, the edgemost syllable will be ignored by stress assignment. Moreover, each language has words with exceptional stress pattern whose "foot structure is marked in its underlying representation. English is a quantity-sensitive, trochaic, right-to-left language with extrametricality and exceptions" (Gussenhoven 2002a).

The phonological[17] domains mentioned above include one another: the IP contains one or more phonological phrases, and the U comprises one or more IPs. The trees they constitute are non-recursive (Selkirk 1984, Nespor – Vogel 1986): no node may dominate another node of the same category, and the requirement is summarised in the constraint called Strict Layer Hypothesis (Selkirk 1984):

(10)   *Strict Layer Hypothesis*

There is a hierarchy of prosodic domain types such that, in a prosodic tree, any domain at a given level of the hierarchy consists exclusively of domains at the next lower level of the hierarchy.

Graphically such relationships can be represented as follows (example 11).

---

[17]Here *phonological* will be used for transparency, due to ambiguity involved in the term *prosodic*.

(11)
```
                 U
          IP    /\
         /\       \
        φ   φ      φ

        A   B      C
```
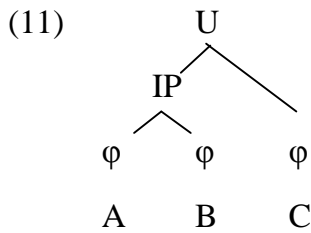

If the prosodic structure takes the form of a hierarchy of constituents, a question arises whether intonational domains are parts of the hierarchy. Gussenhoven (1988) claims they are not, as they cannot be mapped onto prosodic constituents in a consistent manner. As he claims, the location of the boundary tone is determined by the Association Domain.

As discussed in 1.4.1., the pitch accent associates with the accented syllable, while the second tone, namely the phrase accent, spreads. The last tone, the boundary tone, links to the rightmost domain boundary. As for locating the boundaries of the domain, it has been assumed that the Association Domain of intonational tones is a prosodic constituent, namely the IP or ω. However, there seem to be no correspondence between pausally and intonationally defined domains (Gussenhoven 1988: 91), which is the main weakness of the model. Beckman and Pierrehumbert (1986) studied the distribution of tones in the word *California* (example 12).

(12)
```
        L*H    L*H
        | |    |\
        C a l i  f o r n i a        (after Gussenhoven 1988: 91)
```
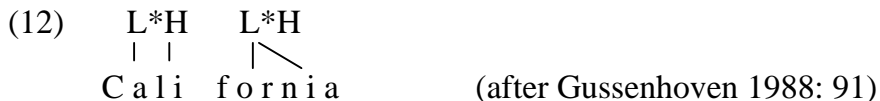

The Association Domain in (19) coincides with the foot; the evidence is supplied by the fact that word-final /i/-tensing (the *happy* vowel) does not occur in *Cali-*, so the domain applies to a smaller constituent. Even if additional elements are introduced, e.g. Pierrehumbert's Intermediate Phrase which is lower in ranking than IP, the relationship between intonational and prosodic constituents is not much clearer. The answer to the problem is that the prosodic hierarchy is independent of intonational domains – they are parallel to one another. The Association Domain begins at the accented syllable and ends at the highest dominating prosodic domain, but does not extend to the next accent (example 13).

(13)
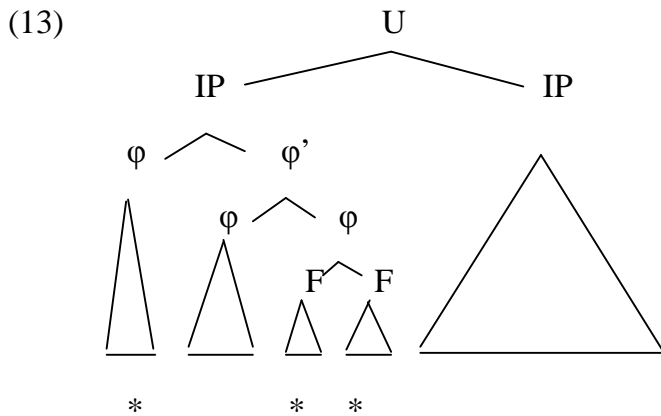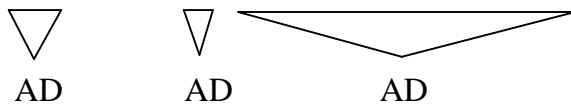
(Gussenhoven 1988: 93)

Although no one-to-one relationship occurs between the domains and the prosodic tree, the tree helps to establish the Association Domains (AD); if, for example, the first domain is extended to *wanted*, an IP boundary would fall after *Mum wanted*, producing an ill-formed phrase. Gussenhoven (1988: 94) thus concludes that "an Association Domain boundary reinforces the prosodic boundary it coincides with". However, as it has been admitted, intonational domains are independent from the prosodic hierarchy, because of their paralinguistic aspects, their limited grammatical functions, as well as attitudinal and discoursal meanings, therefore the prosodic structure may be excluded from a description of intonation.

From the above analysis one can conclude that the tonal structure parallels the segmental structure. However, the autosegmental tier differs from language to language. If tones specify words and morphemes, e.g. Japanese [ha!sõ'] 'chopsticks', [ha'sõ!] 'bridge', [ha'sõ,] 'end', or Mamvu [ma!a'ka'] 'type of seasoning' vs. [ma'a!ka'] 'cat', language belongs to the "tone language" type (Gussenhoven 2001: 15296-7). If tones do not specify morphemes, languages are referred to as "intonation languages". Those two terms do not seem fully relevant, since they suggest that tone languages do not have intonation. In reality, many of these languages use tone intonationally to indicate such discourse meanings as 'finality', 'continuation', 'question', etc. For instance, in Japanese there is a distinction between L-tone in statements from a H-tone in questions utterance-finally. Moreover, speakers of all languages will use pitch to convey universal meanings (for universal vs. language-specific meanings see also chapter 3). There also exist so called "pitch accent languages" which restrict the lexical tonal contrasts to one per word, e.g. the Dutch dialect of Venlo, Norwegian, Swedish, Central Franconian dialect of German, Limburgian dialect of Dutch, Serbocroatian dialects, Basque dialects, and Lithuanian (Gussenhoven 2001: 15296-7).

1.4.8. Discourse intonation

The term discourse, as Crystal (2003: 141) defines it, indicates "a continuous stretch of (especially spoken) language larger than a sentence... a set of utterances which constitute any recognisable speech event." It was not until recently that the study of intonation and discourse has been combined.

Brazil's (1975, 1978) theory of discourse intonation was developed for British English and, although his main concern was to describe functions of intonation without referring to grammar, his taxonomy can be seen as a simplification of O'Connor and Arnold's model. According to the theory, there exist a limited number of contrastive pitch configurations, each with its own meaning. It must be noted, though, that the term "meaning" does not indicate attitudinal or grammatical meanings but refers to interaction, i.e. the speaker's choosing a particular intonational pattern in order to achieve coherence and cohesion in the discourse. Thus, he uses traditional descriptive components to which he applies different meanings and functions.

He does not propose any novel phonological categories, yet uses the traditional notions of tones (Halliday 1963) and key: the general pitch of the sentence or sentence group (Sweet 1890). However, "tone" in his theory indicates the pitch change occurring on the tonic segment in a tone group, while "key" refers to the pitch level of the tone group, i.e. a unit smaller than the sentence. The tones he adopts are the following: falling-rising (referring) "r", falling (proclaiming) "p", rising (marked version of the fall-rise) "r+", rising-falling (marked version of the fall) "p+", and low rising (Brazil 1975).

To show that the grammatical structure of a sentence is irrelevant to intonational meaning, Brazil (1975: 6) gives two syntactically identical sentences (example 14).

(14)   1. // when I've finished $\vee$ Middlemarch // I shall read $\searrow$ Adam Bede //

2. // when I've finished $\searrow$ Middlemarch // I shall read $\searrow\nearrow$ Adam Bede //

1. \/ (fall-rise) = r (referring tone)

2. \ (fall) = p (proclaiming tone)

The tone group in the first sentence which carries the fall-rise (referring tone) contains information shared by both interlocutors. The sentence could be uttered in a situation when the listener already knows that the speaker is reading *Middlemarch* but does not know the speaker's plans for the future reading. The falling tone, by contrast, indicates new information which is "proclaimed". Therefore, in the second sentence the

listener knows what the speaker is going to read and the new information is when it is going to be read (after finishing *Middlemarch*).

The fall and the fall-rise tones possess their marked versions, the rising r+ and the rising-falling p+, which indicate the degree of the speaker's involvement. Compare the following pair of sentences as possible answers to the question *Where are the glasses?* (Chun 2002: 33).

(15) 1. // r in the CUPboard   \/ (fall-rise) = r (referring tone)

2. // r+ in the CUPboard   / (rise) = r+ (intensified referring tone)

The simple referring tone in sentence 1 means 'that's where they always are', while the intensified tone in sentence 2 implies 'why don't you ever remember...?' Capital letters indicate the tonic syllable where the major pitch movement occurs.

The functions traditionally associated with the rising tone are the interrogative ones. A question with a fall-rise (r) is understood as straightforward or neutral. The rising tone (r+) indicates a more insistent question. The choice of the rise-fall (p+) marks an utterance as doubly new, as in example 16, which can be interpreted as 'I also didn't know', 'I'm surprised/disappointed/delighted' (Chun 2002: 34).

(16) // p+ had he READ it? //   /\ (rise-fall) = p+ (intensifying proclaiming tone)

The low-rising tone meaning 'uninvolved', 'careful' or 'patronising' indicates neutral contexts in which the speaker does not proclaim or refer to anything. Thus discourse functions of intonation are interpreted along two independent dimensions: the refer/proclaim dimension and the involved/uninvolved one.

Another important choice to make is the selection of the key (relative pitch) of the tone unit. Brazil distinguishes three keys: high, low and unmarked mid. High key requires raising the pitch of the whole tone group, whereas low key involves its lowering as compared with the speaker's usual pitch level. High key indicates contrast, continuation or the beginning of a new topic. Low key signals the information already known, finality or the completion of an utterance. Mid key does not evoke any specific meanings. That is why the appropriate choice of utterance-final key is vital for an interaction, since it determines not only speakers' expectations but also turn-taking.

Johns-Lewis (1986) is one of Brazil's followers who investigated intonational cues for prominence. According to him, the markers for prominence include segmental

lengthening, which also marks sentence, paragraph and conversation-turn finality, and laryngalisation (creak) before a boundary and pause length indicating boundaries. Topic finality can be also signalled by sentence declination, i.e. the lowering of f_0 towards the end of an utterance. His and Brazil's findings show that intonation functions not only on the sentence level but also at the level of discourse.

Researchers involved in the description of discourse intonation for American English include Liberman – Pierrehumbert (1984), Ward – Hirschberg (1985), Hirschberg – Pierrehumbert 1986, Hirschberg – Litman 1987, and Pierrehumbert – Hirschberg 1990. Hirschberg and Pierrehumbert (1986) suggest a compositional theory of discourse intonation. The theory states that information about the attentional and intentional structures of discourse is revealed through such features of intonation as phrasing (the division of complex utterances), stress (the rhythmic pattern or the relative prominence of a syllable) or accent placement, pitch range (the distance between the highest and the lowest point in the contour) and tune (the abstract source of fundamental frequency patterns). It is the tune (intonational contour) that is the basic unit of analysis. Tunes communicate the relationships between the content of the subsequent utterances and the relationships between the utterance and the shared knowledge of the interlocutors. Pitch accents imply the status of discourse referents and the relationships between accented words. Information conveyed by phrase accents concerns the way intermediate phrases are related, while information provided by boundary tones specifies whether a particular intonational contour is "forward-looking" or not (Hirschberg – Pierrehumbert 1986: 308).

Selting (1988) in her analysis of German, points to the fact that intonation places utterances in context. She distinguishes two types of categories within the system: local categories with semantic functions, e.g. accents, and global categories which have interactive functions, e.g. contour types. She states that intonation is an important device in repairing conversations in a local problem-handling sequence and it influences speakers' cooperation in a global problem-handling sequence.

Yang (1995) shows the role of intonation in conveying emotions and attitudes in Mandarin Chinese, how it contributes to the development of the topic, the management of discourse, e.g. interrupting or agreeing, and signalling discourse structures. Her research reveals that new topics are marked by higher pitch which gradually diminishes towards the end of the topic. Yet if pitch ascends within a topic, the discourse structure undergoes the cognitive building up. She also observes pitch concord in subsequent utterances when the speakers' intonation patterns mirror each other (a similar phenomenon is described in Coulthard – Brazil 1981; Couper-Kuhlen – Selting 1996).

DuBois et al. (1992), although they do not refer directly to the theory of discourse intonation, set their framework of discourse transcription within general theories of intonation and discourse analysis. According to their model, the structure of a conversation includes the turn, which is its basic unit, the intonation unit (IU), and the word unit. The IU resembles Halliday's (1967) tone group, Crystal's (1969) tone unit and Cruttenden's (1997) intonation group, since it contains a prominent meaningful pitch movement placed on the

word with with the primary accent (cf. nuclear accent). They adopt the traditional British five tones: the rise (/), fall (\), rise-fall (/\), fall-rise (\/) and level tones (_). They also attempt to establish the prosodic cues which indicate boundaries between intonational units (Figure 1.10).

1. coherent contour: a unified intonation contour, i.e., one displaying overall gestalt unity
2. reset: a resetting of the baseline pitch level at the beginning of the unit
3. pause: a pause at the beginning of the unit (in effect, between two units)
4. anacrusis: a sequence of accelerated syllables at the beginning of the unit
5. lengthening: a prosodic lengthening of syllable(s) at the end of the unit (e.g., of the last syllable in the unit)

Figure 1.10 The five prosodic cues to IU boundaries (DuBois et al. 1992: 100)

Couper-Kuhlen and Selting (1996) attempt to deal with the issue of the lack of constancy between intonational form and meaning by means of an analysis based on the interaction of prosody with contextual and situational factors. Thus their approach can be contrasted with the structuralist or generativist frameworks, since, instead of finding a one-to-one correspondence between form and function, they base their model on authentic interactions. Intonation constitutes the part of inference-based pragmatic meaning with a contextualizing function: the speakers react to the particular pattern that is employed and cooperate to either avoid conflict or to resolve it.

1.5. Intonational meanings and functions

Recent theories of intonation assume that the melody of speech provides additional cues to the full meaning of utterances. Even though intonation is meaningful, the meanings are hard to define precisely and, as it was mentioned in the previous section, it should be analysed in terms of discourse functions rather than the propositional content of sentences. The functions of intonation are therefore frequently related to the speaker's 'attitudes' or 'emotions', while its grammatical functions or categories are generally disregarded by many theories.

Within each approach to intonation, various intonational functions are recognised, e.g. Crystal (1985) believes that the most vital aspect of intonation is its role of signalling the grammatical structure of sentences, that is contrasting sentence types as well as indicating clause boundaries. Halliday (1967) describes the way intonation shows informational structure, while the attitudinal meaning are the focus of O'Connor – Arnold's (1961) work. Brazil (1974) concentrates on the discourse-level phenomena. Couper-Kuhlen (1986) establishes six of them for English: informational, grammatical, illocutionary, attitudinal, textual/discourse and indexical. Chun (2002) groups the functions into four more general, pedagogically-oriented categories: grammatical, attitudinal,

discourse and sociolinguistic, following learning curricula. None of the authors has managed to avoid certain problems connected with the grouping: categories often overlap, e.g. for some researchers the contrast between new and old information is grammatical in nature, for others it belongs to the discourse. Yet the following description of intonational meanings and functions presents Chun's categorization because of its practical, pedagogical orientation.

1.5.1. Grammatical functions

Establishing grammatical functions of intonation is frequently problematic. First of all, theories emphasizing the link between intonation and grammar can be placed along a continuum, depending on the degree to which the two systems are interrelated. Bolinger (1958: 37), for instance, feels that the relationship between intonation and grammar are "casual, not causal", whereas Halliday (1964) takes the opposite view. The theories on one end of the continuum divide intonational patterns into two tunes which contrast questions and statements, while theories on the other end present how nucleus placement and tonal contrasts mark grammatical functions (Halliday 1963). Furthermore, there is no agreement on what elements constitute the grammatical functions, namely whether one should be concerned with the sentence level or rather on the discourse level phenomena which is the cause of the previously mentioned overlapping of the traditional grammatical and discourse functions of intonation.

At the sentence level, intonational patterns are associated with the following sentence types: statements, yes/no questions, *wh*-questions, commands and exclamations. Smaller units, such as clauses, adverbials, noun phrase subjects, topicalised subjects, parenthetical clauses, vocatives, nouns in apposition and pairs of clauses are also related to particular intonation groups (Cruttenden 1986, 1997).

Associating tunes with sentence types is not always straightforward, since it is difficult to establish the 'marked' and 'unmarked' or 'neutral' tones for sentences. One can consider a rising tone for yes/no questions as neutral, but it is not clear if the 'polite' tone is more unmarked than 'businesslike' one. Recent studies reveal that the choice of the pattern depends more on the type of situation than the actual sentence (Chun 2002).

Apart from associating tones with sentence types, intonation marks the boundaries between and within utterances. Crystal (1969) distinguishes two phonetic factors marking a tone unit boundary: perceivable pitch change and the presence of junctural features at the end of a tone unit which is usually perceived as a pause. However, drawing a clear boundary between intonational groups is not always easy. Cruttenden (1986: 43-44) presents three types of problematic pitch sequences in English.

The first type of problematic pitch sequences concerns a sequence with sentence adverbials, e.g. *He went away unfortunately*. It is not obvious whether the sequence should

consist of one or two intonational groups: if there is a pause, anacrusis on *un-,* lengthening on *-way*, or if *un-* is pronounced at a higher pitch than *-way*, there would be two groups. However, it is likely that none of these criteria would be present and the sequence would be judged to consist of one group only.

The second type involves vocatives and reporting clauses sentence-finally, e.g. *Get a move on, you stupid fool* or *I'll prove you wrong yet, I heard him say.* These should be associated with separate intonational groups, as frequently vocatives and reporting clauses follow a clear pause. However, since they carry no pitch accent, these sequences form one group, regardless of the presence of a pause.

The last type of problematic sequences includes adverbials pronounced on low pitch which may belong, from the semantic point of view, to either the preceding or the following intonational group, e.g. *He went to the States of course he didn't stay very long* (deliberate lack of punctuation after Cruttenden 1986: 44). In such cases additional cues, such as the relative pitch of the unaccented syllables, usually occur.

The examples mentioned above reveal that it is perhaps more relevant to consider intonational groups as independent of syntactic constituents, which in turn proves that finding a link between grammar and intonation is indeed complicated.

1.5.2. Attitudinal/emotional functions

Attitudinal function of intonation is generally considered to belong to its paralinguistic aspect. As Lieb (1980, after Couper-Kuhlen 1986: 174) says, "An angry person does not raise his voice in English or German but simply in anger," so it is necessary to establish the extent to which these functions are part of the linguistic system.

Attitudinal functions have been described in a variety of ways. Sweet (1890) was not very consistent with terms assigned to his five tones, as he used 'attitudinal' labels, such as 'doubt', 'caution', 'warning', 'dogmatic', 'assertion', etc., and labels that could be considered 'grammatical' or 'discoursal', e.g. 'interrogative', 'contrastive', 'expectant', etc. O'Connor and Arnold (1961: 2) admit that intonation is used to express "the speaker's attitude to the situation in which he is placed." For Pike (1945), intonational meanings complement lexical meanings.

Bolinger (1986) based his theory of intonation on the belief in the interrelatedness between intonation and emotion. As he states, facial expressions and body gestures not only match the intonational configurations but they also "operate much of the time in parallel" (Bolinger 1986: 337). However, his assumptions are questioned (Ladd 1990), for if they were valid, the link between intonation and emotions should be universal. According to Ladd (1990), even though intonational contours tend to be similar in many languages, certain prosodic aspects, such as accent placement, remain language-specific; therefore the universality of intonation and emotions is not proved beyond all doubt.

1.5.3. Discourse functions

The discourse functions of intonation refer to the functions which influence continuity and coherence within discourse, such as those marking prominence, focus or the type of information. Other functions include those marking boundaries between sentences, paragraphs, topics and conversational turns. Discourse functions allow to control interactive structure, i.e. replying, continuing a topic or establishing a new one. They help interlocutors to cooperate in developing the discourse, indicate speaker's intentions and expectations, as well as the knowledge shared by both the speaker and the hearer (Chun 2002). Chun classified these functions under the following subcategories: information structure marking, illocutionary/speech act, textual/discourse and interactive/discourse.

The first discourse function of intonation to be discussed here is the one marking information structure. The theory of the division of utterances into parts ("chunks") conveying given and new information originated with the Prague School linguistics and Halliday (1967). Halliday states that each message constitutes a unit of information which can be independent of grammatical clauses but coincides with tone groups. Thus, the number of units of information, or tone groups, depends not on constituent structure but on the information content of the utterance. The units are built on at least one 'point of prominence' which forms the information focus of the unit and "reflect the speaker's decision as to where the main burden of the message lies" (Halliday 1967: 204).

Information focus within a unit of information is realized by assigning prominence in the tone unit, namely by the tonic segment: the nucleus and tail. If a tone group has one focus, the nucleus is simple, but if there are two foci, it is compound, which the following examples illustrate.

(17)   (a) there's a MÀN in the garden //

(b) there's a MÀN // in the GÀRden //

(c) there's a MÀN in the GÁRden //

(d) there's a MÀN [in the GÀRden] //

(Crystal 1975; after Couper-Kuhlen 1986: 123)

The element bearing information focus constitutes information new for the hearer, i.e. information which is factually new or which is not "recoverable from the preceding discourse" (Halliday 1967: 204). In contrast, given information indicates what is recoverable from the preceding discourse or situation or what the speaker believes the listener already knows. Although the given-new dichotomy does not directly depend on grammar, there exist some preferences. For instance, new information is usually conveyed by open-class lexical items, whereas given information is typically carried by grammatical items, such as pronouns, deictic adjectives, adverbs, definite articles, pro-forms, etc. As

for intonation, the distinction is seen in the use of pitch height: high pitch indicates new information, low pitch reflects given information (Couper-Kuhlen 1986).

New information, according to Halliday (1967), is used to contrast with the preceding discourse. This contrastive function can be considered grammatical, though, since it refers to assigning the information focus to the closed-class grammatical items, which normally convey given information, e.g. *HE'S done it* ('not someone else'). Chafe (1976) claims that contrastive information belongs to a separate category, as it is qualitatively different from new items. Both functions are signalled by high pitch, but the exact pitch configurations differ. For instance, the sentence with two foci, *I brought Matthew a book*, has the following patterns.

(18)

(a) I brought MATTHew a BOOK. (new information)

(b) I brought MATTHew a BOOK. (contrastive information)

(after Couper-Kuhlen 1986: 126)

Non-contrastively, both *Matthew* and *book* are relatively high-pitched, with the high pitch beginning on *Matthew* and continuing throughout the rest of the utterance. In the contrastive version, *Matthew* is characterised by a partial drop in pitch which later rises on the word *book*.

Another type of discourse functions of intonation are the illocutionary functions, connected with the Speech Act theory (Austin 1962), according to which utterances are produced in given situations for specific purposes. Acts of speech are communicative activities (a locutionary act) which refer to the intentions of speakers (an illocutionary act) and the effects they have on the listeners (a perlocutionary act). Illocutionary acts have been grouped into the following categories (Crystal 2003: 427):

a) directives, which attempt to get the listeners to do something, such as requesting, commanding or begging;
b) commisives, with which speakers commit themselves to a future course of action, for instance promising, threatening or guaranteeing;
c) expressives, which express the speaker's feelings, such as apologising, welcoming or thanking;
d) declaratives, or utterances which introduce a new external situation, e.g. christening, marrying, resigning;
e) representatives, with which speakers convey their belief about the truth of a proposition, such as asserting, describing, complaining or concluding.

The illocutionary force of utterances is expressed by such factors as syntax and/or lexis, but intonation can also mark illocution. Couper-Kuhlen (1986: 163-164) cites two theories on how the two phenomena are related.

The one-to-one hypothesis states that every distinct illocution corresponds to a distinct intonation marking (Sag – Liberman 1975). Its strong version claims that a particular intonational pattern always occurs when a given illocution is present and never occurs with any other illocution. According to a weaker version, a particular intonational marking is possible in the presence of a given illocution and it can occur elsewhere too. For instance, the illocutionary act of contradiction can be matched with a particular contour, a 'contradiction' one. The contour will always be appropriate regardless of the lexical content of the utterance. However, the theory does not hold for illocutions which are marked performatively, e.g. a questioning contour of *Did you see the murderer?* is out of place in a sentence with the performative verb *ask*: *I (hereby) ask you if you saw the murderer*. On the other hand, the disambiguating hypothesis postulates that intonation can mark illocutionary force only if there exist no explicit performatives. Sag – Liberman (1975) in their analysis of the intonation of indirect speech acts involving interrogatives concluded that in the illocutionary force of suggestion the contour is different than in the same utterance used as a literal question. Thus, "some intonational contours can 'freeze' an utterance pragmatically, i.e. require a literal interpretation, but no intonation can force an indirect interpretation" (Sag – Liberman 1975: 496; after Couper-Kuhlen 1986: 166). However, it has been demonstrated that intonation can indeed force an indirect interpretation (Cutler 1977) which makes the disambiguation hypothesis untenable.

Another subtype of intonational discourse functions is the textual/discourse role of intonation which relate to the content of the speech or discourse. Text refers to any passage, either spoken or written, which constitutes a unified or organized whole (Halliday – Hasan 1976). Texts comprise paragraphs, i.e. conceptual units which are organised around topics. The paragraph is said to be equivalent to the major paratone (Yule 1980; after Couper-Kuhlen 1986) whose beginning is marked obligatorily by high key and its end is signalled by very low pitch and an optional pause. Apart from major paratones, minor paratones can occur in a text. These are units which begin with an unspecified pitch height, i.e. they can begin with either high, mid or low pitch. On the other hand, the final boundary of a minor paratone is specific and obligatory, since it must end with a very low pitch, usually near the bottom of the speaker's pitch range (Couper-Kuhlen 1986).

Textually, minor paratones perform various functions: those which begin high indicate major paratones, thus marking new topics.Those which begin mid signal paratactic additions to or extensions of what precedes them. Finally those which begin low imply hypotactic subordinations to or inclusions in what precedes (Couper-Kuhlen 1986: 193).

The function of intonation as a consolidating device in a chain or sequence of tone groups (compound tone groups) operates through two principles of alternation and succession. The former relates to the alternation of two subsidiary events, e.g. alternating a rise, which demands a 'resolution', with a fall, which provides the resolution and/or

termination, in a compound unit. The latter relates to the creation of a cohesive unit by a succession of two or more events, e.g. either by the reduplication of identical intonational patterns or by the repetition of similar pitch movements in two or more successive tone groups (Couper-Kuhlen 1986).

In phonological texts, intonation helps to establish focus, usually associated with the accent of the greatest pitch height and/or pitch range in a minor paratone. It can indicate the assumed shared knowledge between the speaker and the hearer (falling tones) distinguishing it from new information, depending on the preceding discourse (Couper-Kuhlen 1986).

The textual/discourse function of intonation indicates boundaries within discourse. According to Oreström (1983) the end of turn coincides with the end of the tone unit (TU) carrying a non-level nuclear tone. He claims that prosody, i.e. the end of the TU, syntax, i.e. the end of a sentence, and semantics, i.e. the end of a fully comprehensible stretch, form a major juncture in English. DuBois (1992) adds two more features that are the secondary cues for juncture, namely loudness reduction and pause.

Another textual/discourse function of intonation is marking transitional continuity between intonational units. He distinguishes 'final continuity' which refers to the intonational contours which appear finally in a language, e.g. a fall in English, and 'continuing continuity' referring to the contours which mark continuation, e.g. low rise in English (DuBois 1992).

The final subcategory of intonational discourse functions includes the interactive/discourse ones. They are related to the conversational structure of the discourse, particularly to turn-taking. Interlocutors use intonation to continue with an established topic, to start a new topic, to constrain a hearer to reply, to discourage a hearer from replying, to express a speaker's expectations about a hearer's reply, to show cooperation and/or compliance with the discourse pattern and to facilitate repair when breakdowns or misunderstandings occur (Chun 2002: 64).

As Brown et al. (1980) states that the beginning of a new topic is indicated by a relatively high pitch range. The other speaker, in response to the first utterance, employs a higher pitch too. If the speaker asks a question about the already mentioned topic, s/he can use low pitch and the reply to this question can also be characterized by a low pitch range. To conclude the topic, the speaker's pitch range and pitch amplitude drops, and a pause follows.

Intonation in tag questions reflects the speaker's wish to dominate the hearer. These types of questions can take either a high fall or a low rise, each indicating different meanings. Falling intonation presupposes a slight possibility of the hearer's consent, while a rising contour presupposes a positive reply (Cruttenden 1997).

Prosody can be used to repair errors that could disrupt communication by directing the hearer's attention to them. The errors can refer not only to the lexis but also to the interpretation of the utterance, its semantics and pragmatics (Chun 2002).

1.5.4. Sociolinguistic functions

Sociolinguistic functions of intonation (relating to Couper-Kuhlen's (1986) indexical function) allow to identify speakers in terms of different social groups, according to their sex, age, region or occupation. Intonation also distinguishes the idiosyncratic dialect of the speaker. So far there is a dearth of research devoted to intonation phenomena in dialects of English (e.g. Cruttenden 1997; Lowry 1997; Grabe 2002; 2004; Grabe, Post and Nolan 2000, 2001; Grabe, Kochanski and Coleman 2003) and almost none for Polish (e.g. Dukiewicz 1978; Steffen-Batogowa 1996; Karpiński 2002; Karpiński 2006; for Poznan variety: Grabe – Karpiński 2003; Witaszek-Samborska 1987, 2006).

Grabe et al. (2005) investigate intonation in urban accents of English, namely RP, the accent that is most often taught to foreign learners, as well as the two northern accents of Newcastle and Belfast. For statements, in Cambridge and Newcastle data, no unique patterns were found. The most popular patterns included a fall preceded by one or more prenuclear falls: H*L (H*L) H*L%, or by one or more high prenuclear accents: H* (H*) H*L%. Belfast intonational patterns differed considerably from Cambridge statements, yet share the rise-plateau contour with the Newcastle speakers. In *wh*-questions, Cambridge and Newcastle varieties revealed that the most frequent pattern was the 'flat hat' ('t Hart, Collier and Cohen 1990), also found in Dutch and German. The rise-plateau contour was the most popular pattern in Belfast English. Polar questions are characterized by a greater variety of patterns. Cambridge speakers produced a nuclear fall in most questions, however, rising contours were used too. Belfast and Newcastle shared the most frequent rising contour. Echo questions also revealed variation: a high accent and a rise-plateau predominated in Belfast and Newcastle and a rise in Cambridge data.

1.5.5. Paralinguistic meanings

Gussenhoven (2002b), following Ohala (1983, 1984, 1994) claims that intonational meanings, both universal and language-specific, belong to two components of language, the phonetic implementation and the intonational grammar. Universal meanings which are "based on metaphors of biological conditions" or f_0 (2002b: 47) are expressed by the phonetic implementation, whereas the intonational grammar is the domain of intonational morphology and phonology. He calls the three metaphors "biological codes" which are determined by the biology of human vocal tract. First of all, male and female speakers, as well as adults and children, differ in the sizes of their larynxes which influences

fundamental frequency of their speech. Secondly, speech can be produced with different amount of energy which is detectable in the signal. Finally, the energy used by the speaker coincides with the breathing process. These facts relate to the three codes, the Frequency Code, the Effort Code and the Production Phase Code, or simply Production Code, respectively. The natural meanings the codes include are often grammaticalized, however, as a result of language change, the relations between the meanings and the form they receive can become arbitrary. Moreover, languages differ with respect to the way they employ the meanings included in the codes.

The universal intonational meanings contained in the Frequency Code relate to power relations: smaller larynxes produce higher frequencies and high pitch sounds sound 'feminine', submissive and vulnerable. Therefore utterances that are produced at higher pitch or the ones that end high sound uncertain, depending, appealing, but also friendly and polite. On the contrary, larger larynxes produce lower pitch vocalizations which indicate larger or aggressive individuals. Thus low pitch is perceived as having 'masculine' values; it will sound dominant and protective, while low and low-ending utterances will be perceived powerful, assertive and authoritative. Other interpretations include 'uncertainty' for higher pitch and 'certainty' for those lower, as well as 'questioning' vs. 'asserting'.

The relations indicated by the Frequency Code are supported by experiments with artificial intonation contours of a phrase *for Jane* (Gussenhoven 2002b) in which Swedish and American subjects were to decide whether the contours referred to contours or questions. The contours comprised a single falling-rising peak on *Jane* but differed in peak height and end pitch. The experiment revealed that higher peak was equated with 'Question'. However, native language of the subjects had considerable influence on their answers: Swedish listeners showed greater variability when judging superhigh and high peak, contrary to the Americans because Swedish uses final rises as cues for questions in a different way.

Most languages grammaticalize the informational use of the Frequency Code, since more than 70% of them possess rising contours but rarely for statements. Arbitrary form-function relations that seem 'unnatural' do appear, though. In the case of interrogatives, they include falling intonational contours and, rarely, rising contours for declaratives, as in Chickasaw where the interrogative is H* L%, the declarative H* H%. In the dialect of Roermond falling intonation for questions could have developed as a result of the introduction of a lexical tone in order to preserve the contrast under the interrogative intonation. On the other hand, rising statements can result from truncation[18] of delayed peaks (Gussenhoven 2002b: 50).

The Effort Code relates to the amount of energy expended on speech production. This amount can be varied: if the speaker puts in more effort, it will lead to more canonical and more numerous pitch movements. Meanings that are derived from this phenomenon are, for instance, signalling emphasis: the speaker uses greater force because s/he believes

---

[18]The process of word shortening which is phonologically predictable; truncation can illustrate such processes as template-mapping and prosodic circumscription (Crystal 2003: 477).

a given piece of information is important. Narrow pitch range is interpreted as negation or withdrawal of information. Affective meanings associated with the Effort Code are 'surprise', 'agitation', 'obligingness', or 'helpfulness' – like in speech directed to children. Perception experiments (e.g. Rietveld – Gussenhoven 1985) prove that higher pitch is associated with prominence.

The Effort Code is commonly grammaticalized in the expression of focus in which wide range pitch movements mark focused information. Thus pitch accents associate to syllables in focused parts of the utterance but are removed from the syllables after the focus. Focus, however, is mediated through grammar[19], hence its placement is constrained by language-specific rules (Gussenhoven 2002b: 54).

The Production Code encompasses the generation of energy that is tied to the breathing process, and which is available in phases (cf. Lieberman's breath groups). As a result of the gradual fall in energy towards the end if an utterance, a gradual drop in fundamental frequency, or declination, ensues. In this code high pitch is linked to the beginnings of utterances and low pitch to the ends, e.g. when signaling the end of the speaker's turn in conversation. In addition, high beginnings indicate new topics, while low beginnings signal continuation of topics. As for utterance endings, the opposite holds true: high endings indicate continuation and low endings finality. Common grammaticalization of this phenomenon is found at the ends of utterances: a H% signals continuation, although an initial %H can also signal topic refreshment. The gradual fall in fundamental frequency is grammaticalized as downstep but it can be phonological, hence meaningless (Gussenhoven 2002b: 55).

The physical conditions for the meanings described above do not have to be mentioned to produce the forms. Since the speakers know the form-function relations, they may utter substitute phonetic forms that resemble the targets. First of all, peak delay can be used instead of high pitch, since it takes longer to reach a higher pitch peak than a lower one. Therefore, the meanings that are signalled by high pitch can also be associated with pitch delay. Moreover, the Effort Code makes late peaks be perceived as more prominent. On the other hand, high pitch can substitute wide pitch span. Indeed, studies of a variety of languages seem to prove this. For example, in the Hamburg dialect of German narrow focus, the scope of the Effort Code, is expressed by later peaks which suggests that they are used to signal high pitch (Peters, in press). Substitute variables of the Frequency Code are exemplified by southern varieties of Italian which use a later peak to indicate interrogative intonation (Grice 1992). Japanese categorizes delayed accentual peaks as features of female speech. As for the Production Code, in British English first peaks of intonational phrases including new information were later than other peaks (Wichmann 2000, Gussenhoven 2002b). Figure 1.11 presents the universal and linguistic interpretations of the biological codes.

---

[19]Cf. Selkirk 1995.

Physiological sources	Biological codes	Universal interpretations	Linguistic interpretations
SIZE	**Frequency Code**: small – big → high – low	**Affective**: *submissive – authoritative vulnerable – protective friendly – not friendly* **Informational**: *uncertain - certain*	?
ENERGY (phasing)	**Production Code**: beginning – end → high – low	**Informational**: At beginning: *new topic – continued topic* At end: *continuation – finality*	Question vs. Statement (e.g. H% / L%)
ENERGY (level)	**Effort Code**: less effort – more effort → smaller excursion – greater excursion	**Affective**: *less surprised – more surprised less helpful – more helpful* **Informational**: *less urgent – more urgent*	Polar onset tone (%T) Focus (various)

Figure 1.11 Physiological sources (column 1) of the three biological codes (column 2), some of their interpretations (column 3) and examples of their grammaticalizations (column 4) (Gussenhoven 2004: 95)

1.6. Intonation in English and Polish

The number of possible pitch accents listed for English differs from author to author. Pierrehumbert in the revised analysis enumerates six of them: H*, L*, L+H*, L*+H, H+L*, H*+L. Grabe and Karpiński (2003), on the other hand, list as many as thirteen, including combinations with the mid-level tone and tritonal accents: HL, ML, LL, LH, LM, MH, HH, HM, LHL, LHM, HLH, MLH, MHL. For Polish, Jassem (1961) established a model based on the British school which decomposes intonation into three tones, L, M and H. Demenko (1999) distinguishes five tone heights, xL (extra low), L, M, H and xH (extra high), two prenuclear accents (H, L) and nine nuclear ones: HL, LH, LM, ML, HM, MM, MH xL and LHL. According to Grabe and Karpiński (2003), Polish has a smaller range of nuclear accent types: only six of them, namely HL, ML, LL, LH, LM, MH. Moreover, the rising tone is used in fewer contexts than in English.

The distribution of the tones in the two languages differs also in the absence of tritonal accents in Polish: if they occur, they appear in emotional speech, though rarely. The most frequent final tone in Polish declaratives is ML, while Cambridge English declaratives generally end in HL. Southern varieties of English use predominantly a fall-rise. In rising declaratives English employs a fall-rise and Polish a rise (Grabe and Karpiński 2003).

In both languages *wh*-questions end in falling or rising contours, which variation may be explained by individual speaking styles or various interpretations of utterances (Grabe and Karpiński 2003). According to Steffen-Batogowa (1996), emotional load of the

message can account for such variations in Polish utterances. Karpiński (2002) found two factors that are responsible for the choice of rises in Polish statements, that is "a speaker's uncertainty about what s/he is saying or is about to say" and "a speaker's intention or wish to continue speaking" (Grabe and Karpiński 2003: 1063).

The majority of yes/no questions in Polish end in LH. As Durand et al. (2002: 257) state, in *czy*-questions there is "a double climbing binary foot at the beginning and end of the sentence". Declarative questions in Polish tend to finish with LH or MH, while in Cambridge English the most frequent tone is LH, although HL can also occur.

Grabe and Karpiński's study suggests that in English the widest range of contour types occurs in declarative questions, in Polish it is the opposite: only two contour types are noted. The widest range of contours for Polish can be observed in declaratives which reflects differences in grammatical structure, especially word order: since the sequence of words in Polish sentences is relatively free, Polish speakers may rely more heavily on the use of high pitch in distinguishing interrogatives from other sentence types.

Another difference between English and Polish intonational systems is seen in the timing of peak alignment. Grabe's (1998) contrastive study of English and German reveals that the f_0 peak in English tends to be aligned with the right edge of the stressed syllable. The few studies into peak alignment in Polish (Olivier et al. 2005; Demenko et al. 2007) suggest that it occurs early in the stressed syllable.

CHAPTER 2: Acquisition of L2 intonation

2.1. Introduction

Most of SLA research into phonology has concentrated on segmental phonology rather than the prosodic aspect of language, even though it is intonation that contributes considerably to the so-called 'foreign accent' to such an extent that accurate pronunciation of second language (L2) sounds can be overridden by the faulty choice of pitch patterns (Carmichael 2000).

Recently, interest in the development of prosody during the acquisition of L2 is growing due to a number of reasons. First of all, evidence from studies on first language acquisition reveals that segmental phonology begins to develop later than prosody in infants; therefore a question arises whether there exist separate developmental constraints on segmental and suprasegmental levels of language. Secondly, technical progress in speech science made intonation a vital subject of study. It is hoped that increased knowledge in this aspect of language will lead to the improvement of automatic speech synthesis, speech recognition, language identification and speaker recognition, and human-machine dialogue systems. Finally, previous interest in syntax and its development shifted to the process of communication as a whole, in which intonation plays a crucial role (Vaissière 2005).

2.2. Factors affecting the acquisition of L2 speech

2.2.1. Age of acquisition

Since Lenneberg's (1967) postulated the hypothesis of a critical period (CP)[20] for language acquisition beginning from about age 2 to puberty, that is, the age when the neural plasticity is lost, the question how age influences acquisition has been open to debate[21]. In the case of phonology, a large body of research (e.g. Scovel 1969, 1988; Birdsong 1999; Flege et al. 1995; Piske – MacKay 1999; Carmichael 2000), also studies into bilingual acquisition (e.g. Bergman 1976; Major 1987), suggests that in order to achieve native-like

---

[20]Later, after evidence to the contrary appeared, called 'sensitive period', i.e. a period of time during which mastering a language is most efficient (Lamendella 1977).

[21]It is obvious that a critical period exists for first language acquisition, in the case of second language acquisition, though, its existence is controversial. Many researchers accept now a weaker version of the Critical Period Hypothesis, i.e. the earlier SLA begins, the better (Piller 2002).

pronunciation, a learner must be exposed to L2 before the critical period, or complete biological maturation, finishes.

Research into the subject reveals that there exist independent critical periods for L2 phonology and syntax which is proved by such phenomenon as the Joseph Conrad effect (Scovel 1988) or Henry Kissinger effect (Brown 1987, after Piller 2002). In this case, late L2 learners achieve a native-like command of the L2 but only in the areas of syntax, morphology and lexicon, yet they retain a foreign accent. Seliger (1978) proposed that several critical periods should be recognised, each for a different aspect of language. Other researchers, e.g. Patkowski (1980), agree that puberty marks the age until which good syntax can be achieved. Puberty may be the borderline for pronunciation as well: at about age of 15 the ability to acquire L2 pronunciation diminishes (Patkowski 1990). Flege, Munro and Mackay (1995) found a correlation between foreign accent and the age of arrival in the L2-speaking country. In her paper Ioup (1984) proved that pronunciation is acquired independently of syntax: for native English speakers it was easier to identify non-native users of English by their speech but not their writing. Similar results were obtained by Scovel (1988).

Both short-term (laboratory) and long-term studies have showed superiority of younger learners over older ones[22]. Tahta, Wood and Loewenthal (1981) investigated how five- to fifteen year-old English school children imitated French and Armenian pronunciation. Their findings revealed that this ability declined with age: five- to eight-year-olds replicated foreign intonation accurately, yet in eight- to eleven-year-old children the ability dropped considerably. The conclusion may be that for suprasegmental phonology age-related constraints begin to be set at about age of six, which is earlier than for segmental phonology.

Many researchers (e.g. Scovel 1988) suggested that the critical period for the acquisition of phonology ends between the age of five to seven. According to Asher and Garcia (1969), 93% of L2 learner who were first exposed to L2 at the age of twelve had a clear foreign accent, as compared with those who began learning before seven years of age, among whom 68% were rated as near-native. Fathman (1975) discovered that among 200 eight- to fifteen-year-old children, those aged eleven to fifteen performed significantly better on tasks concerning morphology and syntax, however the younger children outperformed the former group on pronunciation. Other researchers also confirm the superiority of older learners on the syntactic aspect of language, e.g. Erwin-Tripp (1973), Chun (1978), or Krashen, Long and Scarcella (1979). More recent studies (Herschensohn 2000) viewing language acquisition from the perspective of Constructionism agree that native-like pronunciation results from early exposure to the second language, when after a period of feature underspecification a period of building L2 values on other constructions follows. Constructionism concerns morphology and syntax, yet its findings can be translated onto the development of phonology.

---

[22]   There exist highly successful late L2 learners whose pronunciation has been judged as native-like. Their performance is treated either as an exception to the rule, or as an outcome of the general lack of agreement what 'passing for a native' means (see section 2.2.2.).

Despite the vast body of research proving that "younger is better" (Krashen – Long – Scarcella 1979), there have been studies suggesting that adults do possess superior abilities for L2 production (Olson – Samuels 1973). Snow and Hoefnagel-Höhle (1977) proved, though, that the advantage holds for first stages of learning. Bohn and Flege (1990) provided strong evidence proving that adults can master both production and perception of L2 vowels. The mean age of their subjects was 30; therefore they argued that no critical period for learning L2 sounds exists before this age (Leather and James 1991).

According to Piller (2002), many analyses of the ultimate attainment have not exploited linguistic research methods fully. Instead, the methods used have been biased: they have been characterised by monolingual bias, that is, an L2 user's performance have been measured against a native speaker's performance. This can lead to misinterpretation of the facts, as the learner is a multicompetent bilingual; therefore his/her competence will never be identical to a monolingual native speaker's (Cook 1992). Phono-syntactic bias means that most previous research concentrated on two aspects of SLA, namely phonology and syntax, while in reality learners acquire pragmatic and social aspects of communication in L2 as well. Production bias of many previous studies implies the focus on production. Only recently researchers have shifted their interest to perception, i.e. not only how L2 learners speak but also how their production is perceived by the L2 community. Third person bias means that many studies fail to avoid outsider accounts of passing for a native. Finally, lack of ecological validity refers to the overlooking the connection between the cognitive and social processes, therefore testing procedures should be more holistic and qualitative. Only when these postulations are obeyed can research provide valid information on how successful L2 users are.

Those late L2 learners whose pronunciation and syntax are judged indistinguishable from the production of a native speaker are considered exceptional (e.g. Bongaerts et al. 1995), "Olympic high jumpers or opera singers" (Cook 1999: 191). Piller (2002), though, claims to the contrary: in her study she proves that such highly successful L2 users are not that rare. She gathered 38 conversations of bilingual couples where one partner's L1 was English, the other's German. 17 of the conversations could pass as native and 27 out of 73 individuals asserted they were highly proficient in L2 and in certain contexts they could be considered native speakers. What is significant in her study is that the average age at which the subjects first encountered L2 naturally was 20.9 years.

Another reason for conflicting evidence for the existence of age limitations on learning is the fact that in some studies the subjects' ability to imitate L2 sounds was measured, in other a their success in perceptual discrimination tasks. What is more, auditory evaluation of learners' performance often differed from one native judge to another. On the other hand, extensive training of judges to ensure their reliability could disrupt a natural "native-listener" mode of assessing the learner's production (Leather and James 1991). Moreover, when measuring adults' performance such learner factors as motivation, aptitude, etc. must be taken into account, as well as the influence of the circumstances in which learning takes place (Neufeld 1980).

A vast body of research suggests that successful acquisition of L2 phonology must start before the learner is six or seven. As Felix (1978) said, child L2 learners acquire native-like intonation easily. It seems, however, that the question whether older learners can master L2 pronunciation cannot be answered by the age factor alone, since it is not proven that neurological changes indeed limit their learning abilities. Any explanation of the children's and adults' capacities to perceive and discriminate L2 sounds should take into account the differences in their experience in auditory attention to speech. For example, children about the age of 30 months can ignore some sound contrasts since their L1 phonology is still partially developed (Oller 1983). On the other hand, the adults' long experience in a specific linguistic environment limits their discriminatory abilities (Strange – Jenkins 1978). Yet other studies (e.g. Werker – Gilbert – Humphrey – Tees 1981, Werker – Tees 1983) prove that infants can perceive non-native sounds, which ability seems to be, at least partially, retained by 4-year olds.

Other constraints that must be considered when assessing a learner's success or lack of success in acquiring an L2 include the individual and social ones. It is also necessary to decide what it means to acquire a "native-like" competence.

2.2.2. The meaning of "native-like"

The reason for the conflicting findings described in section 2.2.1. is, among others, the lack of a clear definition of a "native-like" pronunciation or possible methodological flaws of many studies. What is more, the term "native speaker" is itself ambiguous. Davies (1991) summarises its definitions in the following way:

a) The native speaker acquires the L1 in childhood.
b) The native speaker has intuitions about their idiolect grammar.
c) The native speaker has intuitions about those elements of the Standard Language grammar which differ from their idiolectal grammar.
d) The native speaker has a unique capacity to produce fluent spontaneous discourse and exhibits a well-developed communicative competence.
e) The native speaker has a unique capacity to write creatively.
f) The native speaker has a unique capacity to interpret and translate into their L1.
It is disputable how close to the "native speaker model" an L2 learner can become.

For sure, criterion (a) cannot be fulfilled: if a learner acquires the L2 in childhood, s/he is going to become a bilingual native speaker. Criterion (d) also tends to be difficult, though not impossible, to achieve, especially if L2 acquisition takes place in a formal setting. The other criteria can be fulfilled, yet some learners will be more successful than others due to both sociolinguistic and psycholinguistic factors: becoming a native speaker is the matter of learning rather than employing the innate capacities.

It is problematic, however, whether a clear-cut distinction between a native speaker and a non-native one can be drawn. Cook (1999) claims that such a judgement can be done on the basis of the speaker's biography, providing that such exceptions as bilinguals, the disabled intellectually or especially gifted learners are excluded. Davies (1999: 8) states: "I believe that [native speaker] membership is largely a matter of self-ascription, not something being given."

Another factor that contributes to the ambiguity of the term "native speaker" is the question what exactly a Standard Language is, which is especially evident in the case of English. There exists the traditional division into British English and American English, each variety consisting of a large number of sub-varieties, some being closer to the standard than others. The 20th century, however, has witnessed the rise of the so-called World Englishes, such as Indian English or African English. The majority of English teachers are people for whom English is not their L1 and the majority of English speakers are multilingual non-native speakers.

The problem has been addressed by many linguists, among whom Singh (1998) argues that American English and British English are different, but nobody would judge any of them inferior. Following this reasoning, Indian English and other World Englishes should be treated as a dialect, not a corrupted performance by foreigners. What is more, Singh (2006, a speech during Poznan Linguistic Meeting) emphasises the disagreement over the status of English in South Asia: some linguists refer to South Asian Englishes, whereas others speak of English in South Asia. The notions "native" and "non-native" varieties, though widely used in pedagogy, are linguistically oxymoronic: South Asian varieties "are just as (non-) native as Texan or Yorkshire English" (Singh 2006). Therefore, given that one cannot enumerate any properties that a "native" and "non-native" varieties do not share and that there exist no psycholinguistic or neurolinguistic differences between the acquisition of monolingual and multilingual competence (Paradis 1998, after Singh 2006), the term "native speaker" should be replaced by a more relevant one, namely "native user"[23]. The fact that English has become a global international language has led Davies (2003) to call the concept of the native speaker a myth and reality.

It is phonology that most acutely reveals the foreign accent in the speech of the learner. Firstly, the "nativeness" of L2 learner's phonology is the easier to asses than, e.g. syntax or pragmatics, since the linguist has the tools necessary to compare the quality of the learner's segments or the shape of intonational contours with the target language model. Secondly, L2 intonation is often considered not teachable, especially in the classroom, as its successful acquisition requires interaction with native speaker (Setter 2005). Therefore, some linguists find the proposal to devise a simplified version of English for foreigners, the Lingua Franca Core (Jenkins 2000), particularly appealing. Jenkins assumes that since the majority of learners of English will communicate mostly with other non-native users of English, there is no need to acquire the native-like accent. According to this view, only those aspects of English phonology which are "crucial for intelligibility"

---

[23] Cf. Davies (1991: 67), "on linguistic grounds Singaporean English does not exist, but of course does British English (…) what does exist is the individual speaker."

(Jenkins 2000:153), the so called phonological core, are important to acquire. In the case of intonation, Jenkins assumes the phenomenon problematic, yet such features as nuclear stress placement and chunking are included in Lingua Franca Core.

Lingua Franca Core is an extreme approach to the problem of reaching native-like competence, and it seems to be misconceived. In the case of intonation, it is necessary to teach the native-like model, be it British English or General American, due to the multitude of its meanings and functions. For instance, the study of Wennerstrom (1998) reveal that non-native speakers fail to use intonation "to signal meaningful contrasts", which may impede communication. Likewise, Goh (2000) while investigating the use of prominence and tone in Singapore English (recognised as a New Variety of English) found that these features "do not always have the same discourse functions as they do in Brazil's model" (Goh 2000: 43), namely referring and proclaiming tones often cease to signal given and new information, or speaker's dominance. Goh emphasises the importance of improving English pronunciation in Singapore "to a level which is internationally intelligible" (Goh 2000: 43). For a detailed discussion of Lingua France Core see, e.g. workshops at Poznan Linguistic Meeting (PLM) 2003, 2004, Dziubalska-Kołaczyk (2005), Dziubalska-Kołaczyk – Przedlacka 2005.

2.2.3. Sociological and psychological factors

There exist various personal and social factors that may motivate the learner either to strive for native-like mastery of L2 speech or to acquire only its minimal command, just to be comfortably intelligible. Learners can be integratively motivated when they wish to identify with the L2 community, or instrumentally motivated if they learn an L2 for practical purposes, such as improving their social status, passing tests, etc. (Gardner – Lambert 1959).

Piller's data show that the age factor can be overshadowed by the learner's motivation and agency, as well as the control over their learning. For instance, one of her subjects described his EFL classes at school in the following way:

> I learnt it more or less in school. Actually I didn't get good marks because I didn't see a point in it.
> [...] I only started to learn some vocabulary when I was about to move to the States (Piller 2002: 188).

Moreover, some of her subjects' speech became indistinguishable from the accent used in the city they live: they acquired the variation used in this particular region. Much ultimate attainment research concentrates on standard variations (Milroy – Milroy 1997), so it is possible that the non-standard pronunciation of some successful learners would not be judged as native-like. As Leather and James (1991: 309-310) put it, it is the variety of the L2 which the learner speaks that makes them sound "acceptable" to a native-speaker.

An important reason why an L2 learner does not acquire native-like pronunciation is their own conscious choice: they may be afraid that if in a conversation with native speakers they will not understand a joke or a remark drawing on some cultural information unknown to the learner, the learner can be seen as stupid. It is safer then to retain a foreign accent to make it clear they are not native. Others may just wish not to be perceived as belonging to a particular nationality, therefore they do not try to speak with a perfect L2 accent. Moreover, their sounding as a native speaker "would negate their achievement in learning an L2 to a very high level" (Piller 2002:195). Bailey (1978) claims that highly successful learners may be rejected by natives, since foreign pronunciation can be expected to mark them as "outsiders".

Another factor influencing the mastery of L2 speech is the attitude of the learner towards the L2-speaking society and culture. Dowd, Zuengler and Berkowitz (1990) prove that learners, regardless of age or level of proficiency, adopt a number of markers in their L2 pronunciation. If the learner's knowledge of L2 marking system is incomplete, s/he must take the risk of making social self-identifications in order to communicate with the L2 culture. Any occurring "mis-markings" possibly result from L1 transfer at a sociolinguistic level. It is assumed that such "social identity" constraint will be most apparent in naturalistic, untutored learning.

2.2.4. Individual factors

Personality of the learner can influence the learner's pronunciation. Such features as empathy, intuition, as well as self-esteem and flexibility of ego boundaries (e.g. Guiora – Brannon – Dull 1972) have been associated with the capacity of acquiring accurate L2 pronunciation. Guiora and his associates claim that as children acquire a general ego, a language ego emerges. Younger learners' ego boundaries are more flexible than those of older learners, which is the reason why children acquire an L2 accent more readily. However, those adult learners endowed with greater empathy have more permeability of language ego boundaries; therefore they can gain an advantage in L2 pronunciation (Larsen-Freeman – Long 1991).

Another factor deciding about the acquisition of accurate pronunciation is the sex of the learner. Eisenstein's (1982) study reveals that women perform better on dialect discrimination tasks, moreover they are more accurate at distinguishing more prestigious dialects from varieties of lesser prestige (Larsen-Freeman – Long 1991). According to Gussenhoven (1979), female learners are more likely to acquire the prestige accent of the L2, which is perhaps a reflection of their tendency towards prestige L1 speech (Leather – James 1991).

Learners can differ in their oral and auditory capacities. During acquisition they need to learn how to shape the oral cavity and control the movements of the articulators. Learners receiving explicit pronunciation training can rely on verbal instructions and

feedback, but those learning naturalistically must find the match between L2 sounds and articulatory configurations. The ability of accurate perception of special configurations within the mouth, i.e. "oral stereognosis", correlates with the ability to acquire L2 pronunciation. Evidence reveals that this capacity increases until mid-teens but later declines with age (Leather – James 1991). As for auditory capacities, learners perform better if the learning conditions suit their individual abilities in auditory discrimination (Leather – James 1991).

2.3. Perception of intonation

There is abundant research into perception of segments (e.g. Goto 1971; Sheldon and Strange 1982; Best 1994, 1995; Flege 1995; Brown 1998, 2000). Still, although intonation has become a fashionable field of study, despite new formal representations of pitch contours and significant advancements in speech technology, no unified model of intonation perception has been proposed. The reason lies in the complex nature of this aspect of speech: the lack of a clear definition, the variety of theoretical approaches, the large number of intonational cues and the lack of standardized research methods (Vaissière 2005).

Definitions and theoretical approaches have been discussed in chapter 1. Intonational cues involve short-ranged local cues (e.g. juncture tone), semi-global cues (e.g. resetting of the baseline in a part of an utterance) and global cues (e.g. controlling declination, pitch range, pitch register and rate of speech in a whole utterance). Local cues are more perceptually significant in languages with strong lexical stress, such as English, while semi-global and global cues are more important in pitch accent and tone languages. Speakers perceive them in an integrated way, which means that they employ more than one property of the pitch contour in recognizing utterance types or in finality judgments (Vaissière 2005).

Apart from pitch, pause duration, intensity and voice quality serve as cues for prosodic contrast. The properties of the pitch contour help to establish whether an utterance is a question or a statement, together with duration it marks phrasing and topical structure of a text. All the cues construct a coherent, context-dependent hierarchy: e.g. the range of f_0 is smaller in the post-focus position, therefore other factors can become leading cues, such as temporal cues for phrasing in French and for stress marking in Swedish[24] (Vaissière 2005).

Another difficulty in studying perception of intonation concerns the non-applicability of many research methods. It is not possible to measure intonation with the methods used traditionally in psychoacoustics. Contrary to segments, the perception of f_0 comprises psychoacoustic level, as well as higher-level cognitive and linguistic processing.

---

[24] Examples of trading relations between parameters.

Phonological distinctiveness and semantic differentiation is not possible in the case of intonation: in affective judgements variations in the range of f_0 are continuous rather than categorial; while it is possible to construct polar contrasts of intonational meanings, the categories will remain gradient which makes the execution and interpretation of perceptual experiments difficult (Vaissière 2005). Another problem to overcome is distinguishing between linguistic and paralinguistic functions of intonation (Crystal 1976). In some languages intonation can replace certain syntactic features, e.g. word order, or it may function as a reinforcement of the feature. Intonation can be the only way of expressing a polar question or a particulat attitude in one language but not in the other. Moreover, due to large variations between speakers, styles or the rate of speech, it is hard to find acoustically well-defined units. Finally, the size of the constituents of the utterance determines its final prosodic organisation (Vaissière 2005).

The perception of intonation depends on context: the intrinsic context includes the timing of f_0 features, loudness, duration and intensity, which requires configurational view of perception; while the discourse context refers to speaker's intentions and listener's interpretation of an utterance. There are differences between adults' and children's abilities to read the intentions: adults rely both on contextual information and intonation, the former being hard for children to decipher (Vaissière 2005). Thus intonational meanings cannot be studied without the pragmatic implicature (Wichmann 2002).

The listener's cultural and linguistic background significantly influences the perception of f_0 contours. Experiments in which subjects' native languages differed proved that "perceptual mapping between the acoustic signal and intonational categories is sensitive to the abstract structural properties of individual phonological systems," (Vaissière 2005: 243), e.g. those by Berinstein (1979) on word stress, Makarova (2001) on prominence and modalities, or Abelin and Allwood (2000), Kim, Curtis and Carmichael (2001) on attitudes and perceived emotion.

Despite the fact that prosody is to a large extent language-specific, it is possible to distinguish some of its universal features of perception between the form and meaning of intonation summarized by a hypothetical psychophonetic code (Vaissière 1995, Gussenhoven 2002).
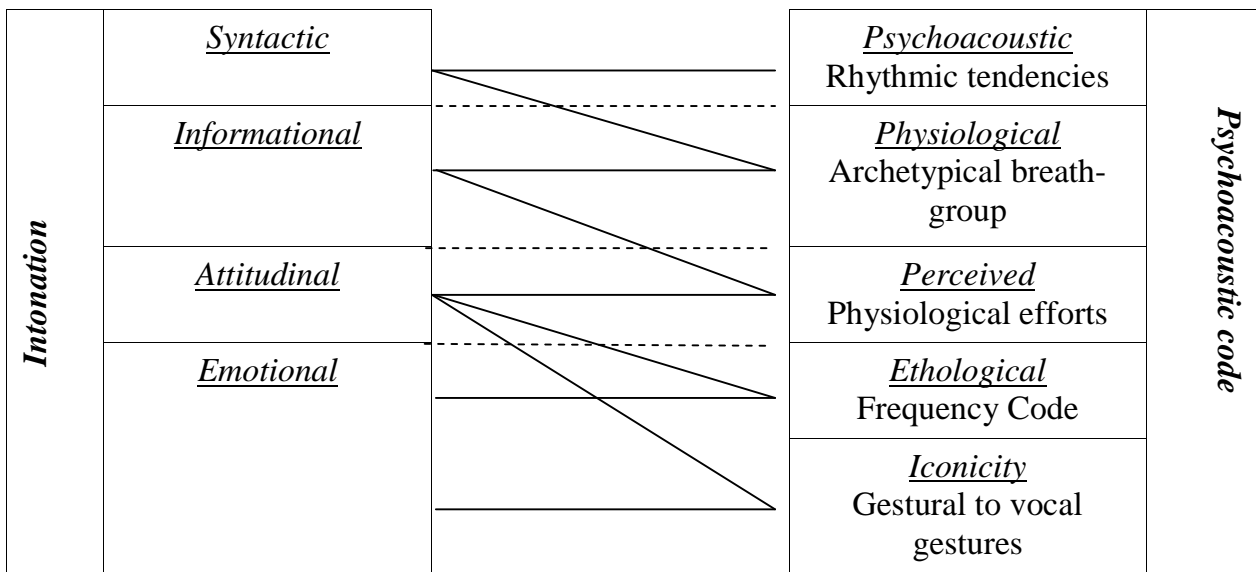
Fig. 2.1 Five elements of the hypothetical psychophonetic code and some functions of intonation (Vaissière 2005: 244).

The hypothetical psychophonetic code involves five elements: psychoacoustic rhythms, the archetypal breath group, perceived expiratory effort, the Frequency Code and iconicity. The basic psychoacoustic rhythmic elements refer to two non-linguistic principles in which the notion of end is associated with lengthening and the notion of beginning with strengthening, both playing role in segmenting speech at the word and phrase level. These tendencies are visible in marking word stress which is the outcome of the grammaticalization of intonation: initial (early) stress, as in English, is marked by extra loudness, as compared to late stress, as in French, which is characterised by extra lengthening. A similar phenomenon is observed at unit boundaries. There is a tendency to lengthen the final element of a unit, e.g. the last syllable in a word or phrase, the final phrase in an utterance and the last utterance in a paragraph. (Vaissière 2005: 244-245). It seems that in the perception of phrase boundaries it is duration and not f_0 that is significant (cf. Lehiste – Olive – Streeter 1976; Price et al. 1991; Verma et al. 1999).

An archetypal, physiologically-based f_0 pattern (Lieberman et al. 1967) which is equivalent to phonation in a single expiration is characterized by a sharp rise and a subsequent fall. With ensuing inhalation subglottal pressure builds up and f_0 values are reset, thus beginning another breath-group. This pattern is universally used for unmarked statements in many languages (Bolinger 1989), therefore any deviations from this pattern will be meaningful. Speakers of different languages focus on various aspects of the archetypal pattern, for instance in English a rapidly falling contour indicates a word-stressed syllable. In contrast, a slow rise and final lengthening is perceptually dominant in French. In Danish, the word-stressed syllable is signaled by low f_0 value. Each language

possesses a prototypical way of combining f_0, duration, intensity and segmental characteristics (Vaissière 2005).

Another element perceived by the listener is the global effort made by the speaker, the Effort Code (Gussenhoven 2002; see section 1.5.5). The timber of the speaker's voice is affected by the degree to which the supraglottal speech organs are strengthened. f_0 contour, voice quality and glottal resistance is influenced by laryngeal effort. Additional activity of the respiratory muscles, which increases the rapidity of glottal opening and intensity, indirectly causes a rise in f_0. Ladefoged, Draper and Whitteridge (1958) correlate nuclear stress, sentence stress and emphatic stress with the sudden extra activity of intercostals muscles which increases subglottal pressure and loudness. On the other hand, focus and emphasis is signaled by expiratory stress. Involvement and arousal are signalled when expiratory effort is maintained, which causes an increase in f_0 range and eliminates declination. Large variations in pitch and high pitch levels express happiness and surprise, while psychological stress and anxiety involve higher pitch and levels of amplitude because of greater respiration rate and a growth in subglottal pressure (Vaissière 2005).

The Frequency Code (e.g. Bolinger 1989, Ohala 1984; discussed in 1.5.5) concerns the influence of the size of the vocal tract on frequency, e.g. lower formant frequencies are associated with larger vocal tract. However, mean formant frequencies, f_0 and breathiness can be partially manipulated by speakers. The Frequency Code is an element of the Iconicity Code, "a gestural-to-lingual code" (Vaissière 2005: 251). In order to read the speaker's intentions, listeners rely not only on vocal gestures, including intonation, but also on facial gestures which are often stronger cues than linguistic ones (Bolinger 1989). For example, when expressing surprise, raised eyebrows often accompany a rise in f_0. Signs expressing attitudes and emotions are claimed to be universal as they are based on an archaic genetic code (Bolinger 1989). The evidence for universality of these intonational meanings is the fact that recognising emotions in a foreign language is often easier than distinguishing types of sentences. The Iconicity Code holds that the signs (instinctive "significants") which were at first used as the means of expressing primary emotions became a part of the linguistic code, i.e. developed into intonational "signifiés". Thus excitement and arousal coincides with higher tension of the vocal folds and higher pitch. Lack of strong emotions is characterised by low f_0 and slower rate of speaking. Greater pitch changes indicate agreeable emotion, while flat intonation signals disagreeable emotion (Fonagy 1981; after Vaissière 2005).

Abundant evidence proves that learners utilise perceptual categories of L1 to process L2 speech sounds. In perception of segments, listeners[25] employ the phonetic categories of their mother tongue following a process of "equivalence classification" during which they project native L1 phonetic categories on L2 sounds. New sounds which have not undergone the equivalence classification will become the basis for a new perceptual category (Flege 1987, 1991). Best (1996) assumes that non-native phonetic segments are perceptually assimilated to native phonetic categories according to their

---

[25] Adults' perception of L2 speech will differ from children's perceptual abilities since their L1 is not fully developed yet.

articulatory similarity. Pisoni (1997) claims that L2 sounds are matched to native exemplars sharing appropriate acoustic parameters.

A similar tendency is observed at the prosodic level. Studdert-Kennedy and Hadding (1973) investigated how different L1 background affects the Swedish and American English subjects' judgments on which resynthesized intonational patterns are appropriate for 'statements' and 'questions'. Cruz-Ferreira (1986; after Leather – James 1991: 312) claimed that "listeners' perception of L2 intonation reflected a combination of intonation transfer (positive or negative) from L1 and universal strategies for intonation interpretation".

Mennen (1999) employs Flege's (1995) Speech Learning Model to account for peak alignment in learner's language. The study reveals that most of the subjects, except one speaker, did not manage to establish a new perceptual category for L2 peak alignment, however, the alignment category in their L1 has apparently been affected and certain distinctions between categories in the L1 and L2 have been lost. The only speaker whose L1 remains unaffected despite her mastery of L2 peak alignment provides counter-evidence to the claim that full mastery of L2 speech indicates a corresponding loss of nativeness in L1 (Major 1990[26]). Abercrombie (1967) poses a question whether there exists "tone deafness", i.e. inability to perceive pitch variations. However, those who consider themselves tone deaf speak their mother tongue as well as other speakers, so they must have heard differences in pitch patterns while acquiring their L1. He explains the difficulty as lying in problems with "adopting an analytic attitude towards something which has become so familiar" (1967: 102-103).

The following subchapters summarise major studies concerning the perception of intonational functions.

2.3.1. Grammatical functions

Since intonation, unlike other prosodic features[27], is inherently meaningful, it conveys additional cues for the hearer to decode messages. First of all, it serves to distinguish sentence types. In their study of the perception of sentence intonation, Gårding and Abramson (1965) determined three categories of intonational contours in American English which hearers described as 'neutral statement', 'yes-or-no question' and 'counting in a series'. Other studies (Hadding-Koch – Studdert-Kennedy 1965) specified such

---

[26] According to Major (1990), one cannot possess native-like phonology in both linguistic systems, as the mutual effects of L1 and L2 seem to interact with one another. A learner can either maintain L1 proficiency but fail to attain L2 proficiency, achieve L2 proficiency but lose native L1 proficiency, or lose native L1 proficiency but also fail to achieve L2 proficiency.

[27] Non-prosodic (segmental) and prosodic features, except intonation, do not possess inherent meaning but rather contribute to distinguishing meaningfully various linguistic elements (Fox 2000).

categories as 'statement', 'question' and 'talking-to-yourself'' for Swedish and American English.

Intonation helps the listener to recognise boundary marking in sentences. Berkovits (1984) found that fundamental frequency is the acoustic cue signalling finality in English and Hebrew. Higher f_0 peaks and smaller f_0 falls marked unfinished sentences, although this feature seemed not to be the only factor at play. She showed that listeners rarely commit mistakes in identifying finished and unfinished sentences.

2.3.2. Attitudinal functions

The interpretation of attitudinal meanings seems to be to some extent language-specific. Luthy (1983) proved that non-native listeners tend to misinterpret the intonational markings of rudeness, politeness, doubt, certainty, surprise or nonchalance. Scherer (1979) concluded that the ability to process attitudinal and emotional signals may innate, since native English subjects mostly agreed on the interpretation of synthesized tones.

2.3.3. Illocutionary functions

Geluykens (1987) in his study of rising intonation in 'queclaratives' claimed that even though felicity conditions help to perceive a declarative with rising intonation as a question, intonation plays a vital role when pragmatic cues are insufficient (Searle 1969).

Research into the perception of sentence and paragraph boundaries reveals that there are several intonational cues at play. Lehiste's (1979) established that timing and fundamental frequency are important markers. The length of an utterance influenced the subjects' judgements, as longer sentences tended to be identified as uttered in isolation rather than within a paragraph. A cue for sentence and paragraph boundaries was the length of the pause, as the perceived paragraph boundaries were signalled by longer pauses. Moreover, sentences with higher f_0 peaks on the first stressed word were recognised as initial in a paragraph, thus marking a paragraph boundary. On the other hand, intonation falling towards the baseline signalled the termination of a unit. Kreiman (1982) proved that non-level intonation patterns, as well as laryngealization, pre-boundary lengthening and pauses indicate sentence boundaries. Paragraph boundaries are additionally marked by cues that are initial and final in a unit, and by "cues that span boundaries and both that one unit has ended and that another has already begun" (Kreiman 1982: 163).

In conversation intonation also serves as a perceptual cue for turn-taking. Schaffer (1984) noticed that turn-beginnings were correctly recognised more frequently than turn-ends. However, syntactic and lexical markings seemed more significant than intonation.

Cutler and Pearson (1986) tentatively concluded that f_0 contour, namely a downstep and an upstep in pitch, is a significant cue for turn-taking; however, their subjects could not determine a considerable number of utterances.

Ford and Thompson (1996) found that the syntactic, intonational and pragmatic completion points converge. The least significant type of completion are syntactic completion points. On the contrary, intonation and pragmatics are the most reliable cues signalling which syntactic completion points are perceived as completed. Intonational completion usually coincides with syntactic and pragmatic completion. On the basis of these three cues interlocutors establish boundaries between turn-constructional units of conversation, the so called complex transition relevance phrases (CTRP).

2.3.4. Sociolinguistic function (perceived foreign accent)

Munro and Derwing (1995) studied the influence of intonation on perceptible foreign accent. They discovered that non-native accent is detectable even if low-pass filtering makes utterances unintelligible. Moreover, it takes longer for native speakers to process utterances produced by non-native speakers than by native speakers.

Studies into the perception and acquisition of segments can suggest how intonation is perceived. Previous research, e.g. by Riney and Flege (1998), revealed that after a period of four years Japanese students in America improved their ability to discriminate between English /r/ and /l/. Lively et al. (1994) showed that after a proper training, Japanese listeners were able to distinguish /r/ from /l/, an ability that did not diminish during the following three months. Thus, the capability to perceive foreign phonetic contrasts resulted from the listeners' selective attention to the acoustic cues of the two sounds. Carmichael (2004) concluded that selective attention could also influence the perception of intonational contrasts in L2.

2.4. Influence of the mother tongue

The influence of the native language on the acquisition of L2 phonology has been generally accepted, despite the debate on the extent to which L1 affected L2 (Dulay – Burt 1972, 1973[28]; Flynn 1987). This makes the acquisition of phonology different from the acquisition of syntax: L1 interference is not that obvious in the area of grammar. According to Richards (1971: 204),

> studies of second language acquisition have tended to imply that contrastive analysis may be most predictive at the level of phonology, and least predictive at the syntactic level (Richards 1971: 204).

---

[28] Dulay and Burt claimed that even if errors seemed to be the result of L1 interference, they were developmental in fact and reflect the growth of an emerging linguistic system.

Moreover, it is possible for a native speaker to recognise non-natives by their pronunciation but doing so on the basis of syntactic evidence alone is not easy (Ioup 1984).

The Contrastive Analysis Hypothesis (CAH) was one of the earliest studies into L1 influence on L2 acquisition. The CAH assumed that for the learner "those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult" (Lado 1957: 2) and that "when learning a foreign language we tend to transfer our entire language system in the process" (Lado 1957: 11). At that time research concentrated on comparing phonemes and the distribution of their allophones in L1 and L2. When faced with L2 sounds that differed from those of L1, the learner assigns two or more allophones in the L1 to different phonemes in the L2. Lado illustrates this claim with the sounds [d] and [ð], which are separate phonemes in English but in Spanish they are allophones of [d]. Thus for Spanish learners assigning the sounds to contrasting English phonemes presents great difficulty. The conclusion was that a thorough comparison of L1 and L2 could allow predicting which L2 structures would be difficult to learn.

Stockwell and Bowen (1965) developed the CAH incorporating ideas of American Structuralism and behavioural psychology (Hockett 1955). They established an eight-level hierarchy of difficulty, based on the classification of L1 and L2 sounds as "null" (non-existent in a language), "optional" (whose distribution is not predictable from the phonological context, so phonemes) and "obligatory" (allophones, as their distribution is based on the phonological environment). Thus they could predict that L2 allophones that are null in the L1 posed maximum phonological difficulty for the learner. The weakness of Stockwell and Bowen's claims was that they were not based on systematically gathered data, yet later studies prove that L1 does influence at least some aspects of L2 pronunciation (cf. Ioup 1984, Eckman 2004).

A number of studies opposed the CAH as they assumed that similarities between L1 and L2 accounted for pronunciation errors. Wode's (1978) Crucial Similarity Measure (CSM) could explain a number of not only phonological but also morphological problems (cf. Young-Scholten 1985). Oller and Ziahosseiny (1970) concluded that while the source of difficulty were the similarities between the writing systems of the L1 and L2, differenced were less likely to cause confusion as they were more salient (after Eckman 2004).

The CAH was eventually refuted by a growing body of research which proved that developmental processes and ease of articulation were more significant than L1-L2 differences. More evidence against the CAH was provided by research into the perception of L2 speech.

L1 transfer can also be affected by the universal typological features of the two languages in contact, as described in the following section.

2.5. Influence of linguistic universals

2.5.1. Interlanguage

Three researchers developed a concept of a "learner's language" independently: Corder (1971) described it as "idiosyncratic dialect", Nemser (1971) referred to it as "approximative system", while Selinker (1972) called the construct "interlanguage". The concept is based on the idea that in the course of acquisition learners create a linguistic system different from either their native language or the target language. Interlanguage (IL) reflects the evolving system of rules which results from such processes as the influence of the L1 ("transfer"), contrastive interference of the L2 and the overgeneralisations of the rules (Crystal 2004: 239).

Corder (1968) postulated that L2 learners might be like L1 learners in that they follow an internally preordained learning sequence (an inbuilt syllabus) during acquisition: certain forms would not be acquired until the learner was "ready" for the acquisition. The learner's systematic behaviour was named a "transitional system", since it was assumed that during SLA the learner formed hypotheses about the target language and tested them in order to confirm or reject them.

The idea of the transitional system was continued by Selinker (1972), who coined the term "interlanguage", referring to the systematic knowledge of L2, intermediate between L1 and L2. He claimed that the L2 learner's mind did not resemble the mind of the L1 learner, which seemed to be confirmed by the fact that only 5% of L2 learners achieved native-like ability in L2. Moreover, Lenneberg's critical period for L1 acquisition also signalled a critical period for the acquisition of any other language, providing more evidence supporting the claim. Among interlanguage processes, such as fossilization, or overgeneralization, L1 transfer, from which some interlanguage rules derive, is mentioned.

Interlanguage Hypothesis deprived L1 transfer of its dominant role. In its extreme version, e.g. in Dulay and Burt's (1974) Creative Construction Hypothesis or Krashen's (e.g. 1981) monitor model, no differences between L1 and L2 acquisition existed and neither did transfer. The moderate version of the Interlanguage Hypothesis, including Markedness Differential Hypothesis (Eckman 1977, 1981), granted L1 transfer a certain responsibility for (un-) successful acquisition of L2 (see section 2.5.2.).

Gass (1988) developed the notion of language transfer, defining it as a phenomenon which involves "the use of native language (or other language) information in the acquisition of a second (or additional) language" (Gass 1988: 387). Her interpretation included, among others, such phenomena as transfer of typological organization, different paths of acquisition, avoidance or overproduction of certain linguistic elements.

Numerous studies revealed that L1 transfer is a complex phenomenon. For example Zobl (1982) found two patterns of L1 influence: the pace at which the learner traverses along developmental sequences and the number of developmental structures within the

sequences. If an L1 structure and the developmental one were similar, the learner tended not only to persist with the developmental structure but also to add an additional stage to the developmental sequence. In this respect similarity between L1 and L2 did not facilitate learning. Other studies, e.g. Kellerman (1978), show that transfer can be utilised as a strategy compensating for the lack of L2 knowledge in all language areas but phonology.

Interlanguage is regarded a key development in SLA theory, since it can help to establish the extent to which the learner's language resembles L1 grammars. Moreover, it can account for utterances whose structures derive neither from L1 transfer nor from L2 input.

2.5.2. Markedness

The concept of markedness was introduced by Nikolai Trubetzkoy (1939) and Roman Jakobson (1941) of the Prague School of Linguistics. Since then, several interpretations of the notion appeared. The earliest one stated that "a sound would be marked if it possessed a certain distinctive feature (e.g. voice), and unmarked if it lacked it" (Crystal 2003: 283). As Larsen-Freeman and Long (1991: 101) observe, "linguistic notions of 'markedness' are usually defined in terms of complexity, relative infrequency of use or departure from something that is more basic, typical or canonical in a language," therefore in such pair of words as *man/woman*, the former is considered unmarked. According to Eckman (2004: 28-29), one member of the binary oppositions, such as voiced and voiceless obstruents, or open and closed syllables, are believed to be privileged in the sense that they have wider distribution not only within a language but also across languages. Those privileged members are considered "unmarked", that is "in some definable way simpler, more basic and more natural than the less widely occurring member of the opposition", the "unmarked" item. Therefore, one can conclude that a falling intonation contour is unmarked while the rising pattern is marked, as the former is apparently more common.

From the typological perspective, markedness relates to the cross-linguistic situation in which the presence of a certain linguistic feature implies the presence of another feature. Gundel et al. (1986: 108) defined typological markedness, which refers the distribution of linguistic representations among the languages of the world, as follows.

> A structure X is typologically marked relative to another structure Y, (and Y is typologically unmarked relative to X) if every language that has X also has Y, but every language that has Y does not necessarily have X.

Generative linguistics formulated a theory of markedness in which unmarked features relate to the general tendencies found in all languages. On the other hand, marked features refer to those properties which are exceptional. Thus a highly unmarked feature is seen to possess universal status, unlike a highly unmarked one. For instance, the CV sequence is considered a phonological universal, whereas the sequences with combinations of consonants and vowels that are different from that pattern are exceptional and, therefore,

more marked. The universal set of linguistic forms is often referred to as the core, whereas the forms departing from the unmarked universals are called periphery.

The notion of typological markedness gave rise to two hypotheses relevant for L2 phonology: the Markedness Differential Hypothesis (MDH) (Eckman 1977) and the Structural Conformity Hypothesis (SCH) (Eckman 1991).

The Markedness Differential Hypothesis (Eckman 1977: 321) states the following:

The areas of difficulty that a language learner will have can be predicted such that

a) Those areas of the target language which differ from the native language and are more marked that the native language will be difficult;
b) The relative degree of difficulty of the areas of difference of target language which are more marked that the native language will correspond to the relative degree of markedness;
c) Those areas of the target language which are different from the native language, but are not marked than the native language will not be difficult.

The MDH, contrary to the postulations of CAH, stated that the differences between L1 and L2 were not sufficient enough to account for learning difficulty. The general claim was that not all differences would appear difficult for the learners. Moreover, a particular L2 item would cause different degrees of difficulty depending on the L1 spoken by the learner. The MDH was addressed in a number of studies (e.g. Anderson 1987 for degrees of difficulty experienced by learners of various L1 backgrounds; Major – Kim 1996, and Yavas 1994 for final devoicing in IL grammars; Carlisle 1997).

Eckman (1991) formulated another hypothesis employing the generalisations which underlie the notion of typological markedness, the Structural Conformity Hypothesis.

The universal generalisations that hold for primary languages hold also for interlanguages (Eckman 1991: 24).

Eckman (1996) argued that the motivation for the SCH is an L2 pattern in which the structures conform to markedness principles, the structures, however, cannot be accounted for by mere differences between L1 and L2. This is the reason why the MDH cannot explain the interlanguage pattern, as it is not L1-like or L2-like, but it is in accord with a kind of universal structure. The SCH was investigated in such papers as Carlisle's (e.g. 1997) and Eckman and Iverson's (1994).

A related hypothesis, the Similarity Differential Rate Hypothesis (SDRH) postulated by Major and Kim (1996) is based on the works by Wode (1976) and Flege (1995). The notion of markedness is merged with the idea that dissimilar sounds may be easier than ones similar to L1 sounds. The SDRH claims that "dissimilar structures are acquired more quickly than similar structures, and that markedness is a mediating factor" (Eckman 2004: 36) and many pronunciation errors can be explained by "rate of acquisition", not "difficulty". The hypothesis is supported by the fact that the learning situation for beginning and advanced learners is opposite: beginners utilize L1 transfer,

therefore similar sounds are easier for them, but not for advanced learners (Major and Kim 1996). The weakness of the hypothesis is that the ideas of "similar" and "dissimilar" are not clearly defined.

2.5.3. The Ontogeny Model and the Ontogeny Phylogeny Model

The Ontogeny Model (OM) (Major 1986, 1987), which was developed into the Ontogeny Phylogeny Model (OPM) (Major 2001), is a significant model of L2 pronunciation that attempts to analyse the interaction of phonological transfer and universals. In OM approach, substitutions of L2 sounds can be accounted for by L1 transfer and L2 development, which are part of universal grammar. The two types of influence on the learner's language change not only over time, but they also evolve according to the formality of the speaking situation[29]. The substitutions caused by L1 transfer gradually decrease in the course of learning. The substitutions increase, however, when the speaking situation becomes more formal. On the other hand, developmental processes first increase with time and with growing formality of the speaking situation, yet later they decrease. The Ontogeny Phylogeny Model adds that interlanguage comprises three parts, the L1, the L2 and universals.

2.5.4. Universal Grammar (generative approach)

The previous subchapters have revised theories that markedness constrains IL phonologies. The following section presents the principles of Universal Grammar (UG) and their influence on IL.

The UG framework claims that IL variation results from different parametrical settings of L1 and L2. The framework provides evidence that IL grammars are systematic, which cannot be predicted on the basis of the analysis of L1 or L2 alone.

Most studies of parametric variation in prosody concern stress. Pater (1997) described the acquisition of English stress by French speakers. The errors the subjects committed involved mis-setting the parameters of Word headedness and Directionality. Neither the L1 values for the parameters were transferred nor were they the same as the ones of the L2. However, the parameters of foot size and foot headedness posed no difficulty. The acquisition of phrasal stress by Polish and Hungarian learners of English was studied by Archibald (1993) whose subjects did transfer their L1 metrical parameters. His findings are in accord with those in other phonological domains, which state that L1 is

---

[29] Cf. Tarone's (1979, 1982, 1983) "continuum paradigm", i.e. the effect of the learner's attention to speech on IL variability.

a vital factor in the acquisition of L2 stress but learners can develop a system that originates from neither the L1 nor L2.

The perception and acquisition of tone and intonation is analysed by Leather et al. (1997). Their subjects were adult speakers of Dutch (a non-tonal language) learning Chinese lexical tone. First one group underwent training in perception of the tones, later they were given a test on their productive abilities. The other group was taught to produce the tones, then their perceptive abilities were tested. The experiment proved that the learning of L2 tones had certain processes that could be encountered L1 acquisition, i.e. learners formed hypotheses about the L2 phonological system, tested and revised them, gradually approximating the target forms. The conclusion was that

> learners did not need to be trained in production to be able to produce, or in perception to be able to perceive, the sound patterns of the target system: training in one modality tended to be sufficient to enable the learner to perform in the other (Leather 1997).

Rintell (1984) studied judgements of the emotional tone in conversations by speakers of different L1 backgrounds. Chinese subjects found recognising emotional states of English speakers difficult, whereas Arabic and Spanish speakers did not.

Zampini (1997), in turn, proved that in the acquisition of Spanish spirantization the L1 rule must be formulated within the domain of the intonational phrase in the prosodic hierarchy.

In the present work, the model of L2 acquisition adopted for the empirical study is based on the Natural Phonology framework, therefore the author chose to present the Natural Approach separately, in the following chapter.

CHAPTER 3: The acquisition of intonation – a model

3.1. Natural Phonology as theoretical framework

In recent decades a considerable number of models of acquisition have been proposed, yet the acquisition of intonation has rarely been approached. The following chapter attempts to integrate theories of L2 learning and models of suprasegmental phonology. Moreover, it aims at discovering the mechanisms behind the acquisition of this aspect of phonology which so far has received not much attention. The model of the acquisition of intonation is therefore based on a theory concerning segmental phonology, the findings of which prove to be applicable to suprasegmentals as well.

An adequate model should incorporate the findings of theoretical linguistics and acquisition studies. One of such attempts concerned the theory of the acquisition of phonology within the framework of Universal Grammar (UG). According to the model, phonological competence included universal principles and language-specific parameters which had to be re-adjusted in the process of acquisition. The weakness of the theory was the unresolved question whether adult learners continued to have access to UG or not (cf. Young-Scholten 1996, James 1996, Birdsong 1989).

Natural Phonology seems to provide the most suitable framework for the construction of a model of the acquisition of intonation, as not only can it provide a formal description of a linguistic system but also the processes taking place during L2 learning. Natural Phonology, however, is not just a descriptive theory, as the generalizations about language behaviour accounted for within its framework take the form of universal or language-specific preferences rather than absolute rules. This makes it is a preference theory rather than a descriptive one and gives it greater explanatory power, since in any analysis of linguistic behaviour the speakers and their active (to some degree) control over language cannot be excluded (Dziubalska-Kołaczyk 2002).

Moreover, the nature of the predictions and explanations provided by Natural Phonology, and Natural Linguistics in general, is both functionalist and semiotic (Dressler 1996, Dziubalska-Kołaczyk 2002). On the one hand, it employs two main functions of phonology by means of which it accounts for the communicative role of language, pronounceability and perceptibility. From this perspective, each linguistic choice of language-users is seen as the consequence of the goal-oriented (functional) linguistic behaviour. On the other hand, semiotics became a metatheory for the natural model linking it with other disciplines, thus allowing a better explanation of language behaviour. Dziubalska-Kołaczyk (2002: 104) illustrated the explanatory system of Natural Linguistics as in Figure 3.1.:

higher principles	non-linguistic
(e.g. the principle of least effort, of cognitive economy)	(cognitive, phonetic, psychological, sociological, etc.)

preferences	linguistic
(e.g. a preference for simple phonotactics, for a CV structure)	
preference parameters (pronunceability, perceptibility)	functional and semiotic

consequences of preferences	linguistic
(e.g. absence of clusters in a language)	

Figure 3.1. Natural Linguistics as an explanatory model (Dziubalska-Kołaczyk 2002: 104)

As Donegan and Stampe (1979: 168) say, even though Natural phonology "lacks any a priori methodology or formalization, is both testable and explanatory. By its nature, it is ultimately accountable for (…) everything language owes to the fact that it is spoken. And by its nature, it must follow from the character of the human capacity for speech."

Such a character of Natural Phonology renders it the most suitable framework for the theory of the acquisition of intonation, as it is conveniently applicable to adult learners of the second language. The natural framework gives a relevant and reliable account of the reorganization of the L2 learner's native phonological system when confronted with L2 requirements (cf. Dziubalska-Kołaczyk 1990). The theory is well-rooted in linguistic research, its beginnings dating back to Plato, who elaborated on natural naming of things in *Cratylus,* and was continued by, among others, Jan Baudouin de Courtenay, Mikolaj Kruszewski, Roman Jakobson, Edward Sapir, Otto Jespersen and Henry Sweet (Katarzyna Dziubalska-Kołaczyk 2002). The tenets of Natural Phonology were formulated by David Stampe (1969, 1979) and further developed by David Stampe and Patricia Donegan (1979), as well as Wolfgang Dressler (e.g. 1984, 1985, 1996). Initially, the model focused on first language acquisition, yet later was extended, among others, to the domain of the second language learning (e.g. Dziubalska-Kołaczyk 1987, 1990a, 1990b; Zborowska 2001, Wrembel 2005).

3.1.1. Basic assumptions

The naturalness of the theory is seen in the way language is characterized: it is "the reflection of the needs, capacities, and world of its users, rather than a merely *conventional*

institution" (Donegan & Stampe 1979: 127). The basic tenet of Natural Phonology is that the sounds of language evolve in individual speakers and over time in a speech community due to the phonetic forces "implicit in human vocalization and perception" (Donegan – Stampe 1979: 126). The forces are demonstrated through universal processes defined as

> mental substitutions which systematically but subconsciously adapt our phonological intentions to our phonetic capacities, and which, conversely, enable us to perceive in other's speech the intentions underlying these superficial phonetic adaptations (Donegan – Stampe 1979: 126).

The above definition of processes stresses their mental status, i.e. they are neither conventional constructs nor peripheral, physical events, although they are physically motivated, therefore they operate in the mind of the language user. The evidence for the mental character is provided by the fact that they are suppressible, e.g. the processes in casual speech concerning style can be suppressed in formal situations (Donegan – Stampe 1979: 136). Otherwise, processes would occur independently of their perceptual or articulatory consequences.

Another important assumption of Natural Phonology is that processes result from the limitations of human vocal and perceptual capacities and can be seen as a natural reaction to the mismatch between the sound intended and actually pronounced, as well as the discrepancy between clarity of perception and ease of articulation (Donegan & Stampe 1979: 130). Due to processes (substitutions), the listener is able to compensate for these limitations and decode the intentions of the speaker. Thus they occur as a response to difficulties in either production or perception of speech. Moreover, it is claimed that when a particular difficult representation is subject to substitution, other representations of the same difficulty will undergo the same process[30]. It follows that representations (sounds) can be categorized into 'natural classes' on which particular processes operate: segments, and "natural prosodic constituents – syllables, accent-groups, words, etc." (Donegan – Stampe 1979: 136). What is more, they undergo implicational hierarchies of applicability, that is, in a language processes apply in accord with a hierarchy of semiotic and functional parameters of naturalness (Dziubalska-Kołaczyk 2002).

Phonological processes begin to operate when the child starts to acquire their mother tongue. The infant has at his/her disposal a universal system of processes which it needs to inhibit in accordance with the requirements of the ambient language. To distinguish the universal processes from those occurring later in acquisition, which are the remnants of the universal system, Dressler (1984) separated 'process types' (universal ones) from 'processes' (substitutions derived from the universals). Although processes are universal, their application differs from language to language and the child needs to learn the constraints imposed by his/her native language on them.

Donegan and Stampe (1979: 142-3) distinguish three types of processes on the basis of their functions in language:

---

[30] Each process results from a particular phonetic motivation, therefore different motivations imply different processes (Donegan & Stampe 1979).

1. Prosodic processes are those which involve such prosodic structures as basic patterns of rhythm and intonation onto which words, phrases and sentences are mapped. Prosody is not provided in the linguistic material itself but rather it is determined by mapping, most relevantly described "as an operation in real-time speech processing of which setting sentences to verse or music are special cases." For Donegan and Stampe prosodic processes are vital for the developing language, since the choice of segmental processes is determined to a large extent by the mapping of segmental representations onto the prosodic structure in speech[31].

The remaining two types of processes apply to segments.

2. Fortition processes (centrifugal, strengthening, paradigmatic) are context-free and make segments more perceptible, e.g. dissimilations, diphthongizations, syllabifications or epentheses in such situations or styles as attentive, formal or expressive.
3. Lenition processes[32] (centripetal, weakening, syntagmatic) are context-sensitive and/or prosody-sensitive, as they make segments easier to pronounce, e.g. in assimilation, monophthongization, disyllabification, reduction or deletion phenomena. They apply in 'weak' positions, such as syllable-final or unstressed positions, in situations and styles in which clarity is of little importance.

Processes cannot be seen as equivalents of rules. The former have synchronic phonetic motivation, contrary to rules, which are the "historical result of conventionalized processes which have lost such a motivation" (Donegan & Stampe 1979: 144) but instead possess semantic or grammatical functions, like umlaut does. What is especially visible in the domain of second language acquisition, processes can be understood as natural responses to innate limitations or difficulties, while rules must be learnt. Processes are involuntary and unconscious but can become evident when confronted with pronunciations not corresponding to the process, whereas rules are created through conscious observation of linguistic differences. Processes cannot be borrowed, unlike rules. Finally, while processes can be either optional (style-dependent) or obligatory (style-independent), rules are invariably obligatory.

3.1.2. First language acquisition

One of the main interests of Natural Phonology is how universal processes (process types) apply during the mastery of the native language.

---

[31] Cf. Keating's (2004) claim to the contrary: segmental and prosodic planning interact in a minor way, at the end of the process of phonetic encoding.
[32] Cf. Dressler's (1984) terms: foregrounding / clarification for dissimilatory processes and backgrounding / obscuration for assimilatory processes.

72

The classical model of Natural Phonology states that the system of natural phonological processes is innate (Stampe 1969). This system is unlimited and unordered, therefore during acquisition the child needs to revise it, so that it approximates the system of adult speakers of the language, by means of three mechanisms: suppression, limitation and ordering. Suppression allows the child to eliminate those processes absent in the mother tongue, e.g. the process of devoicing final obstruents must be suppressed if the ambient language is English. Limitation resembles suppression in that a given process is constrained only in certain contexts, e.g. when the child devoices only dental obstruents. Thus the child, in order to acquire the first language, needs to learn the constraints on processes, not processes themselves.

In the process of acquisition, as the child learns to constrain those processes that do not conform to the native language system, their number is gradually reduced. Yet in the phonological system there remain innate processes which have no manifestation in the first language and therefore have not been activated. It is claimed that although they do not manifest themselves overtly, they reside in the system and can be triggered when an individual acquires a second language (Stampe 1969).

Evidence for the innateness of phonological processes comes from research into acquisition orders, which appear to be similar across studied languages. Moreover, there are parallels between L1 and L2 acquisition (Major 1987). Another piece of evidence is provided by the universal nature of substitutions. It has been proved that certain phenomena undergo similar substitutions during acquisition of various languages, such as substitutions of /w/ for /l/ or /r/, reductions of consonantal clusters or devoicing (Major 1987).

Such view of L1 acquisition poses certain problems, though. For instance, as Dressler (1996) noted, some processes can occur in an irregular way or they never appear as it is expected. Therefore, even if the classical model could account for language universals and certain parallels between emerging phonologies of various languages, it still needed refinement to be applicable to all aspects of language.

The classical model was revised by Donegan (1985) who explained the innateness of phonological processes on the grounds that they are immediate natural responses to the limitations of human articulatory and perceptual capabilities. She claims that the child discovers processes when he/she attempts to use the vocal tract, in other words, processes occur as the child tries to deal with the difficulties that the vocal and perceptual abilities pose. She weakened the claim, though, stating that

> It would not alter the theory of natural phonology substantially to say that processes may be discovered by the child as he learns to use his vocal tract (...). But if processes are learned, they are learned as matters of physical coordination are learned – by *doing* – not by the kind of cognitive processing that is required to learn other components of language, like syntax, morphology, or morphological rules (1985: 26, footnote 5).

Thus the emergence of child's phonology requires physical coordination rather than cognitive processing as in the case of acquisition of other components of language, such as syntax or morphology (cf. Dressler – Karpf 1994).

The constructivist model presents a more accurate attempt to account for language acquisition (Dressler – Dziubalska-Kołaczyk 1997), according to which the child gradually constructs their first language on their own. Contrary to Chomsky's theory, in which acquisition is a passive phenomenon, the child becomes an active participant in the process. The constructivist concept differs also from the Stampean classical model in that phonological processes are no longer believed to be innate and available to the child at birth, they rather arise during acquisition as universal reactions to the articulatory and perceptual limitations the child encounters. As the child's brain receives linguistic information, the process of neuronal specialization begins and, as a result, phonology emerges "as the outcome of the organization and reorganization of processing phonetic information" (Dressler 1996: 48).

The advantage of the constructivist model lies in its potential of explaining why certain phonological processes do not apply and variation in language development among individual children. Moreover, it can be employed to account not only for phonology but also for other components of grammar, thus its scope of application is broader than that of the innateness hypothesis (Stampe 1969, Dziubalska-Kołaczyk 1990a, Dressler 1996, Major 1987). However, as the two models do not stand in contradiction with each other, they can be to some degree integrated in acquisition studies, as in the model of self-organisation (Dziubalska-Kołaczyk 1998, Zborowska 1997 for second language acquisition).

3.1.3. Second language acquisition

Natural Phonology not only accounts for first language acquisition but it can also give remarkable insights into the study of second language acquisition, as in the works of, among many, Abrahamsson (1996), Hammarberg (1988, 1990), Dziubalska-Kołaczyk (1987, 1990a), Dressler – Dziubalska-Kołaczyk (1994), Major (1987), Zborowska (2001). Although it has concentrated mostly on segments and the syllable, it can provide a reliable model for the acquisition of higher levels of prosody.

The problem which SLA has tried to answer is whether and to what extent second language acquisition resembles that of the first language. The question within the natural framework is whether and to what extent the processes taking place in first language phonological acquisition apply when the individual attempts to master a second language. In the previous section it has been said that children need to learn the constraints on universal processes of their mother tongue, revise the set of those processes by means of suppression, limitation and ordering, until their phonological system approximates the system of an adult speaker of the language.

Abrahamsson (1996) believes that the acquisition of the second language follows the same route, except that learners need to learn to ignore the L1-specific phonological rules. As opposed to L1 acquisition, where the child begins with a latent (non-revised) system of phonological processes, during L2 acquisition the learner already has a revised system at his/her disposal.

Major (1987) also claims that L1 and L2 acquisition are the same, since the L2 learner needs to constrain or eliminate "those processes (interference and developmental) which are not characteristic of native pronunciation" (1987: 208). In other words, while the processes that the child needs to successively eliminate in order to arrive at the native set of phonological rules are developmental, the processes which the L2 learner has to constrain are interference ones, resulting from the previously acquired L1, and developmental ones, which begin to operate once interference processes are suppressed. Moreover, the relationship between the two types of processes changes, with interference processes first dominating at the beginning of during L2 acquisition and decreasing with time, and developmental processes increasing first and later diminishing.

Another issue which is vital for SLA concerns the perception-production relationship. Within the natural framework it is believed that in the process of L1 acquisition the child's underlying mental representations resemble adult native speaker's, and more importantly, since they remain unchanged during the process of acquisition, any problems that the child faces are production difficulties. This is not the case in L2 acquisition, as many instances of the learner's failure to attain native-like pronunciation can be attributed to difficulties not only with production but also perception (cf. section 2.4.). Donegan and Stampe (1979) state that from adolescence the production and perception of L2 words is limited by the residual processes. What is more, the failure to constrain L1 processes leads to a regular phonetic change, which most frequently passes unnoticed by the learner, except by listeners. Thus the acquisition of L2 pronunciation must involve the mastery of both abilities. Major (1987) claims that there exists a considerable variability among learners, as some learners are endowed with good perceptual abilities because their mental representations are similar to those of native speakers, while in learners with poor perceptual abilities the representations are closer to their native language[33].

The issue that received a considerable amount of attention in SLA is the problem of interference and "developmental" (however disputable the term) errors reflecting deviation in L2 learner's speech. The causes for each are unquestionable: the former result from the influence of the learner's mother tongue, whereas the latter are a consequence of developmental factors. What is problematic, though, is drawing a clear distinction between them. Within the natural model of acquisition this problematic differentiation is replaced by the concepts of processes and rules, while errors are invariantly seen as the result of the learner's failing to suppress or limit a given process (Abrahamsson 1996).

---

[33] However, the relationship between production and perception is complex, and the priority of perception is not always obvious, cf. the motor theory of speech perception (e.g. Liberman – Mattingly 1985).

The problem of L1-L2 relationship and universals within the Natural Phonology framework is also found in studies concerning markedness (cf. section 2.5.2.). Dziubalska-Kołaczyk (1989) proposed a theory of relative markedness, in which she states that the learner can unsuppress a universal process more easily than suppress a process which earlier eliminated a universal one. Therefore, one can predict that it is easier for a Polish learner of English to acquire aspiration, a foregrounding process suppressed in Polish, than to suppress the process of final obstruent devoicing, which is both universal and unmarked. Moreover, obstruent devoicing in word-final position is backgrounding and serves the ease of articulation, which adds to the difficulty in its suppression.

One can generalize that listener-friendly, foregrounding fortitions are easier for the L2 learner and thus are more likely to be consciously acquired, as compared with speaker-friendly, backgrounding lenitions. Since the latter are grounded on the L1 and universal phonology, they are less susceptible to conscious learning, which hinders acquisition.

The question to what degree the acquisition of phonology is conscious has received a considerable amount of attention. Dziubalska-Kołaczyk (cf. 1990a) assumed that learners in their mastery of L2 pronunciation may follow paths varying between two extremes: entirely subconscious acquisition at one extreme or a purely conscious path on the other. In the former case, in the L2 learner's mind the processes that have been either suppressed or limited during L1 acquisition become activated once again. In the latter situation, the phonological processes have to be learned as rules. The intermediate paths may involve, e.g. conscious learning of those processes that are obligatory, subconscious activation of only a part of the set of latent processes, or the lack of success in the suppression of L1 processes which could explain foreign accent in the learner's speech.

Dziubalska-Kołaczyk (1989, 1990a) formulated a model of phonological acquisition, according to which the L2 learner consciously learns both processes and rules not only on the level of production but also perception. Perception concerns sound intentions, or phonemes, not their surface realizations, i.e. surface phonetic segments. At the initial stages of L2 acquisition, the learner's perceptive abilities are limited to L1-specific sound intentions; therefore he/she needs to learn to perceive L2 surface realizations in order to recognize L2 sound intentions. Such learning can occur as a result of the learner's access to universal processes, and can be significantly facilitated by formal instruction (Dziubalska-Kołaczyk 1989). Subsequent experiments involving natural-setting and formal-setting learners further investigated the influence of the different contexts on the effectiveness of phonological acquisition. The experiments concentrated on phonological processes present in the subjects' performance, especially context-sensitive processes, such as aspiration, word-final devoicing, palatalization or assimilation, and context-free processes, for instance vowel discrimination. The results revealed that those learners who had received formal training not only learned L2 segments more successfully, but also dealt significantly better with the processes notoriously difficult, e.g. final obstruent devoicing (Dziubalska-Kołaczyk 1990a).

Similar results were obtained in an experiment examining phonostylistic processes in casual speech produced by learners of English, as well as their abilities to perceive stylistic variation. Formal-setting learners considerably outperformed those learning in the natural setting both in the production and perception tasks, as the latter could apply some memorized fast speech forms yet did not use casual speech processes productively in new contexts. Furthermore, they appeared to be less sensitive to the differences between styles. The results led to a conclusion that the successful acquisition of L2 phonological processes involves conscious learning for both formal and naturalistic settings, rather than uncontrolled acquisition (Dziubalska-Kołaczyk 1990a, 1990b). Thus, formal instruction enhances the effectiveness of phonological acquisition, especially at its beginning stage when the learner needs to learn to recognize L2 sounds (outputs). However, instruction can also assist at the subsequent stages of phonological acquisition. At the second stage, the learner learns how to use the perceived outputs in order to decode the underlying intentions (inputs), and finally, at the third stage, when the learner begins to associate inputs with outputs and dealing with L2 processes until they are reactivated in their original form, i.e. until they are acquired (Dziubalska-Kołaczyk 1990b).

On the basis of the experiments described above, Dziubalska-Kołaczyk (cf. 1990b) suggested that the order of phonological acquisition in SLA occurs in reverse order as compared with learning L1 phonology. For an infant, it is easier to decode inputs when their distance from outputs is small. Moreover, a child acquires L1 processes through suppression, limitation and ordering of the universal ones, with time gradually acquiring phonostylistic processes and morphological rules. For an L2 learner, though, a greater distance between inputs and outputs is advantageous. S/he first acquires those elements which are easier to perceive, i.e. morphological rules, then progresses to those less and less evident on the surface, i.e. phonostylistic processes, then allophonic and segment-formation processes. Unlike L1 phonological acquisition, the L2 learner needs to cope with a number of sociolinguistic and psycholinguistic factors, e.g. motivation, aptitude, etc., which can either hinder or assist SLA. The model, however, predicts that complete acquisition of L2 processes is possible.

Recapitulating, Natural Phonology, which combines diachronic, synchronic and child language data, provides a sound basis for an account of language acquisition. Even though it mostly concerned segments and phonostylistics, Natural Phonology can account for suprasegmental phenomena; hence it became the basis for the following model of the acquisition of prosody. In the following subsections intonational universals and typology will be discussed first, then the intonational processes taking place in the learner's language and the influence of the extralinguistic factors on the ultimate attainment of L2 learners.

3.2. Intonational universals

Intonation, as all other levels of language, possesses underlying universal rules. According to Vassière (1995), those universals include the biological, psychological and ethological factors which govern fundamental intonational contours (cf. the three biological codes by Ohala (1984)). They are produced naturally due to the perceptual abilities and the phonatory systems of language-users. Because of the influence of these factors, there exists one basic archetypal shape, namely a rise preceding a fall, while the remaining patterns are derived from this archetypal configuration. Moreover, the intonational contours of longer utterances are composed of those derivative patterns occurring recursively. Pitch movements of the basic pattern have become conventionalized as natural 'signifiants' and 'signifiés', thus a rise began to mean the beginning of a semantic unit or an incomplete utterance, a fall, on the other hand, began to signal the termination of the unit (ethological context). Also a 'valley' between a fall and a rise in f_0 track is universally meaningful: when it occurs, it indicates disjunction.

The archetypal intonational shape results from the way utterances are produced. First, the lungs must be filled with air, which is later pushed out during phonation causing the vocal folds to vibrate. The air pressure is greater at the beginning of the utterance, therefore the f_0 rises then and slowly declines towards the end of phonation, which explains the universal declination of the f_0 track and the fact that such a contour universally signals completion. On the other hand, the universal interpretation of the non-fall as continuation or incompleteness derives from the same natural characteristic of the human voice. As Karcevskij (1931) explained, such meaning results from the fact that each global intonational contour has its contrasting counterpart, so if the fall indicates completion, the non-fall must mean the opposite.

Although the archetypal intonational contours depend on human perceptual and phonatory systems, languages grammaticalise those universals in different ways. The way the affective and attitudinal functions of intonation are expressed differ from language to language. Vassière does not include them into her discussions of intonational universals, as she assumes them to be too unpredictable and too language-specific. It seems, however, that languages do constrain their intonation, regardless of the speaker's attitude or emotions, depending on their language typology. The universals can be traced in discourse management and turn-taking rather than in the purely linguistic functions (Välimaa-Blum 1999).

3.3. Intonational typology

Donegan and Stampe (1993) claim that one cannot discuss the intonational structure without reference to morphosyntax, since languages are "not just a collection of autonomous parts, but […] a harmonious and self-contained whole" (Donegan and Stampe

1983: 337). They analysed the interdependence between word order and phrasal accent placement in Munda and Mon-Khmer language families, whose major syntactic structures form mirror opposites. The word order in Munda is operator-first, whereas in Mon-Khmer the order is operator-last. Yet the operand/operator relation does not belong solely to the domain of syntax but is closely linked to the information structure, where the operand relates to the given information and the operator is asserted. The new/given elements are further linked to the phrase accent level, as new information takes the accent vis-à-vis old information[34]. One can hypothesize that this interdependence can be traced in all languages, cf.:

(1)

Her ' feelings were hurt.   It hurt her 'feelings.

(Donegan & Stampe 1993)

In these examples, the word 'feelings' is the operator[35] and bears a phrase accent taken from its operand, even in the case of a changed word order. Since the operator carries new information and thus receives the phrasal accent, the word order which is operator-first bears phrase-initial accent, while operator-last order receives phrase-final accent.

According to Donegan and Stampe phrase accent and word accent is also related. In Munda languages falling phrase accent and falling word accent co-occur and in Mon-Khmer the same relation exists between rising phrase accent and rising word accent. They found this relation also in other languages of India and South-East Asia. A further relationship was revealed, that is the one between falling accent and syllable rhythm, which can be contrasted with the rising accent and word rhythm. Donegan and Stampe hypothesise that such a link can exist in European languages as well. They claim therefore that word order, rhythm and syntax are interdependent and languages should be analysed holistically.

This analysis leads to distinguishing two basic language types on the extremes of a typological continuum:

1. languages with falling initial phrase accent and a predominantly OV order (i.e. dependent-head order) and
2. languages with rising final phrase accent and VO (head-dependent) order, typical of modern Europe.

Intonational systems can also be divided into two basic types: one with a rising and the other with a falling contour. The characteristics of the falling-contour type include word-initial stress and free constituent order, enclisis, grammatical cases, suffixing morphology

---

[34] Cf. Bolinger (1958).
[35] an operator can be replaced with a interrogative word to be questioned vis-à-vis the operand: *What was hurt?* The operand, on the other hand, has no interrogative counterpart (Donegan & Stampe 1983).

and adjectives tend to precede nouns. The rising type is characterized by word-final stress, rigid word order, adjectives following nouns, proclisis, analytic syntax and prefixing morphology (Donegan & Stampe 1983). The following table summarises the holistic organization of the two basic language types.

TYPE	FALLING	RISING
Phrases	Accent initial ("falling")	Accent final ("rising")
Order	Dependent-head order / left branching / OV	Head-dependent order / right branching / VO
Words	Accent relative to beginning (1st or 2nd syllable or mora)	Accent relative to end of word (1st or 2nd syllable or mora from end)
Affixation	Suffixal	Prefixal
Grammar	"Synthetic" (inflecting, free word order)	"Analytic" (non-inflecting, rigid word order)
Timing	Isomoric	Isoaccentual ("stress-timed")
Syllables	CV canon, CV-X internal structure	Accent-dependent canon, C-VX internal structure
Vocalism	Simple nuclei, stable, harmonic tendencies	Accentual nuclei complex and shifting, unaccentual nuclei reductive
Cons'ism	Stable, clusters geminate	Shifting, clusters nongeminate
Tone	Level	Contour
Verse	Quant. metre, "falling" lines, initial rhyme (alliteration)	Accentual metre, "rising" lines, end-rhyme
Music	Modality, unison/monody	Harmony, polyphony
Examples	Proto-Indo-European Munda (India) Tibeto-Burman	Modern Indo-European (excluding India) Mon-Khmer (mainland SE Asia) Sinetic (Chinese)

Table 3.1. Polarity of pansystemic structure and drift, relative to rhythm (Donegan & Stampe 1996[36])

The falling and rising types are extremes on a typological continuum with particular languages being closer to one end or the other. When analysing English and Polish, one can notice the following differences.

3.3.1. English (Välimaa-Blum 1999: 302)

English linguistic system is of a mixed type:

---

[36] The table is based on the material from the Workshop on Typology and Naturalness, SLE 24, Klagenfurt, Austria, September 5, 1996.

1. Content words carry at least one stressed syllable, yet the primary stress placement is not fixed.
2. Function words tend to undergo reduction, unless they are stressed for emphasis.
3. The global intonation is falling, characterized by declination and final lowering, although interrogatives and incomplete utterances can end with a non-fall.
4. *Wh*-questions are formed by morphology and syntax, which makes rising intonation optional. Y/N questions do not contain a question word, so they carry the rise or the fall-rise. In declarative questions, whose constituent order is like in statements, the rise is obligatory.
5. An intonational phrase (IP) can comprise a number of pitch accents, each including either high (H*) or low (L*) tones, phrase accents, such as H⁻ if the pitch accents are rising (e.g. L*+ H⁻), or L⁻ if they are falling (H*+ L⁻), and boundary tones (H% or L%). An IP can also contain one or more intermediate phrases. These make English intonation considerably variegated.
6. English intonation serves as means of expressing the information structure and attitude, as well as conveying those elements of grammatical information which are not signaled by synatx.
7. Syntax is analytic, word order is inflexible, adjectives precede the head noun.
8. Morphology is prefixing and suffixing.

Because of these typological features English possesses a rich variety of intonational shapes, however, they are constrained, as the height of f_0 is not linguistic.

3.3.2. Polish

Polish is closer to the rising type of Donegan and Stampe:

1. Content words normally carry one stressed syllable, although in longer ones there can be secondary stress. Stress placement is fixed.
2. Function words are usually not reduced, although this can occur in casual speech.
3. The canonical intonation is falling, with declination and final lowering. Questions and unfinished utterances can end with a rise or a level tone.
4. Polish Y/N questions can have word ordering like in statements and can include the question word *czy*. In *czy*-questions the rise is optional (to avoid double marking) but when *czy* is not used, the question has a rise. Still, double marking of *czy*-questions is not infrequent – accompanying rising intonation indicates emphasis, involvement or politeness (rising intonation in Polish is perceived as polite).
5. An IP, like in English, can include a number of pitch accents, either high (H*) or low (L*) tones, phrase accents, such as H⁻ if the pitch accents are rising (e.g. L*+ H⁻), or L⁻ if they are falling (H*+ L⁻), and boundary tones (H% or L%). An IP can also contain one or more intermediate phrases.

6. The main function of intonation is to express the information structure and attitude, as well as, though to a lesser degree, grammatical relations.
7. Syntax is synthetic and word order is relatively free: adjectives can go before or after the head noun, which may influence the meaning of the noun phrase. There are 7 grammatical cases and no articles but demonstratives can mark definiteness.
8. Morphology is much richer than in English: prefixing, suffixing and infixing.

These typological features cause Polish inventory of intonational shapes less varied than in English, moreover there exists a noticeable preference for falling and level tones, since a lot of meaning, which in English is expressed by intonation, is carried by other means, e.g. rich morphology, aspect or free constituent order. They influence the information structure, so that new information and focus can be marked by moving a particular element to the beginning or end of a sentence, additionally to accentuation.

The following table summarizing the differences between English and Polish reveals that typologically[37] the two languages are close.

	English	*Polish*
Phrases	Tendency for accent final (close to "rising" type)	Tendency for accent final (close to "rising" type)
Order	Head-first (pre- and post-modifiers but complements always follow the head), right branching / VO	Head-dependent or dependent-head order (with change in meaning) / right branching / VO
Words	Accent relative to end of word, but not fixed (1st, 2nd or 3rd syllable from end)	Accent relative to end of word, fixed (2nd or, rarely, 3rd syllable from end)
Affixation	Suffixal , prefixal	Suffixal, prefixal, infixal
Grammar	"Analytic" (rather non-inflecting, rather rigid word order)	"Synthetic" (inflecting, rather free word order), 7 cases
Timing	Isoaccentual ("stress-timed"), reduction of unstressed function words, feet are quantity-sensitive to the rime, extrametrical syllable on the right	Mixed (with elements of "stress-timing"), occasional reduction of unstressed function words
Syllables	Accent-dependent canon, C-VX internal structure, extrametricality	Accent-dependent canon, C-VX internal structure
Vocalism	Accentual nuclei complex and shifting, unaccentual nuclei reductive	Accentual nuclei complex and shifting, occasionally unaccentual nuclei reductive
Cons'ism	Clusters nongeminate	Clusters can geminate
Tone	Contour: 13 tones (with tritonal accents)	Contour: 6 tones

Table 3.2. The typology of English and Polish

---

[37] From the structural point of view, Polish belongs to synthetic languages, rich in inflections, while English, utilizing relatively few affixes, is close to analytic languages.

3.4. The model of the acquisition of intonation

The proposed model of the acquisition of intonation strives to follow the advice given by Ellis (1994) who stated that a model of L2 acquisition must concentrate on three items: explanation of the relationship between L2 input and knowledge, the representation of L2 knowledge and finally, the relationship between L2 knowledge and output. It also attempts to incorporate theories of intonational phonology in order to provide a comprehensive model of the acquisition of intonational phonology. The findings of the research into L1 and L2 acquisition within the naturalistic framework provide background for the model. Most studies into the acquisition of intonation concentrate on the influence of the native language. The present model concentrates mainly on the interaction between typological preferences and universal processes in the course of mastering L2 intonation, therefore the resulting processes will be accounted for first, following Dziubalska-Kołaczyk's (1990a) framework. The description is accompanied by the account of the extralinguistic component including socio- and psycholinguistic factors, as well as metacompetence, which belongs to the cognitive part of pronunciation acquisition relevant to adult learners (cf. Wrembel 2005).

3.4.1. Processes in the acquisition of intonation

As it was noted in section 3.1., Natural Phonology serves as the basis for the proposed model. The natural framework postulates that during the acquisition of L1 processes occur to help the child to deal with articulatory and perceptual difficulties. The child discovers them unconsciously, without any cognitive processing.

Prosody is the first aspect of the mother tongue to be learned, and possibly the last to be lost in language attrition – its acquisition probably starts already during the prenatal period, as the embryo begins to hear around the fifth month of development and responds to the mother's voice with body movements. At this stage it is clearly the perception of L1 tone system that is acquired. The production of changing pitch patterns begins to be observable from the sixteenth week of the baby's life, i.e. earlier than the babbling stage (starting approximately at the 5-6 month of age) during which a series of consonant-vowel syllables are pronounced. Around the twelfth month of age or later stress and intonation patterns are imposed on babbling which begins to sound like L1. The universal preference to apply simple tones is constrained by the communicative needs of the child: varying pitch patterns help the baby to attract adults' attention and encourage interaction. Adults instinctively respond to babbling with a special kind of talk, i.e. child-directed speech, the characteristic feature of which are adjustments in pronunciation necessary to tune their speech to the perceptive abilities of the child. As a result, adults produce utterances with higher-range, exaggerated and varied ('sing-song') intonation (c.f. Foley & Thompson 2003). By the age of three the universal basic tones grow in complexity as required by the grammatical, attitudinal and discourse functions of the L1. Consequently, the intonational system of the child becomes adult-like (cf. Zharkowa 2002).

The processes that operate during the acquisition of L1 intonation involve, as in the case of segments and phonostylistics, suppression, limitation and reordering. The child needs to suppress the natural tendency of pitch to decline towards the end of an utterance as a consequence of decreasing air pressure, e.g. when learning to apply the rise in questions. Limitation involves constraining processes in certain contexts only, e.g. the child needs to learn to restrict the rise to yes/no questions but not *wh*-interrogatives. If the ambient language lacks additional grammatical markers, e.g. has a poor inflectional system, the two basic patterns (the fall and the rise) are combined into more complex tones to convey the meanings which morphology and/or syntax fail to express: the fewer grammatical markers, the richer tone inventory. The processes are subsequently reordered until the system approximates that of the adult.

The L2 learners, on the other hand, already possess the phonological system of their native language consisting of a limited set of processes, underlying representations and rules. Their access to universal process types is thus limited, so in order to utilize them in L2 acquisition, they need to consciously unsuppress, limit and re-order certain processes (see Figure 3.2., based on Dziubalska-Kołaczyk 1990a, Wrembel 2005).
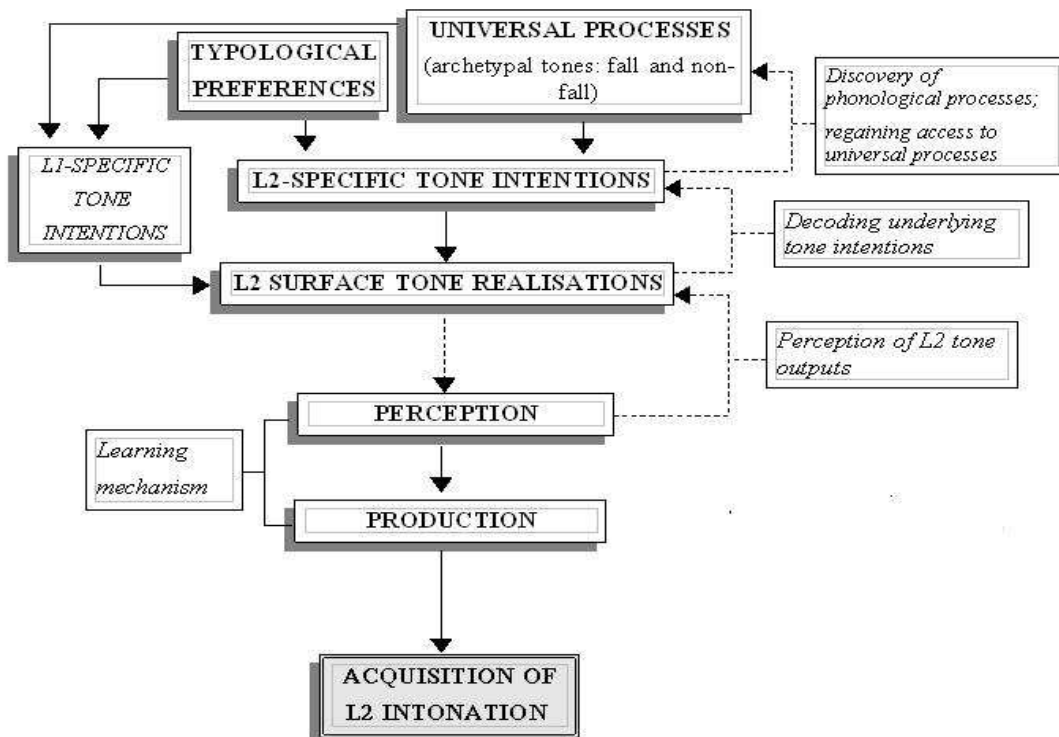

Figure 3.2. A model of the acquisition of L2 intonation (based on Dziubalska-Kołaczyk 1990a, Wrembel 2005)

Perception and production constitute the learning mechanism. The L2 learners are equipped with the residue of universal processes constrained in the process of L1 acquisition which, at initial stages of L2 acquisition, prevents them from perceiving L2 intonational patterns and, consequently, they substitute a difficult representation with a L1 pattern. Therefore, first they need to enhance their perceptual abilities of L2 surface realisations (L2 tonal outputs). For instance, learners whose mother tongue is a language with poorer tonal inventory than the L2 (like Polish), may not perceive the complex f_0 fluctuations in tritonal patterns (e.g. English rise-fall-rise, L+H* L H%) and classify them as a simple rise or a fall-rise. Only then will the learners be able to decode the L2-specific tonal intentions, reactivate universal processes and acquire L2-specific preferences.

According to the tenets of Natural Phonology, in any language preferences are in conflict, yet those which are the most natural, "cognitively simple, easily accessible, elementary and universally preferred, i.e. derivable from human nature" (Dressler 1999) win. Each language differently resolves the conflicts, either by choosing universal or typological preferences. In the case of intonation, which constitutes an integral part of a larger system, both universal and typological processes operate, yet typological preferences seem to be more influential. This is caused by the fact that when the learner's phonological system is confronted with the L2 system, the L1 processes are automatically and unconsciously employed. Thus in interlanguage, the L1 (e.g. the preference for simple tones in Polish) and L2 preferences (e.g. the preference for more complex tones in English) are in conflict resolved mainly by typological and local, language-specific preferences, in addition to universal ones.

As the study described in the following chapter reveals, despite the strong preference for L1 patterns, universal processes (the universal preference of falling tones and narrower pitch span) do operate during the acquisition of intonation, especially at the initial stages. The explanation for this phenomenon comes from the gestural law which holds that speakers favour those combinations which do not involve extensive movements of the articulators[38], here the vibrations of the vocal folds (e.g. Janson 1986 for CV sequences; Ohala 1984, 1990; Gussenhoven – Chen 2000, Gussenhoven 2001, 2002a for the biological codes). Here, the 'more natural' choices mean employing simple tones, that is an unmarked fall and a less marked rise. The more complex, i.e. more marked, tones will appear at the next stages of the process of acquisition – if they appear at all. Although phonetic principles solely are not enough to account for the preference or acquisition processes, which is evident in, for instance, 'unecological' segmental inventories of Caucasian languages, they can be seen as laws complementing the principles of Natural Phonology.

---

[38] Cf. The principle of least effort.

3.4.2. Metacompetence

The learner's success in L2 phonological acquisition, or its lack, is not only determined by socio- and psycholinguistic factors, such as age, attitudes, motivation, etc. (see chapter 2), but also by their metalinguistic awareness (for an extensive study into the role of metacompetence in phonological acquisition see Wrembel 2005).

Wrembel (2005:169-171) advocated that metacompetence serves the learner as a facilitating device. The construct is multilevel, consisting of three components:

1. metalinguistic consciousness which includes language awareness, intentions, knowledge;
2. explicit formal instruction providing theoretical background and ensuring self-monitoring and reflective feedback;
3. L1 competence which helps to counteract L1 interference.

Metacompetence assists the phonological input to become conscious intake by enhancing perceptive skills. Raising the learners' awareness of L2 phonological characteristics helps them to decipher underlying sound and tonal intentions, form suitable L2 realisations and eventually reactivate latent universal processes. Moreover, theoretical knowledge allows the learners to control their production so that they can self-monitor and self-correct their performance.

Metacompetence is especially important in the acquisition of intonation. On the basis of Dziubalska-Kołaczyk's (1990a, 1990b) studies, despite a dearth of studies into the mastery of this element of phonology in natural- and formal-setting learners, it can be assumed that only those whose attention was drawn to prosody can acquire, at least partially, L2 intonation. As Kenworthy (1987) observed, people tend to react to the melody of the unknown language before they learn to distinguish individual sounds. They often judge the foreign speech to be 'melodic' and their native language to be 'flat'. This indicates that intonation operates at the unconscious level, and thus it can be 'overlooked' in L2 learning. Therefore raising the learner's awareness as well as formal training in perception and articulation of intonation seem to be a necessary component of successful phonological acquisition. This issue, however, requires more investigation.

3.4.3. Extralinguistic factors in the acquisition of intonation

The socio- and psycholinguistic factors are considered an important element of any model of phonological acquisition, since language functions in society (cf. Halliday 1973) and its speakers differ in their psychological and emotional features. This area has been widely researched and a number of learner variables have been enumerated. For instance, Ellis

(1991) groups these influences into two categories of personal factors (including group dynamics, attitudes to the teacher and course materials, and individual learning techniques) and general ones (age, intelligence, language aptitude, cognitive style, attitudes, motivation and personality). Leather and James (1996) refer to maturational (age and the stage of cognitive development), individual (motivation, personality variables, sex, oral and auditory capacities) and social constraints (social acceptance and social distance). These factors have been incorporated into several models of SLA, e.g. Ellis's (1991) Variability Competence Model, or Gardner's (1985) Socio-Educational Model.

Gardner's model is particularly interesting, since it concerns also formal language training. Despite the lack of extensive research into the acquisition of intonation in naturalistic settings, one may hypothesise that formal instruction plays a crucial role there for the reasons mentioned above in the discussion of metacompetence and awareness rising. The model postulates an interrelation between four elements of L2 learning which affect acquisition, namely the socio-cultural milieu, individual learner differences, L2 acquisition contexts and linguistic outcomes. The first element, the social milieu, encompasses the learner's beliefs about the second language community and the second language itself. The beliefs directly influence the learner's motivation which, together with aptitude and intelligence belongs to the second influencing factor, individual learner differences. The learner's motivation can be either of instrumental orientation when s/he learns the target language for practical purposes, or of integrative one when the learner wishes to become part of the L2 society. The learner's motivation can change in the course of acquisition and so can his/her attitude. Aptitude and intelligence, however, are invariable, yet affect acquisition in two types of learning contexts, that is formal language training and informal language experience, which ultimately results in second language competence.

Summing up, the present model adopts those factors which are believed to have the greatest influence on the acquisition of L2 pronunciation in general. Individual learner differences and sociolinguistic factors are based on the components of Gardner's Socio-Educational Model, the former including aptitude, intelligence and age, while the latter involves the learner's beliefs and attitudes towards the target language community, as well as their motivation. In the acquisition of L2 intonation, however, motor skills and the learner's auditory abilities which allow them to perceive and mimic L2 intonational patterns cannot be disregarded (Figure 3.3.).
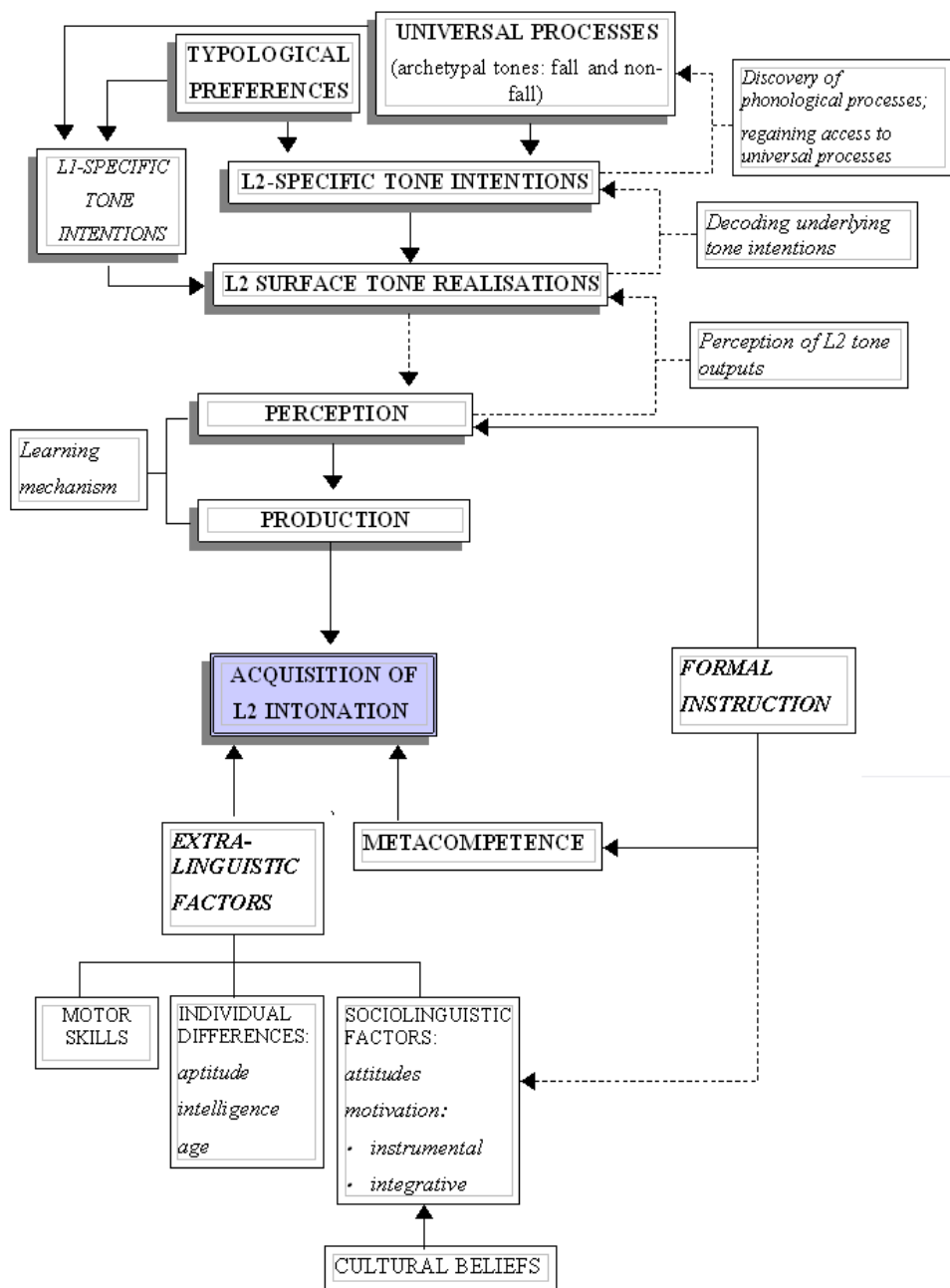
TYPOLOGICAL
PREFERENCES

UNIVERSAL PROCESSES
(archetypal tones: fall and non-fall)

Discovery of phonological processes; regaining access to universal processes

L1-SPECIFIC TONE INTENTIONS

L2-SPECIFIC TONE INTENTIONS

Decoding underlying tone intentions

L2 SURFACE TONE REALISATIONS

Perception of L2 tone outputs

PERCEPTION

Learning mechanism

PRODUCTION

ACQUISITION OF
L2 INTONATION

FORMAL INSTRUCTION

EXTRA-LINGUISTIC FACTORS

METACOMPETENCE

MOTOR
SKILLS

INDIVIDUAL
DIFFERENCES:

aptitude

intelligence

age

SOCIOLINGUISTIC
FACTORS:

attitudes

motivation:
- *instrumental*
- *integrative*

CULTURAL BELIEFS

Figure 3.3. The model of the acquisition of intonation

88

CHAPTER 4: Acquisition of English prosody – the empirical study

4.1. The empirical study

The aim of the research presented in this chapter is to analyse and assess the performance of Polish learners of English as the second language. The subsequent subsections present the analytic procedures which were chosen as most reliable for the description and evaluation of the collected utterances, and the results of their analysis.

4.1.1. Methods of analysis

For the analysis of prosodic phenomena one can choose one of the following analytic techniques: using the f_0 contour as a 'narrow phonetic transcription' followed by drawing stylised contours (Beckman 1995) or a combined auditory/acoustic technique (Grabe 1998). The former can prove useful when the phonologically distinct categories of the language being analysed are already known to the analyst. This approach has not been adopted here for several reasons. Firstly, f_0 cannot be seen as solely acoustic correlate of pitch; length and intensity are important for the auditory perception of intonation as well. Secondly, there exist certain microprosodic variations caused by segmental structure, e.g. by voiceless obstruents. Finally, transcription requires discrete phonetic categories (such as pitch accents), which are hard to be provided solely by the variable f_0. In sum, one can claim that the analysis of intonation should include not only the picture of the f_0 trace, but also of intensity, amplitude and spectrum.

The latter approach, which seems better suited for analysing speech data, constitutes a compromise between the acoustic and auditory techniques, which agrees with Crystal's (1969: 14) view: the analyst should begin the analysis with an auditory method supplementing their judgements with f_0 tracks.

4.1.2. Speech material

The aim of the study presented in this dissertation was to discover the influence of universal and language-specific, determined by language typology, phonological processes on the course of acquisition of English intonational system. In order to collect speech data suitable for the analysis, both the materials and their possible interpretations must be constrained. The same context should elicit a limited number of intonational patterns, which will help to judge whether the appropriate use of a given contour has been acquired

or not. The experimental materials were in English and included a dialogue to be memorised and later acted out by the subjects and a BBC weather forecast to be read aloud (see Appendix A). The dialogue was to ensure a semi-spontaneous, conversational speaking style, which was considered as likely to elicit typically British intonation patterns and livelier, more emotional speech. The weather forecast text provided a different context, more likely to produce a standardised speech style. As many as 1149 intonational phrases were collected, 579 per group A and 570 per group B.

4.1.3. Elicitation

The subjects were recorded in a phonetics laboratory at the Teachers Training College in Nowy Sącz. The recordings were made with the use of a notebook, an MT 382 microphone and the GoldWave program, and saved as .waw files, later analysed by *PRAAT* (Boersma – Weenink 2002). Group A was recorded at the beginning of the academic year before they began their formal education in English phonetics, while group B performed during the last month of the second year of their education in the College when they were already finishing their course in phonetics.

4.1.4. Subjects

There were two groups involved in the experiment: group A consisted of 15 first year students and group B included 15 second year students of English philology. The assumption behind such a choice was that adult learners' pronunciation can benefit from formal instruction and intensive practice, therefore comparison between instructed and non-instructed groups can give some insight into the process of acquisition. The average age of group A was 20.5, the average age of group B was 21. The average age of first exposure to English in group A was 10.4 and of group B: 10.7. In group A 1 student spent a year in England, in group B 3 people lived in England for a short period of time (4 times for 1 month, 3 months, 1 year). The groups did not consist of the same people, as after 4 semesters most 1st year subjects dropped out after their first exam session. This is the reason why in all the statistical calculations the groups were treated as the independent variables. Since intonation of the second language is acquired late in the learning process, only advanced learners were believed to be suitable for the study.

All students came from southern Poland and speak standard Polish. Although there is little information about differences between intonation systems in Polish dialects, yet the relatively small region from which the subjects came from ensures fairly similar variety of Polish they use. Although the materials were designed for the FCE level, for two speakers from group A the texts appeared noticeably difficult, for which reason some utterances

were unintelligible and had to be excluded. For comparison, the same texts were performed by two British speakers, claiming to speak Estuary English.

4.1.5. Categorising and labelling of tones

Individual tones obtained from the recordings were categorised on the acoustic (with the help of *PRAAT*) and perceptual basis. Modern intonational models avoid distinguishing between "low fall" or "high fall" tones, yet for the present analysis such differentiation seems relevant, as the subjects' production of the falling tones differed significantly in the two groups. The intonational contours of the non-instructed group A were detectably flatter than the melody produced by group B, for which single categories "H*LL% (fall)" and "L*HH% (rise)" could not account. Therefore, the following tones were distinguished: L*LL% (low fall) and L*LH% (low rise) for a "small" decrease in pitch (cf. Ladd 1996), in contrast to H*LL% (fall) and L*HH% (rise) for tones traditionally called "high fall" and "high rise" respectively. Moreover, L*L0% (level) for contours with a "flat" change in pitch was used. The decision to which category a given falling or rising tone should be assigned was based not only on the shape of the f_0 contour but also on the perceptual judgements of three speakers: a native speaker of English, one of Polish and the author. The reason behind this decision was that f_0 and pitch are not linearly related and that, in fact, intonation is a perceptual phenomenon, "relating to listener's judgements as to whether a sound is (…) 'higher' or 'lower' than another, and whether the voice is going 'up' or 'down'" (Cruttenden 1986: 4).
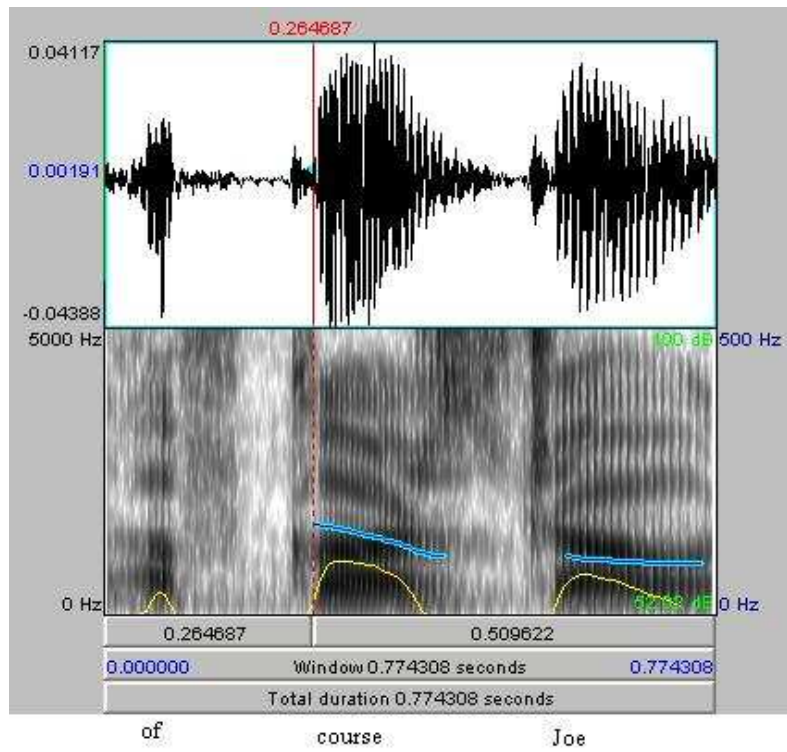
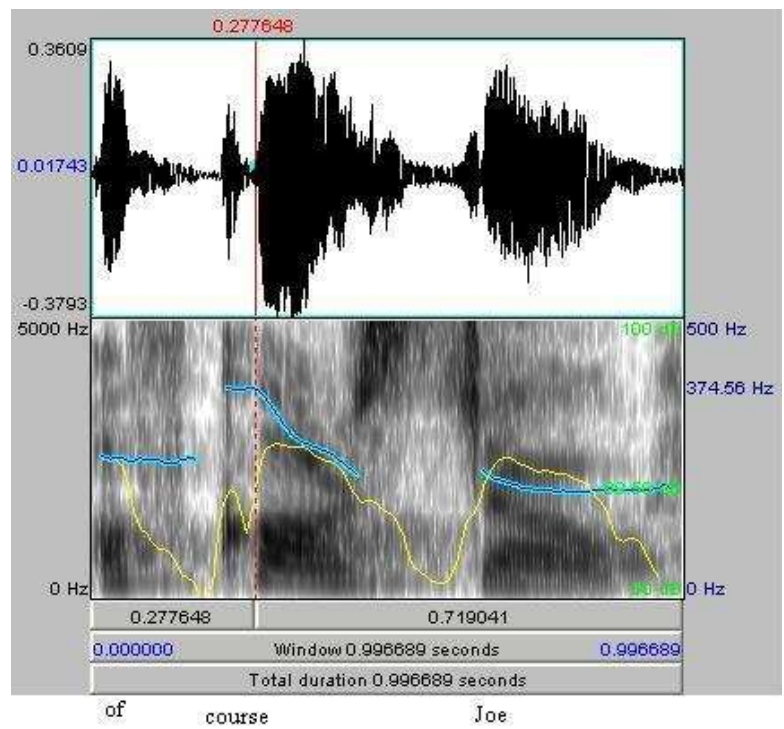Figure 4.1 Example of a low fall (L*LL%)

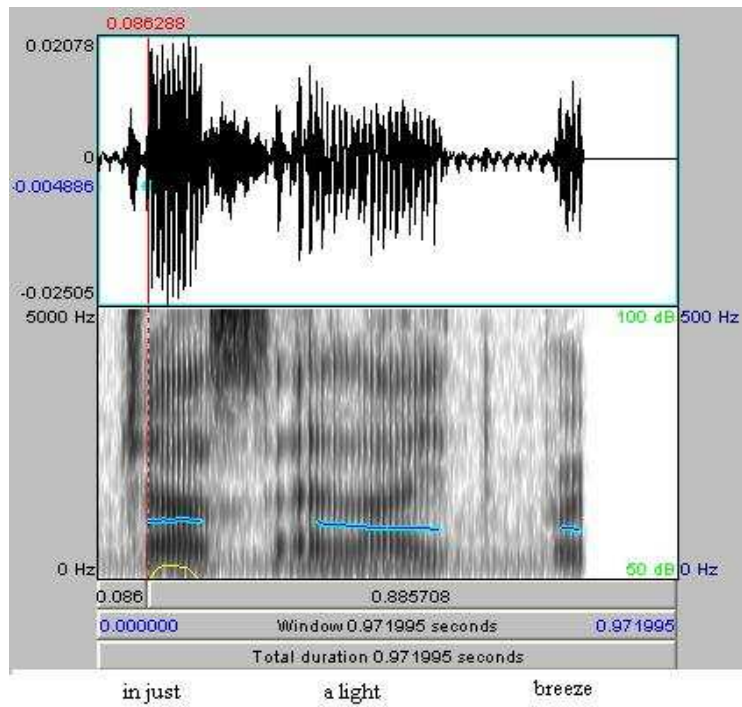
Figure 4.2 Example of a high fall (H*LL%), female voice

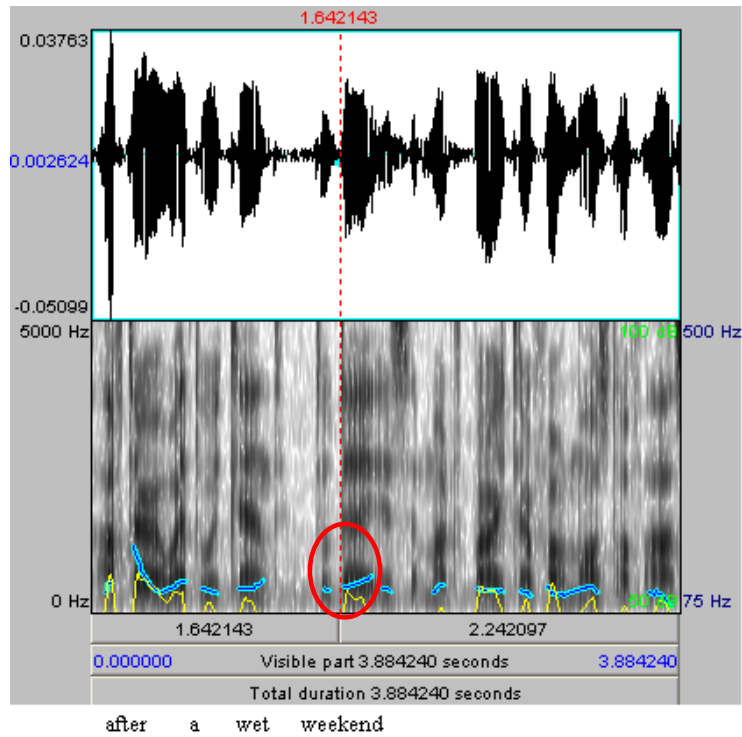Figure 4.3 Example of a level contour (L*L0%), male voice

Figure 4.4 Example of a low rise (L*LH%) at the end of a minor IP, male voice - circled
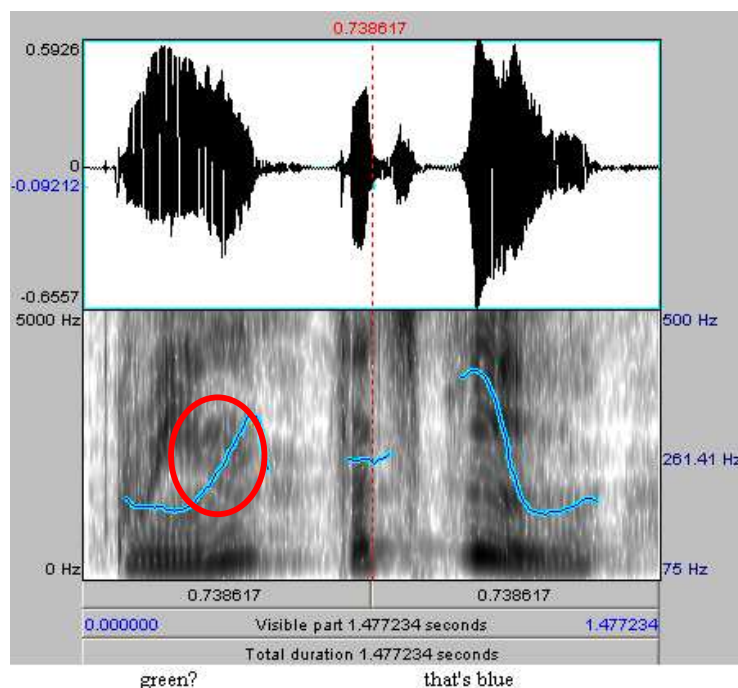
93

Figure 4.5 Example of a high rise (circled), female voice

The labelling of utterances was based on the ToBI system. The auditory impressions of the intonational contours on the tone tier received the following labels:

Pitch accents: **Boundary tones**:

H*, L*, H*+L, L*+H   H%, L%, 0%

The transcription of pitch accents includes two tones accompanied by an asterisk (H* or L*) associated with the accented syllable and a trailing tone of the post-accentual syllable. The label 0% is assigned to the level tone.

Although there exists no single inventory of tones for English or Polish (cf. chapter 1), for the present study 7 tones for Polish and English were distinguished: fall (H*LL%), low fall (L*LL%), level (L*L0%), rise (L*HH%), low rise (L*LH%), fall-rise (H*LH%) and rise-fall (L+H*LL%). Since English and Polish are typologically different languages, their tonal inventories differ significantly. The choice of the above mentioned tones, though, resulted from the analysis of the subjects' speech and is considered to reflect the interlanguage inventory of intonational shapes.

Another step involved the decision which tones would be assumed universal and which language-specific. As it was explained in chapter 3, universal intonational contours are those produced naturally due to the physical abilities of the phonatory systems of

94

language users. Thus the archetypal pitch movements include a rise preceding a fall, parallel to exhaling the air from the lungs. The universal interpretations of the fall is that of completion and the non-fall – of continuation. The archetypal contours (and their combinations) are grammaticalized in a variety of ways by various languages and are assigned various linguistic interpretations, including grammatical, attitudinal, discourse and sociolinguistic ones. In order to distinguish those universal from language-specific tones, Gussenhoven's (2004) universal and linguistic interpretations of the biological codes were employed (cf. chapter 1, Figure 1.11).

Universal interpretations of the archetypal fall include:

a) At the beginning of the IP – continued topic;
b) At the end of the IP – finality (Production Code);
c) Affective interpretation – less surprised, less helpful;
d) Informational interpretation – less urgent (Effort Code).
e) Linguistic interpretation – statement.

Universal interpretations of the archetypal rise include:

a) At the beginning of the IP – new topic;
b) At the end of the IP – continuation (Production Code);
c) Affective interpretation – more surprised, more helpful;
d) Informational interpretation – more urgent (Effort Code).
e) Linguistic interpretation – question.

The language-specific interpretations include the grammatical, attitudinal, discourse and sociolinguistic meanings, typical for English and Polish, described in the following subchapters.

4.1.6. Presentation of data

Various means of presenting intonational contours have been described in previous chapters. It seems, however, that the most representative and objective way of illustrating acoustic evidence is the form of f_0, while for the auditory data stylised contours were chosen. Additionally, an attempt was made to provide more information about the syllables carrying pitch accents by including the spectrum of the analysed utterances. The f_0, spectra and intensity helped in establishing where precisely pitch accents occurred. In the auditory analysis the patterns produced in a given context by the subjects are presented in f_0 diagrams and then compared. The comparison is "cross-speaker", i.e. it involves one particular contour made by the speakers from the two groups in the same context, which allows studying the alignment of the pattern with the segmental structure. Then the choice

of intonational patterns for the same context is described, especially for these contexts which require different contours in English and Polish.

Gussenhoven (2002b, 2004) points to the importance of separating the linguistic from the paralinguistic aspect of intonation. This proved difficult in practice, as in many functions of intonation (e.g. attitudinal, discourse or paralinguistic) the paralinguistic aspect is an inherent element unable to be separated from the phonetic tier. Perhaps phonology of intonation operates on a different level than phonology of segments (cf. Fox 2000, Karpiński 2006).

Another difficulty in the interpretation of the recorded material concerned the gradient character of intonation (Fox 2000). However, despite this feature, intonation is perceived categorically (Gussenhoven 2004). Apart from presenting the distribution of individual tones (falling, rising, etc.), the author employed intonational categories based on grammatical distinctions and some basic dialogue moves (cf. Karpiński 2006).

4.2. The hypothesis

As mentioned in section 4.1.2., it is hypothesized that the acquisition of English intonational system is influenced by universal and language-specific, determined by typology, phonological processes. It is assumed that universal archetypal tones include the following tones, shaped by the tendency for a narrower pitch range (cf. the biological codes):

a) Low fall – a simple tone, requiring little energy, signalling finality.
b) Low rise – a simple tone, requiring little energy, signalling incompleteness.

The following predictions are proposed:

1. Universals and L1 influences will predominate in the non-instructed, less advanced group A.
2. Universals and L1 influences will be overcome by the learned L2 intonation.

Group A will produce significantly more simple low falling and low rising tones than group B, which will reveal the operation of universal processes at the early stage of L2 intonation acquisition. L1 transfer, exemplified by peak alignment, due to the differences between English and Polish, will be more evident in the performance of group A, caused by poorer perception abilities. Moreover, the instructed, more advanced group B will produce more native-like intonational contours, accurate both grammatically and pragmatically.

4.2.1. The influence of universal phonological processes

In order to discover whether universal phonological processes indeed operate in the acquisition of English intonation individual tones were extracted from the two texts performed by the instructed and non-instructed groups. The results are presented in the following charts. Figure 4.6 illustrates the sum of individual tones used in both texts, the dialogue and the monologue. Figures 4.7 and 4.8 depict the distribution of tones in each text.

The non-instructed group A produced more rising contours (116) than the instructed group B (77), however, there occurred more falling tones in the performance of group B (262 tones, summing up the fall and low fall in group B, as opposed to group A's 236 tones in total). The level tone was more frequent in group A (14), unlike the complex tones (fall-rise and rise-fall).
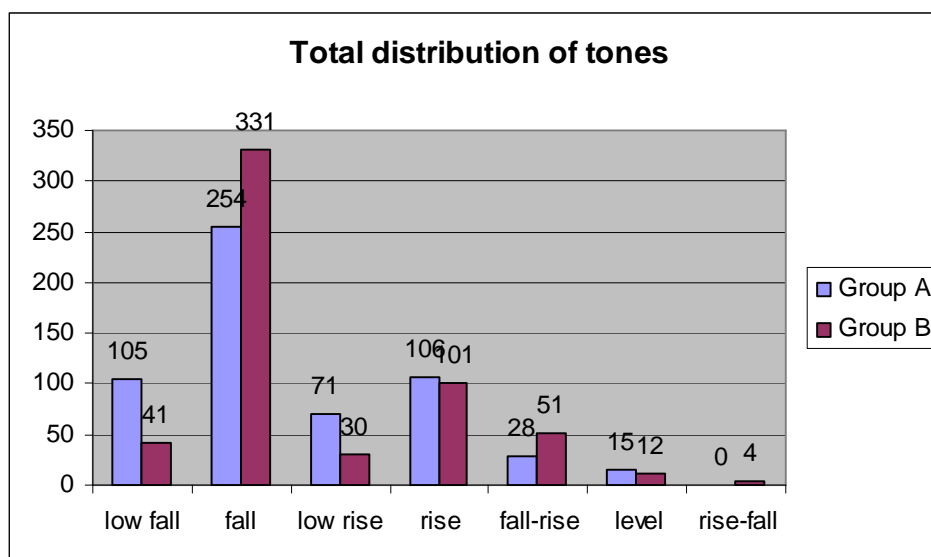

Figure 4.6 The number of tones occurring in both texts
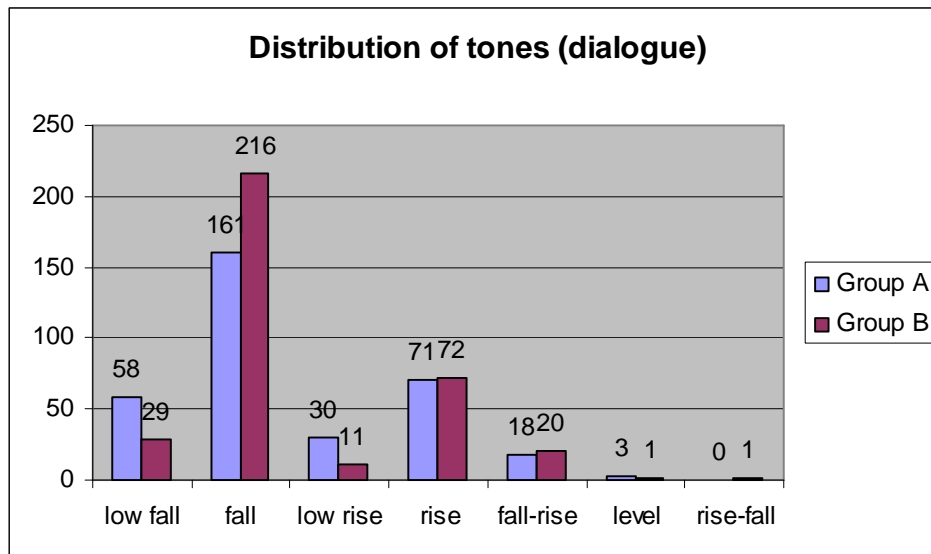
Figure 4.7 The distribution of tones in the dialogue
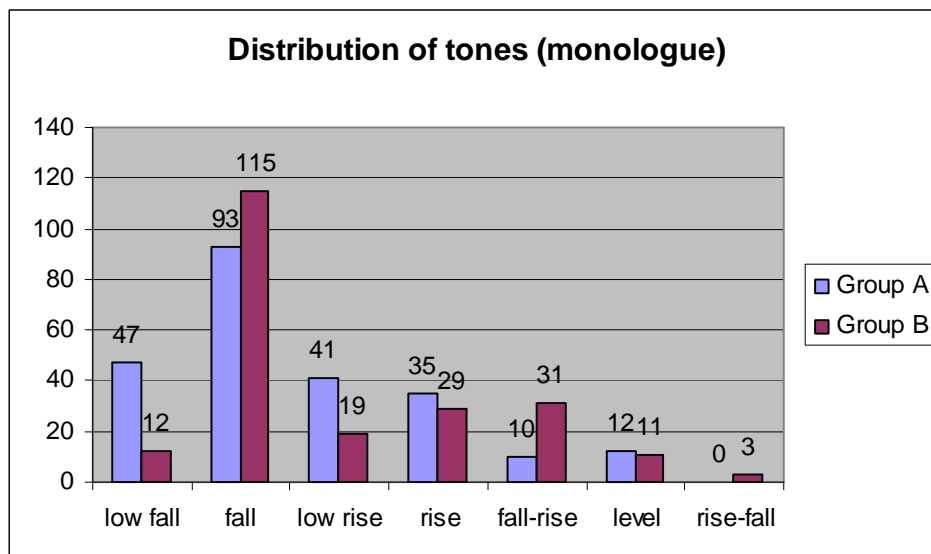
Figure 4.8 The distribution of tones in the monologue

Table 4.1 below presents the more detailed descriptive statistics for the occurrence of particular tones in the production of the two texts by the participants from both groups. The following tables, table 4.2 and 4.3 illustrate the distribution of the tones produced in the dialogue and the monologue respectively. N indicates the number of intonational phrases (IPs) per text, DIAL means "dialogue" and MON indicates "monologue".

IPs: 579	TONE	N (both texts)	mean	SD	variance	SE
Group A, N = 15	Fall (H*L L%)	254	0.439	0.496	0.247	0.0252
	Low fall (L*L L%)	105	0.181	0.385	0.138	0.0189
	Level (L*L 0%)	15	0.026	0.159	0.035	0.0095
	Rise (L*H H%)	106	0.183	0.387	0.15	0.00161
	Low rise (L*LH%)	71	0.123	0.328	0.108	0.0136
	Fall-rise (H*L H%)	28	0.048	0.215	0.046	0.0089
	Rise-fall (L+H* L L%)	0	0	0	0	0
IPs: 570	Fall (H*L L%)	331	0.581	0.493	0.243	0.0207
Group B, N = 15	Low fall (L*L L%)	41	0.072	0.258	0.067	0.0108
	Level (L*L 0%)	12	0.021	0.144	0.021	0.006
	Rise (L*H H%)	101	0.177	0.382	0.146	0.016
	Low rise (L*LH%)	30	0.053	0.223	0.05	0.0094
	Fall-rise (H*L H%)	51	0.089	0.285	0.081	0.012
	Rise-fall (L+H* L L%)	4	0.007	0.083	0.007	0.0035

Table 4.1 The occurrence of individual tones produced by non-instructed students (group A) and instructed ones (group B) in both texts

IPs: 341	TONE	N (DIAL)	mean	SD	variance	SE
Group A, N = 15	Fall (H*L L%)	161	0.472	0.499	0.249	0.027
	Low fall (L*L L%)	58	0.17	0.376	0.141	0.0203
	Level (L*L 0%)	3	0.009	0.093	0.009	0.0051
	Rise (L*H H%)	71	0.208	0.406	0.156	0.022
	Low rise (L*LH%)	30	0.088	0.283	0.08	0.0153
	Fall-rise (H*L H%)	18	0.053	0.224	0.05	0.0121
	Rise-fall (L+H* L L%)	0	0	0	0	0
IPs: 350	Fall (H*L L%)	216	0.617	0.486	0.236	0.026
Group B, N = 15	Low fall (L*L L%)	29	0.083	0.276	0.076	0.0147
	Level (L*L 0%)	1	1	N/A	1	0.0535
	Rise (L*H H%)	72	0.206	0.404	0.163	0.0216
	Low rise (L*LH%)	11	0.031	0.174	0.03	0.0093
	Fall-rise (H*L H%)	20	0.057	0.232	0.054	0.0124
	Rise-fall (L+H* L L%)	1	1	N/A	1	0.0535

Table 4.2 The occurrence of individual tones produced by non-instructed students (group A) and instructed ones (group B) in the dialogue

IPs: 238	TONE	N (MON)	mean	SD	variance	SE
Group A, N = 15	Fall (H*L L%)	93	0.391	0.488	0.238	0.0316
	Low fall (L*L L%)	47	0.197	0.398	0.158	0.0258
	Level (L*L 0%)	12	0.05	0.219	0.048	0.0142
	Rise (L*H H%)	35	0.354	0.354	0.125	0.023
	Low rise (L*LH%)	41	0.172	0.378	0.143	0.0245
	Fall-rise (H*L H%)	10	0.042	0.201	0.04	0.013
	Rise-fall (L+H* L L%)	0	0	0	0	0
IPs: 220	Fall (H*L L%)	115	0.523	0.499	0.249	0.0337
Group B, N = 15	Low fall (L*L L%)	12	0.055	0.227	0.052	0.0153
	Level (L*L 0%)	11	0.05	0.218	0.048	0.0147
	Rise (L*H H%)	29	0.132	0.338	0.114	0.0228
	Low rise (L*LH%)	19	0.089	0.281	0.079	0.0189
	Fall-rise (H*L H%)	31	0.141	0.348	0.121	0.0235
	Rise-fall (L+H* L L%)	3	0.014	0.116	0.013	0.0078

Table 4.3 The occurrence of individual tones produced by non-instructed students (group A) and instructed ones (group B) in the monologue

To examine whether the means obtained in the two groups differ significantly a t-test for independent samples was applied, where the significance level is $p < .05$ (table 4.4).

TONE	t	df-t	p
Fall (H*L L%)	-4.87	1146	0
Low fall (L*L L%)	5.63	1011	0
Level (L*L 0%)	0.56	1139	0.5764
Rise (L*H H%)	0.26	1146	0.7915
Low rise (L*LH%)	4.22	1019	0
Fall-rise (H*L H%)	-2.76	1058	0.0061
Rise-fall (L+H* L L%)	-2.03	569	0.045

Table 4.4 T-test for independent samples (both texts)

As transpires from table 4.4, for the falling tone the difference between 0.439 and 0.581 with standard deviations of 0.496 and 0.493 based on sample sizes of 579 and 570 IPs in group A and group B respectively, was significant at the .05 level. For the low fall the difference between 0.181 and 0.072 with standard deviations of 0.385 and 0.258 based on sample sizes of 579 and 570, respectively, was also significant at the .05 level. For the low rise the difference between 0.123 and 0.053 with standard deviations of 0.328 and 0.223 was significant too, and so it was for the fall-rise, with the difference between 0.048 and 0.089 with standard deviations of 0.215 and 0.285, and the rise-fall, with the difference between 0 and 0.007 with standard deviations of 0 and 0.083.

However, for the level tone the difference between the means obtained by the two groups (0.026 and 0.021 with standard deviations of 0.159 and 0.144) was not significant at the 0.5 level, and neither it was for the rise, where the difference ranged between 0.183 and 0.177 with standard deviations of 0.387 and 0.382.

The t-test revealed the following findings for the dialogue and the monologue (tables 4.5 and 4.6). In the dialogue, the difference between the means for the fall, the low fall and the low rise were significant at the .05 level. In the case of the falling tone the difference was calculated from the means of 0.472 and 0.617 with standard deviations of 0.499 and 0.486 based on sample sizes of 341 and 350 IPs in group A and group B, respectively. For the low fall the difference was calculated from the means of 0.17 and 0.083 with standard deviations of 0.376 and 0.276, while for the low rise the means of 0.088 and 0.031 standard deviations of 0.283 and 0.174 were taken into consideration. Since there was only one occurrence of the rise-fall in group B, no t-test was possible. The groups did not differ considerably in the usage of the fall-rise.

TONE	t	df-t	p
Fall (H*L L%)	-3.87	687	0.0001
Low fall (L*L L%)	3.47	623	0.0006
Level (L*L 0%)	N/A		
Rise (L*H H%)	0.06	688	0.9483
Low rise (L*LH%)	3.20	562	0.0016
Fall-rise (H*L H%)	-0.23	688	0.8177
Rise-fall (L+H* L L%)	N/A		

Table 4.5 T-test for independent samples (dialogue)

In the monologue, there existed a statistically significant difference between the means of the fall, low fall, low rise and fall-rise. For the fall, difference between 0.391 and 0.523 with standard deviations of 0.488 and 0.499 based on sample sizes of 238 and 220 IPs in group A and group B, respectively, was significant at the .05 level. For the low fall the t-value was based on the difference between 0.197 and 0.055 with standard deviations of 0.398 and 0.277, for the low rise – on the range of 0.172 and 0.086 with standard deviations of 0.378 and 0.281, and for the fall-rise – on the difference between 0.042 and 0.141 with standard deviations of 0.201 and 0.348. The difference between the rise, level and rise-fall was not significant for p<.05.

TONE	t	df-t	p
Fall (H*L L%)	-2.86	451	0.0044
Low fall (L*L L%)	4.64	381	0
Level (L*L 0%)	0	453	1
Rise (L*H H%)	0.46	455	0.6428
Low rise (L*LH%)	2.75	436	0.0057
Fall-rise (H*L H%)	-376	344	0.0003
Rise-fall (L+H* L L%)	-1.86	218	0.0748

Table 4.6 T-test for independent samples (monologue)

Therefore, prediction 1 seems to be supported by the evidence collected in the experiment. Group A did produce significantly more simple tones than group B. This reveals that universal processes, rather than those due to L1 influence, indeed predominate at the earlier stages of intonational acquisition. The more frequent occurrence of complex tones later in the acquisition proves that the advanced students have suppressed at least some of the universal processes.

Generally speaking, for the more advanced group B the acquisition of English intonation has not been completed, not even after a two-year course. Of the complex tones, the fall-rise was more frequent in group B, yet it was the only complex tone the use of which was statistically significant. The rise-fall appeared sporadically, although this phenomenon is ambiguous: either the tone has not entered the subjects' interlanguage yet or they consciously decided that the context is not appropriate.

4.2.2. Perception

Perception is one of the crucial factors influencing acquisition. The perceptual abilities of the subjects in both groups were assessed on the basis of a recording consisting of 30 independent sentences (see Appendix F) read aloud by one of the native-English teachers at PWSZ. The test included the following intonation patterns: fall (7 tokens), rise (4), fall-rise (13), rise-fall (5) and level (1). The task was twofold, to recognise the word carrying the nuclear pitch accent and to indicate with a simplified symbol the intonational contour of each sentence. Table 4.7 presents the results of the test.

tone	N	Group A			Group B			t
		%	mean	SD	%	mean		
H*LL%	70	100%	-	-	100%	-	-	-
L*HH%	40	100%	-	-	100%	-	-	-
H*LH%	130	83%	8.31	1.32	92%	9.15	1.34	-5.09
L+H*LL%	50	60%	6	3.16	88%	8.8	1.64	-5.56
L*L0%	10	100%	-	-	100%	-	-	-

Table 4.7 Perception test

In group A, the falling, the rising and level patterns were correctly recognised by all subjects. Of the complex tones, the fall-rise was correctly recognised in 83 % of its occurrences and the rise-fall in 60%.

In group B, the level, falling and rising contours were correctly recognised by all students, while the fall-rise was indicated in 92% of intended contexts and the rise-fall in 88%. The difference between the groups, though, seems to be on the verge of statistical significance, so any conclusions concerning their perceptual abilities should be drawn cautiously.

Therefore, Pearson correlation was performed in order to measure the strength of the relationship between the groups' perception test results (the two complex tones as a whole). The test revealed a strong positive correlation at .05 level: r = 0.974, which means that the perceptual skills of the subjects from the two groups are rather similar.

Both groups experienced greater difficulties with the complex tones, though group B performed significantly better than group A. What is interesting, the fall-rising contour was always categorised as the rise and the rise-fall as the fall. Recognition of complex tones was especially problematic when the word on which the nuclear pitch accent was falling was two- or more syllables longer, as the pattern extends then over the unaccented syllables following the pitch accent, which makes the complex tone more difficult to perceive.

4.2.3. The influence of the L1: peak alignment

The overall shape of intonational contours in Polish roughly resembles the English ones. Therefore, the influence of L1 intonational system can be most vividly observed in peak alignment. In English[39], the peak is reached within the stressed syllable, towards its right edge (cf. Grabe 1998). In Polish the peak is located earlier within the syllable.

In order to provide the means of comparison of the subject's performance, 20 Polish utterances were examined (Appendix G, 2 female native speakers of Polish). The spectrum, pitch and intensity were extracted with the use of *PRAAT* in order to specify the vowel of the nuclear syllable. The peak was indicated by the highest value of the f_0 track and intensity (which was especially helpful when the f_0 was too flat to distinguish its peak), the formants in the spectrum, and the strongest variation in air pressure depicted by the amplitude. The difference in peak timing was calculated with the use of the following formula (*Pt* – peak timing, t_{f0max} – timing of the greatest f_0 value, t_{f0i} – timing of the f_0 value at the beginning of the accented vowel).

$$Pt = t_{f0max} - t_{f0i}$$

The calculated values for the Polish utterances were as follows (Table 4.8). A tentative conclusion is that the timing of peak alignment is later if the vowel is preceded by a liquid or nasal sound, and earlier after a voiceless consonant.

---

[39] Cf. Wichmann's (2000: 44) study of peak alignment in discourse: "the peaks on the first accented syllable of a new topic were consistently later than those at the beginning of a sentence which did not begin a new topic." Moreover, "under some contextual conditions the pitch peak can occur outside the accented syllable" (2000: 46).

UTTERANCE	PEAK TIMING	MEAN	SD
1. Sprawdź za książkami.	0.023	*0.03315*	*0.011089*
2. Co chcesz z nimi zrobić?	0.034		
3. Wyrzucę tę bluzkę.	0.021		
4. Te spodnie są brązowe.	0.021		
5. Fonologia.	0.041		
6. Fonologia?	0.033		
7. Wejdź na górę.	0.043		
8. Zejdź na ziemię.	0.023		
9. Mówił wyraźnie.	0.027		
10. Mówił i mówił...	0.046		
11. Idziesz do domu?	0.042		
12. To jest czerwone.	0.047		
13. Co to jest?	0.052		
14. To jest ta książka?	0.028		
15. Szerokiej drogi.	0.023		
16. Dzień dobry (fall).	0.049		
17. Dzień dobry (rise).	0.025		
18. Tak mi przykro.	0.019		
19. Naprawdę żałuję.	0.023		
20. Możesz mi pomóc?	0.043		

Table 4.8 The timing of peak alignment in Polish utterances

To examine whether peak alignment was acquired by the subjects, a sample of 20 IPs was chosen from the two texts (examples in Figures 4.9 and 4.10). Not only was the performance of the two groups compared against each other, but also against the IPs produced by a native speaker (Figure 4.11). It is assumed that context did not influence the differences in peak alignment, since the tokens were produced in similar situations.
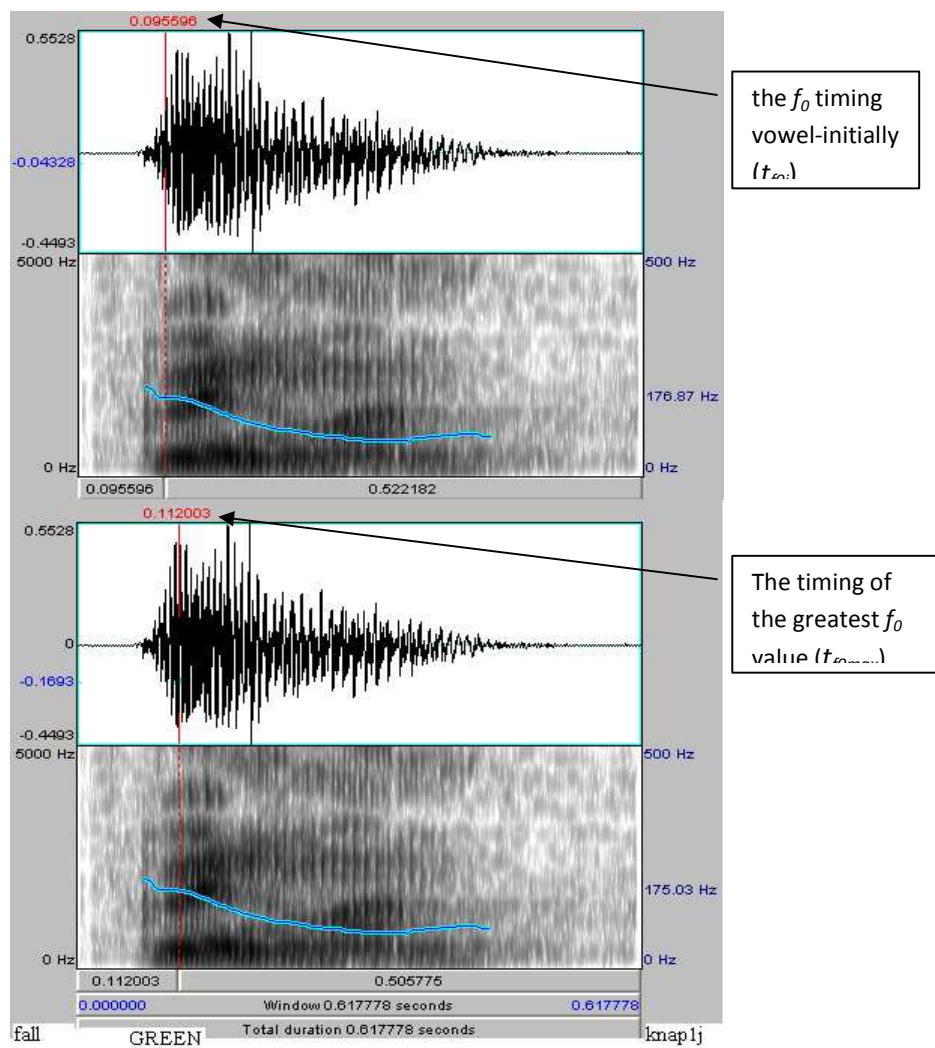
Figure 4.9 Visual representation of a student's IP "Green one" where the nuclear syllable is "green" by a group B subject
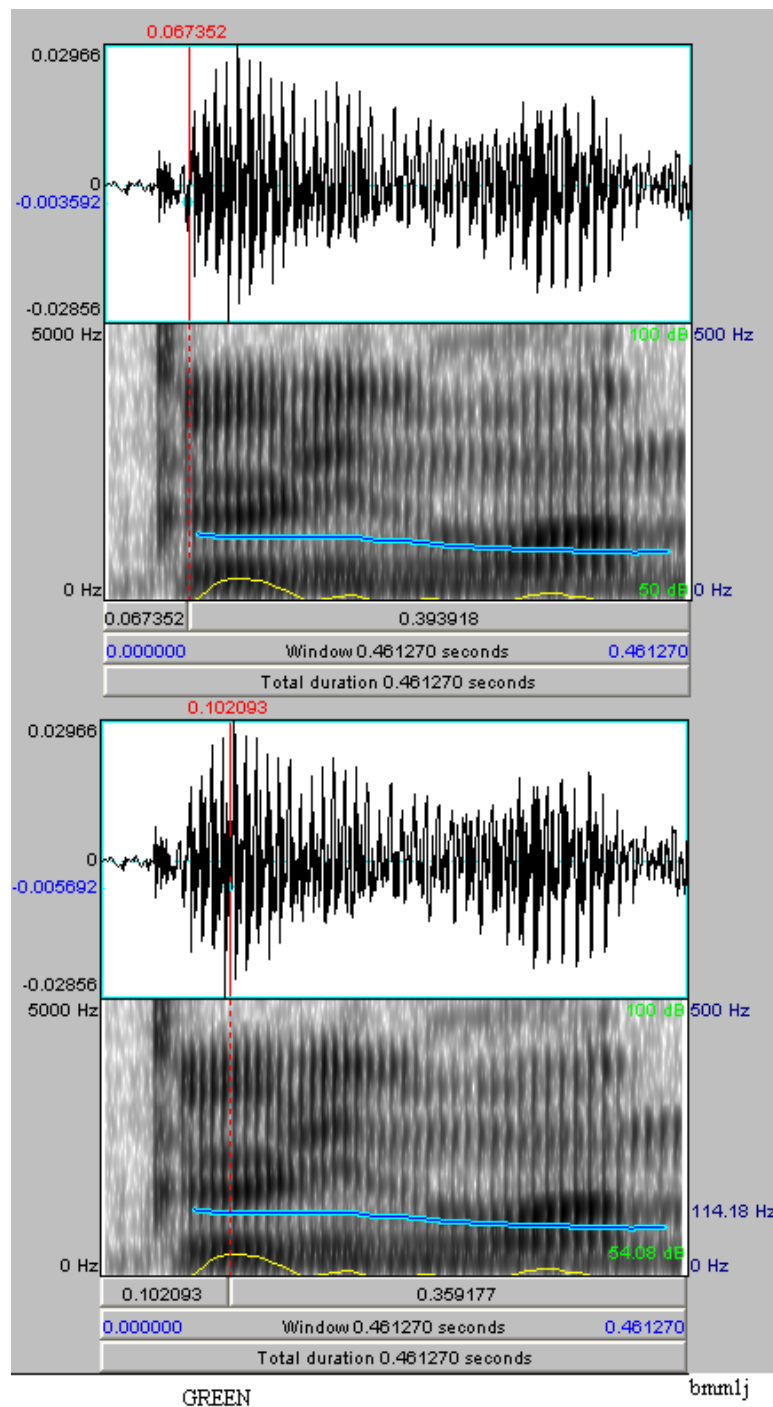
Figure 4.10 Visual representation of a student's IP "Green one" where the nuclear syllable is "green" by a group A subject: the *f0* track (the blue line) is too flat to recognise the peak, therefore intensity (the yellow line) and amplitude are decisive
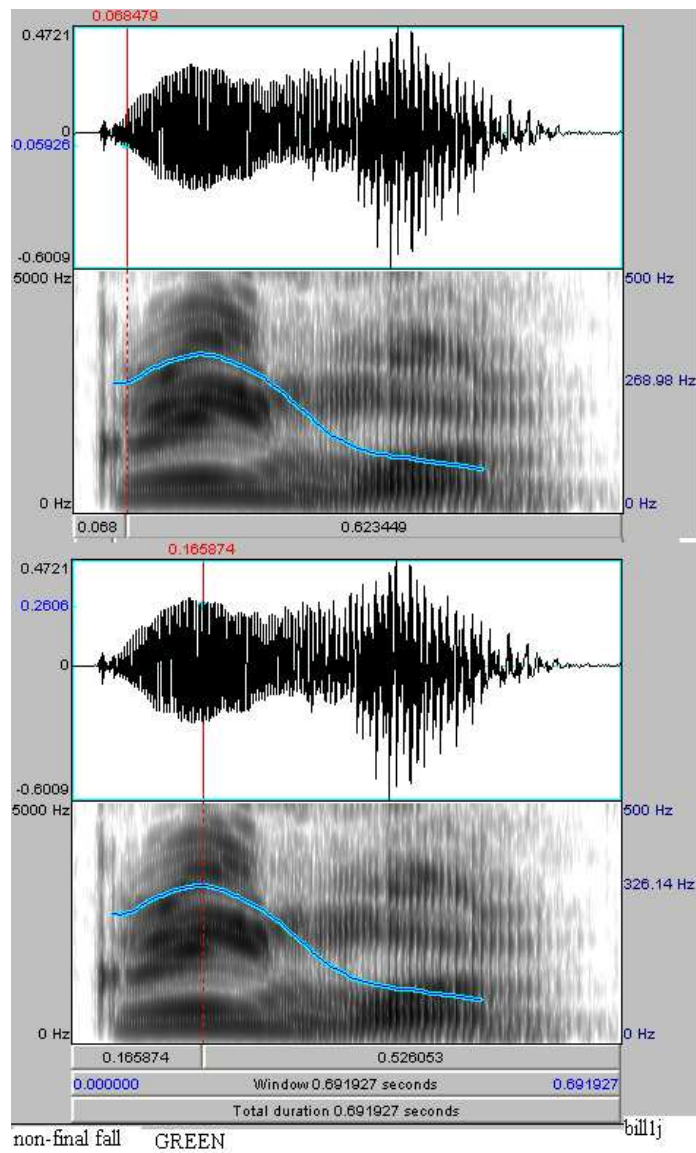
Figure 4.11 Visual representation of a native speaker's IP "Green one" where the nuclear syllable is "green"

The mean distance of the peak from the beginning of the vowel for each student is presented in Figures 4.9 and 4.10. The last bar in each figure depicts the mean of the native speaker. The comparison of the means of both groups, the native speaker and the Polish utterances, is presented in Figures 4.12, 4.13 and 4.14.
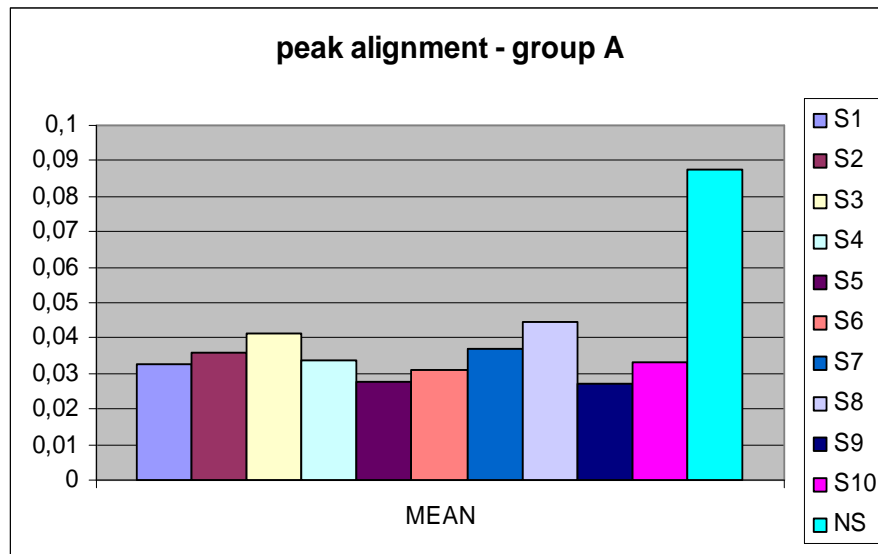
107

Figure 4.12 The mean distance of the peak from the beginning of the vowel in the nuclear syllable in group A
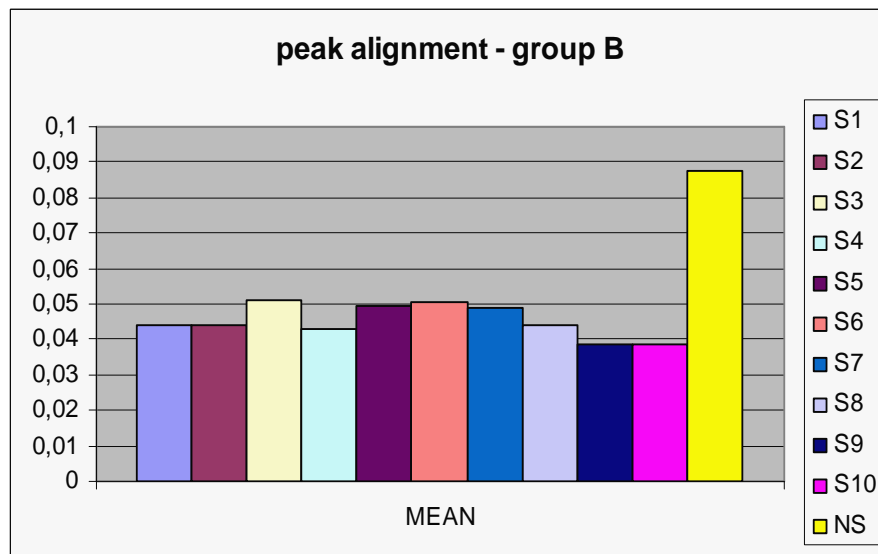(S – student, NS – native speaker)

Figure 4.13 The mean distance of the peak from the beginning of the vowel in the nuclear syllable in group
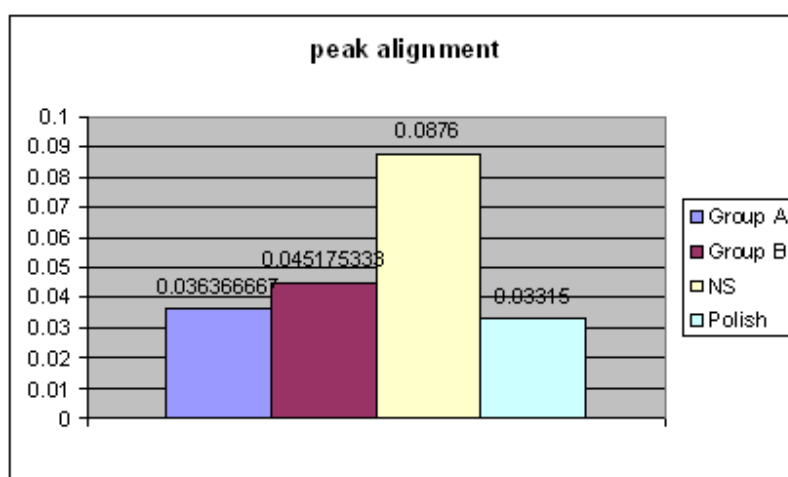B (S – student, NS – native speaker)

Figure 4.14 The comparison of mean peak timing in group A, group B, native speaker and Polish

To test the hypothesis that L1 peak alignment will be more evident in the performance of group A, a t-test for independent samples was run to compare the mean distance of the peak from the beginning of the vowel. The mean for the whole group A is 0.036367, while for group B: 0.045175. The overall means of groups A and B were subsequently compared with the mean of the native speaker: 0.0876 and Polish (Table 4.9).

	Mean	*SD*	*t*	*df-t*	*p*
Group A	0.036367	0.023854	-3.79	396	0.0002
Group B	0.045175	0.022579			
Group A	0.036367	0.023854	-5.78	35	0
NS	0.0876	0.031692			
Group B	0.045175	0.022579	-4.88	34	0
NS	0.0876	0.031692			
Group A	0.036367	0.023854	0.55	26	0.5879
Polish	0.03315	0.011089			
Group B	0.045175	0.022579	2.14	27	0.0403
Polish	0.03315	0.011089			

Table 4.9 T-test for independent samples showing the difference in the performance between the means of peak timing of group A, group B, the native speaker and L1

The t-test revealed that the performance of group A and group B varied from the performance of the native speaker, but it also differed between the subjects. Although the timing of the native speaker was significantly later than the timing of each student, group B's subjects reached the peak significantly later than group A's subjects. In comparison with their L1, in group A's performance the difference between Polish peak timing and their interlanguage timing was not significant (t = 0.55). However, the difference between group B's interlanguage peak timing and L1 timing was significant at the .05 level (t = 2.14).

It can be claimed, though tentatively, that partial acquisition of target-like peak alignment indeed took place and that group B students could develop a "merged" intonational system. This success may be attributed to the previous acquisition of other sublevels of L2 phonological competence, such as vowel length and rhythm, rather than to the postponement of the peak. Thus the prediction that L1 transfer will be less detectable in group B's pronunciation, is supported. Yet the performance of group B did not reach the L2-like peak timing, which contributes to the perceived foreign accent in the speech of the subjects. However, more detailed research is required in order to investigate the factors which may contribute to the difference, for instance, the length of the vowel carrying a pitch accent, the influence of the consonants preceding the vowel and the influence of discourse structure on peak alignment, e.g. the beginning or end of a topic (cf. Wichmann 2000).

4.2.4. The influence of the L1: tonal meanings

The main function of intonation is to convey attitudes of the speaker, mark discourse movements (discourse functions were not analysed in the present study) and indicate the grammatical properties of an utterance. This section presents the analysis of the most transparent instances of L1 influence on the pragmatic and grammatical aspects of the learners' system. The description of intonational meanings for Polish is based on Steffen-Batogowa 1996; Grabe – Karpiński 2003; Karpiński 2006, and the author's judgement, and for English – on Crystal 1969; Brazil 1975, 1978, 1997; Cruttenden 1984; Wichmann 2000 and Wells 2006.

The following tables (Tables 4.10, 4.11) present the differences in the distribution of tones in major sentence types and in chosen semi-ritualized expressions.

	(DIAL)	Group A			Group B			t	df-t
		N	Mean	SD	N	Mean	SD		
Y/N Questions	Fall (H*L L%)	1	1	-	6	0.2	0.4	-	-
	Low fall (L*L L%)	0	0	0	0	0	0	-	-
	Level (L*L 0%)	0	0	0	0	0	0	-	-
	Rise (L*H H%)	23	0.767	0.423	22	0.733	0.442	0.30	57
	Low rise (L*LH%)	3	0.1	0.3	0	0	0	1.826	29
	Fall-rise (H*L H%)	3	0.1	0.3	2	0.067	0.249	0.46	56
	Rise-fall (L+H*L L%)	0	0	0	0	0	0	-	-
Wh-questions	Fall (H*L L%)	22	0.379	0.485	46	0.767	0.423	-4.64	112
	Low fall (L*L L%)	11	0.19	0.392	3	0.05	0.218	2.39	88
	Level (L*L 0%)	0	0	0	0	0	0	-	-
	Rise (L*H H%)	11	0.19	0.392	5	0.083	0.276	1.70	102
	Low rise (L*LH%)	11	0.19	0.392	2	0.033	0.18	2.78	79
	Fall-rise (H*L H%)	3	0.052	0.221	3	0.05	0.218	0.05	115
	Rise-fall (L+H*L L%)	0	0	0	1	1	-	-	-

	Tone								
Echo question	Fall (H*L L%)	0	0	0	0	0	0	-	-
	Low fall (L*L L%)	0	0	0	0	0	0	-	-
	Level (L*L 0%)	0	0	0	0	0	0	-	-
	Rise (L*H H%)	14	0.933	0.249	11	0.733	0.442	1.53	22
	Low rise (L*LH%)	0	0	0	0	0	0	-	-
	Fall-rise (H*L H%)	1	1	-	4	0.267	0.442	-	-
	Rise-fall (L+H*L L%)	0	0	0	0	0	0	-	-
Question tag	Fall (H*L L%)	5	0.333	0.471	11	0.733	0.442	-2.40	27
	Low fall (L*L L%)	0	0	0	0	0	0	-	-
	Level (L*L 0%)	0	0	0	0	0	0	-	-
	Rise (L*H H%)	9	0.6	0.49	4	0.267	0.442	1.95	27
	Low rise (L*LH%)	1	1	-	0	0	0	-	-
	Fall-rise (H*L H%)	0	0	0	0	0	0	-	-
	Rise-fall (L+H*L L%)	0	0	0	0	0	0	-	-
Declaratives	Fall (H*L L%)	108	0.725	0.447	132	0.886	0.318	-3.58	267
	Low fall (L*L L%)	37	0.248	0.432	9	0.06	0.238	4.65	230
	Level (L*L 0%)	0	0	0	0	0	0	-	-
	Rise (L*H H%)	0	0	0	0	0	0	-	-
	Low rise (L*LH%)	0	0	0	0	0	0	-	-
	Fall-rise (H*L H%)	4	0.027	0.162	8	0.054	0.225	-1.19	268
	Rise-fall (L+H*L L%)	0	0	0	0	0	0	-	-
Listing	Fall (H*L L%)	16	0.267	0.442	14	0.222	0.416	0.58	119
	Low fall (L*L L%)	6	0.1	0.3	13	0.206	0.405	-1.64	114
	Level (L*L 0%)	3	0.05	0.218	0	0	0	1.78	59
	Rise (L*H H%)	13	0.217	0.423	27	0.429	0.495	-2.47	119
	Low rise (L*LH%)	15	0.25	0.433	9	0.143	0.35	1.51	113
	Fall-rise (H*L H%)	7	0.117	0.321	0	0	0	2.89	59
	Rise-fall (L+H*LL%)	0	0	0	0	0	0	-	-
Apology	Fall (H*L L%)	9	0.6	0.49	7	0.467	0.499	0.74	27
	Low fall (L*L L%)	4	0.267	0.442	2	0.133	0.34	0.93	26
	Level (L*L 0%)	1	1	-	1	1	-	-	-
	Rise (L*H H%)	1	1	-	2	0.133	0.34	-	-
	Low rise (L*LH%)	0	0	0	0	0	0	-	-
	Fall-rise (H*L H%)	0	0	0	3	0.2	0.4	-0.2	14
	Rise-fall (L+H*L L%)	0	0	0	0	0	0	-	-

Table 4.10 Differences in the distribution of tones in major sentence types and a chosen semi-ritualized expression in the dialogue

	(MON)	Group A			Group B			t	df-t
		N	Mean	SD	N	Mean	SD		
Greetings: "Hello there"	Fall (H*L L%)	8	0.533	0.499	1	1	-	-	-
	Low fall (L*L L%)	3	0.2	0.4	0	0	0	1.94	14
	Level (L*L 0%)	1	1	-	0	0	0	-	-
	Rise (L*H H%)	0	0	0	3	0.2	0.4	-	-
	Low rise (L*LH%)	0	0	0	0	0	0	-	-
	Fall-rise (H*L H%)	3	0.2	0.4	10	0.677	0.471	-2.99	27
	Rise-fall (L+H*L L%)	0	0	0	1	1	-	-	-
Greetings: "Good afternoon"	Fall (H*L L%)	5	0.333	0.471	9	0.6	0.24	-1.95	20
	Low fall (L*L L%)	7	0.467	0.499	1	1	-	-	-
	Level (L*L 0%)	1	1	-	0	0	0	-	-
	Rise (L*H H%)	0	0	0	0	0	0	-	-
	Low rise (L*LH%)	0	0	0	0	0	0	-	-
	Fall-rise (H*L H%)	2	0.133	0.34	0	0	0	1.51	14
	Rise-fall (L+H*L L%)	0	0	0	1	1	-	-	-

Table 4.11 Differences in the distribution of tones in two semi-ritualized expressions in the monologue

Table 4.10 demonstrates that only the difference between the usage of the low rising tone in yes/no (polar) questions (30 samples in group A and 30 in group B) is significant at the .05 level. However, this may be regarded as an instance of universal processes operating. What is striking is the usage of the fall, which group B employed more frequently.

English yes/no questions take the rise and so do Polish equivalent constructions. English speakers use the yes-no fall in special contexts, namely in guessing games, repeated questions and insistent, businesslike interrogatives (Wells 2006). In Polish, the intonation of polar questions is affected by their syntactic structure. Most frequently they tend to carry rising intonation, however, when they contain an initial question word "czy", the rise is optional and other tones can be employed: the fall, fall-rise or level (cf. Karpiński 2006). As it was mentioned in chapter 3, syntax already indicates the question, therefore the rise can be replaced by the less marked, more archetypical fall, with no particular changes in meaning. Though it is hard to decide which factors precisely caused the subjects to employ the fall, one may hypothesise that these were the instances of L1 influence. There were a few examples of the fall-rise in both groups' performance. This is probably an example of transferring Polish emotive features of intonation. Figures 4.15 and 4.16 present a yes/no question with a low rise and a fall.
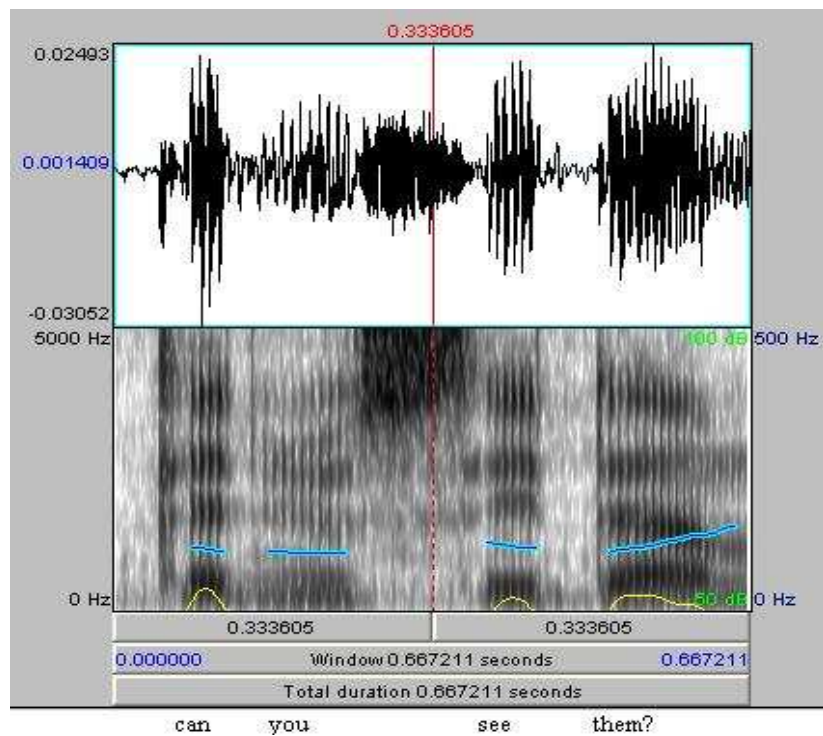
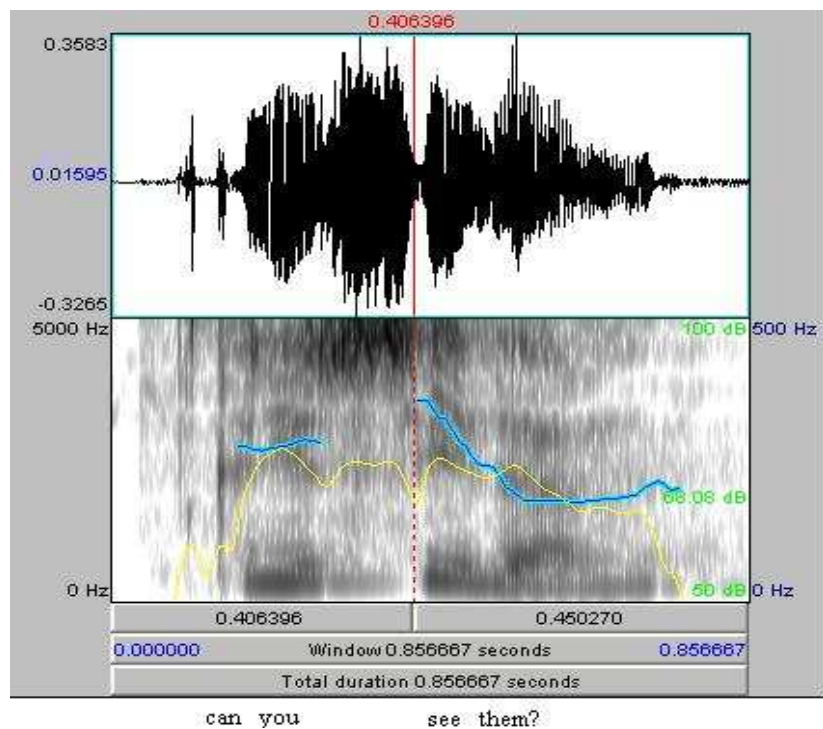Figure 4.15 A yes/no question with a low rise, pronounced by group A male subject

Figure 4.16 A yes/no question with a fall, pronounced by group B female speaker

In both English and Polish, *wh*-questions contain a *wh*-word which is the indicator of the sentence type; therefore the rising intonation pattern is not obligatory. The two

languages, though, differ in the use of the fall and rise in these question types. In English the fall is the most typical tone in *wh*-interrogatives. Other patterns can also appear, yet they affect the meaning of the question. The rise sounds encouraging and may be used to politely initiate a conversation or, rather than demand information, to reassure the listener, especially a child. However, when directed at an adult, it may sound patronising. As for Polish, recent research (Karpiński 2006: 170) reveal that the most typical contour is the rise, although its occurrence in questions of this type is less frequent than in yes/no interrogatives. The falling tone is also used, though its usage does not seem to influence the meaning of the sentence.

The production of *wh*-questions (58 samples in group A and 60 in group B) was more diverse than in the case of yes/no interrogatives. The usage of the falling, low falling and low rising tones differed significantly at the .05 level (the last two low-range tones revealing universal preferences). Group B used significantly more target-like falling tones. Group A employed the rise more frequently than group B, though the difference was not significant.

The use of fall-rise in both groups did not differ to a great extent. The rise-fall was employed once (*And what colour are my TROUsers then?*) by a group B speaker, making the question sound grumpy, which did not match the subject's overall interpretation of the dialogue. The speaker could have been misled by the structure of the sentence (*then* at the end) because of which it resembled a statement. Another possible explanation is the transfer of the emotive load of the Polish rise-fall contour. It is doubtful, therefore, whether the tone was used on purpose.
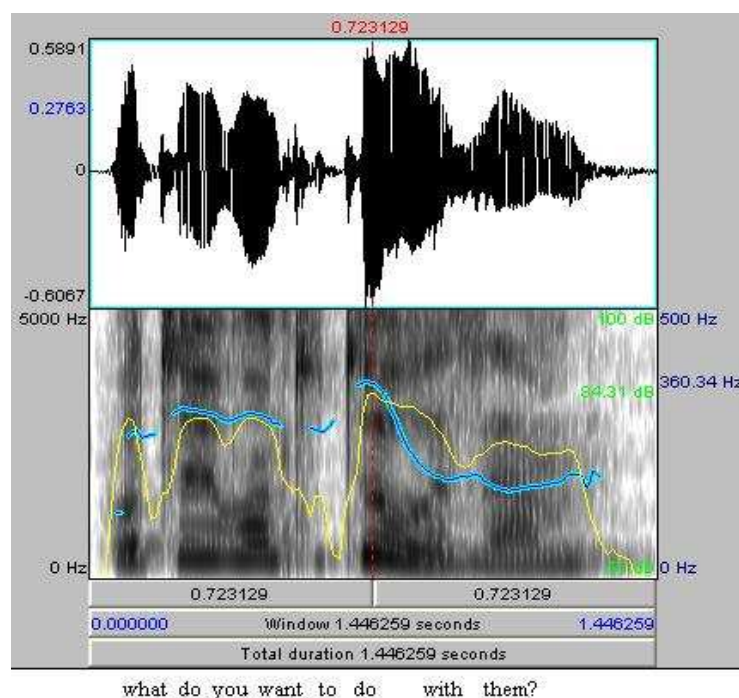

Figure 4.17 A *wh*-question with a fall (group B female speaker)
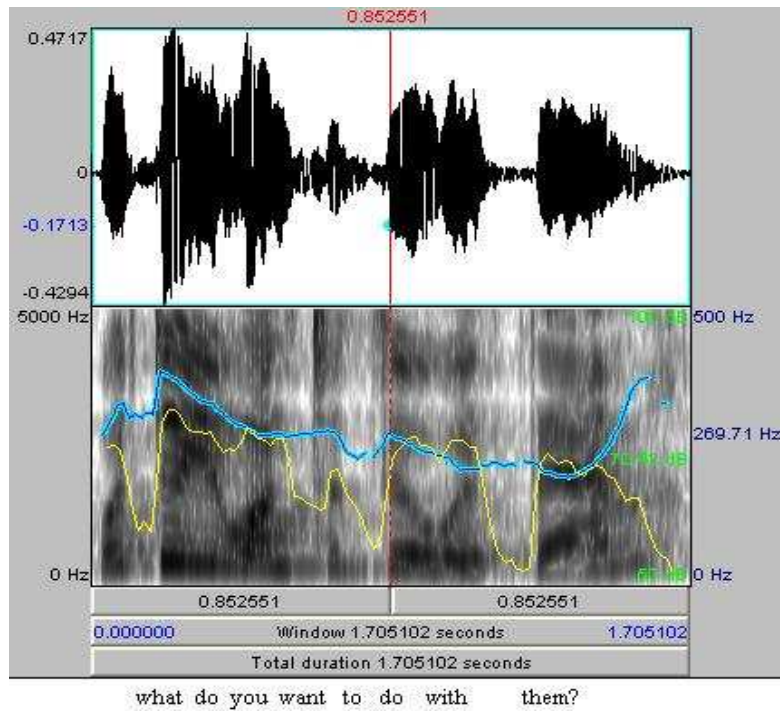
114

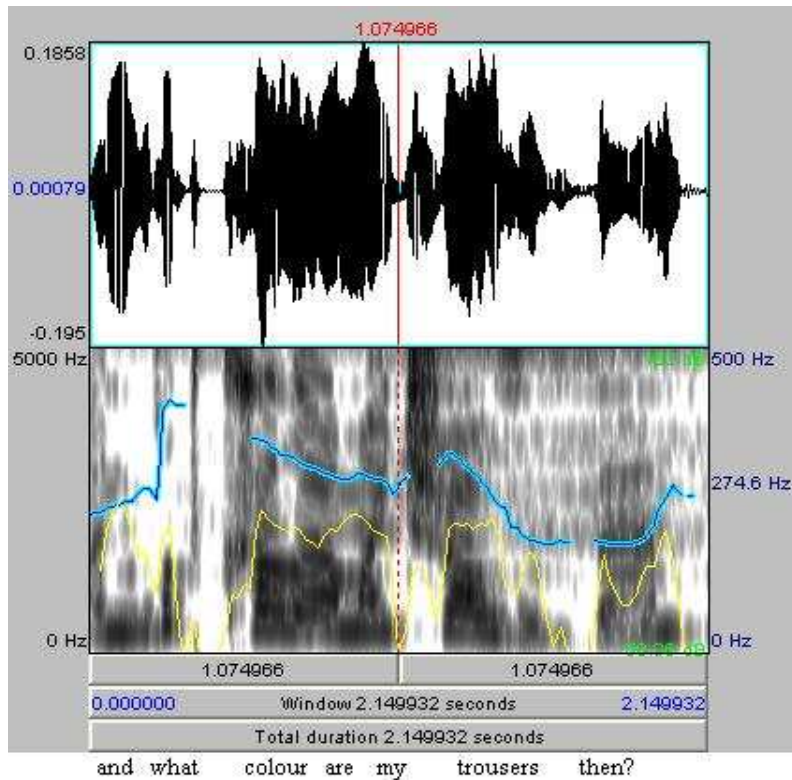Figure 4.18 A *wh*-question with a rise (group A female speaker)

Figure 4.19 A *wh*-question with a fall-rise beginning on *TROUsers* (group A female speaker)

115

The analysed material contained one echo question (15 IPs per group) and one question tag (15 IPs per group, see Figures 4.20 and 4.21). Polish marks echo questions with a high rise, and English either with a high rise or a fall-rise. The groups differed significantly in the use of the fall, where group B employed more appropriate to the context falling tones and group A revealed greater L1 influence. This tendency is confirmed by the author's teaching experience: many quite advanced learners persist in pronouncing all question tags with the final rise, like genuine questions.

Question tags do not correspond directly to any Polish interrogative, although their closest counterparts are the so-called "questions for confirmation", which include a (declarative) question with the falling intonation, as in a statement, followed by a questioning particle, such as *tak? nie? co?* forming a separate IP. Intonation of the particle is of the rising type (Karpiński 2006: 169). On the contrary, English question tags, also pronounced as separate IPs, invite confirmation with the statement when they carry the fall, or verification of the proposition in the statement when the tone is rising (Cruttenden 1986). In the analysed dialogue a falling question tag was more appropriate as an ironic remark about the speaker's inability to name colours (*You should really put your glasses on, shouldn't you?* [H*L L%]).
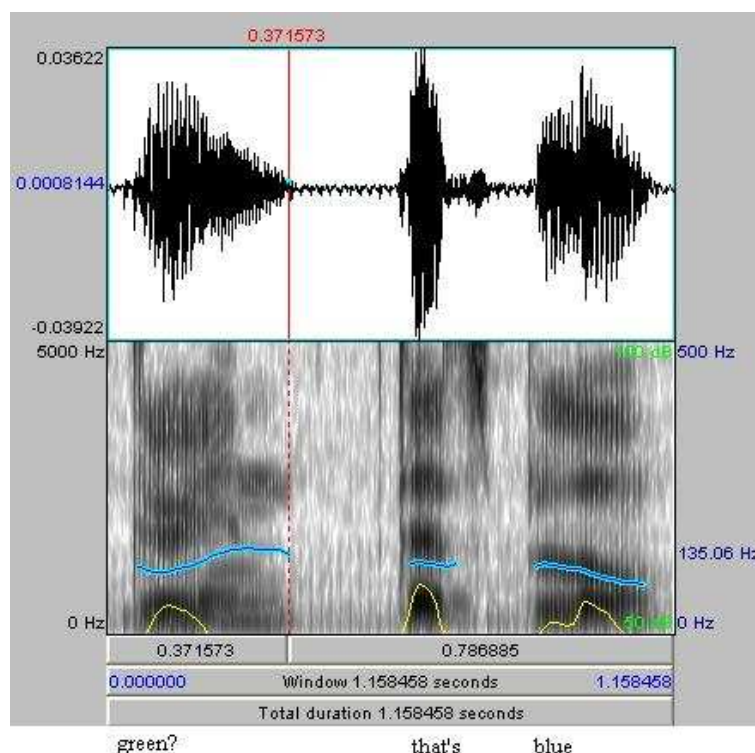

Figure 4.20 An echo question followed by a declarative (group A male speaker)
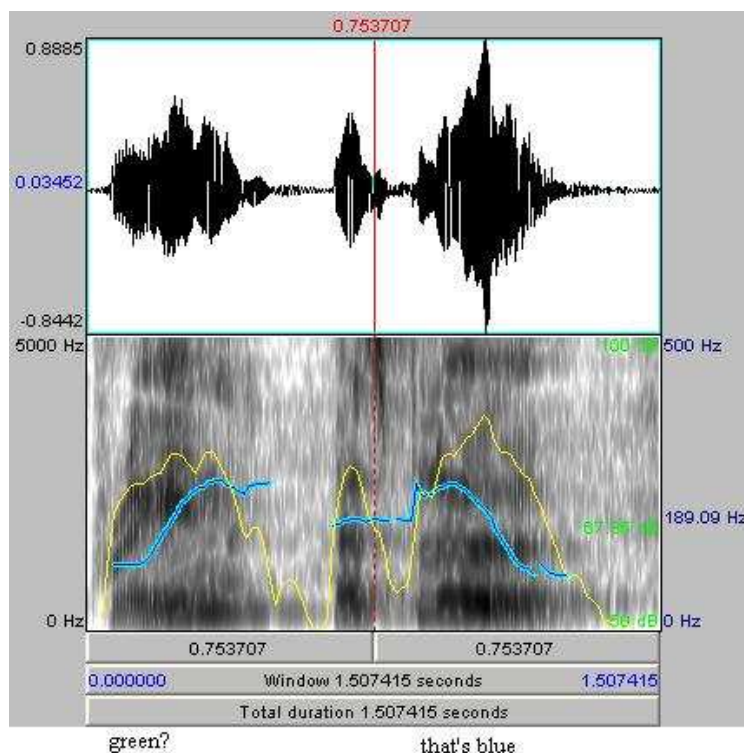
Figure 4.21 An echo question followed by a declarative (group B male speaker)

Declaratives in the dialogue form answers to questions. Karpiński's (2006) study demonstrated that in Polish there exists a strong tendency to indicate positive answers with falling intonation. English does not differ in this respect. In Polish, apart from the fall, the fall-rise is possible; it seems, however, to be used as a discourse-organiser in special contexts, such as encouraging the interlocutor to continue. On the other hand, the fall-rise in English can make a sentence more polite too but it also adds a special meaning to a statement, namely reservations, contradiction or warning (Cruttenden 1986: 109). Polish negative answers carry a more varied inventory of intonational patterns: falling, falling-rising and rising-falling, and the rising melody also possesses discourse-organising function (Karpiński 2006). In English the fall-rise is implicational (Wells 2006) and allows the speaker to contradict the interlocutor without giving offence or to express his/her reservations about what is being said. The rise-fall in English statements adds extra meaning to the utterance, often implying the opposite of the literal meaning.

In the analysed material the use of the fall and low fall (cf. universal preferences) in declaratives as pronounced by the subjects (149 samples in group A and 149 in group B) differed significantly at the .05 level. The difference between the use of the fall-rise by the two groups was not significant, although the tone occurred more often in the performance of group B (8 instances) than in the speech of group A (4 instances). However, instead of conveying the expected L2 meanings, in the performance of both groups the tone always indicated non-finality (e.g. in *Check behind these BOOKS* [H*LH%] followed by *Can you see them?*). This proves that the L2 meanings were not yet acquired.

117

Intonation of lists in English and Polish does not differ considerably and can be considered universal: rising tones indicate incompleteness (the items listed, except the last item), while falling tones signal completeness (the final item in the list). The two groups differed significantly in the use of the rise and the fall-rise: the latter was used more frequently by group A to indicate non-finality rather than L2 meanings.

The most evident language-specific differences were revealed by the intonation of the apology. In Polish, the expression equivalent to *I'm so sorry*, i.e. *Tak mi przykro*, tends to be pronounced with the falling contour, although the rise and fall-rise, making the apology more polite, are possible. In English a sincere apology carries the fall-rise. The fall is considered aggressive and may be intended to elicit an apology from the addressee; therefore it was unacceptable in the context of the analysed dialogue. Both groups used predominantly falling and low falling tones, and the difference between their performance was not significant. The rise and fall-rise were employed sporadically only by group B.
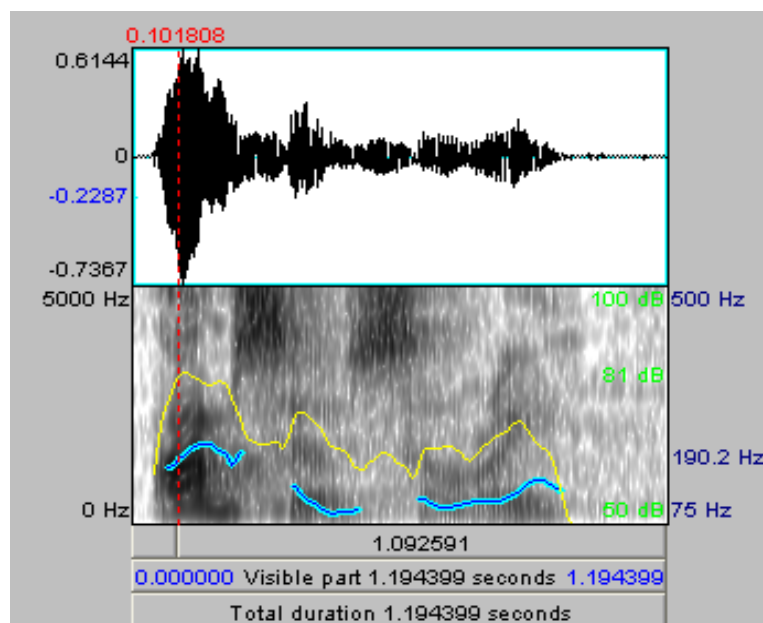

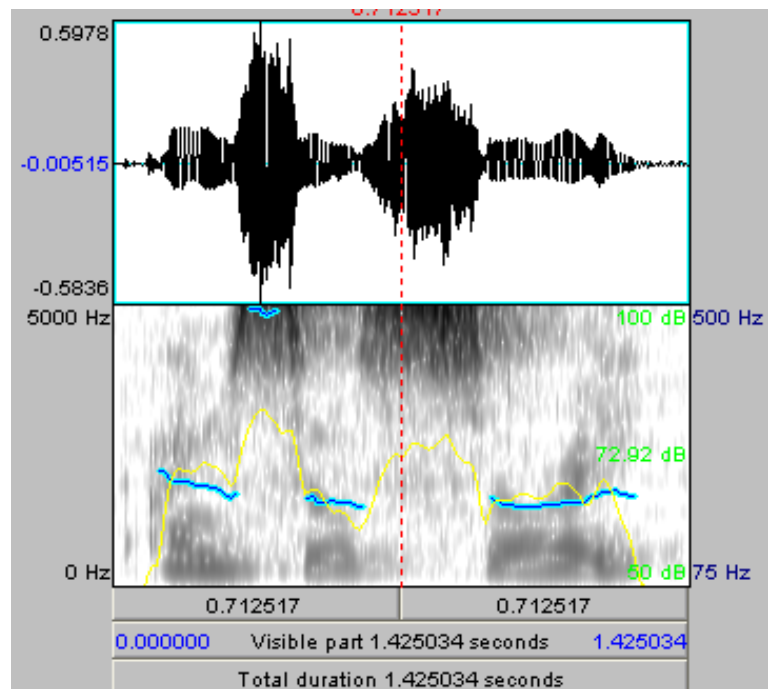Figure 4.22 A fall-rise in an apology (group B male speaker)

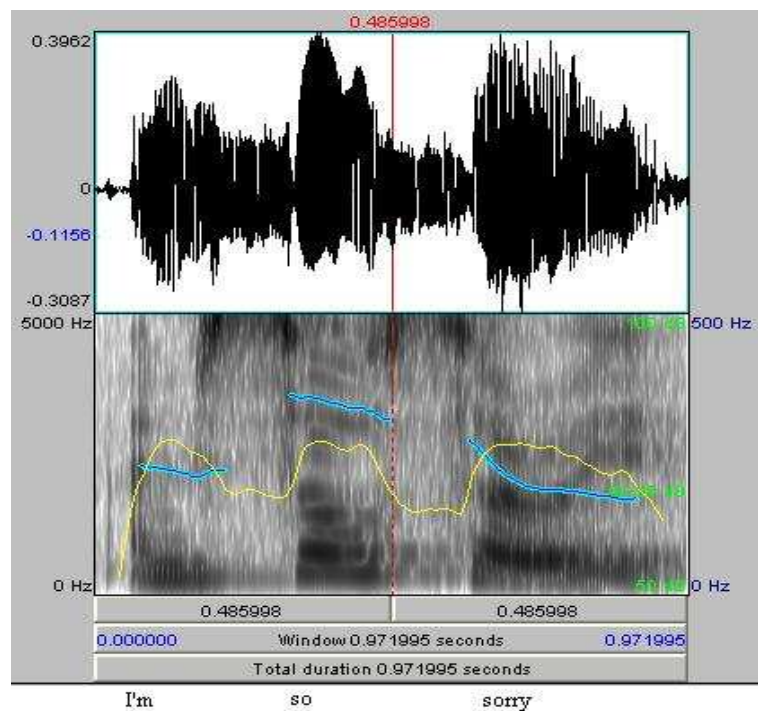Figure 4.23 A rise in an apology (group B female speaker)

Figure 4.24 A fall in an apology (group A female speaker)

In the analysed material there appeared two types of greetings: *hello there* and *good afternoon*. Their Polish equivalents usually carry the falling intonational pattern (relatively

lower than the English fall) or the more polite rising one (e.g. *cześć, dzień dobry*). Likewise, for most English greetings both the fall and the rise are acceptable, the definite fall being more formal and the encouraging rise more personal. However, such greetings as *hello* followed by a vocative are pronounced with a fall on the greeting word and a rise on the vocative forming a separate IP. In Polish, it seems that the vocative may take either the fall or the rise.

The fall-rise in English greetings may give the impression that the speaker is really trying to establish contact with the other person, therefore it is perceived as polite. The fall-rise also indicates that the speaker intends to attract the other person's attention, or when there is an element of surprise in meeting the other person. The rise with low pre-nuclear pitch accents is used when the speaker is answering the telephone. A high rise tends to occur when the speaker wishes to re-establish contact after being kept on hold for a while (e.g. *hello* [L*H H%], *are you there?*).

In the analysed monologue (weather forecast) group A used mostly the fall and the low fall in *HEllo there*. One person pronounced the greeting with the level tone. The fall-rise occurred in 3 instances. Group B used mostly the fall-rise (obviously treating the expression as a single IP) and the rise, yet the difference was not significant. One person employed the fall and one used the rise-fall, the latter of which is not expected in English, indicating the transfer of the Polish emotive function of the tone.

The expression *good afternoon* in both English and Polish can be pronounced with both the fall and the rise in formal, polite situations. As with *hello*, the rise helps to attract attention and engages more directly with the addressee, therefore was more appropriate for the context of the introduction to the weather forecast. The two groups used mostly the fall and the low fall, where the low fall was more frequent in the performance of group A, though the difference was not significant.
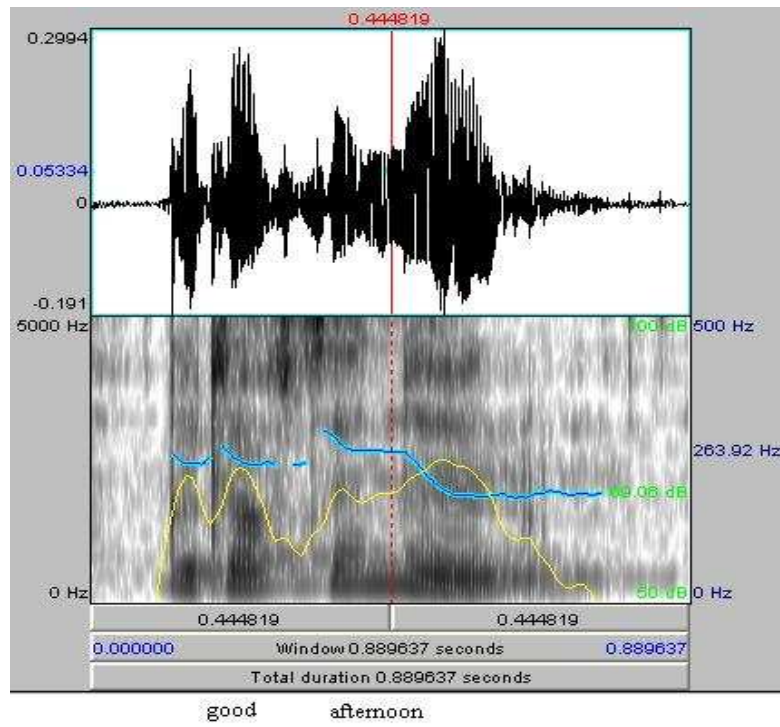
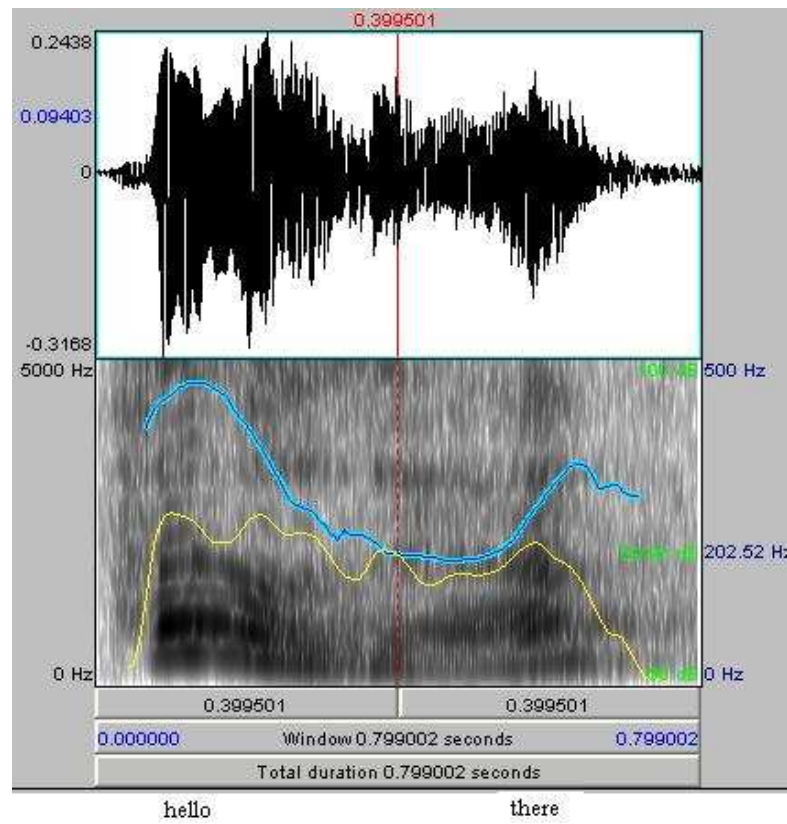Figure 4.25 A fall in a greeting expression (group A female voice)

Figure 4.26 A fall-rise in a greeting expression (group B female voice)

121

The analyses above illustrate general tendencies in the acquisition of grammatical functions of intonation. L2-like intonation patterns were used more frequently by group B in questions. Moreover, significantly fewer question tags were interpreted as genuine questions, which would not fit the context. However, the expressions with relatively fixed intonational patterns proved to be problematic for subjects from both groups indicating that even 4 semesters of phonetics classes may not suffice to successfully acquire the intonation of the second language. Tables 4.12, 4.13 and 4.14 provide the summary of the pragmatic uses of intonation in Polish, English and the learner's system.

Tone	Meaning	Context
Rising / high tones	Non-finality	General questions without *czy*, commands, acceptations (of command), polite statements
		Signalling continuation
Falling / low tones	Finality	Statements, explanations, acceptations (of a statement), *czy*-questions
		Signalling finality
Fall-rise / rise-fall	Specifically Polish	Emotional utterances, the former – positive approbation, the latter – disbelief / surprise

Table 4.12 The use of intonation in Polish (adapted from Karpiński 2006: 187)

Tone	Meaning	Context
Rising / high tones	Non-finality	y/n questions, order, 'gentle' *wh*-questions, tag questions (open to dis-/agreement), 'encouraging' statement
		Signalling continuation, encouragement
Falling / low tones	Finality	Statements, explanations, *wh*-questions, 'insistent' y/n questions, tag questions (seeking agreement), commands
		Signalling finality
Fall-rise	Local	Implicational: contrast, reservations, polite correction (statements, also in negative ones)
Rise-fall	Local	Impressed statements and y/n questions, 'gossip', grammatically marked exclamatives, challenging / sarcastic / ironic statements

Table 4.13 The use of intonation in English (based on Cruttenden 1986 and Wells 2006)

Tone	Meaning	Context
Rising / high tones	Non-finality	y/n questions, commands, most tag questions (esp. earlier stages), *wh*-questions (esp. earlier stages), polite statements
		Signalling continuation
Falling / low tones	Finality	Statements, explanations, acceptations (of a statement), y/n questions (resembling *czy*-questions, esp. earlier stages),
		Signalling finality
Fall-rise	Specifically Polish, non-finality	Emotional utterances
		Signalling continuation
Rise-fall	Specifically Polish	Emotional utterances

Table 4.14 The use of intonation in the Polish learner's system

4.3. Summary

Chapter 3 presented typological differences between English and Polish derived from Donegan and Stampe's (1993) holistic organization of languages, which allowed for the following statements about the intonational systems of the two languages.

a) Richer morphology and more flexible sentence structure in Polish makes intonation less varied than in English: there are fewer tones and they are predominantly simple. More complex tones, including tritonal ones, occur largely in emotional speech. The pitch range is also narrower than in English.

b) English morphology and syntactic structures cause intonation to be more variegated, as one of its functions is to signal certain grammatical relations.

According to Natural Phonology, the L2 learner has to regain access to universal processes in the course of the acquisition of L2 segments. The acquisition of intonation apparently follows a different path, since universal processes seem not to be completely suppressed in the L1. Languages employ narrower pitch range (cf. the Effort Code), as well as archetypal falls and rises (cf. the Production Code) with the universal interpretations. On the other hand, the fall and the rise may be used for different purposes, e.g. they may not signal the polarity between statements and questions in all contexts. Such merging of the paralinguistic and linguistic elements proves that intonation does belong to the "edges of language" and any attempt to separate the paralinguistic component from the linguistic one is, to a large extent, in vain.

What is significant, greater proficiency in English grammar, i.e. the acquisition of the less varied, barely inflected syntactic structures, may not have any impact on the acquisition of intonation. Moreover, neither does the better command of L2 segments: intonation follows its own path, as Wrembel's (2005) and the present study reveal. Finally, the acquisition of intonation starts late and is rarely successful.

The aim of chapter 4 was to verify the model of the acquisition of L2 intonation. As it was hypothesized, group A's performance revealed the existence of phonological processes, which needed to be 'ungrammaticalized' in order to shed their L1 settings and assume those of the L2. L1 influence was most evident in the speech of group A, especially in the timing of peak alignment resulting from deficiencies in the perceptual abilities of the subjects, and the transfer of L1 pragmatic and grammatical functions of intonation.

Group B have suppressed universal processes and L1 influence to a greater degree than group A, but not entirely. Intonational contours were used more appropriately, as long as the context of the experimental texts allowed the author to jugde them as 'appropriate'. It must be noted, though, that the author is aware of the difficulties in deciding which factors exactly guide the speakers in implementing a particular tone. On average, peak alignment was significantly later than in the performance of group A, however, not due to its delaying, but because of the previous acquisition of long vowels and rhythm. Despite this advancement, group B's peak timing was far from being target-like. One cannot forget, though, that the subjects, living in a monolingual country, may not have had enough comprehensive input, therefore none of them has reached native-like alignment. It is also not clear whether formal instruction is capable of affecting the acquisition of L2 alignment.

The conclusions resulting from the perception test remain tentative, due to the limitations inherent in the study of this phenomenon. Unlike segments, intonation resists an easy classification into categories, which was demonstrated by the study by Karpiński and Post (still in progress; after Karpiński 2006): it was possible to establish perceptual categories for patterns performing certain intonational functions (question vs. statement), but the categories proved too broad to draw clear-cut conclusions. The experiment in the present study concentrated on the skill to discern between simple and complex tones. All subjects recognized simple tones correctly, which was not surprising, yet they did not differ much in the ability to recognize complex tones. At the same time, the fall-rise and rise-fall appeared more frequently in group B's utterances. Clearly, a more detailed study is required in order to establish the strength of the relationship between perception and acquisition of intonation.

As it was mentioned in chapter 3, the conflict between the typological and language-specific, as well as universal preferences is resolved to the advantage of the former. Indeed, although universals predominate at the beginning of intonational acquisition, they are soon suppressed, unlike L1 preferences, which is supported by the fact that the number of archetypical simple low falling and low rising tones receded in the performance of group B, but certain L1 patterns persisted, such as the relatively fixed contours of certain expressions (greetings, apologies, question tags, etc.). The more frequent occurrence of the fall-rise also indicates the suppression of the tendency for the universals, however, the L2-specific meanings were not acquired by the more advanced group B. Moreover, peak alignment, although later than in Polish, was still closer to the L1 system.

Whether group B achieved or was "close" to native-like competence depends on which definition of "native-likeness" is adopted. Such aspects as peak alignment, the richness of the tonal inventory or the pragmatically appropriate usage of intonation patterns are the most obvious indicators of foreign accent. It is striking that the range of certain intonational patterns employed by group A and group B did not differ significantly, which indicates that acquisition was not finished, even after four semesters of formal phonetic training. The reason may lie in the learning setting, as well as in the contents of the course: the bulk of course materials concentrated on segments, therefore, successful attainment may require more emphasis on suprasegmentals or, perhaps, different methods of teaching.

More research is needed to confirm the results of the present study. Greater homogeneity of the experimental groups would appear beneficial and would contribute to greater reliability of the experiment. A similar research on less advanced learners could give an answer to the question when exactly the acquisition of intonation is launched, in other words, at which stage of SLA intonation begins to be acquired.

The present study concentrated on the interplay of universal and L1-specific preferences in intonational acquisition. Subsequent research should also investigate the influence of extraneous factors, such as cognitive and psychological ones, age, or gender, which will lead to the better understanding of the phenomenon.

Conclusions

The present dissertation aimed at investigating the nature of the acquisition of English intonation by Polish adult learners. The analysis of the Polish subjects' performance, as well as the research done by other authors, provided an insight into the mechanisms of learning and helped to propose a model of the acquisition of intonation. Moreover, it has been proved that Natural Phonology can provide a valid descriptive framework.

First of all, the research proved that the learners who had not done any formal course in English phonetics but whose command of English was rather advanced employed considerably simpler intonational patterns when speaking English. The learners who had completed the phonetics course produced significantly more complex tones. The findings prove that the acquisition of L2 intonation begins late in comparison with the acquisition of other aspects of L2 phonology, and that during its initial stages universal processes and language-specific preferences are operating. As it was mentioned in chapter 3, the L2 learners are equipped with a residue of universal processes which need to be unsuppressed if the learners want to reconstruct L2 patterns.

This operation needs to be accompanied by the development of L2 perceptual categories. The perception test in the presented experiment suggested that the perceptual categories were already present in the less advanced group: the two groups perceptual abilities seemed not to differ significantly.

Two intonational systems in the learners' minds must come into conflict. According to the assumptions of Natural Phonology, from the conflicting preferences those language-specific are chosen, which is supported by the research presented in chapter 4. Typological differences between English and Polish are depicted in chapter 3, which reveals how the morphology and syntax affect intonational systems of the two languages: contrary to English, Polish rich inflection and flexible sentence structure impoverish the inventory of intonation contours, an obstacle which Polish learners of English must overcome.

The present study revealed that the grammatical and pragmatic aspect of intonation is most strongly affected by L1. They are most clearly visible in expressions with fixed intonational patterns, such as greetings or apologies, and in the local meanings of individual tones, such as the fall-rise. Subjects from both groups committed mistakes by using various intonation patterns in inappropriate contexts, yet the instructed group performed better than the non-instructed one.

It is hoped that the dissertation will contribute not only to the discussion of language acquisition but also to the research into Polish intonation. As it was mentioned by many phoneticians (e.g. Hirst – Di Cristo 1998; Gussenhoven 2004), the prosodic systems of many Asiatic and African languages are well-described, contrary to those of well-known European languages.

Bibliography

Abelin, Å – J. Allwood. 2000. "Crosslinguistic interpretation of emotional prosody." *ISCA Workshop on Speech and Emotion: A conceptual framework for research.* Belfast.

Abercrombie, David. 1965. "Steele, Monboddo and Garrick." In *Studies in Phonetics and Linguistics*, London: Oxford University Press.

Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh University Press.

Abrahamsson, Niclas. 1996. "Natural Phonology and second language acquisition: problems and consequences." In: Kellerman, Eric, Bert Weltens, Theo Bongaerts (eds.). *EUROSLA 6. A selection of papers*. Amsterdam: VU Uitgeverij, 9-21.

Anderson, J. 1987. "The markedness differential hypothesis and syllable structure difficulty." In Ioup, G. – S. Weinberger. *Interlanguage Phonology: The acquisition of a second language sound system.* 279-291. New York: Newbury House/Harper & Row.

Archangeli, Diana – D. Terence Langendoen (eds.) 1997. *Optimality Theory: an overview.* Oxford: Blackwell.

Archibald, John. 1993. *Language learnability and L2 phonology. The acquisition of metrical parameters*. Kluwer Academic Publishers.

Archibald, John. 1998. *Second language phonology*. Chicago: Aldine.

Armstrong, L.E. – I.C. Ward. 1926. *Handbook of English intonation*. Leipzig: Teubner.

Asher, J. – R. Garcia. 1969. "The optimal age to learn the foreign language." *Modern Language Journal*, 53, 5, 334-41.

Austin, J.L. 1962. *How to do things with words*. Cambridge, MA: Harvard University Press.

Backman, N. 1977. Learner intonation: a pilot study. Paper read at Proceedings of the Second Language Research Forum, Los Angeles.

Bailey, C.J.N. 1978. "Native accent and learning English as a foreign language." *IRAL 16*, 229-240.

Beckman Mary E. – Julia Hirschberg. 1994. "The ToBI annotation conventions". mbeckman@ling.ohio-state.edu (date of access: 15 March 2003).

Beckman Mary E. 1995. "Problems of intonation." In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 1, 450-7. Stockholm, Sweden.

Beckman Mary E. – Gayle Ayers Elam. 1997. "Guidelines for ToBI labelling". The Ohio State University Research Foundation. www.ling.ohio-state.edu (date of access: 9 Aug. 2005).

Best, Catherine T.1994. "Emergence of non-native language influences in infants." In Goodman, J. – H. Nusbaum (eds.) *The development of speech perception.* Cambridge, MA: MIT Press.

Best, Catherine T. 1995. "A direct realist view of cross-language speech perception." In Strange, W. (ed.) *Speech perception and linguistic experience.* Baltimore: York Press.

Bergman, C.R. 1976. "Interference vs. independent development in infant bilingualism." In Keller, C. – R. Taeschner – S. Viera (eds.). *Bilingualism in the bicentennial and beyond.* New York: Bilingual Press, 86-95.

Berinstein, Ava. 1979. "A cross-linguistic study on the perception and production of stress." *UCLA Working Papers in Phonetics 47.* Los Angeles: University of California.

Berkovits, R. 1984. "A perceptual study of sentence-final intonation." *Language and Speech* 27, 291-308.

Birdsong, David. 1989. *Metalinguistic Performance and Interlinguistic Competence.* Berlin: Springer.

Bloomfield, Leonard. 1935. *Language.* London: George Allen and Unwin.

Boersma, Paul – David Weenink. 2002. Praat. Doing phonetics by computer. (a computer program, version 4.0.34).

Bohn, O. – J.E. Flege. 1990. "Interlingual identification and the role of foreign language experience in L2 vowel perception." *Applied Psycholinguistics 11*, 303-328.

Bolinger, Dwight L. 1951. "Intonation: levels versus configurations." *Word* 7, 199-210.

Bolinger, Dwight L. 1957. "On certain functions of accents A and B." *Litera*: 4, 80-89.

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

Bolinger, Dwight L. 1986. *Intonation and its parts. Melody in spoken English.* Stanford: Stanford University Press.

Bolinger, Dwight L. 1989. *Intonation and its uses. Melody in grammar and discourse.* London: Edward Arnold.

Bongaerts, Theo – Susan Mennen – Frans van der Slik. (1995). "Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis," in Singleton, David and Zsolt Lengyel (eds.) *The age factor in second language acquisition.* Clevedon, UK: Multilingual Matters, 30-50.

Botinis, A. (ed.). 2000. *Intonation analysis. Modelling and technology.* Dordrecht: Kluver Academic Press.

Brazil, David. 1975. *Discourse intonation I.* Birmingham: English Language Research Monographs.

Brazil, David. 1978. *Discourse intonation II.* Birmingham: English Language Research Monographs.

Brazil, David. 1997. *The communicative value of intonation in English.* Cambridge: Cambridge University Press.

Brown, Gillian – Kirsten L. Currie – John Kenworthy. 1980. *Questions of intonation.* London: Helm.

Brown, C. 1998. "The role of the L1 grammar in the L2 acquisition of segmental structure." *Second Language Research 14*, 136-193.

Brown, C. 2000. "The interrelation between speech production and phonological acquisition from infant to child." In Archibald, J. (ed.) *Second language acquisition and linguistic theory*. Oxford: Blackwell.

Bruce, Gösta. 1977. "Swedish word-accents insentence perspective." *Working Papers 12*. Lund: Lund University, 61-70.

Bruce, Gösta – Eva Gårding. 1978. "A prosodic typology for Swedish dialects." In Bruce, G. – E. Gårding – R. Bannert. *Nordic prosody*. Gleerup.

Carlisle 1997. " The modification of onsets in a markedness relationship: testing the interlanguage structural conformity hypothesis." *Language Learning 47*, 327-361.

Carmichael, Lesley. 2000. *Measurable degrees of foreign accent: a correlational study of production, perception and acquisition*. MA thesis, University of Washington.

Carmichael, Lesley. 2004. "Second language acquisition of suprasegmental phonology."

Chafe, Wallace L. 1974. "Giveness, contrastiveness, definiteness, subjects, topics, and point of view." In: Charles N. Li (ed.), *Subject and topic*. New York: Academic Press, 25-55.

Chomsky, Noam – Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Chun, Dorothy M. 1998. "Signal analysis software for teaching discourse intonation." *Language Learning and Technology* vol. 2, no. 1, 74-93.

Chun, Dorothy M. 2002. *Discourse intonation in L2. From theory and research to practice*. Amsterdam: John Benjamins.

Chun, J. 1978. "A survey of research in second language acquisition." In Croft, W. (ed.) *Readings on ESL: for teachers and teacher trainees*. Cambridge MA: Winthrop Publishers Inc., 181-198.

Cohen, Antonie – Johan, t'Hart. 1967. "On the anatomy of intonation." *Lingua 19*, 177-192.

Cook, Vivian. 1992. "Evidence for multicompetence." *Language Learning*, 42: 557-591.

Cook, Vivian. 1999. "Going beyond the native speaker in language teaching." *TESOL Quarterly*, 33: 185-209.

Corder, Pit. 1967. "The significance of learner's errors." *International Review of Applied Linguistics* 5, 161-169.

Corder, Pit S. 1971. "The elicitation of interlanguage." *IRAL:* special issue for B. Malmberg's sixtieth birthday. Heidelberg: Julius Groos.

Coulthard, M. – David Brazil 1981. "The place of intonation in the description of interaction." In Tannen, D. (ed.) *Analysing discourse: text and talk*. Washington D.C.: Georgetown University Press.

Couper-Kuhlen, Elizabeth. 1985. *An introduction to English prosody*. Max Niemeyer Verlag Tübingen.

Couper-Kuhlen, Elizabeth – M. Selting (eds.). 1996. *Prosody in conversation*. Cambridge: Cambridge University Press.

Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.

Cruttenden, Alan. 1997. *Intonation* (2nd edition). Cambridge: Cambridge University Press.

Cruz-Ferreira, M. 1986. "Non-native interpretative strategies for intonational meaning: an experimental study." In Leather, Jonathan – Allan James. 1986.

Cutler, A. – M. Pearson 1986. "On the analysis of prosodic turn-taking cues." In Johns-Lewis (ed.). *Intonation in discourse*. San Diego: College Hill, 139-155.

Crystal, David. 1969. *Prosodic systems and intonation in English*. London: Cambridge University Press.

Crystal, David. 1976. *The English tone of voice. Essays in intonation, prosody and paralanguage*. New York: St. Martin Press.

Crystal, David. 2003. *A Dictionary of linguistics and phonetics*. 5th edition. Malden: Blackwell Publishing.

Dauer, R.M. 1983. "Stress-timing and syllable-timing reanalysed." *Journal of Phonetics 11*, 51-62.

Davies, Alan. 1991. *The native speaker in applied linguistics*. Edinburgh: Edinburgh University Press.

Davies, Alan. 2003. *The native speaker: myth and reality*. Clevedon: Multilingual Matters.

Demenko, Grażyna. 1999. *Analiza cech suprasegmentalnych języka polskiego na potrzeby technologii mowy* [The analysis of suprasegmental features of Polish for speech technology]. Poznań: UAM.

Demuth, Katherine. 1995. "The acquisition of tonal systems." In Archibald, J. (ed.). *The acquisition of non-linear phonology*. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 111-134.

Donegan, Patricia. 1978 (1985). *On the Natural Phonology of vowels*. New York: Garland.

Donegan, Patricia – David Stampe. 1978. "The syllable in phonological and prosodic structure." In: Boll, A., J.B. Hooper (eds.). *Syllables and segments*. North-Holland Publishing Company.

Donegan, Patricia – David Stampe. 1979. "The study of Natural Phonology." In: Dinnsen Daniel A. (ed.) *Current approaches to phonological theory*. Bloomington: Indiana University Press, 123-173.

Donegan, Patricia – David Stampe. 1983. "Rhythm and the holistic organization of language structure." In Richardson, J. (ed.) *Chicago Linguistic Society*, 337-353.

Donegan, Patricia. 1993. "Rhythm and vocalic drift in Munda and Mon-Khmer." In *Linguistics of the Tibeto-Burman area*, vol. 16:1,1-43.

Dowd, J. – J. Zuengler – D. Berkowitz. 1990. "L2 social marking: research issues." *Applied Linguistics 11*, 16-29.

Dressler, Wolfgang U. 1984. "Explaining Natural Phonology." *Phonology Yearbook,* 1: 29-51.

Dressler, Wolfgang U. – Katarzyna Dziubalska-Kołaczyk. 1994. "Functional analysis in the study of second language acquisition." *Functions of language* 1, 2: 201-228.

Dressler, Wolfgang U. – Annemarie Karpf. 1994. "The theoretical relevance of pre- and protomorphology in language acquisition." *Yearbook of morphology* 1994 (1995): 99-122.

Dressler, Wolfgang U. 1996. "Principles of naturalness in phonology and across components." In: Hurch, Bernard, Richard A. Rhodes (eds.). *Natural Phonology: the state of the art*. Berlin: Mouton de Gruyter, 41-52.

Dressler, Wolfgang U. – Katarzyna Dziubalska-Kołaczyk. 1997. "Contributions from the acquisition of Polish phonology and morphology to theoretical linguistics." In: Eliasson, Stig, Ernst H. Jahr (eds.). *Language and its ecology: essays in memory Einar Haugen*. Trends in linguistics. Studies and monographs, vol. 100. Berlin: Mouton de Gruyter, 379-399.

Dressler, Wolfgang U. 1999. "What is *natural* in natural morphology?" *PragueLinguistic Circle Papers* 3:135-144.

DuBois, J.W. – S. Schuetze-Coburn – D. Paolino – S. Cumming. 1992. *Discourse transcription*. Santa Barbara: Santa Barbara Papers In Linguistics.

Dukiewicz, Leokadia. 1978. *Intonacja wypowiedzi polskich* [Intonation of Polish utterances]. Wrocław, Warszawa, Kraków, Gdańsk: Wydawnictwo PAN.

Dulay, Heidy – Marina Burt. 1972. "Goofing, an indicator of children's second language strategies." *Language Learning 22*, 234-252.

Dulay, Heidi – Marina Burt. 1973. "Should we teach children syntax?" *Language Learning 23*, 254-258.

Dulay, Heidy – Marina Burt. 1974. "Natural sequences in child second language acquisition." *Language Learning* 25: 37-53.

Durand, P. – A. Durand - R. Gubrynowicz – Bogusław Marek. 2002. "Polish: prosodic aspects of 'czy' questions" in *Proceedings of the Speech Prosody Conference*, pp. 255-258.

Dziubalska-Kołaczyk, Katarzyna. 1987. "Phonological rule typology and Second Language Acquisition." In James, Allan, Jonathan Leather (eds.) 193-206.

Dziubalska-Kołaczyk, Katarzyna. 1989. "A natural model of acquisition of second language phonology and the notion of relative markedness." In: Pensado, Carmen, J. Mendez Dosuna (eds.) *Naturalists at Krems*. Salamanca: Ediciones Universidad Salamanca, 43-52.

Dziubalska-Kołaczyk, Katarzyna. 1990a. *A theory of second language acquisition within the framework of Natural Phonology A Polish-English Contrastive Study*. Poznan: AMU Press.

Dziubalska-Kołaczyk, Katarzyna. 1990b. "Phonetic awareness of English in native vs. non-native speakers – a perception study." In: Fisiak, Jacek (ed.) *Further insights into Contrastive Analysis*. Amsterdam/Philadelphia: John Benjamins, 207-217.

Dziubalska-Kołaczyk, Katarzyna. 1998. "Self-organisation in early phonology." In: Puppel, Stanisław (ed.). *Scripta Manent. Publikacja wydziałowa z okazji 10-lecia istnienia Wydziału Neofilologii UAM 1988-1998*. Poznań: Motivex, 99-112.

Dziubalska-Kołaczyk, Katarzyna. 2002a. "Challenges for Natural Linguistics in the twenty first century: a personal view." In Dziubalska-Kołaczyk, Katarzyna, Jarosław Weckwerth (eds.). *Future challenges for Natural Linguistics*. Munich: Lincom: 103-128.

Dziubalska-Kołaczyk, Katarzyna. 2002b. *Beats-and-binding phonology*. Peter Lang: Frankfurt am Main.

Dziubalska-Kołaczyk, Katarzyna. 2003a. "How learners 'repair' second language phonology and whether they may become native speakers." In Sobkowiak, Włodzimierz, Ewa Waniek-Klimczak (eds.), 31-49.

Dziubalska-Kołaczyk, Katarzyna. 2003b. "Speech is in the ear of the listener: some remarks on the acquisition of second language sounds." In: Hales, Kimberli, Angela Terveen (eds.) *Selected papers from the sixth College-wide Conference*

for Students in Languages, Linguistics and Literature 2002. Honolulu: University of Hawai'I at Mānoa, 81-92.

Dziubalska-Kołaczyk, Katarzyna. 2005. "Native or non-native? This is the question: which English to teach in the globalizing world?" *Proceedings of PTLC 2005*, London: Department of Phonetics and Linguistics, UCL, CD-ROM.

Dziubalska-Kołaczyk, Katarzyna – Joanna Przedlacka (eds.). 2005. *English Pronunciation models: a changing scene*. Frankfurt: Peter Lang.

Eckman, Fred R. 1977. "Markedness and the contrastive analysis hypothesis." *Language Learning 27*, 315-330.

Eckman, Fred R. 1981. "On predicting phonological difficulty in second language acquisition." *Studies in Second Language Acquisition* 4, 18-30.

Eckman, Fred R. 1991. "The Structural Conformity Hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners." *Studies in Second Language Acquisition* 13, 23-41.

Eckman, Fred R. – G. Iverson. 1994. "Pronunciation difficulties in ESL: coda consonants in English interlanguage." In Yavas, M. *First and second language phonology*. San Diego: Singular Publishing Group, Inc., 251-266.

Eckman, Fred R. 1996. "A functional-typological approach to second language acquisition theory." In Ritchie W.C. and T.K. Bhatia (eds.) *Handbook of Second Language Acquisition*, pp. 195-211. San Diego: Academic Press.

Eckman, Fred R. 2004. "From phonemic differences to constraint rankings: research on second language phonology." *Studies in Second Language Acquisition* 26, 513-549.

Eisenstein, M. 1982. "A study of social variation in adult second language acquisition." *Language Learning 32*, 376-392.

Ellis, Rod. 1991. *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, Rod. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.

Erwin-Tripp, S. 1973. "Identification and bilingualism." In Dil, A. (ed.). *Language acquisition and communicative choice*. Stanford: Stanford University Press.

Fathman, A. 1975. "Language background, age and the order of acquisition of English structures," in Burt, M., H. Dulay (eds.) *On TESOL'75*. TESOL, Washington DC.

Felix, S. 1978. "Some differences between first and second language acquisition." In Snow, C. – N. Waterson (eds.) *The development of communication*. London: John Wiley & Sons, 469-479.

Flege, James Emil. 1987. "A critical period for learning to pronounce foreign languages." *Applied Linguistics 8*, 162-177.

Flege, James Emil. 1991. "Perception and production: the relevance of phonetic input to L2 phonological learning." In Huebner, T. – C. Ferguson. *Crosscurrents in second language acquisition and linguistic theories*. Amsterdam: John Benjamins.

Flege, James Emil – J.M. Munro - I.R.A. Mackay. 1995. "Factors affecting strength of perceived foreign accent in a second language." *Journal of Acoustical Society of America*, 97:5, 3125-3134

Flege, James Emil – E.M. Frieda – T. Nozawa. 1997. "Amount of native language (L1) use affects the pronunciation of an L2." *Journal of Phonetics*, 25, 169-186.

Foley, Joseph – Linda Thompson. 2003. *Language learning. A lifelong process*. London: Arnold.

Fonagy, Ivan. 1981. "Emotions, voice and music. Research aspects on singing." J. Sundberg, Stockholm. Publication no. 33 of the Royal Swedish Academy of Music.

Ford, C. - S.A. Thompson. 1996. "Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns." In Ochs, E. – E.A. Schlegloff – S.A. Thompson (eds.). *Interaction and grammar.* New York: Cambridge University Press, 134-184.

Fox, Anthony. 2000. *Prosodic features and prosodic structure.* Oxford: Oxford University Press.

Féry, C. 1993. *German intonation patterns*. Tübingen: Niemeyer.

Fry, Dennis B. 1968. "Prosodic phenomena." In B. Malmberg (ed) *Manual of phonetics*. Los Angeles: University of California, 364-410.

Fujisaki, Hiroya – S. Nagashima. 1969. "A model for the synthesis of pitch contours of connected speech." *Annual Report of the Engineering Research Institute.* Tokyo: Faculty of Engineering, University of Tokyo.

Gårding, Eva – A. Abramson. 1965. "A study of the perception of some American English intonation contours." *Studia Linguistica 19*, 61-79.

Gårding, Eva. 1981. "Contrastive prosody. A model and its applications." *Studia Linguistica 35*, 146-165.

Gårding, Eva. 1983. "A generative model of intonation." In Cutler, A – D.R. Ladd (eds.). *Prosody: models and measurements*. Berlin: Springer.

Gardner, Robert C. – W. Lambert 1959. "motivational variables in second language acquisition." *Canadian Journal of Psychology 13*, 266-272.

Gardner, Robert C. 1985. *Social psychology and second language learning. The role of attitude and motivation.* London: Edward Arnold.

Gass, Susan. 1988. "Integrating research areas: a framework for second language studies." *Applied Linguistics 9*, 198-217.

Geluykens, R. 1987. Intonation and speech act type: an experimental approach to to rising intonation in queclaratives. *Studia Linguistica*, 19, 61-79.

Gibbon, Dafydd. 1984. "The role of discourse in intonation theory", in: Dressler, Wolfgang U. et al. (eds.) *Phonologica. Proceedings of the 5th Phonology Meeting.* Cambridge: Cambridge University Press.

Giegerich, Heinz J. 1984. *Relating to metrical structure*. Bloomington: Indiana University Linguistics Club.

Goh, Christiane. 2000. "A discourse approach to the description of intonation in Singapore English." In Brown, A. – D. Deterding – E. Low (eds.). *The English language in

Singapore: research on pronunciation. Singapore: Singapore Association for Applied Linguistics, 35-45.

Goldsmith, John A. 1976. *Autosegmental phonology*. Doctoral dissertation, MIT. Distributed by Indiana University Linguistics Club, Bloomington.

Goldsmith, John A. 1978. "English as a tone language." *Communication and Cognition 11*, 3/4, 453-476.

Goldsmith, John A. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell.

Goodman, Judith C. – Howard C. Nusbaum. 1994. *The development of speech perception: the transition from speech sounds to spoken words*. Cambridge MA: The MIT Press.

Goodman, Judith C. – Lisa Lee – Jenny DeGroot. 1994. "Developing theories of speech perception," in Goodman, Judith C. and Howard C. Nusbaum. *The development of speech perception: the transition from speech sounds to spoken words*. Cambridge MA: The MIT Press.

Goto, H. 1971. "Auditory perception by normal Japanese adults of the sounds /l/ and /r/." *Neuropsychologia 9*, 317-323.

Grabe, Esther. 1998. *Comparative intonational phonology: English and German*. MPI Series in Psycholinguistics 7, Wageningen: Ponsen en Looien.

Grabe, Esther – Maciej Karpiński. 2003. "Universal and language-specific aspects of intonation in English and Polish." 15th ICPhS Barcelona, 1061-1064.

Grabe, Esther. 2004. "Intonational variation in urban dialects of English spoken in the British Isles," in Gilles, P. & J. Peters (eds) *Regional variation in intonation*. Tuebingen: Linguistische Arbeiten.

Grabe, Esther – Greg Kochanski – John Coleman. 2005. "The intonation of native accent varieties in the British Isles – potential for miscommunication?" In Dziubalska-Kołaczyk, Katarzyna, Joanna Przedlacka (eds.), *English pronunciation models: a changing scene*. Linguistic Insights Series, Peter Lang.

Grundel, J. – K. Houlihan – G. Sanders. 1986. "Markedness distribution in phonology and syntax." In Eckman, Fred – E. Moravcsik – J. Wirth (eds.). *Markedness*. New York: Plenum Press, 107-138.

Gussenhoven, Carlos. 1985. "The intonation of 'John and Mildred': post-nuclear generalisations." In Johns-Lewis, C. (ed.) *Intonation in discourse*. London: Croom Helm, 77-121.

Gussenhoven, Carlos. 1988. "Intonational phrasing and the prosodic hierarchy", in Dressler, W. et al. (eds) *Phonologica. Proceedings of the 6th International Phonology Meeting*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos – Aoyu Chen. 2000. "Universal and language-specific effects in the perception of question intonation." *Proceedings of the 6th ICSLP*. 91-94.

Gussenhoven, Carlos. 2001. "Suprasegmentals", in Smelser, N.J., P. Baltes (eds), *International Encyclopaedia of the Social and the Behavioural Sciences*. Oxford: Pergamon. 15294-15298.

Gussenhoven, Carlos. 2002a. "Phonology of intonation. State-of-the-article", in *Glot International 6*, No. 9/10, November/December (271–284).

Gussenhoven, Carlos. 2002b. "Intonation and interpretation: phonetics and phonology." In: *Speech prosody 2002. Proceedings of the 1st International Conference on Speech Prosody.* In Bernard Bel & Isabelle Marlien (eds), Aix-en-Provence: Laboratoire Parole et Language, Université de Provence, 47-57.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

Guiora, A. – R. Brannon – C. Dull. 1972. "Empathy and second language learning." *Language Learning 2.* 111-130.

Hadding-Koch, K. – M. Studdert-Kennedy. 1965. "Intonation contours evaluated by American and Swedish test subjects." In *Proceedings of the 5th International Congress of Phonetic Sciences.* Basel: S. Kargel, 326-331.

Halliday, Michael A.K. 1963. "Intonation in English grammar." *Transactions of the Philological Society*, 143-169.

Halliday, Michael A.K. 1967. *Intonation and grammar in British English.* The Hague: Mouton.

Halliday, Michael A.K. 1973. *Explorations in the functions of language.* London: Edward Arnold.

Halliday, Michael A.K. – R. Hasan 1976. *Cohesion in English.* London: Longman.

Hammarberg, Björn. 1988. "Acquisition of phonology." *Annual review of applied linguistics* 9: 23-41.

Hammarberg, Björn. 1990. "Conditions on transfer in phonology." In: Leather, Jonathan, Allan James (eds.). *New sounds 90.* University of Amsterdam, 198-215.

Hammond, Michael. 1997. "Optimality theory and prosody." In Archangeli, D. & D. T. Langendoen. *Optimality Theory: an overview.* Oxford: Blackwell.

Han, You-Mi – Jean-Yves Dommergues. 1999. "The early acquisition of prosody in bilinguals' weaker language." *ICPhS* San Francisco, 1459-1461.

t'Hart, Johan – Antonie, Cohen. 1973. "Intonation by rule: a perceptual quest." *Journal of Phonetics* 1, 309-327.

't Hart, Johan – René Collier – Antonie Cohen. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody.* Cambridge: Cambridge University Press.

Hayes, Bruce, Aditi Lahiri. 1991. "Bengali intonational phonology." In *Natural Language and Linguistic Theory* vol. 9 no. 1. Kluwer Academic Publishers.

Hayes, Bruce. 1995. *Metrical stress theory.* Chicago: University of Chicago Press.

Herschensohn, J. 2000. *The second time around – minimalism and L2 acquisition.* Philadelphia: John Benjamins.

Hirschberg, J. – Janet Pierrehumbert. 1986. "The intonational structuring of discourse." In Association for Computational Linguistics, *Proceedings of the Twenty-fourth Annual Meeting.* New York: ACL, 136-144.

Hirschberg, J. – D. Litman. 1987. "Now let's talk about 'now': identifying cue phrases intonationally." In Association for Computational Linguistics, *Proceedings of the Twenty-fifth Annual Meeting.* Stanford: ACL, 163-171.

Hirst, Daniel – Albert Di Cristo (eds).1998. *Intonation systems. A survey of twenty languages.* Cambridge: Cambridge University Press.

Hockett, Charles F. 1995. *A manual of phonology*. Baltimore: Waverley Press.

Hyman, L. M. 1975. *Phonology. Theory and analysis.* New York: Holt, Rinehart and Winston.

Ioup, G. (1984). "Is there a structural foreign accent? A comparison of syntactic and phonological errors in second language acquisition" *Language Learning*, 34, 2, 1-17.

Jakobson, Roman. 1941. *Child language, aphasia, and phonological universals.* The Hague: Mouton.

James, Allan R. 1996. "Second language phonology." In: Jordens, P., J. Lalleman (eds.). *Investigating Second Language Acquisition*. Berlin: Mouton de Gruyter, 293-320.

Janson, Tore. 1986. "Cross-linguistic trends in the frequency of CV sequences." *Phonology Yearbook 3*, 179-186.

Jassem, Wiktor. 1961. *Akcent języka polskiego* [Accent in Polish]. Prace Językoznawcze 31, Komitet Językoznawstwa. Kraków: PAN.

Jassem, Wiktor. 1984. "Structural units of intonation in the acoustic speech signal." In Dressler, Wolfgang et al. (eds) *Phonologica. Proceedings of the 5th International Phonology Meeting*. Cambridge: Cambridge University Press.

Jenkins, Jennifer. 2000. *The phonology of English as an international language*. Oxford: Oxford University Press.

Johns-Lewis, C. (ed.). 1986. *Intonation in discourse.* San Diego: College Hill.

Jones, Daniel. 1909. *Intonation curves*. Leipzig: Teubner.

Kahn, Daniel. 1976. *Syllable-based generalisations inEnglish phonology*. PhD dissertation, MIT.

Karcevskij, S. 1931. "Sur la phonologie de la phrase." *Travaux du Cercle Linguistique de Prague 4*, 188-227.

Karpiński, Maciej. 2002. "Polish intonational database PoInt: project report." Available from http://main.amu.edu.pl/~maciejk/idbhome.html [Accessed on 23rd May 2004]

Karpiński, Maciej. 2006. *Struktura i intonacja polskiego dialogu zadaniowego. [The structure and intonation of Polish task-oriented dialogues]*. Poznań: Wydawnictwo Naukowe UAM.

Keating, Patricia A. 2004. "Phonetic encoding of prosodic structure." *UCLA Working Papers in Phonetics* no. 103, pp. 48-63.

Kellerman, E. 1978. "Giving learners a break: native language intuitions as a source of predictions about transferability." *Working Papers on Bilingualism* 15, 59-92.

Kenworthy, Joanne. 1987. *Teaching English Pronunciation*. New York: Longman.

Kim, S. – E. Curtis – Leslie Carmichael. 2001. "Emphatic Korean and neutral Americans?" *Journal of the Acoustical Society of America*, 109, 2474-5.

Kisseberth, Charles. 1970. "On the functional unity of phonological rules." *Linguistic Inquiry 1*, 291-306.

Klinghardt, Hermann. 1927. *Übungen im deutschen Tonfall*. Leipzig: Quelle und Meyer.

Krashen, Stephen – Michael Long – R. Scarcella. 1979. "Age, rate, and eventual attainment in second language acquisition." *TESOL Quarterly 11*, 338-341.

Krashen, Stephen. 1981. *Second language acquisition and second language learning.* Pergamon.

Kreiman, J. 1982. "Perception of sentence and paragraph boundaries in natural conversation." *Journal of Phonetics 10*, 163-175.

Ladd, D. Robert. 1996. *Intonational phonology.* Cambridge: Cambridge University Press.

Ladefoged, Peter – M. Draper – D. Whitteridge. 1958. "Syllables and stress." *Miscellanea Phonetica 3*, 1-14.

Ladefoged, Peter. 1975. *A course in phonetics.* New York: Harcourt, Brace, Jovanovich.

Lado, Robert. 1957. *Linguistics across cultures.* Ann Arbor: University of Michigan Press.

Larsen-Freeman, Diane – Michael H. Long. 1991. *An introduction to second language acquisition research.* Harlow: Longman.

Laver, John D.M. 1994. *Principles of phonetics.* Cambridge: Cambridge University Press.

Lehiste, Ilse. 1970. *Suprasegmentals.* Cambridge, MA: MIT Press.

Lehiste, Ilse – J. Olive – L.A. Streeter. 1976. "Role of duration in disambiguating syntactically ambiguous sentences." *Journal of the Acoustical Society of America, 60*, 1199-1202.

Leather, Jonathan – Allan James. 1991. "The acquisition of second language speech." *Studies in Second Language Acquisition 13*, 305-341.

Leather, Jonathan – Allan James (eds.). 1997. *Second-language speech. Structure and process.* Berlin: Mouton de Gruyter.

Leather, Jonathan (ed.). 1999. *Phonological issues in language learning.* Malden: Blackwell Publishers.

Leben, William R. 1977. "Length and syllable structure in Hausa." *Studies in African Linguistics*, Suppl. 7, 137-143.

Lenneberg, E. (1967). *Biological foundations of language.* New York: John Wiley.

Liberman, A.M. – F. Cooper – D. Shankweiler – M. Studdert-Kennedy. 1967. "Perception of the speech code." *Psychological Review* 74, 431-461.

Liberman, Mark. 1975. *The intonational system of English.* Dissertation, MIT. Distributed by Indiana University Linguistics Club, Bloomington, 1987.

Liberman, Mark – Alan Prince. 1977. "On stress and linguistic rhythm." *Linguistic Enquiry 8*, 249-336.

Liberman, Mark – Janet Pierrehumbert. 1984. "Intonational invariance under changes in pitch range and length." In Aronoff, M. – R. Oehrle (eds.) *Language and structure.* Cambridge, MA: MIT Press, 157-233.

Liberman, A.M. – I.G. Mattingly. 1985. "The motor theory of speech perception revised." *Cognition 21*, 1-36.

Lieb, H. 1980. "Intonation als Mittel verbaler Kommunikation." *Linguistische Berichte 68*, 34-48.

Lively, Scott E. – David B. Pisoni – Reiko A. Yamada – Yoh'ichi Tohkura. 1994. "Training Japanese listeners to identify English /r/ and /l/." *Journal of the Acoustical Society of America*, 96, 4:2076-2087.

Long, Michael H. 1990. "Maturational constraints on language development." *Studies in Second Language Acquisition, 12*, 251-285.

Luthy, M. 1983. "Nonnative speaker's perceptions of English "nonlexical" intonation signals." *Language Learning 33(1)*, 19-36.

McCarthy, John J. – Alan Prince 1993. *Prosodic morphology 1: Constraint interaction and satisfaction*. MS Amherst, MA: University of Massachusetts and New Brunswick.

Magen, H.S. 1998. "The perception of foreign accented speech." *Journal of Phonetics*, 26, 381-400.

Makarova, Veronika. 2001. "Perceptual correlates of sentence-type intonation in Russian and Japanese." *Journal of Phonetics*, 29, 2, 137-54.

Major, Roy C. 1986. "The ontogeny model. Evidence from L2 acquisition of Spanish *r*." *Language Learning 36*, 453-504.

Major, Roy C. 1987. "The Natural Phonology of second language acquisition." In: James, Allan, Jonathan Leather (eds.), 207-224.

Major, Roy C. – E. Kim. 1996. "The similarity differential rate hypothesis." *Language Learning 46,* 465-496.

Major, Roy C. 2001. *Foreign accent: the ontogeny and phylogeny of second language phonology.*" Mahwah, NJ: Lawrence Erlbaum.

Marek, Bogusław. 1984. "The prosodic structure of intonational contours." In Dressler, Wolfgang et al. (eds) *Phonologica. Proceedings of the 5ᵗʰ International Phonology Meeting*. Cambridge: Cambridge University Press.

Mennen, Ineke. 1999. "Second language acquisition of intonation: the case of peak alignment." In Gruber, M.C., D. Higgins, K. Olson and T. Wysocki (eds.) *Chicago Linguistics Society 34II: the panels*. Chicago: Chicago Linguistic Society.

Milroy, James – Lesley Milroy. 1997. "Varieties and variation." In Coulmas, Florian (ed.) *The handbook of sociolinguistics*. Oxford: Blackwell, 47-64.

Munro, M.J. – T.M. Derwing. 1995. "Processing time, accent and comprehensibility in the perception of native and foreign-accented speech." *Language and Speech* 38, 289-306.

Nemser, William. 1971. "Approximative systems of foreign language learners." *International Review of Applied Linguistics 9*, 115-124.

Nespor, Marina – Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.

Neufeld, Gerald – E. Schneiderman. 1986. "Prosodic development." In: Scarcella, Robin C., Stephen D. Krashen (eds.). *Research in second language acquisition*. Rowley, MA: Newbury House, 105-109.

O'Connor, J.D. – G.F. Arnold. 1961. *Intonation of colloquial English*. London: Longman.

Ohala, J. 1983. "Cross-language use of pitch: an ethological view." *Phonetica 40*, 1-18.

Ohala, J. 1984. "An ethological perspective on common cross-language utilization of F0 voice." *Phonetica 41*, 1-16.

Ohala, J. 1994. "The frequency code underlines the sound symbolic use of voice of pitch." In L. Hinton et al. (ed.). *Sound symbolism*. Cambridge: Cambridge University Press, 325-347.

Oliver, Dominika – R. A. Clark. 2005. "Modelling Polish intonation for speech synthesis." *The Proceedings of Interspeech 2005 Conference*, Lisboa.

Oller, John – S. Ziahosseiny 1970. "The contrastive analysis hypothesis and spelling errors." *Language Learning 20*, 183-189.

Oller, John. (ed.). 1983. *Issues in Language Testing Research*. Rowley, MA: Newbury House.

Olson, L. – S. Samuels 1973. "The relationship between age and accuracy of foreign language pronunciation." *Journal of Psycholinguistic research 66*, 263-267.

Oreström, B. 1983. *Turn-taking in English conversation.* Lund: CWK Gleerup.

Paradis, M. 1998. "Neurolinguistic aspects of the native speaker." In Singh, R. (ed.) *The native speaker: multilingual perspective*. New Delhi: Sage.

Palmer, H.E. 1922. *English intonation, with systematic exercises*. Cambridge: Heffer.

Pater, Joe. 1997. "Minimal violation and phonological development." *Language Acquisition 6(3)*, 201-253.

Pastuszek-Lipińska, Barbara. 2004. "An overview of a research project and preliminary results of two experiments on perception and production of foreign language sounds by musicians and non-musicians." *Speech, music and hearing, TMH-QPSR, KTH*, 46: 61-74.

Patkowski, M.S. 1980. "The sensitive period and comprehension of speech." *Working Papers on Bilingualism,* 6, 1-17.

Patkowski, M.S. 1990. "Age and accent in second language. A reply to James Emil Flege." *Applied Linguistics*, 11, 73-89.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation.* PhD, MIT.

Pierrehumbert, Janet – Mary E. Beckman. 1988. *Japanese tone structure.* Cambridge, MA: MIT press.

Pierrehumbert, Janet – J. Hirschberg. 1990. "The meaning of intonational contours in the interpretation of discourse." In Cohen, P. – J. Morgan – M. Pollock (eds.) *Intentions in communication*. Cambridge, MA: MIT Press, 271-312.

Pike, Kenneth L. 1946. *The intonation of American English.* Ann Arbor: University of Michigan Press.

Piller, Ingrid. 2002. "Passing for a native speaker: identity and success in second language learning." *Journal of Sociolinguistics,* 6/2, 179-206.

Piske, Thorsten – Ian R.A. MacKay. 1999. "Age and L1 use effects on degree of foreign accent in English." *ICPhS* San Francisco, 1433-1436.

Pisoni, David, B. 1997. "Some thoughts on 'normalisation' in speech perception." In Johnson, K. – J.W. Mullennix (eds.) *Talker variability in speech processing*. San Diego: Academic Press, 9-32.

Price, P.J. – M. Ostendorf – S. Shattuck-Hufnagel – G. Fong. 1991. "The use of prosody in syntactic disambiguation." *Journal of the Acoustical Society of America, 90*, 2956-2970.

Prince, Alan – Paul Smolensky. 1993. *Optimality Theory: constraint interaction in Generative Grammar*. Piscataway, NJ: Rutgers University Centre for Cognitive Science.

Ramus, Franck – Marina Nespor – Jacques Mehler. 1999. "Correlates of linguistic rhythm in the speech signal." *Cognition 73*, 265-292.

Richards, J. 1971. "A non-contrastive approach to Error Analysis." *English Language Teaching* 25.

Rietveld, A.C.M. – Carlos Gussenhoven. 1985. "'On the relation between pitch excursion size and pitch prominence." *Journal of Phonetics* 15, 273-285.

Riney, Timothy J. – James E. Flege. 1998. "Changes over time in global foreign accent and liquid identifiability and accuracy." *Studies in Second Language Acquisition*, 20, 213-243.

Rintell, E. 1984. "But how do you feel about it? The learner's perception of emotion in speech." *Applied Linguistics 5*, 255-264.

Roca, Iggy – Wyn Johnson. 2000. *A course in phonology*. Malden, MA: Blackwell Publishing.

Sag, I.A. – M. Liberman 1975. "The intonational disambiguation of indirect speech acts." In Chicago Linguistic Society, *Papers of the Eleventh Regional Meeting*, 487-497.

Schaffer, D. 1984. "The role of intonation as a cue to topic management in conversation." *Journal of Phonetics 12*, 327-344.

Scherer, K. 1979. "Acoustic concomitants of emotional dimensions." In Weitz, S. (ed.). *Non-verbal communication: readings with commentary*. New York: Oxford University Press, 249-253.

Scovel, T. 1988. *A time to speak: a psycholinguistic inquiry into the critical period for human speech*. Rowley, MA: Newbury House.

Sheldon, A. – W. Strange. 1982. "The acquisition of /l/ and /r/ by Japanese learners of English." *Applied Psycholinguistics 3*, 243-261.

Seliger, H. 1978. "Implications of a multiple critical periods hypothesis for second language learning," in Ritchie, W. (ed.) *Second language acquisition research. Issues and implications*. New York: Academic Press.

Selinker, Larry. 1972. "Interlanguage." *International Review of Applied Linguistics 10*, 209-232.

Selkirk, Elisabeth. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge MA: MIT Press.

Selkirk, Elisabeth. 1995. "Sentence prosody: intonation, stress and phrasing", in: Goldsmith, John A. *The handbook of phonological theory*.

Selting, M. 1988. "The role of intonation in the organization of repair and problem handling sequences in conversation." *Journal of Pragmatics 12*, 293-322.

Setter, Jane. 2005. "Communicative patterns of intonation in L2 English teaching and lerning: the impact of discourse approaches." In Dziubalska-Kołaczyk, K. – J. Przedlacka (eds.). *English pronunciation models: a changing scene*. Frankfurt: Peter Lang.

Shibatani, Masayoshi. 1973. "The non-cyclic nature of Japanese accentuation." *Language 48*, 584-595.

Singh, Rajendra (ed.). 1998. *The native speaker: multilingual perspective*. New Delhi: Sage.

Singh, Rajendra. 2006. "IN, OF and ABOUT: some unfinished reflections from South Asia." A presentation during *Poznan Linguistic Meeting 2006*.

Snow, C. – M. Hoefnagel-Höhle. 1977. "Age differences and the pronunciation of foreign sounds." *Language and Speech 20*, 357-365.

Stampe, David. 1969. "The acquisition of phonetic representation." *Chicago Linguistic Society* 5: 443-453.

Stampe, David. 1979. *A dissertation on Natural Phonology.* Bloomington: Indiana University Linguistic club.

Stampe, David. 1980. *A dissertation on Natural Phonology.* New York: Garland.

Stampe, David. 1984. "On phonological representations." *Phonologica* 1984: 287-300.

Steele, Joshua. 1775. *An essay towards establishing the melody and measure of speech to be expressed and perpetuated by peculiar symbols.* London: privately printed. Also listed as *Prosodia Rationalis.* Microfiche reproduction published by Scolar Press, London, 1974.

Steffen-Batogowa, Maria. 1996. *Struktura przebiegu melodii polskiego języka ogólnego* [The structure of the Polish language melody]. Poznań: Sorus.

Steffen-Batogowa, Maria. 2000. *Struktura akcentowa języka polskiego: studium statystyczne* [The structure of Polish stress: a statistical study]. Warszawa: Wydawnictwo Naukowe PWN.

Stockwell, Robert P. 1960. The role of intonation in a generative grammar of English. *Language* 36: 360-7.

Stockwell, Robert P. – J. Bowen. 1965. *The sounds of English and Spanish.* Chicago: Chicago University Press.

Stockwell, Robert P. 1972. The role of intonation: reconsiderations and other considerations. In D. Bolinger (ed.) *Intonation.* Harmondsworth: Penguin, pp. 87-109.

Studdert-Kennedy, M. – K. Hadding-Koch. 1973. "Auditory and linguistic processes in the perception of intonation contours." *Language and Speech 16(4)*, 239-313.

Sweet, Henry. 1890. *A primer of spoken English.* Oxford: Clarendon Press.

Tahta, S.M. – M. Wood – K. Loewenthal. 1981. "Age changes in the ability to replicate foreign pronunciation and intonation." *Language and Speech*, 24, 4, 363-372.

Toivanen, Juhani. 2003. "Tone choice in the English intonation of proficient non-native speakers." *PHONUM* 9, 165-166.

Trager, G.L. – H.L. Smith. 1951. *An outline of English structure.* Oxford: Oxford University Press.

Trubetzkoy, Nikolai. 1939. *Prinziples of phonology.* Paris: Klincksieck.

Vaissière, Jacqueline. 1995. "Phonetic explanations for cross-linguistic similarities." *Phonetica*, 52, 123-30.

Vaissière, Jacqueline. 2005. "Perception of intonation," in Pisoni, David B., Robert E. Remez. *The handbook of speech perception.* Malden, MA: Blackwell Publishing.

Välimaa-Blum, Riitta. 1999. "Intonation in English, Finnish and French: ethology versus typology?" *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, California.

Verma, Schivendra – A. Mannering – B.M. Kornell – W.F. Kratz – C. Beach. 1994. "Prosodic cues for phrasal boundaries in productions by children and adults." *Journal of the Acoustical Society of America, 96(5)*, 3308.

Vogel, Irene. 1991. "Prosodic phonology: second language acquisition data as evidence in theoretical phonology", in Huebner, T. & Ch. A. Ferguson (eds.) *Crosscurrents in SLA and Linguistic Theories*. Amsterdam/Philadelphia: John Benjamins.

Ward, G. – J. Hirschberg. 1985. "Implicating uncertainty: the pragmatics of fall-rise intonation." *Language 61*, 747-776.

Wells, J.C. 2006. *English intonation. An introduction*. Cambridge: Cambridge University Press.

Wells, R.S. 1945. "The pitch phonemes of English." *Language 21*, 27-39.

Wennerstrom, Ann. 1994. "Intonational meaning in English discourse: a study of non-native speakers." *Applied Linguistics* 15/4, 399-420.

Wennerstrom, Ann. 2001. The music of everyday speech prosody and discourse analysis. New York.

Werker, J.F. – J. Gilbert – K. Humphrey – R. Tees. 1981. "Developmental aspects of cross-language speech perception." *Child Development 52*, 349-355.

Werker, J.F. – R. Tees. 1983. "Developmental changes across childhood in the perception of non-native speech sounds." *Canadian Journal of Psychology 37*, 278-286.

Wichmann, Anne. 2000. *Intonation in text and discourse. Beginnings, middles and ends*. Harlowe: Longman.

Wichmann, Anne. 2002. "Attitudinal intonation and the inferential process." *Speech Prosody 2002*, 11-5. Aix-en-Provence. http://www.lpl.univ-aix.fr/sp2002/pdf/wichmann.pdf [Accessed on 1st June 2004]

Willems, N. 1982. *English intonation from a Dutch point of view*. Dordrecht: Foris Publications.

Witaszek-Samborska, Małgorzata – Monika Gruchmanowi – Małgorzata Żak-Święcicka. 1987. *Mowa mieszkańców Poznania. [The speech of the inhabitants of Poznan]*. Poznań: Wydawnictwo Poznańskie.

Witaszek-Samborska, Małgorzata. 2006. "Odrębności w polszczyźnie miejskiej Poznania." [Differences in urban Poznan Polish]. *Polonistyka* 8, 20-25.

Wode, Henning. 1978. "The beginning of non-school-room phonological acquisition." *International Review of Applied Linguistics 16*, 109-125.

Wrembel, Magdalena. 2005. *Phonological metacompetence in the acquisition of Second Language phonetics*. Unpublished PhD thesis, Poznan: UAM.

Yang, L. 1995. *Intonational structures of Mandarin discourse*. Unpublished PhD thesis, Georgetown University, Washington D.C.

Yavas, M. 1994. "Final stop devoicing in interlanguage." In Yavas, M. (ed.). *First and second language phonology*. San Diego: Singular.

Young-Scholten, Martha. 1996. "A new research programme for the L2 acquisition of phonology." In: Jordens, P., J. Lalleman (eds.). *Investigating SLA*. Berlin: Mouton de Gruyter, 263-292.

Yule, George. 1980. "Speakers' topics and major paratones." *Lingua 52*, 33-47.

Zampini, M. 1997. "L2 Spanish spirantization, prosodic domains and interlanguage rules." In Hannahs, S.J. – M. Young-Scholten (eds.). *Focus on phonological acquisition*. Amsterdam: Benjamins.

Zborowska, Justyna. 1998. *Self-organisation in second language acquisition. The case of phonostylistics*. Unpublished M.A. thesis. Poznań: Adam Mickiewicz University.

Zborowska, Justyna. 2001. *Rhythmic organisation of language. A Polish-English contrastive study*. Unpublished PhD dissertation. Poznań: Adam Mickiewicz University.

Zharkova, Natalia N. 2002. 'Acquisition of prosody in Russian.' Aix-en-Provence, France. April 11-13.

Zobl, H. 1982. "A direction for contrastive analysis: the comparative study of developmental sequences." *TESOL Quarterly* 16, 169-183.

APPENDIX

Dialogue:

J: Do you remember where I put my glasses, Sally?

S: Of course, Joe. Check behind these books. Can you see them?

J: Yes. I've been looking for them all afternoon. I want to look through my old clothes.

S: What do you want to do with them?

J: I think I'll throw away that tattered T-shirt, that jumper, those trousers and trainers.

S: Which jumper?

J: Green one.

S: Green? That's blue.

J: Oh. And what colour are my trousers then?

S: I'd call the trousers brownish. You should really put your glasses on, shouldn't you?

J: But where are my trainers?

S: I'm so sorry. I must have thrown them away.

The monologue:

Hello there, good afternoon. Well, after a wet weekend, today is much brighter for most of us, although there are still some showers in parts of Scotland. But let's start across South-East England and East Anglia, where most places will be dry with sunny spells this afternoon, although there will be quite a lot of clouds at times. Top temperature 23 degrees in London, very pleasant here, in just a light breeze.

Possible interpretations

Dialogue (reading 1)

J: Do you remember where I put my glasses, Sally? (L+H* H H% high rise on *glasses*, *Sally* – a vocative, de-accented)

S: Of course, Joe. (declarative statement: H* L L%, a fall on *course*) Check behind these books. (imperative: a fall H* L L% on *books*) Can you see them? (inversion question: a high rise L+H* H H% on *them*)

J: Yes. I've been looking for them all afternoon. I want to look through my old clothes (three declaratives, a fall H* L L% on *yes, afternoon, clothes*)

S: What do you want to do with them? (wh-question: high fall, H* L L% on *do*)

J: I think I'll throw away that tattered T-shirt, that jumper, those trousers and trainers. (a list: first two items: a rise L* L H%, penultimate item, *trousers*, a fall-rise: H* L H%, final item: high fall, H* L L%)

S: Which jumper? (wh-question: high fall H* L L% on *jumper*)

J: Green one. (high fall H* L L% on *green*)

S: Green? (exclamation/surprised echo question: fall-rise H* L H%) That's blue. (declarative: high fall H* L L%)

J: Oh. (exclamation: high fall H* L L%) And what colour are my trousers then? (wh-question: high fall H* L L% on *trousers*)

S: I'd call the trousers brownish. (high fall: H* L L% on *brownish*) You should really put your glasses on, (low fall L* L L% on *really*, low key) shouldn't you? (tag question, a fall H* L L%)

J: But where are my trainers? (wh-question: high fall, H* L L% on *trainers*)

S: I'm so sorry. (apology: fall-rise H* L H% on *so*) I must have thrown them away. (declarative, a fall H* L L% on *away*)

Dialogue (reading 2, when love breaks down)

J: Do you remember where I put my glasses, Sally? (high fall H* L L% on *glasses*)

S: Of course, Joe. (grumpy: a rise-fall L+H* L L% on *course*) Check behind these books. (imperative: a fall H* L L% on *books*) Can you see them? (ironic inversion question: a high rise L+H* H H% on *see*, extra long vowel in *see*)

J: Yes. I've been looking for them all afternoon. I want to look through my old clothes (three declaratives: high fall H* L L% on *yes, afternoon, clothes*)

S: What do you want to do with them? (wh-question: high fall H* L L% on *do*)

J: I think I'll throw away that tattered T-shirt, that jumper, those trousers and trainers. (four IPs: a high fall L* L H%, on *away*, *trousers*, *jumper*, *trousers*, *trainers*)

S: Which jumper? (angry wh-question: high fall H* L L% on *jumper*)

J: Green one. (a fall H* L L% on *green*)

S: Green? (rise L* L H%) That's blue. (high fall H* L L%)

J: Oh. (ironic exclamation: rise-fall L+H* L L%) And what colour are my trousers then? (ironic wh-question: rise-fall L+H* L L% on *trousers*)

S: I'd call the trousers brownish. (high fall H* L L% on *brownish*, extra long vowel in the last word) *J:* You should really put your glasses on, (high fall H* L L% on *really*) shouldn't you? (tag question, high fall H* L L%)

J: But where are my trainers? (wh-question: high fall H* L L% on *trainers*)

S: I'm so sorry. (high fall H* L L% on *sorry*) I must have thrown them away. (ironic rise-fall L+H* L L% on *away*)

The monologue

He<u>llo</u> there, good after<u>noon</u>. (casual, mid-level L* L L%) Well, after a wet week<u>end</u>, (fall-rise H* L H%) today is <u>much</u> brighter for most of us, (fall+rise with nucleus on *much* H* L H%) although there are <u>still</u> some showers (high fall H* L L%) in parts of Scotland. (high fall H* L L%) But let's start across South-East <u>Eng</u>land and East <u>Ang</u>lia, (fall H* L L%) where <u>most</u> places will be dry (fall H* L L%) with sunny spells this after<u>noon</u>, (fall-rise H* L H%) although there will be <u>quite</u> a lot of <u>clouds</u> at times. (fall H* L L%) Top <u>temperature</u> (low rise L* L H%) 23 degrees in <u>Lon</u>don, (low rise L* L H%) very <u>plea</u>sant here, (low rise L* L H%) in just a light <u>breeze</u> (high fall H* L L%).