

Univerzita Karlova
Filozofická fakulta

DIPLOMOVÁ PRÁCE

2023

Bc. Veronika Vonzová

Univerzita Karlova
Filozofická fakulta
Ústav anglického jazyka a didaktiky

Diplomová práce
Bc. Veronika Vonzová

**Melodické průběhy v politických debatách: srovnání britské a americké
angličtiny**

**Melodic contours in political debates: comparison of British and American
English**

Poděkování

Ráda bych poděkovala všem vyučujícím Ústavu anglického jazyka a didaktiky za jedinečné roky naplněné inspirativními setkáními, za jejich přístup, zájem, neutuchající podporu a vstřícnost při studiu. Chtěla bych také poděkovat své rodině, kolegům a kamarádům a Alence, kteří mě celé studium podporovali, jak se dalo, pomáhali s mnohdy složitým časovým rozvrhem a vycházeli ve všem vstříc. Bez vás bych to studium zvládala mnohem hůř.

Mé velké díky patří doc. Mgr. Radku Skarnitzlovi, Ph.D., který mi byl nejen skvělým vedoucím této diplomové práce, ale i mentorem na celé cestě studiem. Jeho neochvějná důvěra, rady, podpora a ochota mi umožnily dostudovat s radostí.

Prohlašuji, že jsem diplomovou práci vypracoval/a samostatně, že jsem řádně citoval/a všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V dne

Podpis

Souhlasím se zapůjčením diplomové práce ke studijním účelům.

I have no objections to the MA thesis being borrowed and used for study purposes.

Abstrakt

Tato diplomová práce se zaměřuje na rozdíly charakteristik prozodické fráze v britské a americké angličtině (délka fráze, intonační pohyb, intonační rozpětí). Snaží se tím zaplnit mezeru v komparativním výzkumu, která spočívá v nedostatku studií porovnávajících výše zmíněné variety angličtiny v uvedených aspektech suprasegmentální úrovně. Teoretická část popisuje prozodii a její funkce v jazyce, také intonaci, jednotlivé typy intonačních pohybů a jejich funkce. Zaměřuje se také na rozdíly mezi britskou a americkou tradicí zápisu intonace v průběhu analýzy a na rozdíly a nesrovnalosti v jejich terminologiích. Nakonec také uvádí nový termín, inspirovaný českou terminologií, tzv. FMU, který v anglické terminologii chybí a který pomáhá při rozlišování funkce a formy jednotlivých intonačních pohybů. FMU odpovídá českému termínu melodém (tj. „funkce“ intonačního pohybu, např. klesavý melodém) a stojí proti konkrétnějšímu typu průběhu melodému, kadenci („forma“, konkrétní realizace intonačního pohybu, např. klesavo-rovná kadence). Empirická část práce nabízí studii provedenou na vzorku 800 frází (400 britských a 400 amerických), které jsou porovnány v následujících oblastech: délka fráze, intonační pohyb a intonační rozpětí mluvčích. Studie je zakončena diskuzí výsledků a krátkým přehledem doporučení pro další výzkum v podobné oblasti.

Klíčová slova: prozodie, intonace, fráze, délka, intonační pohyb, intonační rozpětí, FMU, terminologie

Abstract

The thesis focuses on the differences between British and American English in terms of their phrasal characteristics (length of phrases, nuclear pitch movement, pitch range). By doing so, it aims to fill a gap in research which lies in a lack of comparative studies of the two mentioned varieties of English on the suprasegmental level. The theoretical part describes prosody and its functions in language, intonation, and individual types of nuclear pitch movements and their functions. The paper further concentrates on the differences between the British and American schools of intonation notation and also the differences and discrepancies between their terminologies. Finally, it introduces a new term inspired by Czech terminology, FMU, which is not present in English terminology and helps to differentiate between the form and function of the nuclear pitch movements. FMU corresponds to the scope of the Czech term “melodém” (the function of the pitch movement, e.g., the falling FMU), as opposed to the realization (Czech “kadence”, the concrete realization of the FMU, e.g., the level-fall). The empirical part presents a study carried out on a sample of 800 phrases (400 British and 400 American), which are compared in the following areas: the length of phrases, type of the nuclear pitch movement, and speakers’ pitch range. The study is concluded with a discussion of the results and with a short overview of the implications for further research.

Key words: prosody, intonation, phrase, length, nuclear pitch movement, pitch range, FMU, terminology

Table of contents

List of abbreviations	9
List of figures	10
List of tables	12
1. Introduction	13
2. Prosody	14
2.1. Functions of prosody	17
2.1.1. <i>Lexical function</i>	17
2.1.2. <i>Grammatical function</i>	18
2.1.3. <i>Accentual function</i>	21
2.1.4. <i>Discourse function</i>	22
2.1.5. <i>Affective and indexical function</i>	23
3. Intonation	23
3.1. Intonation: the form and function of individual nuclear pitch movements	26
3.2. Intonation in British English.....	31
3.3. Intonation in American English.....	32
4. The British and American traditions and terminology	34
4.1. Lacking English terminology in intonation analysis	40
5. Research questions and hypothesis	44
6. Material and method	44
6.1. Material.....	44
6.2. Method.....	46
7. Results	47
7.1. Prosodic phrases in numbers	48
7.2. The types of the nuclear pitch movements in phrases	50
7.3. The pitch range of the whole phrase.....	53
7.4. The pitch ranges of the individual nuclear pitch movements	55
7.4.1. <i>The pitch range in falls</i>	58
7.4.2. <i>The pitch range in rises</i>	60
7.4.3. <i>The pitch range in rise-falls</i>	62
7.4.4. <i>The pitch range in fall-rises</i>	64
8. Discussion	66
9. Conclusion	68
10. References	69
11. Resumé	75
12. Appendix	81

List of abbreviations

FMU	Functional melodic unit
FSP	Functional sentence perspective
GA	General American English
HRT	High rising terminal
RP	Received Pronunciation
SBE	Southern British English

List of figures

Figure 1: The prosodic hierarchy	16
Figure 2: a use of pitch for a lexical distinction between two words in Kono	17
Figure 3: Visual representation of a phrasal structure.....	27
Figure 4: wh-question, accompanied by different pitch movements	27
Figure 5: a change in sentence type caused by a different pitch movement	28
Figure 6: The same utterance pronounced with	28
Figure 7: Low-rising intonation, with low pitch accent for “I don’t want to talk to anyone”. 33	
Figure 8: intonational idiom in AmE I.....	33
Figure 9: intonational idiom in AmE II.....	33
Figure 11: AmE notation of intonation	35
Figure 10: AmE notation of intonation	35
Figure 12: AmE notation of intonation, the ToBI system.....	36
Figure 13: the difference in pitch movement progression.....	41
Figure 14: an analysed phrase with marked boundaries and nuclear pitch movement in Praat	47
Figure 15: the comparison of individual speakers in terms of the number of syllables in a phrase between AmE (left) and BrE (right)	49
Figure 16: the frequency of each nuclear pitch movement in AmE (left) and BrE (right)	50
Figure 17: the comparison of individual speakers’ pitch ranges of the whole phrases in AmE (left) and BrE (right)	53
Figure 18: the pitch range of the individual nuclear pitch movements in AmE (left) and BrE (right).....	55
Figure 19: the original phrase with wrongly computed pitch range in the analysis process (91 Hz to 88 Hz)	56
Figure 20: the 80-perc range in falling nuclear pitch movement in AmE speakers (left) and BrE speakers (right)	58
Figure 21: the falling nuclear pitch movement in speaker AM-M2 (a drop from 179.8 Hz to 121 Hz; 6.8 ST)	59

Figure 22: the 80-perc range in rising nuclear pitch movement in AmE speakers (left) and BrE speakers (right)	60
Figure 23: the pitch range of speaker BR-F2’s rising nuclear pitch movement (rise from 158.3 Hz to 211.8 Hz)	61
Figure 24: the 80-perc range in rise-falling nuclear pitch movement in AmE speakers (left) and BrE speakers (right).....	62
Figure 25: rise-fall in speaker AM-M2, reaching 14.8 ST	63
Figure 26: rise-fall in speaker BR-M4, reaching 3.8 ST (113 Hz to 91 Hz).....	63
Figure 27: the 80-perc range in fall-rising nuclear pitch movement in AmE speakers (left) and BrE speakers (right).....	64
Figure 28: fall-rise in speaker AM-M4 with pitch range reaching 13.1 ST (201 Hz to 91 Hz)	65
Figure 29: fall-rise in speaker BR-F3 with pitch range reaching 2.4 ST	65

List of tables

Table 1: an overview of quantitative data for BrE and AmE (mean number of syllables and words in a phrase, articulation rate)	48
--	----

1. Introduction

While the segmental differences between individual English varieties have been described in many studies, books and textbooks, comparative studies of the British and American English varieties in the realm of suprasegmental features are relatively scarce to this day. The aim of the current thesis is therefore to fill this gap by comparing British and American prosodic phrases in terms of their length, the type of nuclear pitch movement employed and the speakers' pitch range.

Firstly, the theoretical part maps the field of prosody in part 2, describes what it is and what functions it serves in language. Secondly, it concentrates on intonation, while it describes the form and function of intonation and the individual nuclear pitch movements in section 3. Thirdly, the theoretical part summarizes the differences between the British and American traditions of notation (section 4) in the process of intonation analysis and explains how it poses a problem in comparing these two varieties in terms of their intonation structures. Last but not least, it also discusses the insufficient English terminology when it comes to the form and function of nuclear pitch movements (subsection 4.1). The paper offers a solution in the form of a newly coined term, the functional melodic unit (FMU), which is equivalent to the Czech term "melodém", but which has not been used in the English terminology as a separate term.

The empirical part of the thesis presents a study which focuses on the differences between British and American English in the characteristics of a prosodic phrase (the length of phrase in syllables and words, the rate of speech, the types of nuclear pitch movement, the pitch range of the whole phrase and the pitch range of the nuclear pitch movement). The material consists of 800 phrases (400 British and 400 American) which were obtained from political speeches in public debates and compared in the previously mentioned areas. The results are presented in section 7, accompanied by an analysis of the revealed differences, and they are followed by a final discussion in section 8 and a conclusion in section 9. The paper also briefly discusses the implications for further research.

2. Prosody

It is not what you said, it is how you said it, is a commonplace, but “a common-sense saying” (Levis, 2013, p. 5). It captures the essence of a very powerful component of language: prosody.

English prosody has been studied for centuries now, with the first attempts to describe it dating back to 1569, the beginning of the Modern English era (Ward, 2019, p. 62). Human communication stretches far beyond mere words, phrases, and grammatical rules of a language. If we stripped language of the prosodic layer, what would be left is a mere robotic stream of words, surprisingly difficult to decode by other human beings. Conversations would become just alternating short speeches (Ward, 2019, p. 1) with no rhythmical structure and no clear peaks created by accents or intonation. We can imagine prosody as an umbrella that covers the segments and gives them a specific form and, in colloquial terms, colour. The prosodic features help speakers navigate the meaning of spoken words, while they serve as “road signs” (Gilbert, 2008, p. 2): the speakers use them to make their thoughts understandable, and listeners follow these cues to decode the meaning of the words they hear. Mainly with the arrival of new technologies (like automatic speech synthesis, or speech recognition), it has become clear that prosody, specifically intonation, will have to be given a much greater deal of attention for the synthesised text to be more or less effortlessly understandable (Hirst, 1998, p. 2). That realization triggered more research in the realm of prosody, which brought many interesting findings and accentuated the importance of prosody for language users.

Ohala and Gilbert (1981), for example, described that speakers were capable of distinguishing the language they heard based on its prosody alone. Another study done by Mehler et al. (1988) showed that already four-day-old infants were capable of distinguishing the prosody of their native language from the prosody of other languages (Mehler et al. 1988, cited in Hirst, 1998, p. 2). Linguists have also discovered that when children get to their babbling stage, they start to produce “utterances” that have characteristic prosodic melodies of the language they are exposed to (Schreiber, 2009, p. 159). It has also been proven that prosody, especially its melodic and temporal aspects, is crucial in the communication process (Skarnitzl & Hledíková, 2022, p. 2). Firstly, it impacts speakers’ intelligibility and comprehensibility; Intelligibility is the actual degree of comprehension, in other words, it measures to what extent the message was understood as intended, and comprehensibility is the amount of effort one has to exert to understand what is being said (Derwing & Munro, 2015, p. 5). Prosodic features can

influence these two spheres to a great extent (by unnatural pausing, variable tempo, etc.). Secondly, prosody affects what impression the speakers make on listeners (Skarnitzl & Hledíková, 2022, p. 2). The impression the sent message leaves on the other participant of the discourse, the listener, may be influenced by a specific use of prosody; In some cases, for example, people mean the opposite of what they are saying with the actual words, and they make use of prosody to make other people understand that (Wharton, 2012, p. 98). Such strategies are quite common in communication.

However, delineating more specifically what “prosodic features” encompass is far from clear-cut. Linguists differ in their approaches to prosody and the definitions also vary across linguistic fields. The number and type of the particular features that prosody covers are variable, and one of the main factors that influences the delineation is the degree of abstraction of the term prosody itself. Shattuck-Hufnagel and Turk (1996) present prosody in quite an abstract manner, when they explain that it “refers to the phonological organization of segments into higher-level constituents and to the pattern of relative prominences within these constituents” (Shattuck-Hufnagel & Turk, 1996, p. 196). Similarly, other linguists may use the term prosody as a term for “the structure that organizes sound” (Cutler, Dahan & van Donselaar, 1997, p. 142). These broad types of definitions regard prosody as properly linguistic (Wharton, 2012, p. 99), but they do not mention any specific aspects which are covered by the term prosody nor do they specify anything about the realization (Cutler et al., 1997, p. 142). Due to the abstract, linguistic nature of these definitions, their adherents do not include “questions of speaker identification” (Cutler et al., 1997, 142), such as the speaker’s emotional state, attitude, or identity, as being relevant to and channelled through language prosody (Mennen & de Leeuw, 2014, p. 184).

Other linguists rather focus on the pragmatic aspect of prosody; they make the definition more specific by using prosody as a synonym for suprasegmental features of speech, that is the realization itself (Cutler et al., 1997, p. 142). That, by referring to anything that exists above the individual segments, covers quite a large scale of phenomena, like intonation, tempo, loudness, stress or rhythm. Bolinger, a proponent of this second definition, also stresses the important role prosody plays in expressing emotions and their degrees of intensity (Bolinger, 1983a, p. 98), an aspect that the proponents of the broader definition discard as being irrelevant in prosody studies. He also complements the broader definition by saying that prosody assists grammar but that it is not ultimately grammatical in its core, because it has roots in the “natural

behaviours from which it evolved”, meaning for example eyebrow movement or nodding a head (Bolinger, 1983a, p. 106, cited in Wharton, 2012, p. 100). The supporters of viewing prosody in terms of suprasegmental features might diverge further; they differ in the number of suprasegmental features they consider important for their purposes, or if not the number, then the degree of importance the individual suprasegmentals play in the language structure.

The most common and broadly applied definition can be found somewhere between the two already mentioned. It views prosody as “the linguistic structure which determines the suprasegmental properties of utterances”, in other words, abstract structure “coupled to a particular type of realization” (Cutler et al., 1997, p. 142). It also has been proposed that prosodic effects exist along a continuum from the previously mentioned “natural” to language-specific (Gussenhoven, 2002) and many phonologists then claimed that prosody encodes both linguistic and paralinguistic meaning (Wharton, 2012, p. 100).

It follows naturally that proponents of each of the streams look at prosody differently and employ different terms. Also, not only personal preferences, but also the type of research is an important factor in the definition of prosody; no one definition is valid for all research (Cutler et al., 1997, p. 142). Whatever delineation one chooses to follow, it is necessary to mention that prosody cannot be separated from language, it cannot be taken away, and it is present in any utterance we produce. The prosodic structure is closely related to the syntactic structure, and they interact, even though they are not always in agreement. A lot of theoretical work about prosody is based on a so-called *prosodic structure theory*, which regards the prosodic organization of sentences as being a hierarchically ordered structure distinct from syntactic structure (Elfner, 2018, p. 2). The prosodic structure belongs to phonological grammar and is somewhat of a mediator between syntax and the phonetic output of the speech stream (Elfner, 2018, p. 3).

The prosodic organization of a sentence has long been perceived as universal across languages (Elfner, 2018, p. 3). A simple hierarchy is presented in figure 1 below.

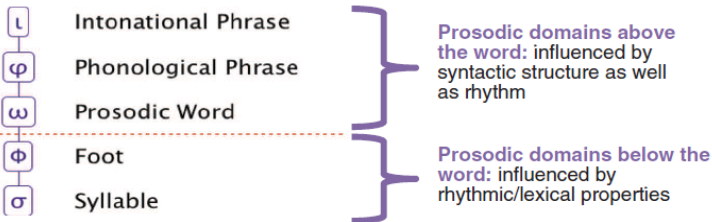


Figure 1: The prosodic hierarchy (Elfner, 2018, p. 3)

Research has, however, proposed including additional prosodic categories to this hierarchy specific to individual languages and their needs (Elfner, 2018, p. 3). Elfner mentions Japanese or Basque, which, as lexical pitch accent systems, require some kind of intermediate categories such as *Minor/Accentual Phrase* and *Major Phrase*, replacing the phonological phrase in the figure above (Elfner, 2018, p. 3). Even the system of description of intonation has proven to be subject to the specific needs of a particular language; the ToBI system, for example, which was created as a universal framework for intonation description across languages, has now been fragmented into various versions which can more aptly suit individual languages (see section 4 for details on the differences in the notation systems). This individualization serves the needs of individual languages, but it also decreased the universality of prosodic terminology which therefore became more non-uniform (see chapter 4 for more details on the non-uniformity of terminology in the intonation analysis).

It has been briefly outlined what prosody means and how it can be viewed. Let us now turn to how it is related to language and what specific functions it performs or channels.

2.1. Functions of prosody

2.1.1. Lexical function

The functions of prosody differ across languages. Some use prosody (mainly intonation) lexically (Wells, 2006, p. 3). Such languages are called “tone” or “tonal” and we put Chinese, Thai or some African languages, like Kono, in this category. In these languages, using certain prosodic cues incorrectly might change the very meaning of the word, which can result in a breakdown of communication. In figure 2 below, an example of a difference between two words in Kono is displayed; here it is the difference in pitch level (high vs. low), which causes the lexical-semantic difference (Roach, 2009, p. 122).

High level	$\bar{b}\epsilon\eta$ ('uncle')	$\bar{b}uu$ ('horn')
Low level	$_b\epsilon\eta$ ('greedy')	$_buu$ ('to be cross')

Figure 2: a use of pitch for a lexical distinction between two words in Kono

On the other end of the spectrum, we would find languages like Czech or English, the so-called “non-tonal” languages, which do not use prosody (specifically, intonation) in the same way Chinese, Thai or Kono does. Changes in prosody do not result in a complete change of lexical meaning as we have seen above, some prosodic aspects are, nevertheless, obligatory for a successful message transmission; for example, lexical stress placement (Frazier, 2006, p.

245). Changes in such respects then cause certain changes in meaning. That is one of the important functions prosody serves in English. Since the thesis is overall concerned with the English language, the following parts will focus on English prosody only.

English uses prosody to distinguish word classes; by shifting lexical stress in a word, the word changes its word class, as in for'bear x 'forebear (Cutler et al., 1997, p. 150) or 'import (noun) x im'port (verb). Nevertheless, such cases which only require a stress shift to change a word class are rare in languages. Usually, such a shift of stress placement is accompanied by a change in vowel quality, as in 'contest ['kɒntest] and con'test [kən'test], per'fect ['pɜ:fɪkt] and 'perfect [pə'fekt] etc., which leads to the vowels in the unstressed syllables being weakened (Mennen & de Leeuw, 2014, p. 186; Cutler et al., 1997, p. 150). That means that the distinction between the word classes is not dependent on the stress placement alone but is also supported by the quality changes in several vowels of the word (Wells, 2006, p. 3).

2.1.2. Grammatical function

Secondly, prosody is a tool for meeting our communication needs on levels above words, too. The analysis on the higher level is usually built around the system of phrases (these phrases exist under different terminology, the most commonly used terms are intonational or prosodic phrases, prosodic units, but other terms, like tone groups, may also be encountered. In this thesis, they are going to be addressed as prosodic phrases, but the original terminology is going to be preserved in citations). This function of prosody is called prosodic phrasing and it relates to the way speakers group words together within an utterance. These groups of words cohere semantically, and they express the preferred rhythmic and intonational patterns of a given language (Cole, 2010, p. 1142). The arrangement of language chunks into phrases is also known as tonality, a term which started to be used in the 60s by Halliday. Tonality reflects the speaker's perception of the number of units they deliver in their speech (Tench, 2020, p. 250) and it is the first part of a rather subconscious process of linguistic planning.

Prosodic phrases are variable in terms of their length. Crystal (1969) noticed in his corpus that the phrases (which he calls intonation units) include five words on average and 80% of them are shorter than eight words (Crystal, 1969, p. 256). It has been observed that stretches of speech longer than that are usually broken up into two or more units (Hirst, 1998, p. 69). The only exception are WH-questions, which “appear to impose greater restrictions on the possible intonation breaks.” (Hirst, 1998, p. 69).

Formerly, researchers assumed that prosodic phrases are identical to syntactic phrases. The early studies in this realm, dating back to the 60s and 70s, showed that major acoustic phenomena, such as intonational boundaries, tend to occur at major syntactic boundaries (Shattuck-Hufnagel & Turk, 1996, p. 196). These studies supported the assumption of a direct link between prosody and syntax and their mutual dependence. It was backed by further evidence, as, for example, the fact that some syntactic ambiguities can be disambiguated by placing a boundary in speech, either on the phrase level, as in “old men and women”: (old men) (and women) vs. (old) (men and women) or on the level of sentences, as in “When you learn gradually / you worry more” vs “When you learn / gradually you worry more” (Shattuck-Hufnagel & Turk, 1996, p. 196). Furthermore, some utterances strictly require a particular prosodic structure: That is the case of the end of an initial subordinate clause, which requires a prosodic break in a place where a comma appears in notation, as in “After it rained,...” or in the case of an appositive structure: “Lance Armstrong, the cyclist,...” (Frazier, 2006, p. 245). There are other syntactic structures which also block the possibility to choose where to insert the prosodic boundary, as “George/ and Mary give blood (Shattuck-Hufnagel & Turk, 1996, p. 197). The sign “/” marks the place where a prosodic boundary cannot be located. This example is another indication that surface syntax imposes constraints on the prosodic organization (Shattuck-Hufnagel & Turk, 1996, p. 197).

However, later research carried out on larger corpora revealed that traditional morphosyntactic boundaries do not always coincide with the prosodic structure of an utterance (Shattuck-Hufnagel & Turk, 1996, p. 198) and quite on the contrary, there are usually many prosodic possibilities how to utter a sentence, some of which violate their syntactic structure (Shattuck-Hufnagel & Turk, 1996, p. 199). As a result, modern approaches to meaning now do not assume any direct causal link between intonation and grammar (Levis, 2013, p. 5).

The question arises of where the boundaries of prosodic phrases lie. There are no rigid criteria which could be used as guidelines for reliable detection of the boundaries of prosodic phrases. Some linguists suppose that prosodic phrases should be defined with reference to prosodic features (Himmelman, 2022, p. 718). The prosodic phrases are most commonly characterized as “the domain of a perceptually coherent intonation contour” (Shattuck-Hufnagel & Turk, 1996, p. 210). Some linguists also add a rhythmical aspect to the definition, specified as “an interruption of the rhythmic delivery by a pause, lengthening of the last segment at the end of a unit and/or increased speed rate at the beginning of a new unit” (Himmelman,

2022, p. 718). But in reality, it happens more often than not that these cues are ambiguous or not even present in the utterance (Cruttenden, 1997, p. 29). Determining the prosodic boundaries in phrases is hindered by phenomena often occurring in spontaneous speech, like hesitation, false starts, incomplete sentences etc. (Cruttenden, 1997, p. 29).

Despite everything that has been mentioned so far about boundary detection in phrases, the tendency for us to find a syntactic unit within a prosodic unit is still quite strong and the boundaries of syntactic and prosodic units typically align (Himmelman, 2022, p. 721). The listener must use his or her judgement about where most of the external criteria meet to “make the assignment of a boundary relatively certain” (Cruttenden, 1997, p. 29).

As a second step in the prosodic phrasing process (i.e., after the speaker divides the speech into prosodic phrases), he decides which piece of information in the phrase is the most important one; that word, called the nucleus, is then assigned as the carrier of the nuclear syllable, the most prominent stress in the whole phrase. The nuclear syllable is accented by rhythmic and pitch prominence (Wells, 2006, p. 7). With an 80% probability, this prominence is going to be placed on the final lexical word of the unit in English (Tench, 2020, p. 250). That is also interconnected with the information structure in the English syntax; English positions new information at the end of a sentence, therefore the intonation highlights the rheme of the utterance (Levis, 2013, p. 5). Therefore, if the speaker chooses to emphasize a different part of the structure, the nuclear syllable is going to move with the stress. The nuclear syllable is also where the so-called nuclear pitch movement starts. The nuclear pitch movements will be dealt with separately in the following chapter, n. 3.

The part of the prosodic phrase which follows the nucleus, if it is present at all, is usually known as the “tail”. This part may include stressed syllables, but no accented syllables (Wells, 2006, p. 8), meaning no other nuclear stress. If a phrase contains another stressed syllable before the nuclear tone, we call it the “onset”, adopting a term used by Crystal (1969). The part between the onset and the nucleus is the “head” and all the unstressed syllables preceding it are called the “prehead”, so the structure of the whole phrase could be described as follows: prehead – head – nucleus – tail (Levis and Wichmann, 2015, p. 140). A phrase must include a nucleus, but no other parts have to occur in the phrase, they are optional. Their presence depends on the length and the form of the whole prosodic phrase.

Another grammatical function of prosody is the possibility to change sentence types. There is a difference between saying “Mile End is in London” with a falling pitch,

conventionally marking a statement, and “Mile End is in London” with a rising pitch movement, which, on the contrary, often indicates a question (Mennen & de Leeuw, 2014, p. 185). Nevertheless, this categorization is a highly simplified statement. The prosody of questions has been proven to be highly variable, as no reliable connections between pitch rises and questions have been found so far (Ward, 2019, p. 63). This topic is further developed in chapter n. 3.

Last but not least, prosody may assist subordination, too: “Jerry, who I did not know then, came in the room.” Our pitch range in the subordinate clause gets compressed, the speech rate might be faster, and we may pronounce the clause more quietly. By such variation in the prosodic patterns, we let the listener know that it is a complementary piece of information.

2.1.3. Accentual function

In a prototypical English sentence, governed by the FSP (functional sentence perspective) principle, new pieces of information are positioned at the end: “Mile End is in London.” (Mennen & de Leeuw, 2014, p. 185). Such an arrangement means that “Mile End”, as the theme of the sentence, has either already been mentioned in the conversation before, or its meaning is clear to the discourse participants from the context of the given situation. “London”, the rheme of the sentence, is the new piece of information, it is positioned at the end and as such it will naturally carry the sentence prominence. Once the speakers need to highlight any other part of the utterance, they can either use syntactic strategies to achieve that (e.g., introducing a cleft sentence), or they can make use of the prosodic cues (the upper case letters symbolize the important, stressed part of the message in the following examples): “MILE End is in London” specifying the correct name of the neighbourhood as in “MILE End, not Upper End/ “Mile End IS in London.” When arguing whether Mile End is or is not in the capital city; or even “Mile End is IN London.” – it is not outside the city. The prominence can be added even to the already existing prominence on the word “London” in the original sentence guided regularly by the FSP principle: in such a case, the prominence would indicate contrasting London with a different city, such as “Mile End is in LONDON, not Birmingham” (Mennen & de Leeuw, 2014, p. 185). Using the prosodic properties this way, almost any part of a sentence can be emphasized without the need to change its syntactic structure.

Even though many findings have shown that grammar is not directly reflected by prosody, in other words, that the syntactic structure cannot be truly mapped onto the acoustic features (Cutler et al., 1997, p. 143), the prosodic functions help us express and form many syntactic structures. Prosody not only supports the speaker in conveying the meaning of what

he or she is encoding in words, but it is also crucial for the listener to be able to decode that meaning again. This important role of prosody has been discovered through many studies, one of the most ground-breaking ones being a finding by Epstein at the beginning of the 1960s (Epstein, 1961, cited in Cutler et al., 1997, p. 143), who discovered that a string of nonsense syllables was recalled better if it was accompanied by sentence morphology, as in: “meeving gups keebed gompily”, but the string also had to be spoken with sentence prosody, otherwise the advantage was lost (Cutler et al., 1997, p. 143). Another study by Speer, Crowder, and Thomas (1993) showed that if participants heard nonsense utterances spoken with the same prosody on the second presentation as on the first one, they recognized the structures more accurately the second time around than when spoken with different prosody (Speer et al., 1993, cited in Cutler et al., 1997, p. 144).

2.1.4. Discourse function

Prosody is also an indispensable tool in managing conversation, helping the participants with focusing on the sentence level (we might stress what is less predictable in a sentence, see section 2.1.3. above for examples), turn taking and structuring the dialogue. For example, the falling nuclear pitch movement, accompanied by other suprasegmental cues, such as the lengthening of the final syllable of the phrase, may send a signal to the listener that the speaker has finished talking and the listener may take his or her turn (Mennen & de Leeuw, 2014, p. 185). The receiver of the message is guided in his interpretation of the semantic contents by the prosodic structure, which is encoded in its phonetic form (Cole, Mo & Baek, 2010, p. 1142). At its very core, the transmission of a message can be described in quite a straightforward way: the prosodic structure of a spoken utterance is interpreted in a phonetic implementation shaping the articulation, which results in specific acoustic patterns; those patterns then encode the prosodic elements marking prominence and phrasing (Cole et al., 2010, p. 1142). These acoustic patterns are those “road signs” that have been already mentioned, and they help the listener interpret the related syntactic and semantic properties (Cole et al., 2010, p. 1142). It is important to note that this description is a laboratory-like depiction of how such a transmission works. In real-life communication, many more factors come into play. Some are purely linguistic (e.g., syntactic, or semantic factors), and others follow from the speakers’ affective state, intent, rate of speech, etc. (Cole et al., 2010, p. 1142). Factors from the latter category may divert the course of the transmission, change the meaning of what the speaker is saying or distort the message

completely. In such cases, a misunderstanding may arise even between two people speaking the same language.

2.1.5. Affective and indexical function

Prosody also has an affective function: it is a tool for conveying our emotions, attitudes, feelings, interpersonal stances, etc. Expressing anger might be accompanied by short duration and loudness, a shy person might speak quietly and rather slowly; we may express irony, sarcasm, or politeness; the possibilities are plentiful.

Last but not least, prosody can be a marker of “personal or social identity” (Wells, 2006, p. 12): people can be recognized as members of a group thanks to their characteristic use of certain prosodic features (e.g., typical nuclear pitch movements, phrasing, etc.). These groups might be divided on geographical grounds, based on age, gender and many more. We call this function of prosody the indexical function.

The two above-mentioned functions are not significant in connection to the subject of the thesis itself, therefore they will not be developed further. We will now move to one particular aspect of prosody which is the subject of the whole thesis, intonation.

3. Intonation

Defining suprasegmental features of speech might be difficult because some terms overlap to a great extent; how is intonation different from prosody? Crystal, for example, decided to separate those two phenomena based on form: in his words, intonation is “the distinctive use of patterns of pitch” (Crystal, 1997b, p. 202) and prosody are “variations in pitch, loudness, tempo and rhythm” (Crystal, 1997b, p. 313). Some authors use the terms interchangeably and others perceive intonation to be a subcomponent of prosody (Wakefield, 2020, p. 10). Similarly to prosody, definitions of intonation vary in literature based on what linguists regard as important and worth studying in their research. It has both a linguistic and a paralinguistic dimension in which it operates; the linguistic one concerns aspects like what information is new in the utterance or if it is a question or a statement, and the paralinguistic one relates to the speaker – their state of mind, degree of politeness, degree of dissociation from the listener etc. (Tench, 1996, p. 2). Some linguists, like Wakefield, work with the term “intonation” only in connection to its linguistic functions and discards its paralinguistic functions outside grammar, such as expressing emotions and attitudes, which he believes to lie solely in the field of non-linguistic prosody (Wakefield, 2020, p. 7-8). Others do not separate

the paralinguistic functions from intonation at all, for instance, Pike (1945), Gussenhoven (2004), Brazil (1997) or O'Connor and Arnold (1961), who believe that expressing the speaker's attitude is the main contribution of intonation (Chun, 2002, p. 17).

Intonation may be characterized broadly, as roughly corresponding to prosody itself, in other words, the suprasegmentals in language. That reflects the fact that we tend to perceive intonation simply as “the way someone says something” (Levis, 1999, p. 38). In such a case, intonation refers to the combination of several acoustic parameters, which include duration, intensity, and fundamental frequency F0 (Levis, 2013, p. 1). These parameters increase and expand, and they are perceived by listeners in terms of length (long-short), loudness (loud-soft) and pitch (high-low) respectively (Grice & Baumann, 2007, p. 27). These two domains, the acoustic and the perceptual one, are not linearly related to each other; for example, what listeners hear twice as high tends to be a much larger difference in acoustic terms (Cruttenden, 1997, p. 4). The terms “intonation” and “prosody” or “suprasegmentals” are then often used interchangeably. If we perceive intonation in this broad sense, it also covers aspects like melody or rhythm and is “produced by tonal height and depth along with stress, volume and varying lengths of pause” (Allen, 1971, p. 74).

In other cases, the scope of intonation is narrowed down to a mere description of pitch and its movement; generally, the rising versus falling pitch, and the accented vs non-accented syllables; that is, locating the nuclear syllable in the phrase/sentence (Levis, 1999, p. 39). For the purposes of this thesis, we will now treat intonation as a system of pitch variations following the narrower sense of the word “intonation”. Aspects like rhythm or loudness will not be described.

Pitch, as the perceptual parameter of F0, is customarily described using the terms “high” and “low”. Speakers differ in pitch height, which is usually connected to and influenced a lot by their physical structure (Roach, 2009, p. 119). The average F0 is 220 Hz for women and 180 Hz for men (Cruttenden, 1997, p. 3). Only some changes in pitch are relevant for intonation analysis. We, as speakers, can intentionally control our pitch to some extent; we may choose to speak with a lower or higher pitch than usual in certain contexts, which is linguistically significant (Roach, 2009, p. 119). There are situations in which our pitch may change and fluctuate as we speak, and it will usually happen due to some extra-linguistic circumstances (e.g., fast changes in pitch due to riding on a horse), but it is not regarded linguistically significant for analysis (Roach, 2009, p. 119). In many areas of research, for example in turn-

taking, researchers often put emphasis on the pitch while they do not study other aspects of the speech signal in such detail (Walker, 2017, p. 3). The boom in intonation studies came in the 70s when the first automatic pitch trackers appeared and became available, which made research in the realm of pitch much easier for phoneticians and linguists (Ward, 2019, p. 67). Unlike other suprasegmental features, like loudness or lengthening, the pitch is also not so sensitive to extra-linguistic factors during measuring (Ward, 2019, p. 67), which made this particular aspect more attractive for researchers. Loudness, lengthening and other such aspects are dependent on the speaker's thought processes, their distance from the microphone, noises in the background and other circumstances (Ward, 2019, p. 68), which might make the conclusions of the experiments potentially unreliable.

Analysing pitch gives us valuable information not only about the language we study but also about the individual speakers themselves and the differences between them. It helps the listeners categorize the speakers in terms of their social status, social role, or geographical location – it serves as a differentiator of accents (Mennen et al., 2007, p. 1769).

The span between the bottom and upper limit of a person's pitch is referred to as the pitch range. Some languages are flatter in terms of their pitch range and some, like English, expand their range much more, sometimes even on a very short portion of speech. A detailed description of the characteristic features of intonation in a particular language is a demanding task, which resulted in only a few linguists undertaking comparative studies of intonation systems between more than two languages until the 90s (Hirst, 1998, p. 2). In the following years, the situation changed for the better with many more studies occurring in this realm, but there are still surprisingly few systematic comparisons of the pitch range between speakers of different languages (Mennen et al., 2007, p. 1769). In a study from 2015, Volín, Poesová and Weingartová compared radio broadcasters, Czech and English native speakers, in terms of their pitch range. They discovered that the pitch range was 2 ST narrower in Czech than in English (Volín et al., 2015). A study with a comparable aim was done in 2011 by Busà and Urbani, only this time with American and Italian speakers, both reading English sentences. The results showed that the American speakers spoke with a wider pitch range than Italians and there was also a greater intonational variation in American English than in Italian (Busà & Urbani, 2011, p. 383). Another similar comparison was made between British English and German in 2007 when Mennen et al. discovered that British English speakers in the experiment had a wider pitch range than German speakers (Mennen et al., 2007, p. 1771). The lack of comparative

studies in terms of intonation and nuclear pitch movements also concerns British and American English, which is the subject of the empirical part of this thesis.

The width of a speaker's pitch range may also be influenced by other factors, such as their emotional state. Ward mentions that a speaker's pitch span becomes narrower in complaints or in the so-called "grudging admiration" (which is a term for conveying praise that is not easy for us to express) (Ward, 2019, p. 15-16). The emotional state of a speaker is, however, not the subject matter of this thesis, therefore this topic is not going to be developed further.

The differences in pitch realize the pitch movements. Their kinds and functions are going to be the subject of the following subchapter 3.1.

3.1. Intonation: the form and function of individual nuclear pitch movements

The functions of intonation greatly overlap with the functions of prosody mentioned in chapter 2. Since already mentioned functions are performed not by intonation only but by other suprasegmental features as well (such as rhythmical properties, loudness, tempo, etc), they were included in the chapter on prosody. In this chapter, the emphasis is put on the grammatical function of intonation. Specifically, it focuses on the description of the form of the individual types of the nuclear pitch movement in connection to their function in language.

The nuclear pitch movement starts on the nuclear syllable. The most commonly used terms for the shapes of the pitch movements in English are falling, rising, fall-rising, rise-falling and level (Levis & Wichmann, 2015, p. 140), which is a set originally proposed by Sweet (1890) and later adopted by other linguists, for example, Brazil (1997). There have been many suggested classifications of these movements or alterations of the proposed set since then; Armstrong and Ward (1926), for example, saw the fall-rise as a mere emphatic form of the rise, Brazil (1975), in opposition to Armstrong and Ward, asserted that it is the rising pattern which is an emphatic form of the fall-rise (Hirst, 1998, p. 69). Halliday (1967a) worked on further, finer differentiation of the movements based on the pitch height at which the movement begins: he introduced the high-rise and low-rise (Hirst, 1998, p. 69). Before him, O'Connor and Arnold (1961) already classified the falling movements in the same way: the high fall and the low fall. These terms are still used to this day. The level movement is often not even present in the basic set of pitch movements; it is up to individual authors to decide which are going to be included. Hirst does not suppose there is a sufficiently convincing way to find reliable proof for any of these analyses (Hirst, 1998, p. 69).

We will now return to the phrasal structure described earlier (section 2.1.2.), but this time focusing on the nuclear pitch movement. In figure 3 below, Wells’ visual representation of the phrase “We’re planning to fly to Italy” is depicted (Wells, 2006, p. 9). The large dots symbolize stressed syllables, and the smaller ones represent unstressed syllables. At the end of the phrase, in the “tail”, the lower position of the dots indicates the falling nuclear pitch movement:

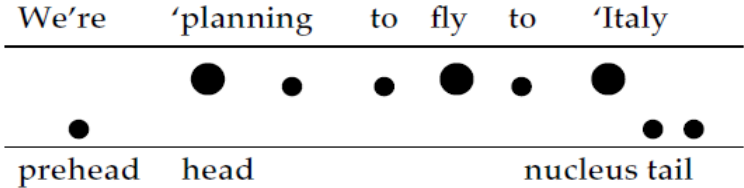


Figure 3: Visual representation of a phrasal structure (Wells, 2006, p.9)

It is generally accepted that, under no special circumstances, the fall is associated with complete statements, exclamations, commands, and wh-questions (Wells, 2006, p. 10).

The rising pitch movement, on the contrary, has usually been thought to be connected to yes-no questions. In terms of intonation, the wh-questions are then much more similar to statements than to yes-no questions (Hirst, 1998, p. 26). Nevertheless, corpus studies in the 90s showed that yes-no questions are more commonly accompanied by falling pitch movements than by rising movements (Bartles, 1999, p. 7). It has been discovered that both the falling and the rising movements are possible to use with yes-no questions, but each of the movements mediates a different message, which is demonstrated by two examples in figure 4 below, taken from Bartles (1999). Sentence a), pronounced with a rising pitch, is a genuine question, while question b) is rather a reaction to something that has already happened:

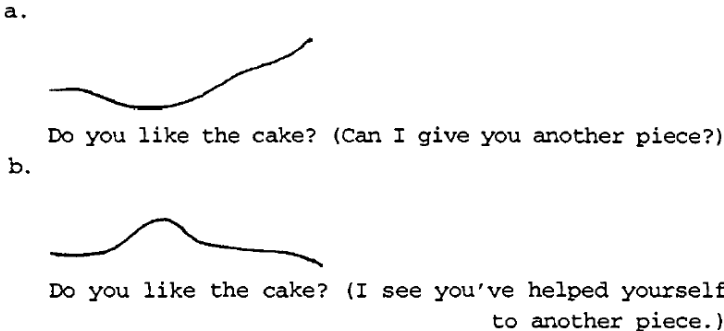


Figure 4: wh-question, accompanied by different pitch movements (Bartles, 1999, p. 7)

Sometimes, if there are no strong contextual cues which would block the effect, a final rise in phrases or sentences can be sufficient to change a declarative sentence to a question without changes in syntax (Bartles, 1999, p. 6), as demonstrated in figure 5 below:

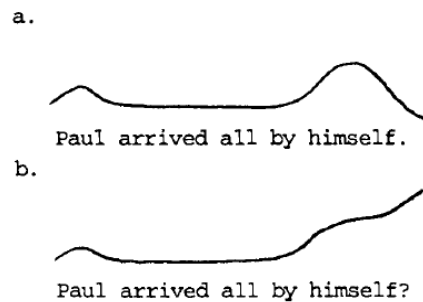


Figure 5: a change in sentence type caused by a different pitch movement (Bartles, 1999, p. 6)

Furthermore, rising pitch movements are customarily connected to uncertainty and a speaker's unfinished, incomplete utterance; rising intonation leaves the impression that something more will follow (Tench, 2020, p. 251). We may compare the two identically worded utterances in figure 6 below, each of which is accompanied by a different pitch movement:

(13) | this is my first visit to \Cardiff |

(13a) | this is my first visit to /Cardiff |

Figure 6: The same utterance pronounced with two different pitch movements (Tench, 2020, p. 252)

While the first one (13) indicates definiteness and the end of the speaker's turn, the second sentence (13a) will probably be followed by some more details.

Moreover, Tench adds that falls are usually associated with major information, while rise often indicates some kind of incomplete information (Tench, 2020, p. 252). We can observe what effect a chosen pitch movement in a reply has on the flow of a conversation, too. If a person asks: "Do you know John Smith?", the positive reply "yes" may occur with both a falling and a rising tone (Roach, 2009, p. 124). The falling movement gives the feeling of finality or the end of the conversation, while the rise invites the other person to continue (Roach, 2009, p. 124). The pitch movement the speaker chooses a) is influenced by the extralinguistic context and b) affects the other participants of the discourse and their decisions in the conversation.

In the 70s, rising intonation slowly started to be used with declarative clauses in sentence-final positions, too. It has been described as “question-like intonation with declarative utterances” (Warren, 2016, p. 23). The speaker usually uses it with the intention to say, “I am checking that you are following me” (Cruttenden, 1997, p. 129). This trend, usually addressed as “uptalk” or “high-rising terminal” (HRT), originated in Australia and has spread to other English-speaking countries quite quickly. It is usually typical of particular groups only; teenagers, working-class speakers, and female speakers (Cruttenden, 1997, p. 130). HRT may also point towards uncertainty on the speaker’s part. The use of uptalk is rather stigmatized and typically associated with being “wealthy, white, young and female” (Tyler, 2015, p. 289).

There are two additional movements, fall-rise and rise-fall. Their structure is more complex in terms of the intonation contour progression since there are two movements inside of each of them; the tail of rise-fall begins with a significant upward movement before the falling, while fall-rise falls first after the nuclear syllable and then starts to rise (Wells, 2006, p. 217). Fall-rise may indicate that “something is known, but there is some doubt or reservation”, or it can be associated with implication or contrast (Herment & Tortel, 2021, p. 8). Trench adds that in colloquial speech, when fall-rise is used to imply disagreement with room for further debate, it is often preceded by the word “Well”, which helps to “reduce the sense of total disagreement” (Trench, 1996, p. 79).

Rise-fall is described as a movement which raises doubt “in order to dismiss it” (Chun, 2002, p. 19), it can also express irony or surprise (Herment & Tortel, 2021, p. 8). Not all linguists consider them to be individual types. Nevertheless, since their complex contour starts on the nuclear syllable as well as the simple rising and falling movements, they are considered equal to the rest of the set in this thesis.

The last movement, level, is produced by the vocal folds vibrating at a constant rate (Wells, 2006, p. 3), which means that it does not involve any significant changes in pitch. What a significant change is, however, tends to be defined differently by various linguists. A study by Hancock et al. (2014), for example, defines level intonation as a change in pitch smaller than 2 ST (Hancock et al., 2014, p. 204). Halliday and Greaves mention that it is difficult to find movement that is actually level in pitch and that there is “almost always a final rise”, even though it is sometimes barely perceptible (Halliday & Greaves, 2008, p. 114). Level is used when the speaker is suggesting rather than asserting, when the sentence or thought is incomplete, or when the speaker wishes to “hold the floor” (Halliday & Greaves, 2008, p. 114).

All of these situations are likely to occur regularly in spontaneous speech or conversation. The level movement is therefore also included in the analysis in the empirical part of this thesis because it was expected to occur often with regard to the nature of the material (spontaneous, unprepared speech).

Linking the pitch movements to interpersonal stances they might express is not as difficult as trying to categorize them in terms of the sentence types they are connected to. Decades of research have shown that a strict categorization of nuclear pitch movements in this respect cannot be reached. The system is too complicated to be summarized dogmatically. As Bolinger comments, “There are both questions without rises and rises without questions.” (Bolinger, 1978, p. 474). Bartles raises the fundamental question if one can even associate a pitch movement “at some level of abstraction with the same interpretational feature across all occurrences, independent of lexical context and situational context” (Bartles, 1999, p. 4). That is a very fair point, given that the choice of pitch movement is actually influenced by several other factors, for example, the context of the whole utterance. Pike rejected the idea of direct connections between nuclear pitch movements and their grammar function already in the 40s, saying that such an error is “the easiest to commit” in intonation analysis (Pike, 1945, p. 23). According to his words, it is not possible to select a grammatical construction and claim that a certain contour indicates this construction at all times (Pike, 1945, p. 23). Grabe and Post (2002) also mention in their study that “the mapping between grammatical structures and intonational form is dialect-specific” and that the “change in grammatical function can be associated with the production of a different pattern in one dialect but not in another” (Grabe & Post, 2002, p. 4).

Nevertheless, Levis mentions that the somewhat simplified categorization forms useful bases for certain areas of research, e.g., large corpus studies focused on improving speech technology (Reed & Levis, 2015, p. 177). Human beings do not usually experience trouble assessing the intention of the other speaker’s utterance – if it is a request, question, statement etc. – but machines are not so well equipped to do that (Reed & Levis, 2015, p. 177), which reveals how complex language actually is. Reed and Levis mention a study by Shriberg et al. (1998) in which the authors tried to distinguish backchannels, e.g., “uhuh”, from agreements, like “yeah”, using automatic analysis (Reed & Levis, 2015, p. 177). The problematic part was that both of these short utterances could perform both functions under certain circumstances, which posed a challenge for automatic recognition (Reed & Levis, 2015, p. 177).

The multi-functionality of some language structures uncovers the many layers of a language, which has a direct influence on how variedly intonation is used. In English, even some short utterances are a great example of this matter, e.g., a simple “thank you” routine. Aijmer claims that there are various intonational patterns used while thanking, depending on the size of the favour or the way gratitude is expressed (Aijmer, 1996, p. 41). The falling movement is usually reserved for expressing real gratitude for greater favours, unlike the rising pitch movement, more often associated with casualness and routine (Aijmer, 1996, p. 41). The phrase can, however, be also a demonstration of dismissiveness (“I can do it myself thank you”), irony (“thank you, that’s all I needed”), and many other attitudes and emotions (Reed & Levis, 2015, p. 177).

Defining and understanding intonational patterns in general terms has proven to be, to say the least, complicated. Looking individually at different English accents makes the matter even more complicated. Although researchers have discovered that speakers of different languages vary in their use of pitch, the studies which would analyse differences in intonation between individual varieties of English are, still to this day, relatively scarce (Fuchs, 2018, p. 1). There is a need to continue this line of research to not only shed light on the linguistic differences at the suprasegmental level but also to help speakers understand other varieties or languages so that they can successfully engage in a conversation. We will now turn to the description of intonation in the British and American English varieties, which were used for the study in the empirical part of this thesis.

3.2. Intonation in British English

There is considerable regional variation across the British dialects, as far as intonation is concerned (Grice et al, 2019, p. 289). Hence it is important for research to clearly define the area of interest. Received Pronunciation (RP), the former standard used as the target dialect of British English in research, has lately been replaced by the Southern British English dialect (SBE). SBE shares the qualities of previously used RP, and therefore it seems “safe to assume” that their intonation systems are going to overlap (Hirst, 1998, p. 56). SBE is therefore the standard this paper adheres to both in this chapter and the later, empirical part.

In terms of the differences in segmental phonetics, Southern British English (SBE) has been extensively compared to other British dialects, but the variation in intonation has received much less attention (Grabe & Post, 2002, p. 1). Falling movements are usually associated with finality and completeness in SBE, which indicates that “neutral complete statements are usually

pronounced with a fall” (Herment & Tortel, 2021, p. 7) unlike other British dialects; Belfast speakers, for example, tend to produce both declarative sentences and questions with rising intonation, according to the study by Grabe and Post (2002). Their next experiment showed that SBE speakers used two intonational movements for read statements; fall and fall-rise, with fall being present in a vast majority of the cases (94%) (Grabe et al., 2005, p. 3). In questions, speakers relied mainly on fall (in 61 % of the wh-questions and 44 % of polar questions), fall-rise and high rise (Grabe et al., 2005, p. 3).

A very recent study by Hudson et al. (2019) compared British English and Hong Kong English speakers and among other phenomena, they evaluated their pitch movements at the end of sentences. The numbers for individual types of movements in British English are very similar to the ones obtained in Grabe’s research task, with statements and wh-questions being mostly accompanied by falling intonation, (93 % of all falls in both categories) and yes-no questions by fall-rising intonation (in 64 % of all the cases) (Hudson et al., 2019, p. 322).

In a very recent study by Herment and Tortel, native English female speakers participated in a reading task aiming to discover which intonation contours prevail in read SBE speech. Similarly to the previously mentioned studies, the results showed that the most commonly used movement for intonation units was the fall (71%), regardless of its position; i.e., if it was at the end of a final or a non-final unit (Herment & Tortel, 2021, p. 12). The even more surprising result is that fall was present in 68% of the non-final intonation units, contrary to the belief that non-finality is usually expressed by rise, potentially by level (Herment & Tortel, 2021, p. 10).

3.3. Intonation in American English

One of the intonational movements that seems to be typical for American English is the low-rise. Low-rise is a type of movement which does not include pitch rise at the stressed syllable, on the contrary, the syllable has a low pitch accent and then it is followed by a moderate rise (Levis, 2002, p. 56). This type of rising differs from the high-rise in that the high-rise already starts high and climbs even higher. To imagine what the low-rise looks like, a figure is attached below. The L* marks the beginning of the nuclear pitch movement.

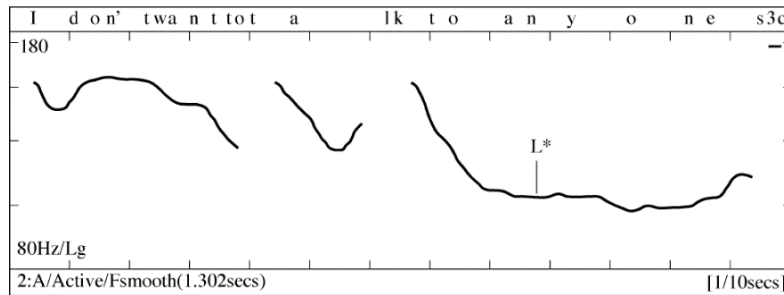


Figure 7: Low-rising intonation, with low pitch accent for
 “I don’t want to talk to anyone” (Levis, 2002, p. 57)

The claims about the frequency of this pattern in British and American English are rather confusing. Levis mentions that both American and British linguists have claimed the low-rise to be more typical in their variety than in the other (Levis 2002, p. 56). In American English, it has been compared, in terms of meaning, to the falling contour by Gunter (1974) or to fall-rise by Bing (1985) (Levis, 2002, p. 57). According to Pierrehumbert and Hirschberg, low-rise has been perceived as the standard contour for the yes-no questions in the USA (Levis, 2002, p. 57). There have even been conflicts over whether the low rise serves as a question intonation movement at all (Levis, 2002, p. 57). Cruttenden, for example, wrote in his book that the use of low-rise in questions may sound patronizing to Americans and that they are more likely to accompany their questions with high rise (Cruttenden, 1997, p. 88). There are no other reported differences between the use of individual pitch movements or their frequency of use between American and British English.

To specify the nature of American English, Bolinger states that there are many intonational idioms, which we can imagine as “stereotyped connections between intonation and particular locutions” (Bolinger in Hirst, 1998, p. 46). These connections are illustrated in the below-attached figures 8 and 9:

Big
de a!

Figure 8: intonational idiom in AmE I
 (Bolinger in Hirst, 1998, p. 46)

I’ll
show him a thing or t^wo!

Figure 9: intonational idiom in AmE II
 (Bolinger in Hirst, 1998, p. 46)

The first idiom, according to Bolinger’s explanation, downplays the importance of something, while the second idiom should express some kind of “vengeful threat” (Bolinger in Hirst, 1998, p. 46). Even though Bolinger specifies these idioms for American English, they probably do

not demonstrate a unique American English feature in speech and British English could show a similar tendency.

Southern British English and American English share, with minor differences, the same intonation system (Bolinger, 1998, p. 45). These differences do not lie in the configurations, but in frequency or pragmatic choice (Bolinger, 1998, p. 45). Bolinger mentions that what makes the American variety unique is its iconic character because gestures and facial expressions are especially important for American speakers (Bolinger, 1998, p. 45). Southern British English and American English share the same set of nuclear pitch movements that can be identified at the end of phrases, but the comparison in the frequency of their use is not supported by much data.

Even papers from the last decade still mention that the number of comparative studies of English dialects and regional varieties is unsatisfactory and insufficient (see e.g., Clopper & Smiljanic, 2011). The reason for the scarcity of comparative studies might lie in the different notational systems of intonation in both varieties. The next chapter is going to briefly introduce and compare these two systems.

4. The British and American traditions and terminology

The earliest attempts to represent the intonational movements on paper occurred already in the eighteenth century when the system for musical notation found its use in the description of intonation as well (Levis & Wichmann, 2015, p. 139). Although it was convenient in the beginning stages of linguistic description, it was later found unsuitable because the fixed position of the staves rarely corresponded to the values of intonation (Levis & Wichmann, 2015, p. 139).

In the past, researchers in applied linguistics could not agree on whether the components of intonation should be represented statically or dynamically; in other words, if the contours are made up of pitch levels or should rather be regarded as holistic configurations (Levis, 2013, p. 2). Two distinct approaches gradually developed, which have been influential the most until the present day: the British and the American. These schools share some features and differ in others. They both work with the melody of an entire phrase in the analysis, and they also work with compositional building blocks of intonation – pitch levels in the American tradition, and heads, pre-heads and contours in the British one (Levis, 2005, p. 341). In the previous chapters, various figures were included describing the pitch movements in either British or American

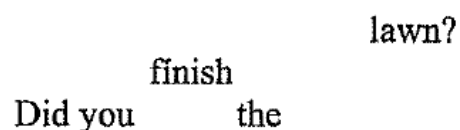
English. Almost each of these movements was depicted differently, using various strategies and different terminology which stem from these two different traditions. At first sight, it might seem that the differences in intonation transcription are not the problematic part of the intonation analysis; however, any transcription we encounter is a direct reflection of what seemed important to the author in the analysis and what did not (Levis, 2005, p. 348), which will transpire below when discussing the specific differences between the two traditions of notation.

There are moving type formats, dots and lines, tiers and many more ways how linguists capture speech prosody on paper. Very broadly put, the British tradition usually favours contours and dots, while the American school rather works with the level system which mirrors the pitch levels. The basic difference might be observed in figures 10 and 11. The examples are borrowed from Grabe et al. (2005) in fig. 10 and Levis (2005) in fig. 11.



It's YOUR fault!

Figure 11: AmE notation of intonation
(Grabe et al., 2005, p. 3)



Did you **finish** the **lawn?**

Figure 10: AmE notation of intonation
(Levis, 2005, p. 248)

An American linguist, Kenneth Pike, was immensely influential in the development of the notation system on the American continent. Already in 1945, he suggested a four-level system to describe intonation in American English. It has become a basis for many of his successors who have been building on his theory up until today, even though it has not been actively used in applied linguistics anymore, having been replaced by Pierrehumbert's two-level approach in the 80s (Levis, 2013, p. 3). Nevertheless, it is still sometimes used in American materials for teaching English until the present day (Levis, 2013, p. 3). Pierrehumbert's work was groundbreaking in the realm of the notation systems of intonation. She laid the foundation of the "most widespread phonological framework for representing intonation", known now as, according to Ladd's terminology, the "autosegmental-metrical" framework (Grice and Baumann, 2007, p. 43). One of such frameworks widely used today is called the ToBI system ("Tones and Break Indices"), originally developed for American English intonation. It has been gradually becoming a general framework for other intonation systems, not only in English but in other languages, too, e.g., Standard German, marked as (G)ToBI (Grice & Baumann, 2007, p. 43).

We have already mentioned in chapter 2 that the ToBI system is being adapted to the needs of the particular language where it is used.

Briefly explained, in ToBI, linguists work with only two kinds of tones: low and high tones (abbreviated as L and H). The system centres around pitch accents (marked with *, e.g., H*), phrase accents (marked with -, e.g., H-), and boundary tones (marked with %, e.g., H%). The system also recognizes the so-called “intermediate phrases”, which we can imagine as a finer phrasing distinction. Intermediate phrases are shorter than prosodic phrases and they exist within the prosodic phrases, separated by a subtle but audible break, very often where a comma would occur in notation. Intermediate phrases do not occur obligatorily, usually, they can be found in a little more complicated phrases. Figure 12 depicts the ToBI system used on an example. The sentence was taken from Beckman and Elam (1997).

Will you have marmalade, or jam?
L* H- L* H-H%

Figure 12: AmE notation of intonation, the ToBI system (Beckman & Elam, 1997, p. 9)

There are two pitch accents (L*), two intermediate phrase accents (H-) and a boundary tone (H%) in the phrase (Beckman & Elam, 1997, p. 9). The end of an intermediate phrase has to be involved in the notation, even when it follows naturally from the marking of the end of the whole prosodic phrase; that is why there is H-H% at the end of the phrase above.

The Break indices rank from 0 to 4 and they indicate the strength of the boundary between individual words. The system was developed to serve a wide scale of users and not to be limited to a small number of experts (Beckman et al., 2005, p. 13). That is probably the reason why linguists started to use it as a basis for intonation transcription in other languages as well; it provides the researchers with a unified and solid system of description, which did not exist up to that point in time. Grice and Baumann mention that the advantage of the American approach over the British one lies in the fact that “the tonal information can be precisely localised on single syllables and/or at the edges of phrases”, while the British school only makes the connection between text and tones occur on the nucleus (Grice & Baumann, 2007, p. 44). The placement of the nuclear syllable is not indicated in ToBI, since “the word with nuclear stress is defined positionally; it is the last accented word, or the accented word (if there is only one in the phrase)” (Beckman & Elam, 1997, p. 11).

Unlike the American school, the British tradition treats intonation as a unified whole (Levis, 2013, p. 4), in terms of holistic contours, as shown in figure 10 above; that particular system (dots and lines) is sometimes called the “tadpole notation” (Grice & Baumann, 2007, p. 43). The contour depicts what levels represent in American tradition. The British approach, sometimes addressed as configurational, centres around nuclear stress and pitch movement. Apart from the visual representation that could be seen in many of the figures throughout this thesis, the movements are also often depicted iconically inside of the text in the British tradition, using - for level, \ for fall, / for rise, and their combinations for combined intonational variants: ∨ for fall-rise and ∧ for rise-fall (Levis & Wichmann, 2015, p. 140). These symbols are then inserted before the syllable, which marks the beginning of the nuclear pitch movement as in “I’d like to \thank you” (Levis & Wichmann, 2015, p. 140).

The development of the British tradition dates back to the work of the German scholar Hermann Klinghardt in the 20s (Fox, 2000, p. 278). He represented pitch movements as “a string of dots of different heights, one for each syllable” (Fox, 2000, p. 278). His approach was further developed and elaborated by other linguists, like O’Connor and Arnold in the 60s, whose work reached a scope comparable to Pike’s in American English (Levis, 2013, p. 4).

The two notation systems are applied to the phrase below to illustrate the differences between the systems and also how they capture differences in pitch movements. The ToBI notation is always first, and the British configurational approach follows after.

1) Jenna came home with Toby.

H* !H* H-L%

Phrase number 1 is a simple prosodic phrase. We start with a high tone, and we end with a low tone. It follows the FSP principle, so the stress lies on the last word, as if in an answer to the question “*Who* did Jenna come with?” When there are more H* accents, the F0 of the phrase tends to fall as the phrase progresses. The exclamation mark signals that the tone is still high, but it is lower in comparison to the previous one.

1) `Jenna came `home with \Toby.

The sentence is transcribed using the British approach to intonation notation. “Jenna” and “home” are marked as stressed, and the intonation falls on “Toby”, which is indicated by the backslash before the word. In the British approach, we do not mark “home” as lower than “Jenna”, so the ToBI transcription is a little more detailed in this respect.

2) Jenna came home, together with Toby.

H* H*L- H* L-L%

Phrase number 2 is a little bit more complex because it includes an intermediate phrase boundary. It starts high on the word “Jenna” as well as phrase 1, then it falls on the word “home”. An intermediate phrase boundary follows (L-). The last word, “Toby”, carries stress again and it starts high, then it falls down to a low tone. Both the end of the intermediate phrase and the end of the whole prosodic phrase are marked on the last word in the notation (H-L%).

2) `Jenna came \home, | together with \Tobby ||

If we want to mark the phrase boundaries in the British notation system, a vertical line | can be used to indicate the intermediate phrase boundaries, double vertical line || is then an indicator of the boundary of the whole phrase, as we can see in the sentence above.

The intonation progression in the British tradition is sometimes also depicted as a horizontal line, or “pitch tracker”, above (or below) the phrase, which offers a better and more precise idea about how the intonation contour develops. It might serve as a quicker and more effective way to teach learners, unlike the little bit more complicated ToBI system, which needs to be studied in depth to be understood. The disadvantage of such a simple depiction lies in the fact that the line might not indicate so clearly on which syllable exactly the potential changes in pitch happen.

The different notation conventions are not the only troublesome part of the potential comparative analysis; many of the problematic differences between the two traditions lie in the terminology. A lot of the terms are usually used only in one of the schools, and if they are used in both, they do not always mean the same (Levis, 2005, p. 344). Therefore, researchers may face a challenge when trying to compare the two systems, the potential comparative research suddenly becomes much more time-consuming.

The pitch movement is regularly called the “tone” in the British tradition, while American linguists will rather refer to “intonation” (Levis, 2005, p. 346). Both of these words are misleading in some respects: “Intonation” might be a confusing term to use because of the broader meaning of the word, which does not refer to the nuclear pitch movement only (see discussion of meaning in chapter 3). However, “tone” might be problematic as well because of its use in tonal languages, where “tone” has a meaning-altering function and it is “a feature of the lexicon” (Cruttenden, 1997, p. 8). Therefore, it differs immensely from what it should

represent in terminology in non-tonal languages. The term “tonic”, used for the syllable where the nuclear pitch movement starts, is misleading for the same reasons, which is why it seems advisable to avoid its use as well.

Some differences between the terms stem from a different perception of their components, as is the case with nuclear syllables. Levis mentions that while the British tradition considers the nuclear syllable (also called tonic, nucleus or nuclear tone) a part of the intonation system, the American linguists rather classify it as a part of the stress system and use such terms as “sentence stress” or “phrase stress” instead (Levis, 2005, p. 349), unlike their colleagues from Great Britain. This discrepancy reflects the fact that for many, stress and intonation cannot be separated from each other entirely during the analysis; for example, pitch accents are usually accompanied by syllable lengthening, which lies in the domain of stress, or rhythmical patterns (Levis, 2005, p. 349). Stress itself is also referred to confusingly in literature; some authors tend to refer to the stressed syllables as accented syllables (see Wells 2006), while others use the word accent only in connection to the main stress of the phrase. Levis (2005) offers a good, brief overview of the differences between the terms used on the suprasegmental level of language in both the American and the British tradition (Levis, 2005, p. 347). Nevertheless, it neither does nor can cover all the differences found in the literature on intonation, which condemns every reader to study the terms, their scopes and meaning carefully to avoid confusion and wrong deduction.

Another troublesome part comes with the different scope of the terms; Levis mentions that “some terms refer to the acoustic patterns of pitch change (e.g., nucleus), while other terms used for the same phenomena refer to the semantic effects (e.g., focus or highlighting).” (Levis, 2005, p. 346). Fox warns about this problematic part of intonation analysis when he mentions that “the different analyses are not entirely equivalent, since each framework provides a different range of possibilities and allows different generalizations to be made” (Fox, 2000, p. 297).

It is not only the question of what tradition a particular study adheres to; even individual authors differ in their attitude towards the categorization of terms and phenomena in the study of intonation. What one author considers to be one category might be treated separately by others, terms used in one study often cannot be found in the next one. The will to compare language dialects or varieties across different traditions of notation might be thwarted by the various attitudes towards terminology we might encounter and have to unravel. Navigating the

terminological chaos is time-consuming and might potentially discourage or slow down further comparative research unnecessarily. Therefore, the meaning of the terms always has to be carefully examined and “translated” to avoid making wrong assumptions.

To avoid confusion, some theoretical work on intonation is prefaced with a specification of the scope of the applied terminology by stating what the individual terms are comparable to in other studies or books. Pierrehumbert, for example, states at the very beginning of her chapter on intonation what her term “intonation phrase” corresponds to in other works on a similar topic (Pierrehumbert, 1980, p. 64). She mentions four different terms which are identical in their scope to the term “intonation phrase”, one of which only “*appears* to be the same” (Pierrehumbert, 1980, p. 64). Such a gesture is helpful for other researchers, however, the explanation of the scope of terminology and its comparison to similar terms elsewhere is not present in all studies and books.

For the reasons stated above, I have decided to use those terms which seem to cause the least confusion, and which are not ambiguous. I work with chunks of language which I call “prosodic phrases”. I intend to avoid using the term “tonic” because of its distinct meaning in tone languages and instead, I prefer the term “nuclear syllable” for the syllable where the movement starts. I describe those final movements of pitch, which start at the nuclear syllable, as “the nuclear pitch movements” because it is very specific, it matches the “nuclear syllable” in the name and hence it leaves little room for confusion. For the sake of transparency, I also use the term “nuclear stress” as a term for the main stress in the phrase. As for the concrete nuclear pitch movements, the five following are going to be considered, corresponding to Brazil’s classification: fall, rise, fall-rise, rise-fall, and level (Brazil, 1997, p. 10).

4.1. Lacking English terminology in intonation analysis

The proposed set of nuclear pitch movements is mostly sufficient for what we need to describe in the intonation analysis of English. However, there are finer details in the process which do not seem to be covered by the current terminology.

To clarify this issue, it is important to start by pointing out certain differences in the general progression of a pitch movement. The majority of the cases that were shown in this thesis so far are “gliding”, in other words, we can hear the smooth progression from one level of the pitch to another since it usually happens between syllables of a single word. However, the transition might also be realized as a step between two levels of pitch, without the audible

glide. Cruttenden (1997) illustrates this difference in pitch movement progression on two examples of short words, which I included in figure 13 below:

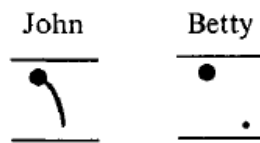
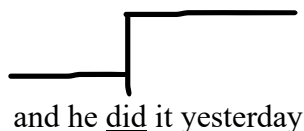


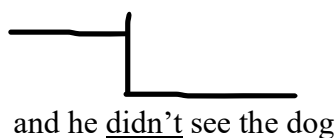
Figure 13: the difference in pitch movement progression (Cruttenden, 1997, p. 4)

While “John” is pronounced with the gliding type of pitch movement, illustrated by the black line, “Betty” is pronounced in a step-like intonation, represented by two black dots. The author even mentions that “some people actually consider the ‘essence’ of the pattern to be a sequence of high and low tones, rather than considering it a fall” (Cruttenden, 1997, p. 5), which is common in the American way of perception (in contrast to the British holistic approach to intonation analysis). The type of progression is also influenced by other factors, like the number of syllables.

The problematic area lies in the way we describe the progression of an intonation contour in longer chunks of language than individual words, specifically, in phrases. The following illustrations of two phrases will assist in explaining this issue further. The underlined word “did” and “didn’t” carry the nuclear stress of the whole phrase.

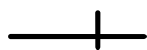


If the pitch movement in a phrase rises on the nuclear syllable in steps from lower pitch to higher pitch but continues in a level movement (represented by the horizontal line after the vertical line on “did”), most of the literature on intonation would consider the nuclear pitch movement to be level, since the movement starting at the nuclear syllable is level. An analogous example follows, this time with falling intonation.



I believe that even though the last portion of the phrases, starting at the nuclear syllable, certainly is level in both cases, the overall “function” of the pitch movement, including the

onset and the head, is rising (in the first phrase) and falling (in the second phrase), even when we cannot hear the “glide”. We can compare the two phrases with the following one, which demonstrates a plain level:



I like rice

Since I suppose that the stepping progression of rise-to-level/ fall-to-level is different from a usual “gliding rise/gliding fall” and it is also different from plain level in the third phrase, it would be convenient if the terminology provided the opportunity to reflect this difference. We are looking for a term for the overall rising and falling scheme in the first two examples of phrases, the “function”. Unfortunately, any such distinction does not seem to be covered by the terminology in the literature on intonation. To shed some more light on the lacking terms in English, we will briefly compare the English system with the Czech intonation terminology.

Czech uses two terms: “melodém”, which is described in Czech literature as a “funkční melodické schéma” (Palková, 1994, p. 306) and which we might imagine as a higher-level term, describing the more general progression of an intonation contour (the “function”). The second term is “kadence”, which is a term for a concrete realization of melodém (the concrete “form” or “realization”). This distinction is more or less analogous to the distinction between a phoneme and (allo)phone on the lower level in phonetics. So, for example, the rise-falling “kadence” may be a realization of a conclusive falling “melodém”. This differentiation between concrete form and more general function captures the finer differences in the realization of pitch movements in Czech; unfortunately, I have not encountered any such distinction in the English terminology, or at least not a widely accepted and used one.

It is possible to use the word “realization” for the concrete type of movement (which roughly corresponds to “kadence”), but we lack an analogous term for “melodém”, the higher-level scheme with a specified function in the language. It is not possible to use the word “function” because of its established different meaning in this area; it would only lead to more confusion. Because my thesis supervisor and I believe that such distinction might be helpful in the process of intonation analysis and it might provide us with more detailed results about the types of movements that are used in English, we propose to coin a new term, which would help differentiate between the two scopes, similarly to Czech intonation terminology: we are proposing the term “functional melodic unit” (abbreviated as FMU), corresponding to the scope of the Czech term “melodém”. Just as we differentiate between fall and rise-fall and we

acknowledge them to be two distinct movements in this thesis, we believe we should make finer distinctions in the terminology, too, to capture the details and differences between form and function. Two additional movements are therefore added to the set of the previous five movements: level-rise (illustrated by the phrase “and he did it yesterday”) and level-fall (illustrated by the phrase “and he didn’t see the dog”).

The term FMU is going to be used as the greater scheme term (the “function”, i.e., level, rise or fall) in relation to the newly observed, lower-scheme terms, the concrete realizations of FMU (“the form”, level-fall and level-rise). Even though these are treated as subtypes, they will be depicted separately in the graphs and in the analysis to illustrate properly how big a share they form in level FMU and rise/fall FMU and thus if they are worth any attention as separate subtypes.

5. Research questions and hypothesis

The first aim of the current thesis is to analyse and compare English phrases, obtained from 8 British and 8 American speakers, in terms of their length (in syllables and words) and articulation rate (in syllables per second).

Secondly, the study aims to discover the most commonly used nuclear pitch movements in each variety. We aim to assess if there are any differences between the two English varieties in terms of the frequency of the five FMUs (level, fall, rise, fall-rise and rise-fall). A null hypothesis is employed, according to which no significant differences are expected to be found between the British and American varieties in the above-mentioned respects. We are also going to be observing the frequency of the newly established “stepping” realizations, level-rise and level-fall, and what proportion they form within the level, fall (in case of level-fall) and rise (in case of level-rise) FMUs.

Thirdly, the research is going to determine whether there are any differences in the width of the British and American speakers’ pitch range of a) the whole phrase and b) the final portion of the phrase (the nuclear pitch movement).

Apart from the comparison of the two English varieties, the analysis is also going to show and discuss the features of the individual speakers’ performances and potential within-group differences between the eight speakers.

6. Material and method

6.1. Material

The material used for the analysis consists of prosodic phrases, which were extracted from the political speeches of 8 British speakers and 8 American speakers (4 men and 4 women in each variety). We obtained around 3-4 minutes of speech for each participant. From this material, 50 phrases per speaker were extracted. It amounts to 800 phrases in total: 400 phrases for British English and 400 phrases for American English. The material was downloaded from online websites containing freely accessible political speeches. Political speeches were chosen as a source for the analysis for several reasons.

Firstly, political debates are an easily accessible source which contains a large amount of material suitable for comparison. To our knowledge, there is no other similar easily accessible source which offers this quantity of authentic spoken language for analysis, and which is, at the same time, of sufficiently high quality (e.g., no background music, noises, no

extensive overlap of speakers, etc). Secondly, such a source can satisfactorily meet most of the requirements that we need to take into consideration when trying to make any kind of tentative conclusions about language prosody: the discourse is authentic, spontaneous and the undesirable influences in the form of stress or nervousness are less likely to interfere since all the speakers are public figures who are expected to be used to talking publicly in front of a camera and a microphone. Last but not least, it would be extremely difficult in our environment to obtain a similar amount and type of data by means of interviews.

The British political debates come from the website called Westminster Hour accessible at <https://www.bbc.co.uk/programmes>, a political discussion program run by the BBC station. The American debates all come from a website called C-span, which belongs to the American C-span cable and satellite television network, accessible at <https://c-span.org>. All of the participants were engaged in a debate with other speakers, meaning that they had to react to each other and therefore their speech could not be prepared in advance, which we believe could, to a certain extent, influence the choice of intonational patterns. The first minute of each of the participant's speeches, in which they introduced themselves and their work, was not included in the analysis, since it was concluded that this particular portion of their speech could be memorized and would not be suitable for comparison with spontaneous (i.e., unprepared) speech.

40 suitable political debates were downloaded in total. Manual selection of the final sample followed which resulted in the choice of 8 politicians for each variety. The selection was based on the following conditions: for British English, only SBE speakers were selected. This narrow regional selection was necessary in order to prevent possible undesirable infiltration of other accents in the analysis which differ in their prosodic patterns, and which might therefore skew the final results.

For American English, General American English (GA) speakers were selected. This part was a little bit more problematic with regard to the rather vague definition and delineation of what GA is. Nevertheless, authors in the literature mention that the regional differences in American English are expected to be relatively minor (Grice et al., 2019, p. 6). Therefore, we narrowed it down to the northern part of the United States; 6 out of 8 speakers come from the Midwest and two from the Northeast. All the speakers that showed any typical regional features in their language were excluded from the sample. The maps with the locations the speakers come from can be found in the appendix (section 12).

Moreover, since all the debates were downloaded from the internet, the recordings had to be distinguished by a good sound quality of speech to ensure the best possible conditions for the subsequent automatic computation of data.

6.2. Method

We obtained 3-5 minutes of usable material for each speaker, which were transcribed in the online tool Beey by Newton Technologies. For the following analysis, the Praat programme (Boersma and Weenink, 2019) was used.

Firstly, a textgrid was created for every sound file. A manual adjustment in Praat followed, during which the potential errors in the positions of word and sound boundaries were corrected. The first fifty phrases, which were taken from spontaneous responses to other participants of the debate, were chosen for the analysis of each speaker.

Secondly, a prosodic segmentation had to be carried out, which was a result of a subjective listening task that included establishing prosodic boundaries of all the phrases in the participant's speech. Because a lot of two-word-long phrases were incomplete statements, repetitions, false starts etc. and they could not be considered full prosodic phrases, they were excluded from the analysis, and we only worked with phrases that contained at least three words.

Thirdly, the nuclear syllable was marked in each phrase, which carries the main stress in the phrase and serves as the starting point for the nuclear pitch movement. For the measurement to be done properly, the point was inserted within the boundaries of the first vowel of the nuclear syllable.

Finally, the type of nuclear pitch movement was defined and assigned one of the following abbreviations (F=fall, R=rise, FR=fall-rise, RF=rise-fall, L=level for FMUs, L-f=level-fall, L-r=level-rise for the two newly established realizations). All the participants were listened to at least three times, with periods of about three weeks between the individual listening sessions to try to minimize errors in the analysis. This thesis was originally supposed to be a part of a larger research project in which one more student was supposed to participate, which would provide two listeners and assessors for the analysis. Since the project was not realized in its entirety, all the potential uncertainties regarding the boundaries of the prosodic phrases or the identification of the nuclear pitch movement types were discussed and subsequently settled with my thesis supervisor, doc. Mgr. Radek Skarnitzl, Ph.D. A fully analysed phrase in Praat is depicted in figure 14 below:

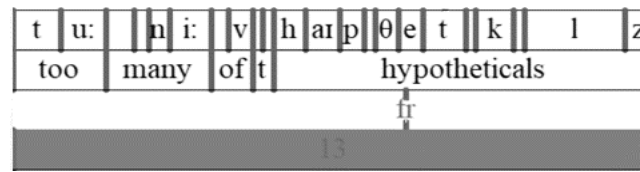


Figure 14: an analysed phrase with marked boundaries and nuclear pitch movement in Praat

In the first two tiers, the individual words and phonemes can be seen. The third tier contains a vertical point, which is placed on the nuclear syllable, and it marks the beginning of the nuclear pitch movement of the phrase. It is labelled “fr” (fall-rise), an abbreviated term for one of the five FMUs mentioned above.

A Praat script was then used to extract all the following information from the acquired data:

1. the number of words per prosodic phrase
2. the number of syllables per prosodic phrase
3. articulation rate in syllables per second
4. the type of the nuclear pitch movement
5. the pitch range of the whole phrase (in ST)
6. the pitch range of the nuclear pitch movement (in ST)

All the data were then transferred to an Excel table and afterwards to the Rstudio programme. A script was used in Rstudio to generate all the graphs displaying the results, which we can see in section 7.

7. Results

The results section comments on the tendencies in British and American English in regard to the six areas mentioned above, it analyses potential similarities or differences between the varieties, and also between individual speakers within the variety. Some graphs show a comparison of the British and American varieties only, other graphs include all the speakers individually. In such cases, speakers are coded according to their variety (BR= British, AM=American), gender (F=female, M=male) and their random order (1-4 for each gender). The final label might then, for example, look like this: BR-M2.

7.1. Prosodic phrases in numbers

The analysis showed that the two English varieties do not differ almost at all in terms of the mean number of words and syllables contained in a single prosodic phrase, or in terms of the mean speakers' speech rate. A simple comparison is provided in table 1 below.

category/variety	British English	American English
syllables	6.3	6.6
words	4.6	4.5
rate (syll/sec)	5.4	5.7

Table 1: an overview of quantitative data for BrE and AmE (mean number of syllables and words in a phrase, articulation rate)

An American English phrase contains 4.5 words on average, a British English phrase 4.6 words. These numbers are approaching Crystal's observation that phrases include 5 words on average (Crystal, 1969, p. 256). He adds that 80% of the phrases in his material were shorter than eight words (Crystal, 1969, p. 256). Crystal's observations are based on data from his collected corpus, which includes "informal discussions and conversations" of people speaking "educated English", i.e., people with a university degree (Crystal, 1969, p. 12). For comparison, the speakers in the current study produced phrases which counted less than 8 words in 95% of all cases. The results in the current study, being higher than Crystal's by 15%, may be influenced by several factors, one of which is the data the analysis is based on. Braga (2004) asserts in her research paper that "in the political debate, there is a more frequent division of the speech in prosodic groups (...) with the purpose of keeping the audience's attention" (Braga, 2004, p. 2). Therefore, the percentage of shorter prosodic phrases may be higher in the current thesis, analysing political speeches, than in Crystal's observations from informal discussions.

The numbers of syllables in the phrases are also almost identical; an American English phrase contains 6.6 syllables, and a British English phrase contains 6.3 syllables on average. 85.5% of all the British English phrases and 81.2% of the American ones contained 8 or fewer syllables.

The articulation rate was calculated from the duration of the phrase and the number of syllables in the phrase. The results showed that the mean rate is slightly higher in the British variety, in which it reaches 5.7 syllables per second, while the mean rate in American English

is 5.4 syllables per second. For comparison, Skarnitzl and Hledíková (2022) conducted a comparative study on Czech and American TED talk speakers (good speakers), in which they compared the speakers' rates as well. The numbers in the current study tally with their results, as they showed that American speakers reach the articulation rate of 5.2 syllables per second in their speech (Skarnitzl & Hledíková, 2022, p. 7). The highest mean rate of speech was measured in British male speakers, who reached 5,9 syll/sec, while the lowest rate of 5,3 syll/sec was reached by American female speakers.

The box plots in figure 15 below show the differences in syllable counts between the two English varieties, but the differences transpire between individual speakers rather than between the varieties generally. Some phrases, symbolized by dots in the plot, deviated from the mean value by as many as 4-9 syllables. These outliers occurred in both American English and British English, and they were present at least once in most speakers, as the box plots below demonstrate.

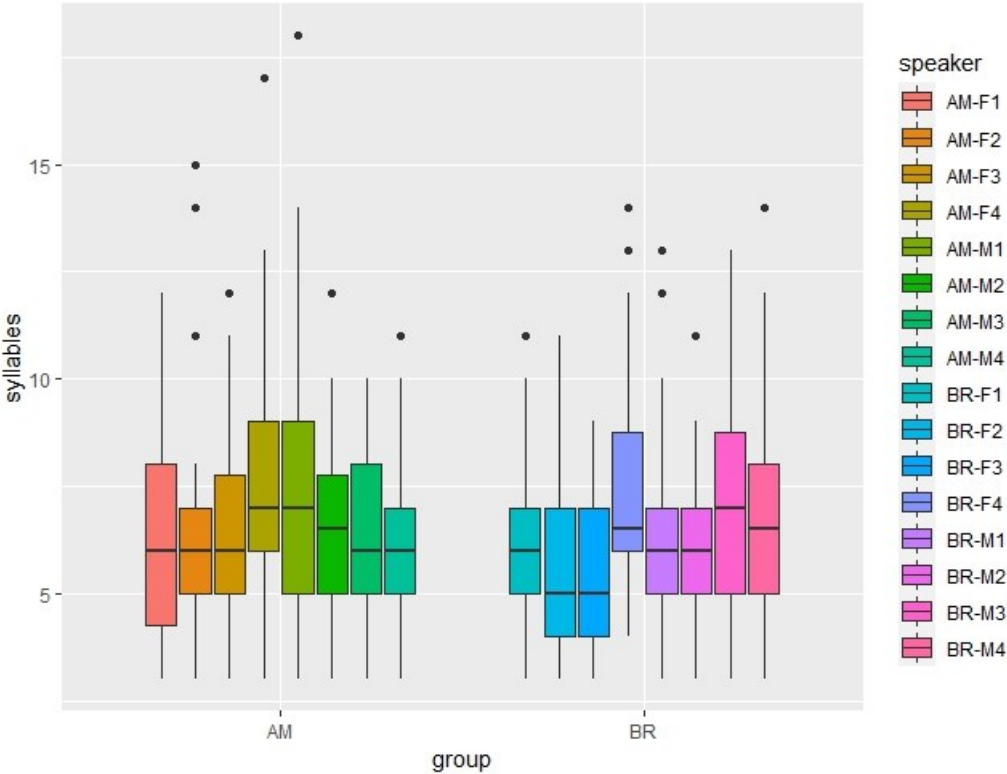


Figure 15: the comparison of individual speakers in terms of the number of syllables in a phrase between AmE (left) and BrE (right)

Since the numbers of words in phrases in British and American English are also comparable, it does not seem necessary to provide any graphs or plots for illustration in that area. The next section, therefore, moves to the types of nuclear pitch movement in the phrases.

7.2. The types of nuclear pitch movements in phrases

British and American English differ more in the area of nuclear pitch movements. An overview of the frequency of the individual movements is depicted in figure 16 below.

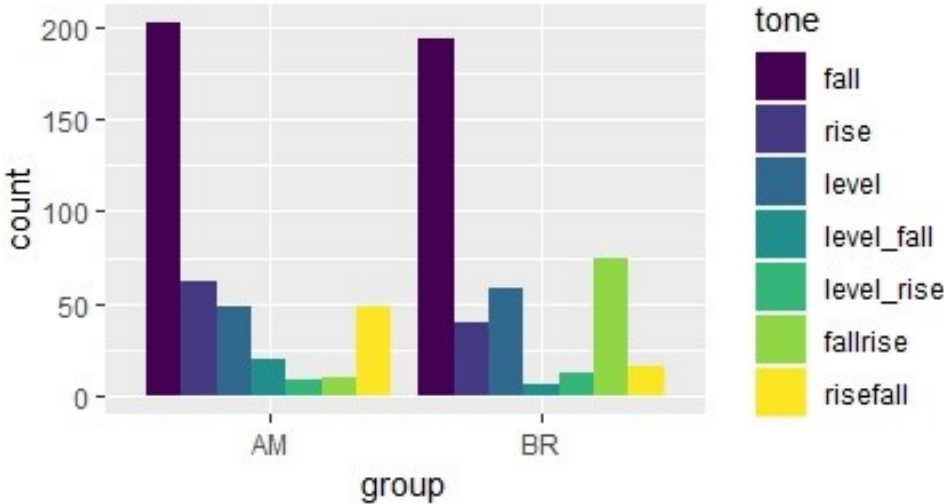


Figure 16: the frequency of each nuclear pitch movement in AmE (left) and BrE (right)

The most commonly used movement by far is the fall, which forms almost 50% of all the movements in each variety. This finding might be surprising given the fact that we analysed spontaneous speech which is usually abundant in incomplete statements, false starts, and potential interruptions by other participants of the discourse. The speaker might also want to “hold the floor” in these debates, which would be more likely accompanied by level intonation (Halliday & Greaves, 2008, p. 114). Therefore, such a high share of the falling movement was not expected in the results of the present research task. Sophie Herment (2021), who conducted a study on read material in English, actually yielded similar, surprising results, when she also discovered falling contours to be most commonly employed in the participants’ performances. Fall was present in 71% of all, what she calls, intonation units (Herment and Tortel, 2021, p. 11).

The frequency of use of the falling intonation contour in the current thesis is slightly higher in American English, but the difference is practically negligible (50.5% in AmE and 48.25% in BrE).

There is a considerable gap in frequency between the first and the second movement in both varieties. Rise is the second most used movement in American English. Nevertheless, it occurred only in 62 phrases, forming 15.5% of all the movements. The second most used movement in British English is the fall-rise, which forms 18.5% of all the analysed phrases. It was mentioned in the theoretical part that some linguists consider fall-rise to be a subtype of a rise, but this study treats them separately; the nuclear syllable is a starting point for two consecutive movements, which makes the structure more complex.

Due to various influences typical for unprepared, spontaneous speech named above, a higher share of the plain level intonation was expected. It was employed by British speakers in 14.5% of the phrases and 12.3% of the phrases by American speakers, which means that it is the third most commonly used type of movement in both varieties. It is interesting to compare the newly established realizations, i.e., the concrete forms of level, rise and fall FMUs, level-rise and level-fall (see section 4.1. for a more detailed explanation). In American English, a higher share of the level intonation is constituted by level-fall than by level-rise; it occurred in 20 phrases, which forms 5% of the 400 phrases, in comparison to 9 instances of the level-rise realization (2.3%). On the contrary, British English phrases provided a slightly higher number of level-rise occurrences (13 instances; 3.25%) and fewer level-fall realizations (6 instances; 1.5%).

One of the questions for which we were seeking an answer was how big a share the newly established and observed realizations, level-fall and level-rise, form within the level FMU and also within fall and rise FMUs. It has been discovered that level-rise forms 16.8% and level-fall forms 7.7% of all level FMUs occurring in British phrases. In American English phrases, level-rise forms 11.5% of all the level FMUs and level-fall forms as much as 25.6% of the level FMUs, which is every fourth instance.

As we have outlined in section 4.1, we believe that the overall function of these new realizations is fall in the case of level-fall and rise in the case of level-rise. In British English, fall is in 97% formed by true fall and in 3% by level-fall. Rise is formed in 75.5% by true rise and in 24.5% by level-rise. In American English, fall is formed by true fall in 91% of the cases, and by level-fall in 9%. True rise constitutes 81.3% of the rising movement, level-rise forms 12.7%.

We believe such results indicate that these newly established realizations are worth being noticed and distinguished in the intonation analysis. This proposed finer distinction in

terminology and in the scope (FMU and its realizations) might be useful in three ways: Firstly, it could provide a better idea about the distribution of all nuclear pitch movements in English. Secondly, it might alleviate the burden for linguists in comparative research, since some languages, for example, Czech, do differentiate between the FMU (in Czech, “melodém”) and its realizations (in Czech, “kadence”; see section 4.1). The above-proposed distinction might be a step towards the unification of the varied terminology. Thirdly, it could be helpful to students of English when studying intonation of English and learning how to shape their intonation contour in speech and what possibilities there are.

The biggest difference between the two varieties, regarding the frequency of use, can be seen with the compound movements: fall-rise and rise-fall. Fall-rise is employed much more often in British English (forming 18.5%) than in American English, in which speakers only used it in 2.5% of all instances. Rise-fall, on the other hand, was more common in American English; it was used in 12% of all 400 phrases. British speakers used rise-fall only on 16 occasions, which forms 4% of the British data set.

The following chapters present the results from the area of pitch range and the differences between varieties and individual speakers. Firstly, the range of the whole phrase is going to be analysed in subchapter 7.3. Secondly, the pitch range of the individual nuclear pitch movements is going to be presented in subchapter 7.4.

7.3. The pitch range of the whole phrase

The following figure, n. 17, offers a comparison of both the varieties and the individual speakers' pitch ranges in the whole phrase. Precisely speaking, all the following sections which concern pitch range depict the results of measuring the f0 range, as an objective quantity.

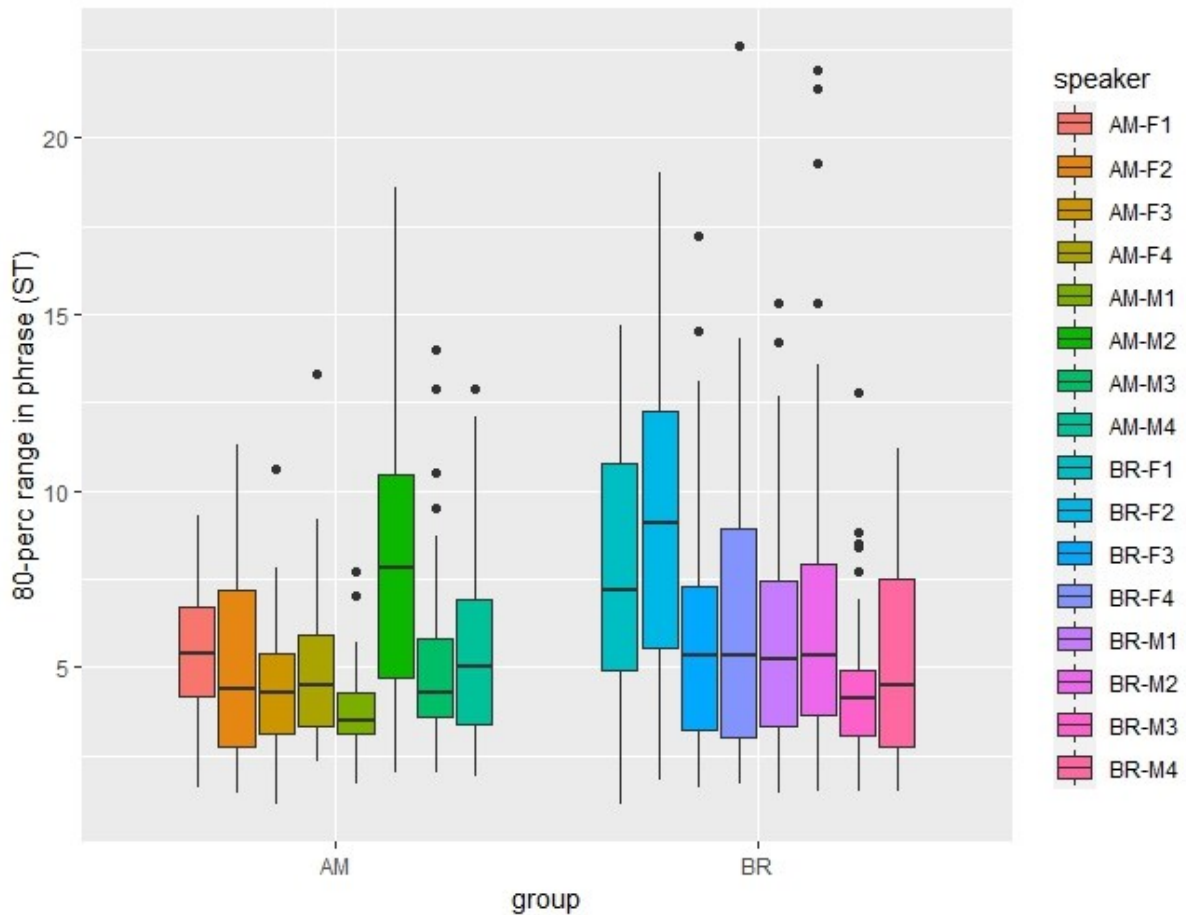


Figure 17: the comparison of individual speakers' pitch ranges of the whole phrases in AmE (left) and BrE (right)

Firstly, the mean pitch range of a phrase was calculated for the whole variety, and it has been discovered that British speakers' pitch ranges are wider by 1.2 ST; the mean range of the whole phrase is 6.4 ST for British English and 5.2 ST for American English.

Secondly, it may be concluded that there is greater variability in the degree of pitch range expansion in individual British speakers than in American speakers. That is illustrated by a taller interquartile range and longer whiskers in most speakers of British English.

The performances of individual British speakers are comparable with one another in terms of their pitch range. The sole exception is represented by speaker BR-M3, who shows the least variability across his phrases in this respect. His mean range reaches 4.4 ST, which is 2

ST narrower than the mean value of the whole group. Speaker BR-F2 shows the greatest variability of all speakers, with a mean pitch range of 9.12 ST. More cases of extreme values, which deviated from the average, and which are represented by the black dots above the whiskers of the box plots, occurred in British English than in American English.

The differences between speakers within the American English group are more noticeable. The most striking one occurs between speakers AM-M1 and AM-M2. Speaker AM-M1 has a mean range of 3.7 ST, while speaker AM-M2 has a mean range of 8.2 ST. The interquartile range of the AM-M1 speaker's phrases (the middle 50% portion of the box plot) is lower than the range of speaker AM-M2. Over 75% of the AM-M1 speaker's range lies within the lowest 25% of the speaker AM-M2's range.

7.4. The pitch ranges of the individual nuclear pitch movements

The results show that there are differences between the varieties in the realm of the speakers' pitch ranges of the nuclear pitch movements (in other words, in the “tail” – a part of a phrase from the nuclear stress to the end of the phrase). We may compare the pitch ranges of the individual pitch movements in the form of box plots in figure 18 below:

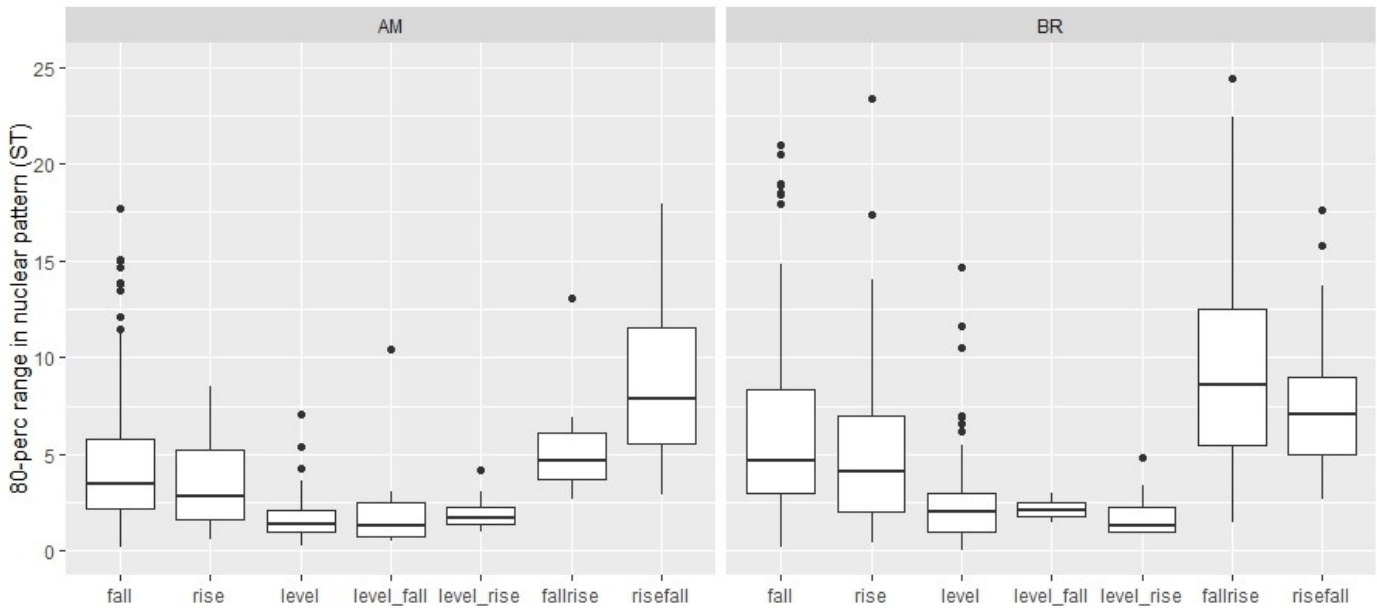


Figure 18: the pitch range of the individual nuclear pitch movements in AmE (left) and BrE (right)

The falling intonation seems to evince slightly higher variability of the pitch range in the British variety, ranging from 0.9 ST to 21 ST. American speakers range from 0.2 ST to 17.7 ST. Rising and fall-rising movements show a similar tendency. The varieties differ most when fall-rise is employed: American speakers range between 3-6 ST in most cases (with one exception of 13.1 ST) in the fall-rising contour, while British speakers range between 1.5-22.4 (with one exception of 24.4 ST), which points to greater variability in the speakers' performances. The middle portion of the fall-rise box plots (50% of all fall-rise cases) is wider in British English. In rise-falls, the variability between groups is comparable, while the pitch ranges between 2.5 ST and 17.6 ST in both varieties, only the IQR (the middle portion of the box plot) is wider in American English.

The plots also demonstrate that speakers expanded their pitch range most when they used one of the two complex movements, fall-rise, or rise-fall, situated on the very right of the respective varieties in figure 18. That is illustrated by their higher position in the graph.

Level FMU, and the two realizations, level-fall, and level-rise, range approximately from 1 ST to 2.5 ST in both varieties. Determining the plain level FMU in the process of the analysis posed a challenging listening task. Identifying the nuances in intonation and whether the movement is still level or whether it falls or rises proved to be difficult and required many listening sessions. The results represented by the box plots prove that most of the nuclear pitch movements identified as level were determined correctly if we follow Hancock’s (2014) assertion that level intonation is a change in pitch smaller than 2 ST (Hancock et al., 2014, p. 204) and also if we consider that level, level-fall and level-rise are the lowest situated box plots in the graph. Nevertheless, even after a successful identification of the type of movement and potential discussion of some problematic cases, we can still see some extremes in the box plots whose range does not align with level intonation. Namely, the British variety in the “level” box plot contains some of those extreme results, with an intonation range as high as 15 ST, which naturally would not indicate the employment of a level movement. Therefore, those extreme cases underwent further inspection in Praat to see where the mistake has occurred. It has been discovered that such examples are usually phrases ending in a creaky or breathy voice, which was probably the reason for an error in the computation process. One of such examples is depicted in figure 19 below.

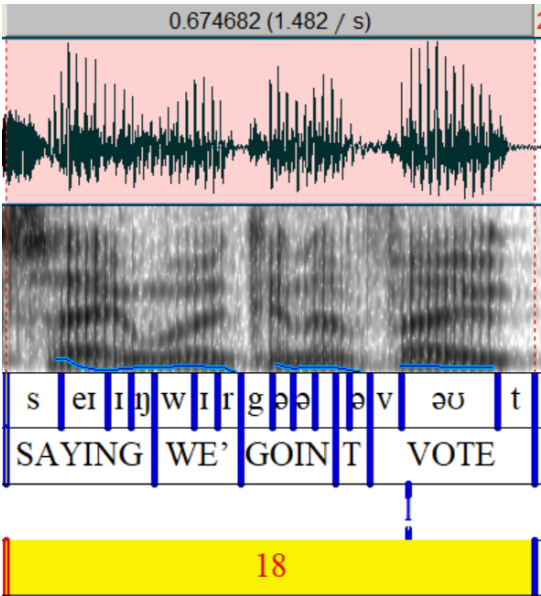


Figure 19: the original phrase with wrongly computed pitch range in the analysis process (91 Hz to 88 Hz)

Even though the blue line, which represents the F0 of the phrase, is not always a reliable indicator of how the pitch movement develops, it is in accordance with the real shape of the

intonation contour in this case. Speaker BR-M1 starts the nuclear pitch movement at the word “vote” at 91 Hz and ends at 88 Hz, which is a difference of about 0.5 ST, but the error results in detecting an almost 15-semitone difference. Another listening inspection then confirmed that there is no such significant fall or rise, but it revealed that the phrase ends in a creaky voice, which creates a misleading environment for automatic computation. It then leads to such high numbers in the results. Fortunately, such cases were rare in the analysis process and did not affect the overall results of the research task in any significant way.

The mean range in the nuclear pitch movement is 6.1 ST in British English, while it is almost 2 ST lower in American speakers, specifically 4.3 ST. 5 out of 8 British speakers have a wider pitch range than all the American speakers, with a mean range of approximately 7 ST.

The following parts are going to present some of the nuclear pitch movements individually in the form of box plots for a finer and more demonstrative comparison of the two groups. With respect to some of the errors in the calculation of the level movements discussed above, the plain level intonation is not depicted graphically in the results section in the same way the other movements are, because the box plots naturally reflected the occasionally skewed pitch range data, too.

7.4.1. The pitch range in falls

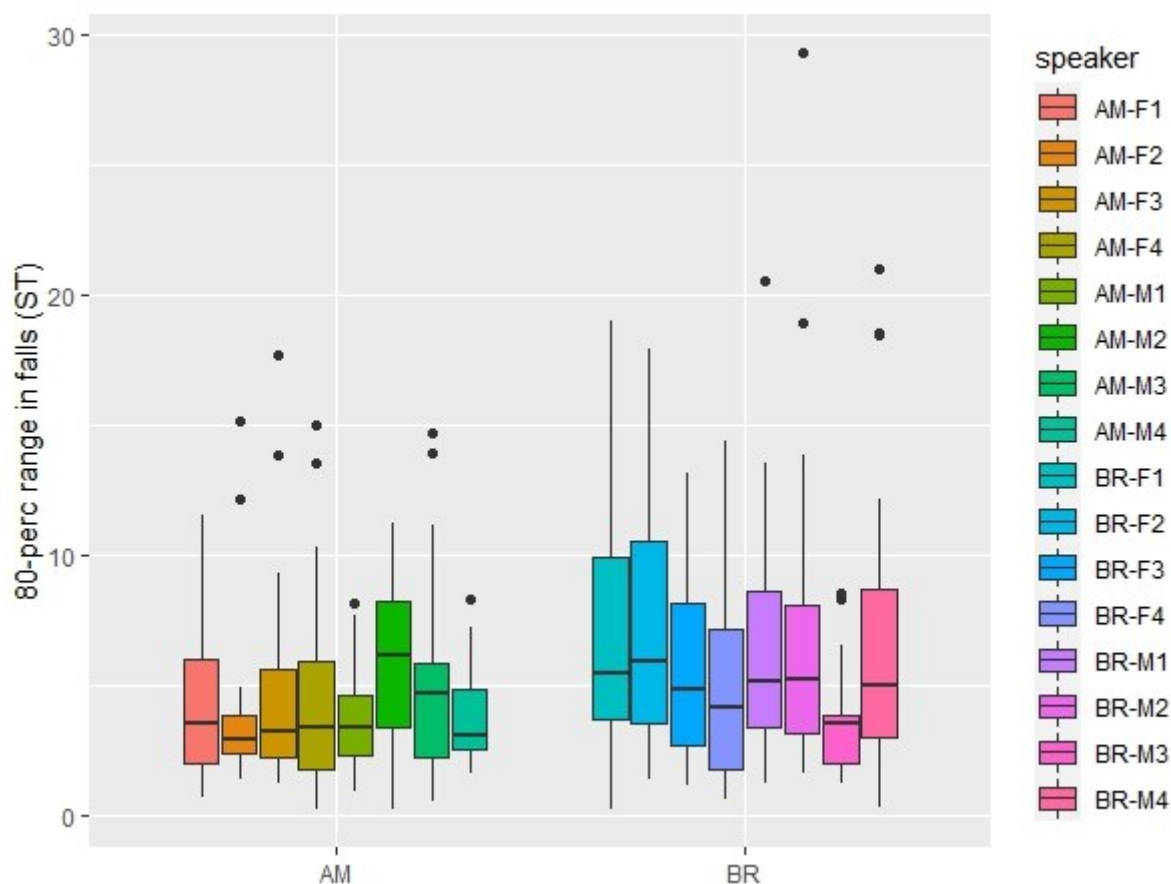


Figure 20: the 80-perc range in falling nuclear pitch movement in AmE speakers (left) and BrE speakers (right)

The falling contour was the most frequently used pitch movement, and it was also used by all 16 speakers. The mean pitch range of the falling movement is 4.4 ST in American speakers and 6.1 ST in British speakers. It shows that British speakers expand their range more in their speech than American speakers. The middle portions of the box plots of British speakers are also wider in most cases, which points to a greater variability within their performances, although the variability is not as high as in the other types of movements.

The American speakers' performances exhibit less variability, demonstrated in large measure by narrower middle portions of the box plots and shorter whiskers. Nevertheless, fall is the movement in which American speakers expand their range most of all movements. Figure 21 below shows an example of a falling movement in speaker AM-M2. The beginning of the movement, positioned at [ɪ], has a value of 179.8 Hz and it falls to 121 Hz at [aʊ] which corresponds to a difference of approximately 6.8 ST.

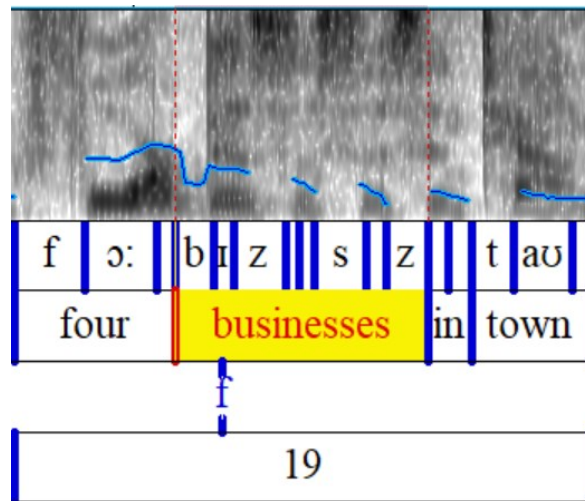


Figure 21: the falling nuclear pitch movement in speaker AM-M2 (a drop from 179.8 Hz to 121 Hz; 6.8 ST)

On the blue line, we can notice a slight “drop” in the f_0 tracker on the first sound of the word carrying the nuclear stress, [b]. That is typical for voiced obstruents, in which we do not perceive the pitch height at all. It only shows the importance of assigning the first vowel to be the beginning of the nuclear pitch movement, not the first sound.

7.4.2. The pitch range in rises

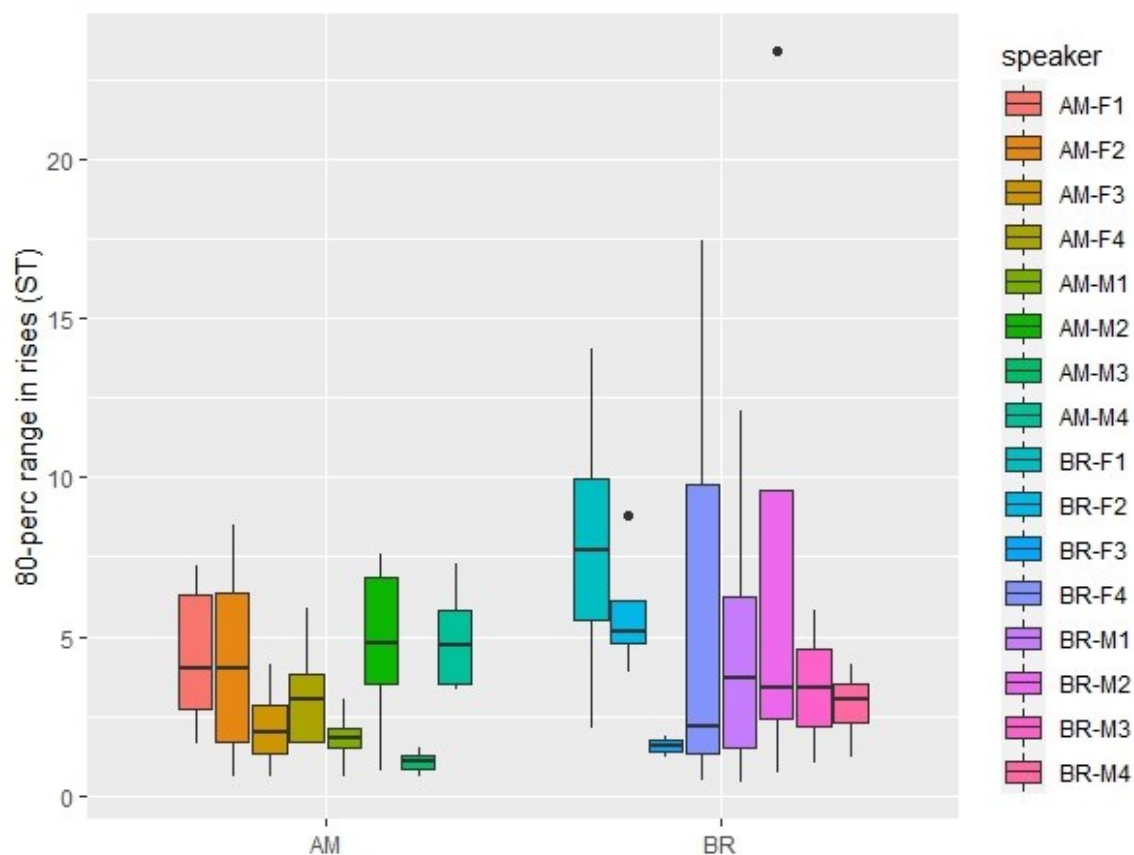


Figure 22: the 80-perc range in rising nuclear pitch movement in AmE speakers (left) and BrE speakers (right)

The rise, although much less frequent, was also used by all 16 speakers. The box plots for rising intonation in figure 22 above demonstrate greater variability than the falling intonation both between groups and between individual speakers, too. The mean range for rises in British English is 5.4 ST, whereas the mean range of American speakers is 3.4 ST, 2 ST narrower.

Some British speakers exhibit considerable variability in their speech, mainly BR-F4, whose pitch range covers around 16.5 ST across her performance, the values oscillating between 0.5 ST and 17 ST. It is interesting to mention that the speaker only used rising contour in her speech three times, once with the range of 0.5 ST, the second time with the range of 2 ST and the third time with the range of 17 ST. On the contrary, speaker BR-F3, for example, showed almost no variability in her performance, while all of her rising movements ranged around 1-2 ST.

The within-group differences in American English are not that striking, but they do occur. Speaker AM-F2 exhibits the greatest variability, while her pitch ranges between

0.6-8 ST. On the contrary, speaker AM-M3 was the most consistent one in his performance of the rising contour, with a range of 0.6-1.5 ST. Figure 23 below shows one of the phrases of speaker BR-F2: the nuclear pitch movement starts at the frequency of 158.3 Hz and ends at 211.8 Hz, which corresponds to a pitch range of 5.2 ST. This value approximates the mean value of 5.4 ST in the rising movement in the British variety.

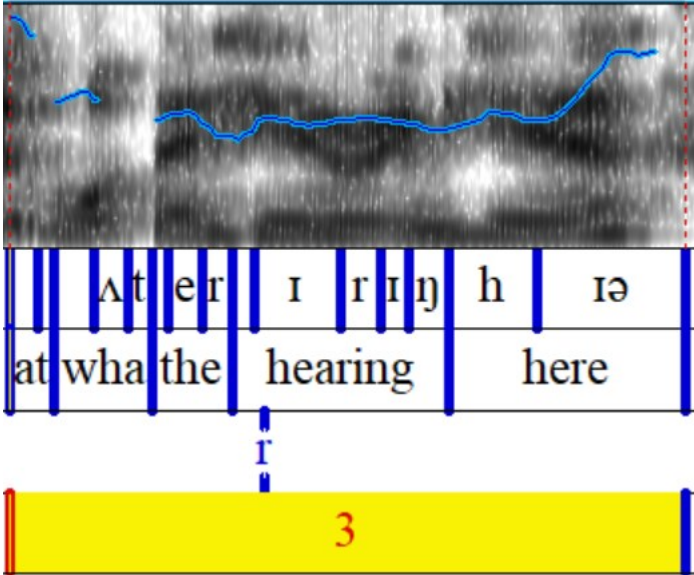


Figure 23: the pitch range of speaker BR-F2's rising nuclear pitch movement (rise from 158.3 Hz to 211.8 Hz)

7.4.3. The pitch range in rise-falls

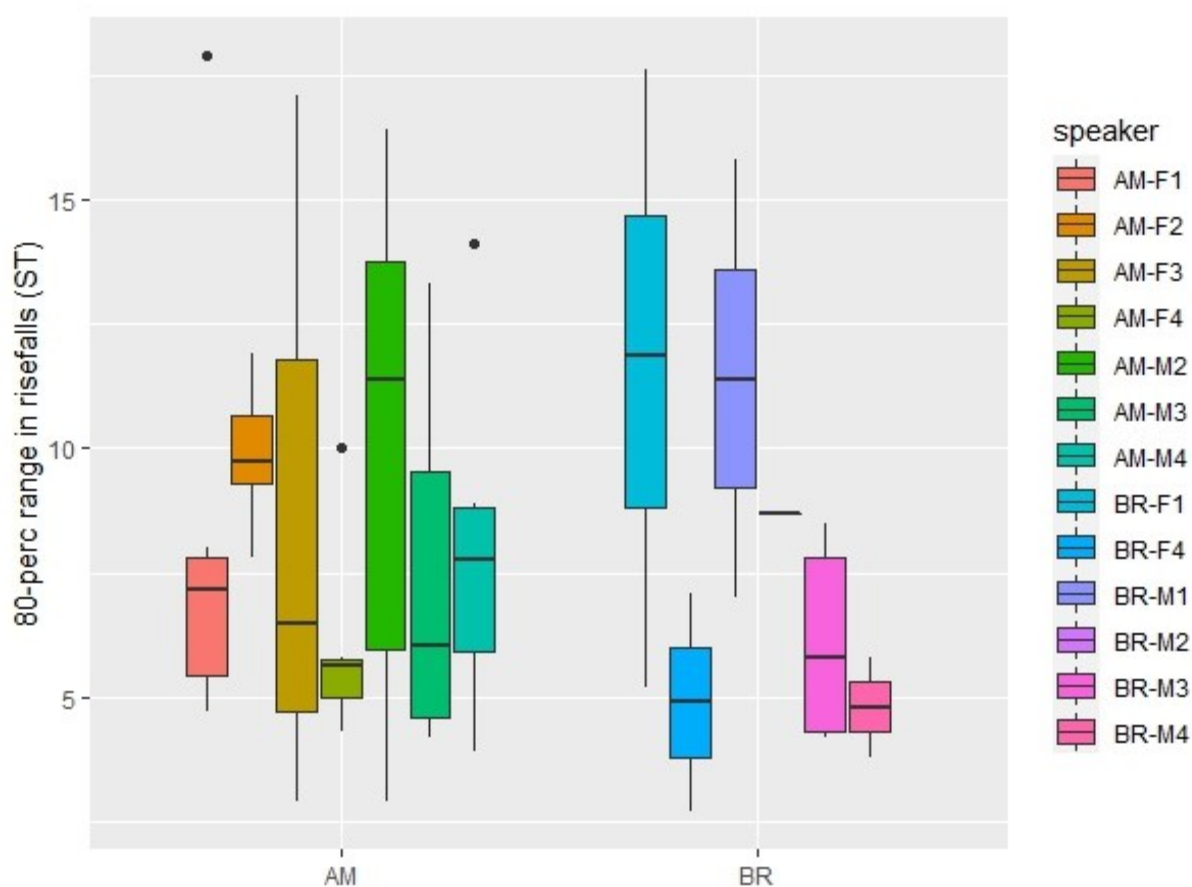


Figure 24: the 80-perc range in rise-falling nuclear pitch movement in AmE speakers (left) and BrE speakers (right)

Rise-fall is much more common in the American variety than the British variety, but it was only present in 7 out of 8 American speakers. In British English, 2 out of 8 speakers did not use it in their phrases at all.

The mean pitch range is quite similar in both varieties; it is 8.7 ST for American speakers and 8 ST for British speakers. Speakers of British English exhibit big differences between one another in terms of their pitch range; BR-M4's mean pitch range is 4.8 ST, while BR-F1 has a mean pitch range of 11.6 ST, which is a difference of 6.8 ST. American speakers vary a little less; AM-F4, a speaker with the least variability, has a mean range of 6 ST, in comparison with speaker AM-M2 with his mean pitch range of 10.2 ST. The difference between the means of those two American speakers is 4.2 ST. Figure 25 below shows speaker AM-M2 and one of the widest pitch ranges in the results section of rise-falls, reaching 14.8 ST.

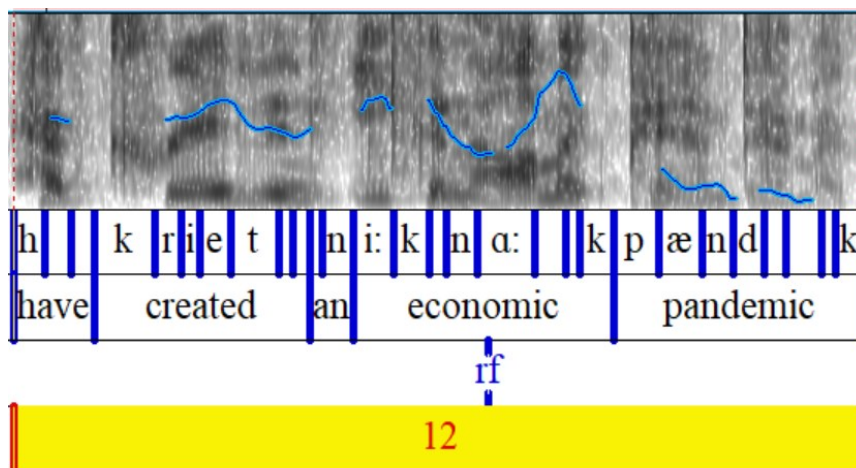


Figure 25: rise-fall in speaker AM-M2, reaching 14.8 ST (99 Hz to 232 Hz)

The changes in pitch are substantial in this case, in comparison with the following example of speaker BR-M4's rise-fall, where the movement is much more moderate; there is only a subtle rise followed by a slight fall. The pitch range of the rise-falling movement in speaker BR-M4 is 3.8 ST.

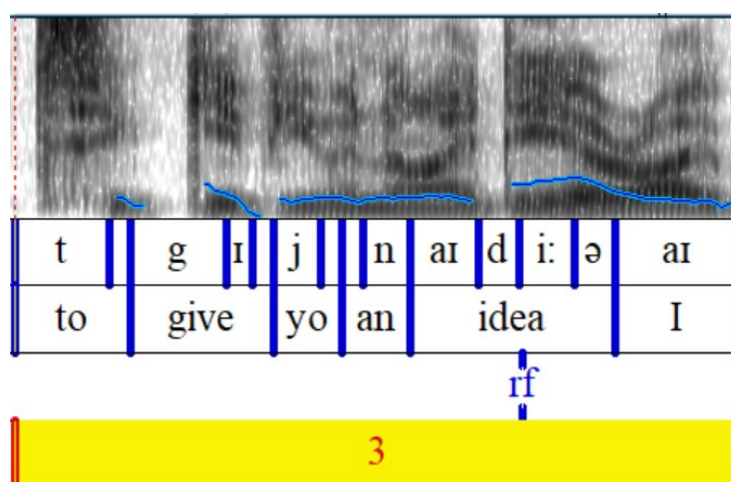


Figure 26: rise-fall in speaker BR-M4, reaching 3.8 ST (113 Hz to 91 Hz)

7.4.4. The pitch range in fall-rises

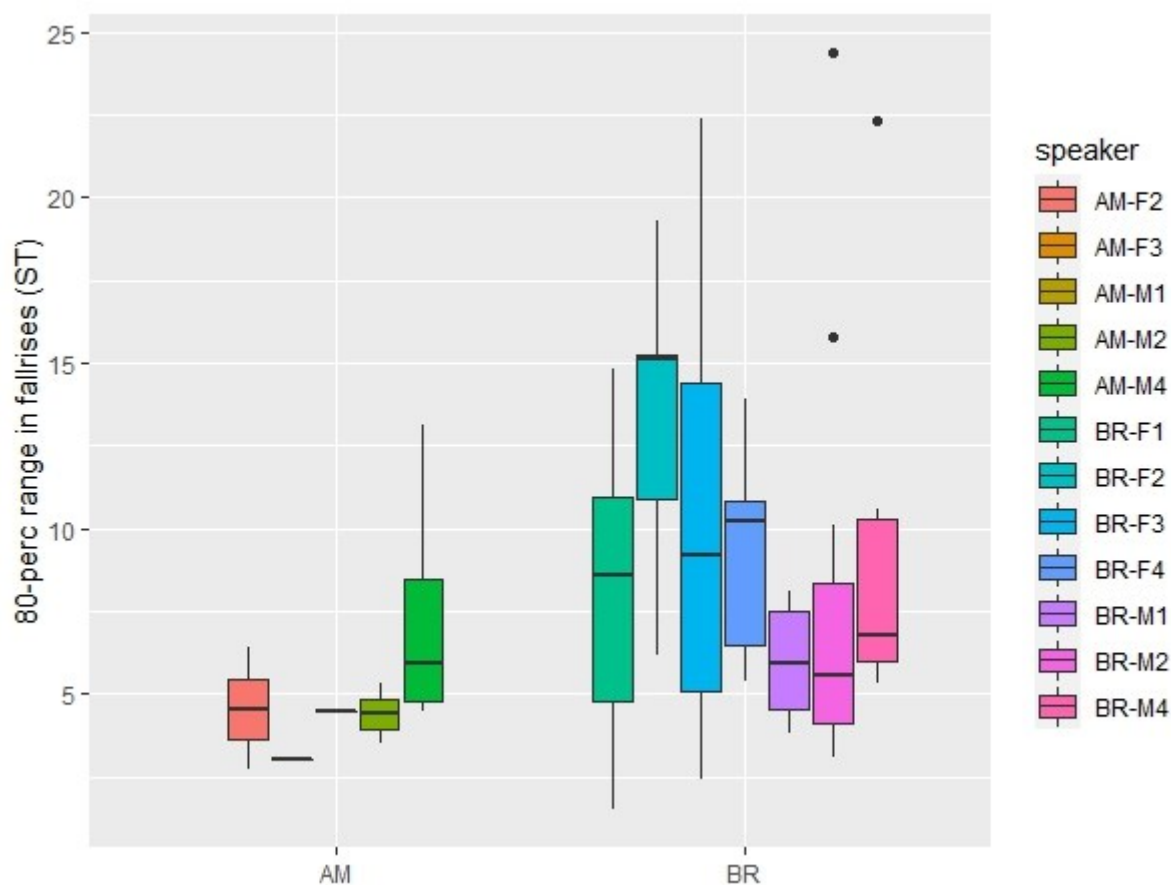


Figure 27: the 80-perc range in fall-rising nuclear pitch movement in AmE speakers (left) and BrE speakers (right)

The fall-rise was used only by five American speakers, two of whom, speakers AM-F3 and AM-M1, used this pattern merely once in their speech. The frequency of the fall-rising pattern is much higher in British English, which could be already seen in figure 16 in section 7.2. The mean value of the British speakers' pitch range is 9.4 ST, while it is only 5.5 ST for American speakers. That shows that British speakers expand their pitch range more by 4.4 ST in comparison with American speakers. Furthermore, fall-rise is the movement in which British speakers expand their ranges most of all movements.

The British speakers also show greater variability in their performances, for example, speaker BR-F3, who has the lowest value of 2 ST and the highest value of 22 ST. With the sole exception of speaker AM-M4, who expands his range from 4.5 to 13.1 ST, American English does not show such big variability within the speakers' performances. Figure 28 below shows an example of a fall-rise performed by speaker AM-M4.

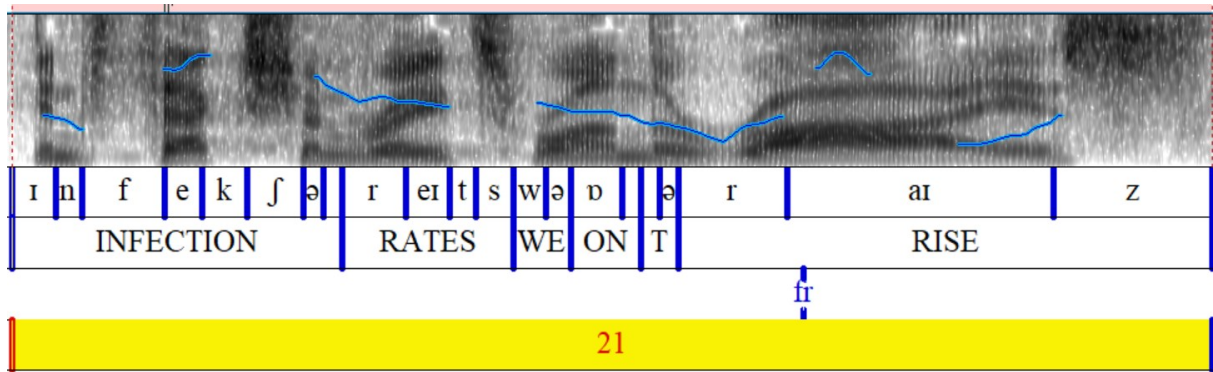


Figure 28: fall-rise in speaker AM-M4 with pitch range reaching 13.1 ST (201 Hz to 91 Hz)

The pitch range of speaker AM-M4 reached 13.1 ST in the nuclear pitch movement, which is the highest value in the American variety. When we compare it with the mean value of the American speakers, AM-M4 represents an exception in the group.

Figure 29 offers a phrase from speaker BR-F3, in which the fall-rise is much less prominent. The pitch range in the nuclear pitch movement in this case is 2.4 ST.

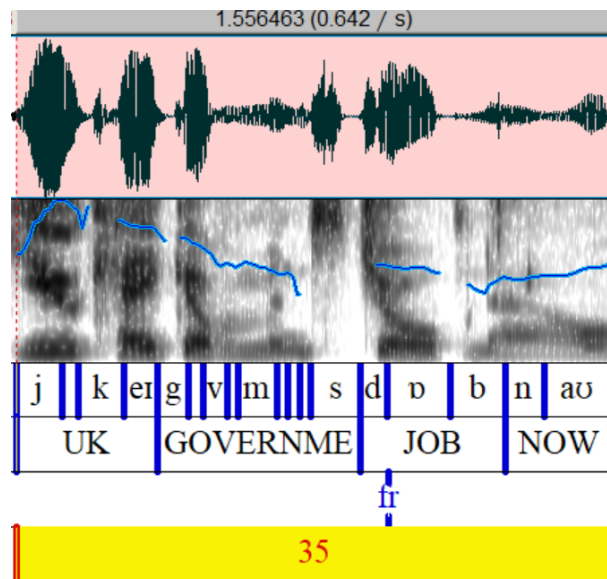


Figure 29: fall-rise in speaker BR-F3 with pitch range reaching 2.4 ST

8. Discussion

The present thesis provided an analysis of 800 English phrases, 400 British and 400 American, to discover the tendencies in both varieties in the intonational aspect of language. The material was extracted from British and American political debates (for a detailed description of the criteria and method of the material extraction see chapter 6).

The main objectives of the empirical part were numerous. Firstly, it aimed to discover potential differences between two English varieties, the British and the American, in the following respects: the number of syllables and words per phrase, the speech rate, the type of nuclear pitch movement, and the pitch range of the whole phrase and the nuclear pitch movement. These areas are scarcely covered in comparative research of American and British English. For the lack of previous evidence for potential dissimilarities between the varieties, we did not assume to find any substantial differences between them, and a null hypothesis was employed.

The results showed that there are no significant differences between the two varieties in the quantitative respect; namely, in the number of syllables or words per phrase. An American phrase contains 4.5 words and 6.3 syllables on average, while a British phrase contains 4.6 words and 6.6 syllables on average. The articulation rate was slightly higher in British English and the results for the rate in American English tally with the numbers in a recent study by Skarnitzl and Hledíková (2022).

As for the type of nuclear pitch movement, the results were somewhat surprising. The analysis showed that the most frequently used movement at the end of phrases is the fall, employed in almost 50% of the cases in both British and American English. Such an outcome was unexpected due to the nature of the research, which included phrases extracted from spontaneous, spoken production. With regards to the frequently higher share of incomplete statements, false starts, or potential interruptions by other participants of the discourse, fall was not expected to be represented in such high numbers due to its customary association with definiteness and complete statements. These results, however, correspond to the discovery of Sophie Herment and her recent study (2021), in which she yielded similar results, only in a reading task.

The second most used nuclear pitch movement is the rise in American English (15% of all the movements) and the fall-rise in British English (18.5%).

The varieties differ most in the frequency of the fall-rising contour. It was scarcely present in American English, forming only 2.5% of all the movements, which made it the least-used pitch movement in American English.

Level was the third most used movement in both varieties, with a comparable frequency of use; it occurred in 14.5% of all phrases in British English and 12.3% of American phrases.

Rise-fall was more frequently used by American speakers as it occurred in 12% of all the phrases, while British speakers only used rise-fall in 4% of the cases. In British English, it was the least used one from the set of the five established FMUs.

Furthermore, the thesis shows how big a portion of the level movements is formed by two newly coined realizations, the level-rise and level-fall. The results showed that level-rise is slightly more prominent in British English, in which it forms 16.6% of all level movements, while American speakers used it in 11.5% of the sample. Level-fall is more used in American English, forming 25.6% of all level intonation occurrences (every fourth instance), while British speakers used it in 7.7% of the phrases. The level-rise and level-fall realizations are also related to the rise and fall FMUs respectively in both varieties. In British English, fall is in 97% formed by true fall and in 3% by level-fall. Rise is formed in 75.5% by true rise and in 24.5% by level-rise. In American English, fall is formed by true fall in 91% of the cases, and by level-fall in 9%. True rise constitutes 81.3% of the rising movement, level-rise forms 12.7%.

The thesis offers three reasons why it might be beneficial to differentiate nuclear pitch movements in greater detail (FMU and its realizations): Firstly, it could provide a better idea about the distribution of all nuclear pitch movements in English. Secondly, it might make the comparative research easier by applying the same hierarchy and by introducing terms which already exist in other languages (FMU: Czech “melodém”, and its realizations: Czech “kadence”; see section 4.1). The above-proposed distinction might be a step towards the unification of the varied terminology. Thirdly, students of English could benefit from a more unified system when studying intonation and learning how the intonation contour in speech can be shaped and what possibilities there are.

It has been concluded that the pitch range of the whole phrase is 1.2 ST wider in British English than in American English. The result section then provided results for the pitch range of each nuclear pitch movement, showing that British speakers have a wider pitch range than American speakers in most movements, specifically in fall, rise, and fall-rise. American speakers have a wider pitch range only in rise-fall.

The results also showed that American participants have the widest pitch range in the falling movement (mean value of 8.7 ST) and the narrowest in the rising movement (3.4 ST). The British speakers expand their range most in fall-rise (mean value of 9.4 ST) and least in rise (5.4 ST), similarly to their American counterparts.

9. Conclusion

The current thesis reports on the findings regarding differences between British and American phrases. Some differences have been found between the two English varieties, mainly in the types of nuclear pitch movements at the end of phrases, but some were also identified in the area of the width of speakers' pitch ranges. By comparing the two varieties, the thesis aimed to fill the gap in the area of comparative research, which has only rarely studied the differences between British and American English in the respects mentioned above. The results correspond to the scarce data obtained by some researchers in the recent past. Moreover, the thesis introduces a new term for a finer distinction in intonation analysis, FMU, corresponding to the Czech term "melodém", which, to the best of our knowledge, the English terminology lacks. Incorporating this term in the analysis process could help paint a more accurate picture of the types and subtypes of nuclear pitch movements in English. We believe that this distinction might contribute positively to the intonation comparative studies.

The thesis could serve as a source and basis for further research in the area of suprasegmental features of English, be it a comparative study between different languages or between different varieties of English. Future studies could examine other potential influences and variables; for example, the effect of gender or age on some of the suprasegmental aspects of language, like the speakers' articulation rate or the width of their pitch range. The findings are valuable not only for linguists and other researchers but also for learners, who could utilize gained knowledge about the particularities of individual English varieties to engage successfully in communication with native speakers of the given variety.

10. References

- Aijmer, K. (1997). *Conversational Routines in English: Convention and Creativity*. New York: Routledge. <https://doi.org/10.4324/9781315845128>
- Allen, V. F. (1971). Teaching Intonation, from Theory to Practice. *TESOL Quarterly*, 5(1), 73-81. <https://doi.org/10.2307/3586113>
- Bartels, C. (1999). *The Intonation of English Statements and Questions: A Compositional Interpretation*. New York: Routledge. <https://doi.org/10.4324/9781315053332>
- Beckman, M. E., & Elam, G. A. (1997). *Guidelines for ToBI Labelling*. Version 3. Ohio: Ohio State University. May 2012. http://www.ling.ohiostate.edu/research/phonetics/E_ToBI/.
- Beckman, M.E. et al (2005). The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 9-54). Oxford: Oxford University Press.
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.41, retrieved 1 January 2022 from <http://www.praat.org/>.
- Bolinger, D. L. (1978). Intonation across languages. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcik (Eds) *Universals of Human Language, Phonology*, 2, 471-524. Stanford: Stanford University Press.
- Bolinger, D. L. (1983a). The inherent iconism of intonation. In J. Haiman (ed.), *Iconicity in syntax* (pp. 97-109). Amsterdam: John Benjamins.
- Bolinger, D. L. (1998). Intonation in American English. In D. Hirst, & A. D. Cristo (eds.), *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.
- Braga, D., & Marques, M.A. (2004). The Pragmatics of Prosodic Features in the Political Debate. Conference Proceedings, Speech Prosody.
- Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge: Cambridge University Press.
- Busà, M. G., & Urbani, M. (2011). *A Cross Linguistic Analysis of Pitch Range in English L1 and L2*. International Congress of Phonetic Sciences.

- Büring, D. (2013). Syntax, information structure and prosody. In M. Den Dikken, (Ed.), *The Cambridge Handbook of Generative Syntax*, 860-896. doi:10.1017/cbo9780511804571.029
- Chun, D. M. (2002). Discourse intonation in L2. *Language Learning & Language Teaching*. Amsterdam: John Benjamins. doi:10.1075/llt.1
- Clopper, C., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of phonetics*, 39, 237-245. doi:10.1016/j.wocn.2011.02.006.
- Cole, J., Mo, Y., & Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and cognitive processes*, 25, 1141-1177. doi 10.1080/01690960903525507.
- Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139166973
- Crystal D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press. Accessed through <https://archive.org>.
- Crystal, D. (1997b). *English as a global language*. Cambridge: Cambridge University Press.
- Cutler, A., Dahan, D., & Donselaar, W. (1997). Prosody in the Comprehension of Spoken Language: A Literature Review. *Language and speech*, 40(2), 141-201. doi:10.1177/002383099704000203.
- Derwing, T. M., & Rossiter, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Elfner, E. (2018). The syntax-prosody interface: current theoretical approaches and outstanding questions. *Linguistics Vanguard*, 4(1), 20160081. <https://doi.org/10.1515/lingvan-2016-0081>
- Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. New York: Oxford University Press.
- Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in cognitive sciences*, 10, 244-249. doi:10.1016/j.tics.2006.04.002.
- Fuchs, R. (2018). Pitch Range, Dynamism and Level in Postcolonial Varieties of English: A Comparison of Educated Indian English and British English. *Proceedings of Speech Prosody 9*, 893-897. 10.21437/SpeechProsody.2018-180.

- Gilbert, J. M. (2008). *Teaching Pronunciation Using the Prosody Pyramid*. New York: Cambridge University Press.
- Grabe, E. & Post, B. (2002). *Intonational Variation in the British Isles*. In B. Bel & I. Marlin (Eds.), *Proceedings of the Speech Prosody 2002 conference* (pp. 343-346). Aix-en-Provence: Laboratoire Parole et Langage.
- Grabe, E., Kochanski, G., & Coleman, J. (2005). The intonation of native accent varieties in the British Isles: Potential for miscommunication. *English Pronunciation Models: A Changing Scene*, 311-338.
- Grice, M., & Baumann, S. (2007). An introduction to intonation – functions and models. In J. Trouvain & U. Gut (Ed.), *Non-Native Prosody: Phonetic Description and Teaching Practice* (pp. 25-52). Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110198751.1.25>
- Grice, M., German, J. & Warren, P. (2019). Intonation systems across varieties of English. In C. Gussenhoven, and A. Chen (Eds), *The Oxford Handbook of Language Prosody, Oxford Handbooks* (pp. 285-302). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198832232.013.18>, accessed 2 Jan. 2023.
- Gussenhoven, C. (2002). Intonation and Interpretation: Phonetics and Phonology. *Conference proceedings, Speech Prosody, 2002*. Available at https://www.isca-speech.org/archive_open/sp2002/sp02_047.pdf
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Greaves, W. S., & Halliday, M. A. K. (2008). *Intonation in the grammar of English*. London: Equinox Publishing.
- Hancock, A., Colton, L., & Douglas, F. (2014). Intonation and gender perception: Applications for transgender speakers. *Journal of Voice*, 28(2), 203-209. doi:10.1016/j.jvoice.2013.08.009
- Herment, S., & Tortel, A. (2021). Chapter 8. The intonation contour of non-finality revisited. *AILA Applied Linguistics Series*, 176-195. doi:10.1075/aals.19.08her
- Himmelmann, N. P. (2022). Prosodic phrasing and the emergence of Phrase structure. *Linguistics*, 60(3), 715-743. doi:10.1515/ling-2020-0135

- Hirst, D., & Cristo, A. D. (1998). *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.
- Hudson, T., Setter, J. & Mok, P. (2019). Nuclear Tones in Hong Kong and British English. In S. Calhoun, P. Escudero, M. Tabain, and P Warren. (Eds), *Proceedings of the 19th International Congress of Phonetic Sciences*, (pp. 320-323). Canberra: Australasian Speech Science and Technology Association Inc. Available at <https://centaur.reading.ac.uk/80883>.
- Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly*, 33(1), 37-63. doi:10.2307/3588190
- Levis, J. M. (2002). Reconsidering low-rising intonation in American English. *Applied Linguistics*, 23(1), 56-82. doi:10.1093/applin/23.1.56
- Levis, J. M. (2005). Comparing apples and oranges? Pedagogical approaches to intonation in British and American English. In Dziubalska-Kolaczyk, K. and Przedlacka, J. (Eds) *English pronunciation models: A changing scene* (pp. 339-366). Bern: Peter Lang. Available at: http://lib.dr.iastate.edu/engl_pubs/138
- Levis, J. M. (2013). Suprasegmentals: Intonation. In C.A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 5443-5449). Oxford: Blackwell/Wiley. doi:10.1002/9781405198431.wbeal1124
- Mennen, I., Schaeffler, F. & Docherty, G. (2007). Pitching it differently: a comparison of the pitch ranges of German and English speakers. *16th International Congress of Phonetic Sciences*, 1769-1772.
- Mennen, I., & de Leeuw, E. (2014). *Beyond segments: Prosody in SLA*. Cambridge: Cambridge University Press.
- O'Connor, J. D., & Arnold, G. F. (1961). *Intonation of colloquial English* (2nd ed., 1973). London: Longman.
- Ohala, J. J., & Gilbert, J.B. (1981). Listeners' ability to identify languages by their prosody. In P. Léon and L. Rossi (eds.), *Problemes de prosodie, Vol. 11: Experimentations, modeles et fonctions* (pp. 123-131). Ottawa: Didier.
- Palková, Z. (1994). *Fonetika a Fonologie češtiny: s obecným úvodem do problematiky oboru*. Praha: Karolinum.

- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Cambridge: Massachusetts Institute of Technology. PhD thesis, MIT. Distributed 1988, Indiana University Linguistics Club.
- Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Reed, M., & Levis, J. M. (2015). *The handbook of English pronunciation*. Hoboken, NJ: Wiley Blackwell.
- Roach, P. (2009). *English phonetics and phonology* (4th ed.). Cambridge: Cambridge University Press.
- Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice*, 30(3), 158-164. doi 10.1080/00405849109543496
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247. doi:10.1007/bf01708572
- Skarnitzl, R., & Hledíková, H. (2022). Prosodic Phrasing of Good Speakers in English and Czech. *Frontiers in psychology*, 13, 857647. <https://doi.org/10.3389/fpsyg.2022.857647>
- Tench, P. (1996). *The intonation systems of English*. London: Cassell.
- Tench, P. (2020). Intonation in semantic system networks. In G. Tucker, G. Huang, L. Fontaine, & E. McDonald (Eds.). *Approaches to systemic functional grammar* (pp. 246-268). UK: Equinox Publishing.
- Tyler, J. C. (2015). Expanding and mapping the Indexical Field. *Journal of English Linguistics*, 43(4), 284-310. doi:10.1177/0075424215607061
- Volín, J., Poesová, K., & Weingartová, L. (2015). Speech melody properties in English, Czech and Czech English: Reference and interference. *Research in Language*, 13(1), 107-123. doi:10.1515/rela-2015-0018
- Wakefield, J. (2020). *Intonational morphology*. Singapore: Springer.
- Walker, G. (2017). Pitch and the projection of More talk. *Research on Language and Social Interaction*, 50(2), 206-225. doi:10.1080/08351813.2017.1301310

- Ward, N. G. (2019). *Prosodic patterns in English conversation*. Cambridge; New York: Cambridge University Press.
- Warren, P. (2016). *Uptalk: the phenomenon of rising intonation*. Cambridge: Cambridge University Press.
- Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge: Cambridge University Press.
- Wharton, T. (2012). Prosody and meaning: Theory and practice. In J. Romero-Trillo (Ed.). *Pragmatics and Prosody in English Language Teaching* (pp. 97-116). Madrid: Springer.
- Wichmann, A. & Levis, J. M. (2015). Chapter 8. English Intonation: Form and Meaning. In M. Reed, & J. M. Levis (Eds). *The handbook of English pronunciation* (pp. 139-155). Malden, MA: Wiley Blackwell.

11. Resumé

Tato práce se zabývá porovnáním britské a americké angličtiny z hlediska prozodických frází a jejich charakteristik na suprasegmentální úrovni jazyka. Cílem výzkumu bylo zjistit, zda a případně jak se obě variety od sebe liší v oblasti délky fráze, intonačního pohybu na konci fráze a intonačního rozpětí mluvčích.

Teoretická část začíná v části 2 přehledem o tom, co je to prozodie, jaký význam v jazyce má a k čemu se v lidské komunikaci využívá. Objasňuje komplikace při vymezení toho, co do prozodie patří a zdůrazňuje, že jevy, které se do prozodie zahrnují, se liší od autora k autoru a jsou odrazem toho, co je do konkrétního výzkumu potřeba zahrnout. Prozodie se odráží na všech úrovních, od slabik až po celkový diskurz. Prozodická organizace věty byla dlouho považována za univerzální ve všech jazycích, nicméně postupem času bylo zjištěno, že každý jazyk oplývá specifickými prozodickými rysy a v každém se prozodie chová jinak. Notační systémy, které se prozodii jazyků snaží popsat na papíře a které byly nejdříve vytvářeny s ideou univerzálně použitelného rámce (např. systém ToBI), se nyní fragmentují a vznikají podtypy, které mají více vycházet vstříc individuálním potřebám určitých jazyků. Práce zdůrazňuje, že prozodie je nedílnou součástí jazyka a nelze ji oddělit či nepoužívat.

Část 2.1. popisuje funkce prozodie a postupuje od úrovně slov směrem k vyšším celkům. Nejprve se v krátkosti zaměřuje na lexikální funkci prozodie; a to prve v tónových jazycích, jakými jsou například Africký jazyk Kono nebo čínština. Tyto jazyky používají prozodii lexikálně, tedy změnou v prozodii dochází ke změnám ve významu slov. Na druhé straně spektra stojí například angličtina nebo čeština, které nejsou tónovými jazyky a prozodie v nich nemá takto přímý vliv na významy jednotlivých slov. Pro angličtinu má nicméně prozodie význam ještě trochu jiný než pro češtinu. Jisté prozodické jevy, konkrétně přízvuk, se používá na lexikální úrovni k rozlišení slovních druhů (for'bear x 'forebear). To je vždy ovšem doprovázeno také změnou v kvalitě vokálu v daném slově.

Část 2.1.2. se dále zabývá gramatickou funkcí prozodie. Začíná prozodickým frázováním. Zaměřuje se na stavbu anglické fráze a tonalitu, neboli na to, jak mluvčí člení svou výpověď, a na detekci hranic frází. Crystal (1969) uvádí, že prozodické fráze v korpusu, se kterým pracoval při vypracovávání své studie, obsahovaly průměrně pět slov. Dříve bylo předpokládáno, že prozodické fráze mají hranice identické s hranicemi syntaktických frází. To bylo podpořeno mnoha jevy v jazyce, které této skutečnosti nasvědčovaly. Postupem času výzkum prokázal, že tyto hranice se sice velmi často překrývají, ale neexistuje pravidlo, které

by mohlo být na základě toho formulováno. Neexistují žádná kritéria, na základě kterých by bylo možné rigidně ustanovit pravidlo pro detekci hranic prozodických frází, či žádný univerzální návod. Prozodické fráze tedy definujeme zkrátka jako doménu intonační linky, kterou vnímáme uceleně, jako ukončenou. Někdy je dodáván do definice také rytmický aspekt, který říká, že fráze jsou často odděleny pauzou nebo ukončeny prodloužením posledního segmentu fráze. Je důležité zmínit, že ve skutečnosti je detekce hranic frází velmi složitá a nejednoznačná. Je tedy nutné použít i vlastní úsudek a intuici při identifikaci předělu frází. Prozodie je z hlediska gramatické funkce také prostředkem k rozlišení větných typů. Stejná věta může být vyslovena vícero způsoby (s klesavou intonací – značící tvrzení, oznamovací větu, se stoupavou intonací – značící otázku). Práce ovšem zdůrazňuje, že toto rozdělení je jen velké zjednodušení toho, jak ve skutečnosti prozodie, potažmo intonace, v jazyce funguje.

Větná stavba v angličtině je do určité míry ovlivněna aktuálním členěním větným (AČV). Nová informace se nachází přirozeně na konci věty, kam se přesouvá i hlavní větný přízvuk. Pokud mluvčí potřebuje zdůraznit jinou část věty, může k tomu využít buď syntax (např. vytýkáci vazbu), nebo prozodii. Kombinací důrazu, hlasitosti a výšky tónu můžeme dosáhnout důrazu na jiném slově, než které by bylo zdůrazněno v lineárním průběhu podle AČV. Prozodie slouží nejen mluvčím, kteří ji využívají k výše zmíněným manipulacím jazyka, ale také posluchači, který prozodii využívá k dekodování zprávy. To bylo potvrzeno několika výzkumy, např. Epstein (1960) odhalil při své studii, že slova, která nemají žádný význam, si posluchači lépe zapamatovali a vybavili, pokud byla tvořena podle morfologických pravidel daného jazyka a doprovázena jeho typickou prozodií.

Poslední funkcí zmíněnou v práci je funkce diskurzivní. Prozodie pomáhá mluvčím členit delší úseky, jako je celá výpověď, na kratší, navíc je nástrojem pro strukturu konverzace a výměny replik mluvčích. Do průběhu této funkce ovšem vstupují další faktory, které dekodování zprávy mohou ztěžovat: například emoční stav mluvčího, nálada, záměr, rychlost řeči apod. Ty mohou konverzaci stočit jiným směrem a ztížit porozumění ostatních účastníků, což může vést k nedorozumění v komunikaci.

Práce se nadále zabývá především jednou určitou složkou prozodie, kterou je intonace (kapitola 3). Definovat, jak přesně se intonace liší od samotné prozodie, je velice složité. Někteří autoři oba termíny zaměňují, někdo vnímá intonaci jako podkategorii prozodie. To je součástí takzvané širší definice prozodie, která v podstatě říká, že intonace je „způsob, jak něco říkáme“. V takovém případě intonace ztělesňuje kombinaci akustických jevů (délka, intenzita

a základní frekvence) a je velmi nespecifická. V užším slova smyslu se intonace soustředí pouze na výšku tónu a změny v této výšce. Další části práce se víceméně soustředí právě na výšku tónu a změny ve výšce tónu v souladu s užším vymezením termínu intonace. Výška tónu je ovlivněna faktory jako je tělesná stavba nebo pohlaví mluvčího, může být ale také manipulována vědomě. Rozmezí nejnižšího a nejvyššího tónu se nazývá intonační rozpětí. Některé jazyky, například angličtina, mají větší intonační rozpětí než jiné, například čeština. Ta je v porovnání s angličtinou intonačně mnohem plošší.

Sekce 3.1. rozebírá formu a funkci intonace (především jednotlivých intonačních pohybů), zaměřuje se hlavně na gramatickou funkci, tedy jak intonace funguje v kombinaci s větnými typy. Intonační pohyb začíná ve slově nesoucí hlavní větný přízvuk. Nejběžněji používané intonační pohyby zahrnují klesavý, stoupavý, klesavě-stoupavý, stoupavě-klesavý a neukončující (často realizován rovnou kadencí), což je set pěti pohybů původně navržený lingvistou Sweetem již v roce 1890. V angličtině se tyto pohyby nazývají v literatuře zabývající se intonací různě, tato práce dává přednost názvu „nuclear pitch movement“, v češtině se používá termín melodém. Klesavý melodém se většinou používá ve větách oznamovacích, rozkazech a otázkách uvozených tázacími zájmeny (wh-questions). Stoupavý melodém se naopak používá ve zjišťovacích otázkách. Sekce 3.1. ovšem také rozebírá detailnější rozdíly v použití různých typů intonace u otázek. Stoupavý melodém se také používá ve spojitosti s nejistotou či s nedokončenými či neúplnými výroky. Taktéž velmi krátké odpovědi typu „Ano/Ne“ mohou být doprovázené klesavým i stoupavým melodémem, v každém případě ale předávají jinou informaci – klesavý melodém vyvolává pocit definitivy, zatímco stoupavý melodém spíše vyzývá druhého účastníka hovoru, aby pokračoval. Důležité je zmínit, že desetiletí výzkumu ukazují, že nelze typy melodémů striktně rozdělit podle toho, k jakému větnému typu se pojí a nelze tedy takové rozdělení předpokládat.

Stoupavý melodém našel své využití i mimo výše uvedené případy. V 70. letech 20. století se začal objevovat v Australské angličtině i u deklarativních vět, s úmyslem udržet si posluchačovu pozornost. Tento trend, nazývaný v angličtině „uptalk“ či „high-rising terminal“ (HRT), se velmi rychle rozšířil do dalších zemí anglofonního světa, typický je pro mladé ženy. Kromě tohoto nešvaru, jak je fenomén často nazýván, se stoupavý melodém běžně pojí i s pocitem nejistoty na straně mluvčího. Melodém neukončující (v angličtině level) je melodém, který vzniká, pokud hlasivky vibrují na víceméně stálé frekvenci (tj., ve výšce tónu nejsou rozdíly větší než zhruba 2 půltóny, přestože definice tohoto melodému z hlediska intonačního

rozpětí se v literatuře buď nevyskytuje, nebo je ono rozpětí definováno pouze nárazově jednotlivými autory pro účely jejich studie). Tento melodém může být realizován rovnou kadencí. Většinou se spojuje s nedokončenými výroky, například pokud chce mluvčí vyslat signál, že ještě bude pokračovat v promluvě. Tento typ melodému nachází proto své uplatnění především ve spontánní konverzaci Sekce 3.1. také rozebírá stoupavě-klesavý a klesavě-stoupavý melodém, které jsou z hlediska své struktury komplexnější než předchozí uvedené melodémy.

V kapitole 3.2. se práce zabývá Britskou angličtinou a jejími specifiky z hlediska intonace a intonačních pohybů. Britskou angličtinou se v této i v empirické části rozumí „Southern British English“ (SBE), neboli angličtina z jižní části Anglie, která jako standard nahradila předchozí takzvanou „Received Pronunciation“. Z hlediska suprasegmentálních jevů v jazyce nebyl tento dialekt porovnáván s ostatními britskými dialekty či s jinými varietami angličtiny tolik jako v případě segmentálních jevů. Jedni z mála autorů, kteří se tomuto srovnání věnovali, byli Grabe a Post (2002). Ve svých studiích zjistili, že nejčastěji používaným melodémem ve čteném zadání je klesavý melodém (71%), dokonce i ve chvíli, kdy se jednalo o intonační pohyb jinde než na konci vět. Sekce 3.3. navazuje představením Americké angličtiny, pro kterou opět neexistuje příliš mnoho popisu z hlediska suprasegmentálních jevů. Nedostatečné množství studií porovnávajících Britskou a Americkou angličtinu z hlediska suprasegmentálních jevů (především intonace) je v práci připisováno mimo jiné velmi rozdílným notačním systémům obou variet. Následující sekce 4 tyto systémy představuje a porovnává (str. 22).

Nejprve se zaměřuje na stručnou historii vývoje obou systémů zápisu intonace (Britské a Americké školy), poté popisuje rozdíly mezi nimi. Britská tradice vnímá intonaci více holisticky, tedy zapisuje intonaci pomocí kontur nebo ikon v textu, Americká škola vychází ze práce lingvisty Kennetha Pikea a jeho systému čtyř úrovní výšky tónu, který se později přetvořil do binárního systému vysokých a nízkých tónů (H a L) v auto-segmentální metrické teorii (nejrozšířenějším systémem tohoto druhu je systém ToBI). Práce nabízí příklady frází, na kterých jsou uplatněny oba přístupy pro jejich vzájemné porovnání. Rozdílné zvyklosti v zápisu intonace ovšem nejsou jediným problematickým bodem v komparativní analýze variet angličtiny: dalším kamenem úrazu je nesjednocená terminologie v analýze intonace. Mnoho termínů se liší v jednotlivých tradicích (Br x Am), ale často se v užívání termínů neshodnou ani jednotliví autoři – používají jeden termín pro rozdílné jevy, nebo pokud jimi nazvou stejné jevy,

mohou se lišit v rozsahu onoho fenoménu, který termín pokrývá. Chaos v terminologii může mařit nebo minimálně ztěžovat potenciální snahy o komparativní analýzy a srovnání jazyků či jejich variet.

Sekce 4.1. navazuje na nesjednocenou terminologii a dodává, že v anglické terminologii chybí ekvivalent českého termínu „melodém“. Tento termín je tedy zaveden pod zkratkou FMU (Functional melodic unit) a použit pro pět základních melodémů, které se v práci vyskytují: stoupavý, klesavý, klesavě-stoupavý, stoupavě-klesavý a neukončující. Současně jsou představeny dvě nové kadence, „level-fall“ a „level-rise“. Tyto kadence mají trochu jiný průběh: intonační pohyb (konkrétní realizace), který začíná na slabice nesoucí větný přízvuk, je neukončující, tedy neklesá ani nestoupá. Ovšem jeho celková „funkce“ je stoupavá (část fráze před větným přízvukem je intonačně položená níže). Angličtina toto nerozlišuje a považuje takový průběh za rovnou kadenci. Proto práce vysvětluje, že by bylo výhodné mít k dispozici termíny pro rozlišení prostého neukončujícího melodému, oproti stoupavému/klesavému melodému realizovanému pouze neukončující, rovnou kadencí. Zároveň uvádí důvody, proč je výhodné terminologii sjednotit, respektive vytvořit k existujícím českým termínům jejich anglické ekvivalenty: takový krok nabízí detailnější přehled o typech a distribuci intonačních pohybů v angličtině, také usnadňuje komparativní studie díky sjednocené terminologii, také může sloužit ku prospěchu lidem, kteří se angličtinu učí.

Sekce 5. je první kapitolou empirické části práce. Uvádí výzkumné otázky a hypotézy. Cílem výzkumu bylo zjistit, zda a případně jak se britská a americká varieta angličtiny od sebe liší v oblasti délky fráze, intonačního pohybu na konci fráze a intonačního rozpětí mluvčích. Materiál a metoda jsou popsány v části 6. Pro vypracování výzkumu bylo použito 800 prozodických frází od celkem 16 mluvčích (8x50 Br frází a 8x50 Am frází), kteří se účastnili politických debat. Tyto debaty byly staženy z volně dostupných internetových zdrojů. Metoda popisuje postup při výběru mluvčích, způsob manuální segmentace frází v programu Praat a následné určení typu intonačního pohybu na konci fráze. Následně byl použit v Praatu skript pro získání potřebných dat.

Sekce 7 poskytuje výsledky výzkumu ve formě tabulek, grafů a stručných analýz. Nejprve jsou v podkapitole 7.1. rozebrány výsledky, které se týkají délky frází. Bylo zjištěno, že obě porovnávané variety angličtiny jsou z tohoto hlediska téměř identické (britská fráze obsahuje průměrně 6,3 slabiky, americká 6,6.) Též tempo je srovnatelné, lehce vyšší v americké angličtině (5,4 slabik/s v britské angličtině, 5,7 slabik/s v americké angličtině). Sekce 7.2. se

stáčí k typu intonačního pohybu na konci frází. Bylo zjištěno, že nejčastěji byl použit klesavý melodém (tvořil téměř 50 % v každé varietě). Takový výsledek byl překvapivý vzhledem k povaze materiálu – spontánní řeč, která je bohatá na nedokončené věty, zaváhání, případný překryv mluvčích, apod. Tyto výsledky se ovšem shodují s poznatky Sophie Herment (2021), jejíž studie ovšem byla provedena na čteném materiálu. Druhým nejčastějším melodémem je stoupavý v americké angličtině (15,5 %), klesavě-stoupavý v britské angličtině (18,5 %). Neukončující melodém byl očekáván ve větším zastoupení vzhledem k povaze materiálu, vyskytl se ovšem pouze ve 14,5 % u britských mluvčích a v 12,3 % u Amerických mluvčích. Nově zavedená kadence „level-fall“ byla častější v americké angličtině: tvořila 25,6 % neukončujícího melodému, oproti 7,7 % neukončujícího melodému v britské angličtině. Z hlediska podílu, který tvoří v rámci klesavého melodému, je to 9 % v americké angličtině a 3 % v britské angličtině. Level-rise tvořil 16,8 % neukončujícího melodému v britské angličtině a 11,5 % v americké angličtině. Z hlediska podílu v rámci stoupavého melodému to bylo 12,7 % v americké angličtině a 24,5 % v britské angličtině.

Nejvíce se obě variety lišily ve frekvenci užití komplexních melodémů: stoupavě-klesavého a klesavě-stoupavého. Stoupavě-klesavý se se více vyskytoval v americké angličtině (12 % všech frází oproti 4 % všech britských frází). Klesavě-stoupavý melodém se vyskytoval více naopak v britské angličtině, ve které tvořil 18,5 % všech melodémů. Američtí mluvčí ho využili v pouhých 2,5 % všech případů.

Intonační rozpětí se ukázalo být větší v britské angličtině (v průměru o 1,2 půltónu). Konkrétně využili britští mluvčí větší intonační rozpětí v klesavém, stoupavém a klesavě-stoupavém melodému. Američtí mluvčí měli větší rozpětí než britští mluvčí pouze v případě stoupavě-klesavého melodému. Podkapitoly v rámci sekce 7.4. nabízí srovnání jednotlivých intonačních pohybů v obou varietách i u jednotlivých mluvčích v podobě grafů a komentářů. Sekce 8 a 9 shrnují poznatky a uzavírají celou práci doporučeními a implikacemi pro další výzkum v podobné oblasti.

12. Appendix

The appendix offers two maps with marked locations where the speakers come from. The first picture shows the map of British speakers, the second picture provides a map of the American speakers. Due to the limited space, the map of SBE speakers is complemented with a caption explaining under what number the speakers can be identified in the map.



a map of the British speakers

Speakers are coded in numbers:

- | | |
|----------|----------|
| 1= BR-F1 | 5= BR-M3 |
| 2= BR-M4 | 6= BR-M2 |
| 3= BR-M1 | 7= BR-F2 |
| 4= BR-F4 | 8= BR-F3 |



a map of the American speakers' origin