2022 ACTA UNIVERSITATIS CAROLINAE PAG. 97–113
PHILOLOGICA 1 / PHONETICA PRAGENSIA

DESCRIPTION OF F0 CONTOURS WITH LEGENDRE POLYNOMIALS

MICHAELA SVATOŠOVÁ, JAN VOLÍN

ABSTRACT

Phonetic research has developed both impressionistic and more objective means of describing the basic units of intonation. The quantification involved in the approaches based on acoustic measurements provides more detail and it is a necessary prerequisite for the comparability and replicability of the results of different studies. In addition to having these characteristics, a proper description of intonation should be comprehensible and meaningful. This article presents a method for describing melodic contours using Legendre polynomials, which yields a few coefficients that capture the basic properties of the analysed contour (e.g. level or slope). This approach thus connects objectivity and quantitative precision with common linguistic concepts. The article also proposes the use of Legendre polynomials for the description of traditionally recognized Czech *melodemes* through the analysis of schemes reported in the literature. Further research on real material could verify the validity of these categories and the usefulness of the method itself.

Key words: intonation, fundamental frequency, Legendre polynomials, polynomial modelling

1. Introduction

Despite differing substantially, various models of intonation share a common goal. They aim to simplify the enormous variability of melodic contours produced by speakers into a limited number of perceptually distinctive categories. This effort involves two tasks – identifying the relevant categories and characterising them appropriately in phonetic terms. The first accounts of intonation were based on careful listening, which is an accessible method that considers perceptually relevant changes in F0. Nevertheless, the impressionistic descriptions formulated as verbal labels (e.g. *fall*, *rise-plateau*) or autosegmental labels (e.g. H*, L+H*) suffer from subjectivity and vagueness.

When instrumental measurements of fundamental frequency became available, new methods emerged that attempted to quantify melodic patterns objectively, e.g. by stating the size of a melodic step in semitones. Experiments have shown that listeners perceive intonation only in the central parts of vowels (Hermes, 2006: 13–15), allowing for one value to represent the pitch of a short syllable. The reduction of a contour into

https://doi.org/10.14712/24646830.2022.31

a set of points connected by lines was licensed by the close-copy stylization approach ('t Hart et al., 1990). On the other hand, the models produced by Fujisaki (1983), Taylor (1994) or Hirst et al. (2000) used complex equations in order to reconstruct F0 contours. Previous studies have also approximated F0 contours with polynomial equations (Andruski & Costello, 2004; Volín & Bořil, 2014). These approaches usually model the original contours more accurately at the expense of interpretability. A compromise is therefore sought that adequately captures the data but remains easily understandable.

This article presents a method for the description of melodic patterns with Legendre polynomials, which provide a quantification that is more linguistically meaningful than conventional polynomial approximations. This approach was already used for the analysis of British nuclear tones (Grabe et al., 2007) and it was applied also to German (de Ruiter, 2011) and Czech (Volín et al., 2017). Furthermore, the research on Czech has exploited Legendre coefficients in the field of automatic speech processing, where they were shown to effectively parameterize the nuclear patterns and improve the prosody of the TTS synthesis (Matura & Jůzová, 2018). Section 2.1 introduces the first four Legendre polynomials (as a subset of the whole Legendre polynomial family) and their coefficients, which relate them to complex curves. The practical steps constituting the process of obtaining the coefficients from the F0 contour are outlined in Section 2.2. The meaning of the coefficients is discussed using real examples in Section 2.3. The following Section 3 then suggests the application of Legendre coefficients in the description of Czech nuclear contours. This demonstration with schematic patterns could serve as a starting point for other studies that could test this approach on real material.

2. Legendre polynomials

2.1 Mathematical basis and modelling of curves

Polynomials are mathematical functions that can be used for describing curves. Legendre polynomials are named after the French mathematician Adrien-Marie Legendre, who discovered them in 1782. They are defined in the interval $[-1, 1]$ and normalized to $L_n(1) = 1$. The first four polynomials are shown in Figure 1 (all figures in this article were created with the R packages *tidyverse*, *grid* and *gridExtra* (Auguie, 2017; R Core Team, 2022; Wickham et al., 2019)). The equation of the Legendre polynomial of the n-th degree can be derived from the general formula given in (1). The degree of a polynomial expresses how many changes it can describe. The first polynomial in Figure 1 is constant and it has a degree of 0. The rising line of L_1 already captures one change (from low to high values), which corresponds to the first degree, while the subsequent polynomials of higher degrees turn their directions more times.

(1) $$L_n(x) = 2^n \sum_{k=0}^{n} x^k \binom{n}{k} \binom{\frac{n+k-1}{2}}{n}$$

These polynomials represent basic building blocks for creating more complex curves. Combining pairs of polynomials (by adding together values that correspond to each other on the x-axis) yields curves shown in Figure 2. The panel a) illustrates the combination

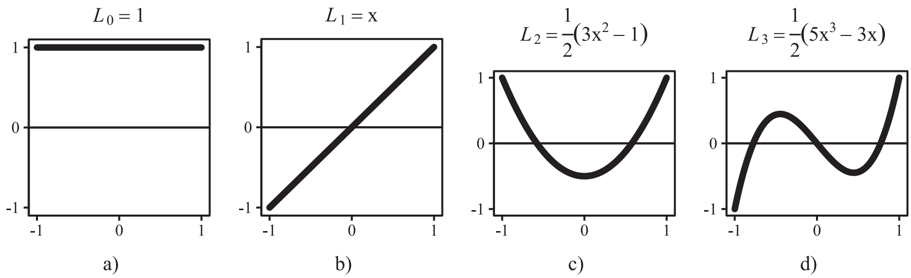

Figure 1. The first four Legendre polynomials (L_0–L_3) with their equations.

of the first two polynomials. While the values of L_1 in its basic form range from −1 to 1 (on the y-axis), the addition of L_0 makes them range from 0 to 2, because L_0 has a constant value of 1 in the whole interval. Adding L_0 to any other polynomial would also shift the given polynomial on the y-axis, but its shape would remain the same. The curve in the second panel retains the cup-shape of L_2, but it also has a clear rising tendency overall due to the presence of L_1. In the panel c), the direct rise of L_1 is modified by the wave shape of L_3. Finally, the last panel d) shows that summing the nearly opposite values in the first halves of L_2 and L_3 produces values around zero, while their similarly rising shape toward the end results in a more prominent rise.

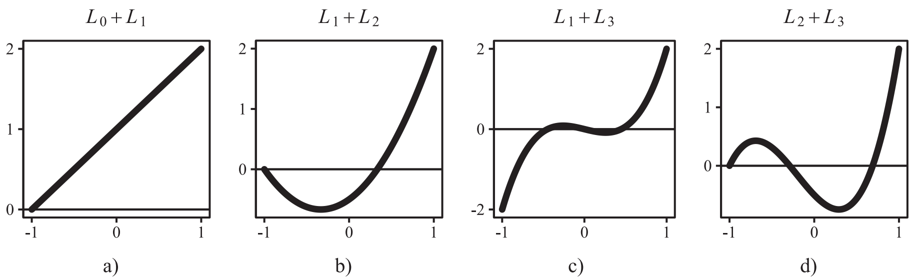

Figure 2. Combinations of pairs of Legendre polynomials.

Each polynomial L_n can be multiplied by the coefficient c_n. The basic polynomials in Figure 1 are not accompanied by any number, which implicitly refers to $c_n = 1$. A different value of c_0 simply makes L_0 represent a different constant, as shown in the panel a) of Figure 3, where $c_0 = 1.8$. The other coefficients affect the span of their respective polynomials. Lowering c_1 to 0.5 produces halved values across the whole interval of L_1, as illustrated in the second panel. Multiplying the polynomials by negative numbers creates curves that are mirror-shape images (according to the x-axis) of their counterparts with positive coefficients. Negative values of c_2 therefore lead to dome-shaped curves instead of the cup-shaped ones that were presented so far. This is shown in the panel c), where the coefficient is not only negative, but also of higher absolute value ($c_2 = -2$, compared to $c_2 = 1$ in Figure 1), which is reflected in the wider range of its values. Similarly, a negative c_1 would produce a falling line and a negative c_3 corresponds to a falling-rising-falling shape.

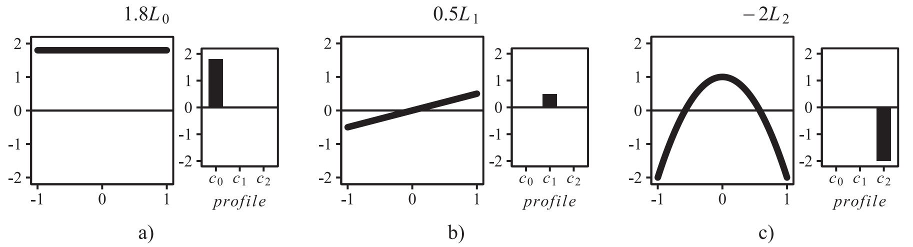

Figure 3. The first three Legendre polynomials multiplied by different coefficients. The *profiles* (on the right side of each panel) show the coefficient values.

The two variations just presented imply the possibility to multiply each basic polynomial with a specific coefficient and add them together. Since the precise values of individual coefficients are no longer easily recognizable from the complex curve, they can be summarized in a *profile* accompanying the curve, as will be done in the rest of the figures in this article. Figure 3 contains three simple *profiles* that graphically depict the respective coefficients. The coefficients of polynomials that are not part of a given curve equal zero. It is therefore sufficient if the *profile* includes coefficients from c_0 up to the last coefficient with a nonzero value. Figure 4 illustrates some of the curves that can be modelled using only two polynomials (L_1 and L_2), but in different ratios. The panel a) starts with a simple fall that corresponds to the polynomial L_1 multiplied by a negative coefficient ($c_1 = -1$). In addition to L_1, the curves in the following panels also include the polynomial L_2 modified by negative values of the coefficient c_2 (these produce dome-shaped curves as in the panel c) of Figure 3). As the relative magnitude of c_2 compared to c_1 gradually increases in panels a) – e), the falling L_1 transforms into the dome-shaped L_2. The curves in panels f) – i) contain a positive c_1, making the overall slope rising. Mirror-shaped images of the first eight curves could be modelled using opposite values of c_1 and c_2 (as indicated in the last four panels), forming a transition from a rise through a cup-shaped parabola to a fall.

Figure 4 shows that curve shapes are determined by the ratios between L_1 and L_2 (and possibly other higher coefficients). In order to compare various curve shapes, relative coefficients (rc_n) can be calculated from the raw ones (c_n) by the formula given in (2), where N stands for the highest degree of a polynomial with a nonzero coefficient. It transforms all coefficients with nonzero values to make the sum of their absolute values equal 1. The first coefficient (c_0) is excluded from this conversion, because it shifts the curve on the y-axis, but it is not related to its shape. The relative coefficients do still express the positive or negative orientation of their respective polynomials. However, they also indicate to what extent do the individual polynomials contribute to the overall shape of the modelled curve. The *profiles* in Figure 4 actually present relative coefficients (the sum of their absolute values equals 1).

(2) $$rc_n = \frac{c_n}{\sum_{i=1}^{N} |c_i|}$$

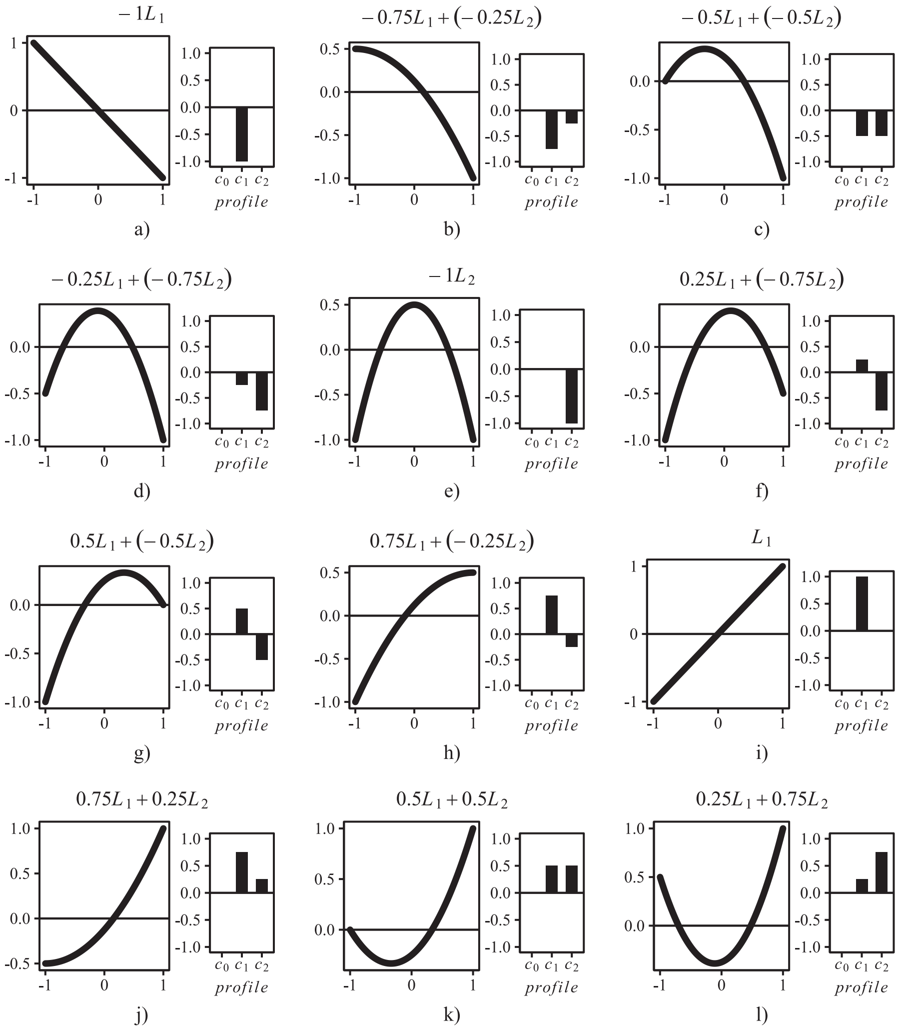

Figure 4. Curve shapes produced by combinations of the second (L_1) and third (L_2) Legendre polynomial. Each polynomial is multiplied by a different coefficient (their values are shown in the *profiles* on the right side of each panel).

An important feature of Legendre polynomials is their orthogonality. In mathematical terms it means that the inner product of each two polynomials equals zero. It refers to the fact that each polynomial captures a property that is not explainable by any other polynomial. The average value of a given curve can be described only using L_0, since all the other polynomials have an average of zero. Similarly, the linear slope is only reflected in L_1, because the linear regression of the other polynomials equals zero. The polynomials of higher degrees describe unique characteristics in the same manner. Thanks to orthogonality, any curve in the [−1, 1] interval can be formed by summing polynomials of different degrees, each multiplied by a specific coefficient, although the modelling of more complex curves (containing numerous or abrupt changes) requires using polynomials of higher degrees. From the opposite perspective, it is possible to decompose any curve (e.g. an interpolated F0 contour) into a set of Legendre polynomials multiplied by different coefficients (having different "amplitudes"). This is analogical to Fourier analysis of a sound wave, which works with cosine functions of various frequencies instead of Legendre polynomials. Identifying the values of these coefficients lies at the core of the analysis presented in Section 2.2. Since the inner product of any pair of Legendre polynomials equals zero, the coefficient of each polynomial can be calculated from the inner product of that polynomial and the analysed curve. Another consequence of orthogonality is the independence of the coefficients, which means they are not correlated and can be statistically evaluated as separate variables within one analysis.

The first few polynomials are sufficient for the analysis of intonation, because they capture the main melodic movements and ignore microprosodic effects. In other words, they model the relatively simple underlying shape of the F0 contour. Specifically, in nuclear patterns that usually span over a few syllables, a higher number of changes is not expected. For convenience, the first four coefficients are referred to as AVERAGE (c_0), SLOPE (c_1), PARABOLA (c_2) and WAVE (c_3) following Grabe et al. (2007). Their further advantage over other methods of polynomial modelling (e.g. least squares approximation) is that they can be interpreted in linguistic terms. As mentioned earlier, the AVERAGE has a special position, because it expresses the mean value of the curve, while all the other coefficients are related to its shape. The meaning of AVERAGE and units in which it is expressed depend on the method of normalization of the original F0 contour (described in detail in Section 2.2), but it is connected to the position of the nuclear pattern in the pitch range. The relative ratios of SLOPE, PARABOLA and WAVE affect the shape of the curve (as illustrated in Figure 4), while the absolute values of these coefficients reflect its span (compressed or expanded). A feature that is not inherently captured by this method is the temporal dimension, since each contour needs to be transformed to the interval $[-1, 1]$.

2.2 Analysis of F0 contours

This section describes the necessary steps that have to be undertaken in order to obtain the Legendre coefficients of a given melodic contour. Some of them are common in intonation research, while others are required specifically by this type of analysis.

First of all, the analysis domain has to be chosen. Stress-groups usually consist of a few syllables bearing relatively simple melodic movements. These can be adequately captured by a few coefficients and therefore seem as a reasonable choice. The decision about the appropriate domain should be guided by the research question and by the findings from the previous research on the given language. For example, it might be useful to include the pre-stressed syllable, if its relative pitch plays a role in distinguishing various patterns, as was done for German in de Ruiter (2011). However, longer units such as prosodic phrases could be modelled as well, if the differences in interpretation are taken into account. One caveat arising from this approach might be the diversity of attested patterns. The range of possibly distinct contours expands with every additional syllable of the analysed domain, which also affects the amount of data necessary for their adequate description.

The present analysis is based on the measurement of F0 alone. It differs in this respect from the method utilized by Grabe et al. (2007), which additionally required the parameters of intensity and periodicity. These were used for weighting the importance of particular parts of the F0 contour. The idea underlying their approach derives from the finding that listeners do not pay equal attention to variations of pitch in different segments. More sonorant phones like vowels appear to form the basis of the perceived melodic patterns, while the pitch in voiced obstruents is ignored with regard to intonation (Hermes, 2006: 13–15). The method described here proposes an alternative way of reflecting this knowledge through the exclusion of all F0 values in the irrelevant regions from the analysis. Although it provides only a categorical distinction (values are either used or deleted) instead of a gradual scale, this approach avoids further errors that are related to the measurements of other parameters.

Since listeners' sensitivity to pitch variations seems to be language-specific, the choice of particular parts of the F0 contour for the analysis should be theoretically grounded. Generally, the F0 values would be retained in vowels and discarded in consonants, although some languages might exploit also the regions occupied by sonorants. This elementary distinction could be further refined to eliminate some of the microprosodic effects. This means extracting only certain parts of each vowel, defined either absolutely (e.g. starting 10 ms after the beginning of the vowel) or relatively (e.g. using the middle third of its duration). Comparing both approaches might show if the simpler method is robust enough or whether the further adjustments need to be made. It is obvious that in any case, the F0 contour should be annotated at the level of segments.

From the practical point of view, the F0 contour in the domain of interest has to be extracted and corrected for errors like octave jumps and missing values, which is a standard procedure for studies concerning pitch. The values calculated in Hertz are then commonly converted to semitones (ST), which applies also for the present analysis, because this unit is perceptually more relevant. Semitones are used (instead of octave ratios as in Grabe et al., 2007) for two main reasons – their values are easier to interpret and they are well known. Nevertheless, the two units are mutually convertible (for comparative reasons) by simply dividing the values in semitones by 12 and vice versa, which holds for the coefficients as well.

Furthermore, the contours can be normalized to allow comparisons between speakers and utterances. The interpretation of AVERAGE (the first coefficient) follows directly from the chosen reference. Subtracting the mean F0 of each speaker implies that the coefficient is to be understood in relation to it. For example, if the contours of a speaker are expressed in semitones with the reference of 100 Hz (his mean F0), then the AVERAGE of 1.5 corresponds to 109 Hz, which is 1.5 ST above 100 Hz. Depending on the research question, an alternative reference for the normalization might be chosen (e.g. the mean F0 of the given utterance).

Finally, the coefficients are calculated using the method implemented in the *rPraat* package for the R software (Bořil & Skarnitzl, 2016; R Core Team, 2022). As already mentioned, the procedure differs from the one described in Grabe et al. (2007), but it derives from the orthogonality of Legendre polynomials. The algorithm takes the adapted F0 contour (normalized and including only the relevant values) and performs the following operations on it. First, the contour is linearly interpolated into 1000 points to ensure equivalent sensitivity in the whole analysed domain. Secondly, the time scale is transformed into the interval [−1, 1] in which Legendre polynomials are defined. This means that the coefficients alone are unable to capture the stretching or compressing of an identically shaped contour in the temporal dimension. Lastly, the coefficients are calculated from the inner products of the transformed contour and the respective polynomials.

The outlined method is summarized in these four steps:

1. selection of the analysis domain and its relevant parts
2. extraction of the F0 contour (in semitones)
3. normalization
4. calculation of the coefficients

An example of a practical application of this procedure is provided here, using a short polarity question [ˈr̝ɛknɛtɛ jɪm ˈt͡so sɪ ˈmɪsliːtɛ] (“Will you tell them what you think?”) produced by six speakers. The nuclear contour spanning the last stress-group [ˈmɪsliːtɛ] was chosen as the analysis domain, limiting the relevant parts only to vowels, which form the nuclei of the three syllables. F0 estimates in 10 ms intervals were extracted in Praat (Boersma & Weenink, 2022) using the autocorrelation method with standard settings and then manually corrected for octave jumps and missed voicing regions. All values in Hertz were converted to semitones with the reference of 1 Hz. The speakers’ means (obtained from a collection of their utterances) were then subtracted from the respective contours.

Figure 5 shows all six nuclear contours. The black points represent the extracted F0 values, while the interpolated values (also used for the analysis) are coloured in grey. The Legendre coefficients (c_0–c_3 in this case) are presented in the *profiles* on the right side of each panel. The black curves are models based solely on these four coefficients. Their values are summarised in Table 1 together with their relative counterparts and explained in the next section.

2.3 Interpreting the coefficients

So far, only coefficients relating to isolated curves were discussed. However, working with real data usually requires comparing multiple contours. This section therefore provides a more detailed description of the relationship between Legendre coefficients and the curves they represent. It also explains how similarities of contour shapes translate into the coefficient values.

The speaker S1 realized the analysed nuclear pattern as a *rise* (with the main melodic step located between the first and the second syllable), as shown in the top left panel of Figure 5. The accompanying *profile* reveals that its most prominent coefficient is the positive SLOPE, followed by the negative PARABOLA, which reflect the rising and dome-shaped appearance (see Figure 3 above for the individual polynomials and Figure 4 for their combination). The minor AVERAGE indicates that the whole contour is located 0.1 ST below the speaker’s mean pitch.

The coefficients of S1’s contour can be compared to those of another speaker, S2. Turning to both top panels of Figure 5, it can be seen that the extracted F0 values

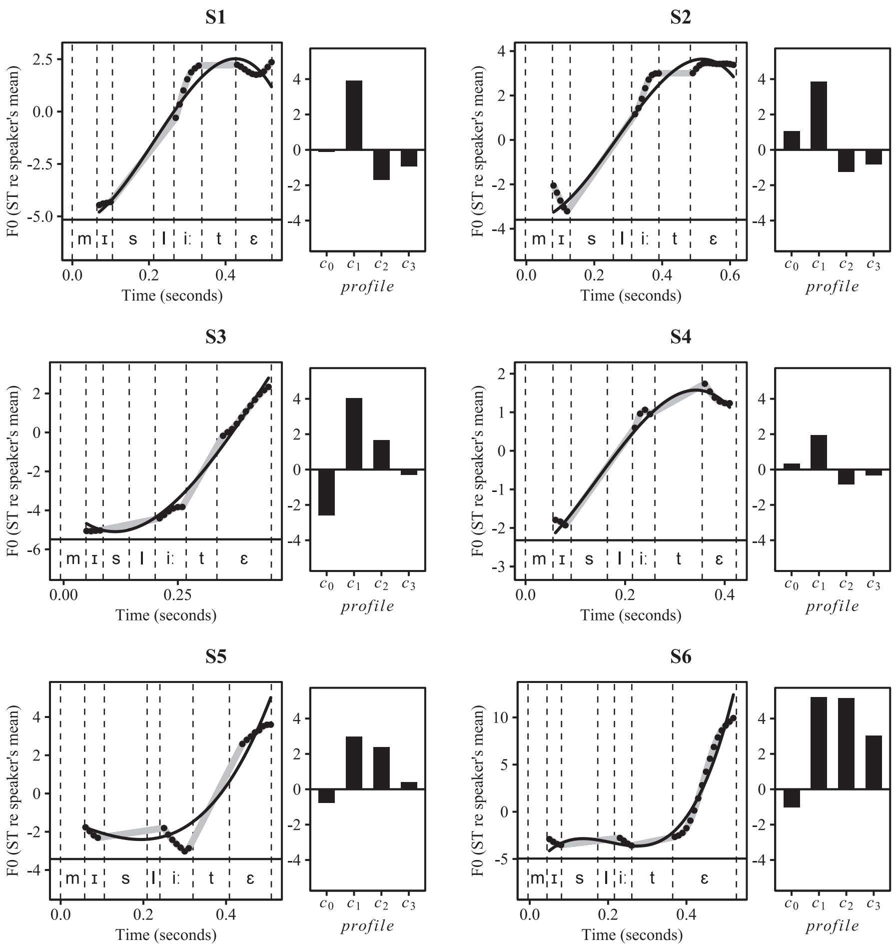

Figure 5. The nuclear contours on the stress-group [ˈmɪsliːtɛ] produced by six speakers. Each panel includes the extracted F0 values (black points), the interpolated values (in grey) and a curve constructed from the first four Legendre coefficients (these are shown in the *profiles* on the right side of each panel).

resemble each other a lot. Conveniently, the similarity is preserved in the coefficients of these contours. Excluding the AVERAGE that is not related to the shape, but rather signals the position of the nuclear pattern in the pitch range (1.1 ST above S2's mean pitch), the coefficients do not differ from each other by more than 0.4. In contrast, S3 produced a pattern that could be called a *late rise*, realizing the main melodic step between the second and the third syllable. It results in a considerable change in PARABOLA, which switches to a positive value. It captures the cup-shaped pattern that is present in the contour, although reduced by the more prominent SLOPE. It also becomes clear that the coefficients should not be interpreted in isolation. The similar SLOPE of the first three contours does not imply their resemblance in the overall shape. Nevertheless, it still holds that they all include overall rising, which is exactly what SLOPE represents. However, the type of rising is specified only in combination with higher coefficients.

The contour of S3 also features a lower AVERAGE. While this coefficient is autonomous in the sense that it merely distinguishes between realizations of an identical shape on different levels in the pitch range, it also interacts with the shape in one respect. The first and third contour begin and end with comparable F0 values of approximately −4.5 and 2 ST, yet their values of AVERAGE differ by 2.5 ST. The reason is the shape of the contour (here specifically the position of the second syllable), because all interpolated points in the analysed interval contribute equally to the value of AVERAGE. The AVERAGE will always be lower for contours with a low middle part than for those with a high one, even though they have the same values on the edges. It remains to be tested whether it reflects the fact that listeners perceive more low or high values in the whole contour. However, this effect should be taken into account in the interpretation.

Table 1. The Legendre coefficients (raw on the left, relative in italics on the right) of the six nuclear contours from Figure 5.

	AVERAGE	SLOPE	PARABOLA	WAVE	*SLOPE*	*PARABOLA*	*WAVE*
speaker	c_0	c_1	c_2	c_3	rc_1	rc_2	rc_3
S1	−0.1	3.9	−1.7	−0.9	0.60	−0.26	−0.14
S2	1.1	3.9	−1.3	−0.8	0.65	−0.21	−0.14
S3	−2.6	4.0	1.7	−0.3	0.67	0.28	−0.05
S4	0.3	1.9	−0.8	−0.3	0.63	−0.27	−0.10
S5	−0.8	3.0	2.4	0.4	0.52	0.41	0.07
S6	−1.0	5.2	5.2	3.1	0.39	0.38	0.23

Despite the fact that the contour produced by S4 follows the same pattern as the first two, its coefficients (especially SLOPE) substantially differ. This is due to the narrower span it covers, because the *rise* stretches only around 3.5 ST compared to approximately 6.5 ST in the previous contours. The relative coefficients in Table 1 clarify the similarity of S4's contour to those of S1 and S2. It lies in the approximately 63% share of positive SLOPE and 23% share of negative PARABOLA. The inclination of S3's contour to the shape of a *late rise* is therefore mainly caused by the positive PARABOLA, although it is present in a similar ratio. The relative coefficients also show that SLOPE forms the most important component in these contours, leaving only half as much space to PARABOLA. Notice that the ratios between coefficients can be to a certain extent visually assessed from the *profiles* alone, even if their *y*-axes have the same range and the raw coefficients differ in magnitude.

S5 realised a contour with a similar appearance to the one produced by S3 (a *late rise*), which is reflected in the comparable relative coefficients of these two contours. The prominence of the SLOPE is reduced in S5's contour (although it is still the strongest component), while the PARABOLA has a greater share than is S3's pattern. This change in the ratios of both coefficients relates to the position of the second syllable. The third panel resembles a straight rising line, while the fifth is more bent (compare with panels j) and k) in Figure 4). An important feature for the differentiation of curve shapes is the polarity of the most prominent coefficients, which applies for SLOPE and PARABOLA here. The contrast of positive and negative WAVE does not affect the shapes dramatically, because its relative values are close to zero.

The last panel of Figure 5 shows another *late rise*. However, the speaker S6 did not produce the melodic step between the second and third syllable (as did S3 and S5), but compressed this movement into the final syllable, which starts at a low pitch and ends high. The modelling of this abrupt change requires the presence of WAVE. Its relative value is three times higher for S6 than for S5 and at the same time the highest of all the contours in Figure 5. Besides allowing for the steep rise, the WAVE also makes the first half relatively flat. Without it, the combination of SLOPE and PARABOLA would result in a fall in that part of the curve (as in the panel k) in Figure 4). In fact, any curve with a steeper or sharper shape than those in Figure 4 necessarily includes WAVE (or other higher coefficients) in a non-negligible ratio. The same tendency can be observed even in the contours of S1 and S2. Their values of relative SLOPE and PARABOLA lie somewhere between those from panels g) and h) in Figure 4, but they rise in a straighter manner due to 14% of negative WAVE.

3. Modelling Czech nuclear patterns with Legendre polynomials

3.1 Connecting Legendre polynomials to linguistic categories

Previous sections have described the method which allows for the simplification of a F0 contour into a few Legendre coefficients that capture its basic properties. Nevertheless, these individual numbers do not represent the ultimate goal of phonetic research. Returning to the introduction, the aim of the analysis is to describe intonation patterns generally, which involves classification into various categories. Each category includes a range of possible realizations, while remaining distinct from other categories in various ways. The difference between declarative and interrogative utterances is a commonly mentioned one, but contours of different types can among other things also signal the speaker's dialect, thus expressing the indexical function.

The most thoroughly studied unit within the intonation of Czech is the nuclear pattern (*melodeme*), which is considered the most information-laden. Three functionally distinct categories are distinguished – conclusive, interrogative and continuative patterns (Daneš, 1957: 38-54; Palková, 1994: 307–315), each consisting of several contour types (*cadences*). Their traditional descriptions were based on auditory assessment, resulting in schematic stylizations using four pitch levels (Daneš, 1957: 53–54). It would be desirable to experimentally test their perceptual validity. However, the controlled design of test items requires quantitative characteristics of these types. Legendre coefficients might serve as a convenient tool in this respect. A given category can be described with a model contour created from the average coefficients of the contours belonging to that category. For illustration purposes, the following sections present an experiment indicating approximate coefficient values which could be expected for some of the traditionally reported contour types.

3.2 Method

The three categories of *melodemes* were represented by schematic patterns taken from Palková (1994: 309–315) and spanned over a three-syllabic stress-group, which allows for a greater variability of contour shapes. The patterns are summarised in Table 2. In order to analyse them with Legendre polynomials, F0 contours based on these schemes were created in Praat. The four levels were set to 305, 265, 230 and 200 Hz, which yielded equidistant pitch levels after the conversion to semitones (approximately 2.4 ST apart). The values of the highest and lowest level were chosen to reflect a possible human pitch range, which produces coefficient values comparable in magnitude to those that could be obtained from real recordings. The target values were placed in the centres of vowels in a simple CVCVCV template (assuming the same duration for all segments) and then interpolated quadratically with the built-in function in Praat. The edge values in the first and last vowel were adjusted to produce a mean F0 (in these vowels) equivalent to the desired levels. The analysis domain thus corresponded to a stress-group and the relevant parts used for the analysis were limited to the vowels.

Table 2. Schemes of Czech nuclear patterns based on Palková (1994: 309–315). Number 1 represents the highest pitch level, number 4 the lowest; * marks the stressed nuclear syllable and the number in brackets denotes the level of the pre-nuclear syllable.

	conclusive patterns (CCL)	interrogative patterns (INT)	continuative patterns (CNT)
1	(1) 2* 3 4	(2) 4* 4 2	(3) 4* 3 2
2	(2) 3* 2 4	(2) 4* 1 2	(4) 2* 1 2
3		(4) 1* 1 2	(4) 1* 1 3

The pitch of the pre-nuclear syllable was chosen as the reference value for normalization for two reasons. First of all, the mean F0 of a speaker or utterance could not be used, since the contours were created artificially in isolation. Secondly, the relative position of the stressed syllable and the preceding (pre-nuclear) syllable is argued to differentiate various patterns and therefore represents a relevant component of the whole contour (Daneš, 1957: 51). Legendre coefficients were calculated in *rPraat* with the method explained in detail at the end of Section 2.2. Their relative counterparts were obtained using the formula presented in Section 2.1.

3.3 Interpretation

Figure 6 illustrates all analysed contours. Similarly to Figure 5, it contains the original F0 values (black points) and the whole interpolated contours (grey lines). The black curves represent the models based on the first four Legendre coefficients, which are shown in the profiles on the right. Both raw and relative coefficients are summarised in Table 3. However, the specific values should not be taken literally, since a few arbitrary decisions had to be made when transforming the schematic representations into analysable F0 contours.

The first two contours are conclusive and they both have a negative SLOPE. It indicates an overall fall, which is a typical property of this category. However, they differ in some respects. The fall in CCL-2 is milder (it has a smaller absolute SLOPE) and it is complemented by a rising-falling element, which is captured by the negative PARABOLA. As can be seen from the relative coefficients, the parabolic shape in fact contributes to the whole contour to a greater extent than SLOPE. On the other hand, the first two interrogative contours are rising, since they both contain a positive SLOPE, although it is not the only polynomial present in them. Nevertheless, considering just the first four contours, the negative or positive SLOPE seems to differentiate between the conclusive and interrogative types.

Table 3. The Legendre coefficients (raw on the left, relative in italics on the right) of the contours from Figure 6. The sum of the absolute values of relative coefficients does not equal 1 in some rows due to rounding.

	AVERAGE	SLOPE	PARABOLA	WAVE	*SLOPE*	*PARABOLA*	*WAVE*
contour	c_0	c_1	c_2	c_3	rc_1	rc_2	rc_3
CCL-1	−4.9	−3.0	0.0	0.2	−0.94	0.00	0.06
CCL-2	−2.2	−1.5	−2.9	0.1	−0.33	−0.65	0.02
INT-1	−3.4	3.0	1.9	−0.2	0.59	0.37	−0.04
INT-2	−0.5	3.0	−3.9	−0.2	0.43	−0.55	−0.03
INT-3	6.6	−1.5	−0.9	0.1	−0.60	−0.37	0.04
CNT-1	0.0	3.0	0.0	−0.2	0.94	0.00	−0.06
CNT-2	5.8	0.0	−1.9	0.0	0.00	−1.00	0.00
CNT-3	5.8	−3.0	−1.9	0.2	−0.59	−0.37	0.04

The advantage of Legendre coefficients over verbal labels manifests itself in the comparison of CCL-2 with INT-2 and CNT-2. They could all be called *rise-falls* based on the relative positions of the three syllables, despite the fact that they are visually and perceptually distinct. A longer specification is then required to capture the different magnitudes of the melodic steps between syllables. The descriptions become much more concise when translated into Legendre coefficients. The most prominent element is the negative PARABOLA, corresponding to the rising-falling skeleton shared by all three

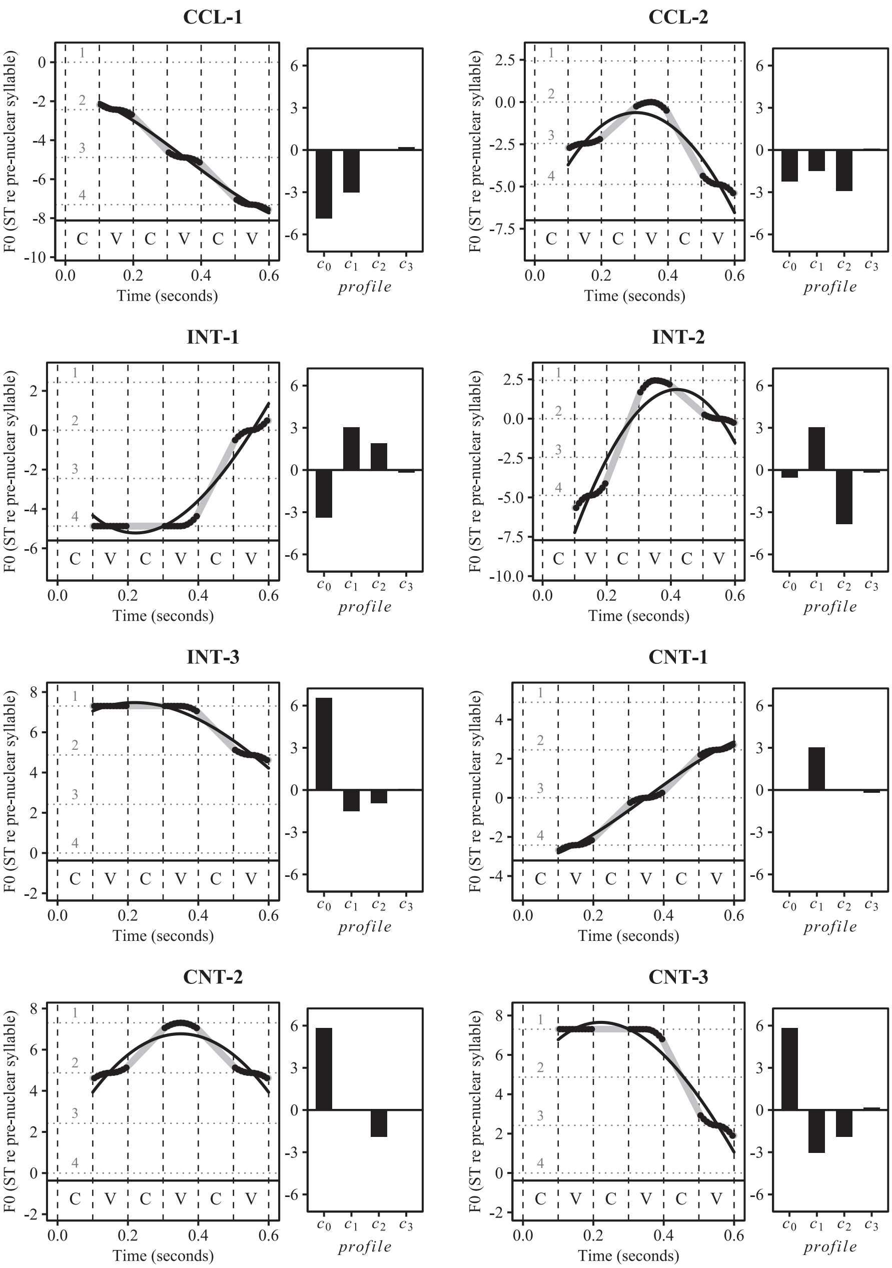

Figure 6. The modelled contours based on the schematic patterns from Table 2. Each panel includes the F0 values (black points), the interpolated values (in grey) and a curve constructed from the first four Legendre coefficients (these are shown in the *profiles* on the right side of each panel). The horizontal dotted lines indicate the four pitch levels.

contours. It is the only constitutive component of the model in the case of CNT-2, but it is complemented by SLOPE in the other two contours. These are distinguished by the values of SLOPE – CCL-2 has a falling tendency, while INT-2 is rising. The SLOPE is thus indirectly signalling the ratio between the two melodic steps.

The first interrogative contour resembles the second one in the rising aspect, but it has a positive value of PARABOLA, which on its own means a fall-rise. However, the relative coefficient shows that it amounts to approximately one third of the whole contour, while the SLOPE has a greater share. This combination lies halfway between the shapes j) and k) in Figure 4 and leads to a plateau (rather than a fall) between the first two syllables followed by a rise. The third interrogative contour diverges from the general pattern, since it is falling, although not prominently in absolute terms (small absolute value of SLOPE). This opposite tendency is compensated by a high AVERAGE that strongly contrasts with the negative or zero values present in the contours discussed so far. The specific values are arbitrary, because they result from the F0 levels chosen during the modelling of the contours, but the ratios between them hold true. The interpretation of AVERAGE depends on the current reference, which is the F0 level of the pre-nuclear syllable here. While the first four contours are located at about the same pitch or below the previous syllable, INT-3 lies higher. This might serve a similar function as the overall rising, since both strategies end at a relatively high pitch.

Interestingly, the same pattern is present in the continuative contours. The first one has a positive SLOPE, which is also dominant relatively. On the contrary, CNT-2 and CNT-3 contain a high AVERAGE, although their SLOPE is zero or even negative. In fact, CNT-3 closely mirrors the relative coefficients of INT-3. Figure 6 shows that the two shapes are alike, but the two contours differ in the level of the last syllable (see Table 2). In other words, CNT-3 covers a wider span. This difference is normalised in the relative coefficients, but retained in the raw coefficients, which are halved for INT-3. The level and span of nuclear patterns might play an important role for the listeners when distinguishing the categories. Although INT-3 and CNT-3 seem to differ only in the span, the present analysis is strongly limited by the four-level schematization and a proper description would require real data.

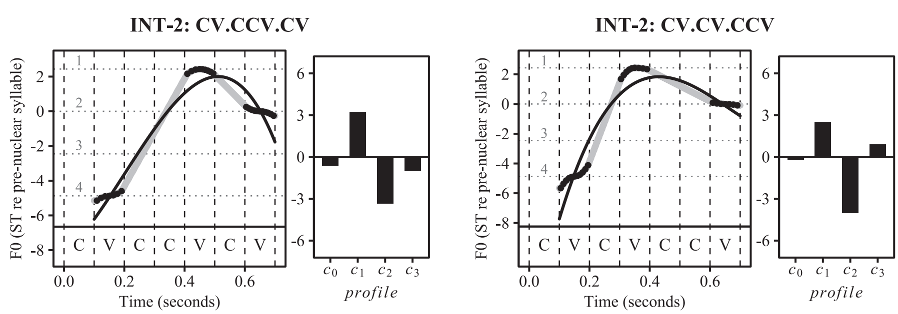

Figure 7. The modelled contours of the INT-2 pattern as combined with two different syllabic templates. Each panel includes the F0 values (black points), the interpolated values (in grey) and a curve constructed from the first four Legendre coefficients (these are shown in the *profiles* on the right side of each panel). The horizontal dotted lines indicate the four pitch levels.

Finally, it can be seen at first glance that WAVE is only marginally involved in all contours, probably due to the regular temporal distribution of the three melodic targets. For comparison, Figure 7 simulates the presence of a consonant cluster before or after the second vowel in contour INT-2. The relative position of the peak is therefore shifted towards the beginning or end of the contour. It leads to four times higher relative values of WAVE (−0.13 and 0.12, respectively) compared to the rc_3 of the original INT-2 contour. These account for the steeper falls at the edges of the modified contours.

4. Conclusion

The article presented a method for the description of F0 contours using Legendre polynomials. The main melodic movements are converted into a few (usually four) coefficients that are linguistically interpretable, while remaining quantitatively precise. They therefore combine the advantages of simple verbal labels and complex mathematical equations. The coefficients capture the elementary properties of the analysed contours and ignore microprosodic effects. Both dimensions of the pitch range are referred to in this approach, since the AVERAGE (the first Legendre coefficient, c_0) relates to the level and the absolute values of the other coefficients reflect the span. Different contour shapes can be easily compared using the relative coefficients, which inherently express the internal temporal distribution of the pitch targets in the analysed contour. However, the total duration of the analysed unit is not accounted for due to the normalization that is required for the calculations. Relating the coefficients to speech tempo thus remains one of the questions for further research.

Section 3 suggested the application of Legendre coefficients in the description of Czech nuclear patterns. However, it only outlined the procedure that should be repeated with natural material in order to explore the differences between the nuclear pattern categories and also the specifics of their subtypes. These studies could compare their results with those observed here for the traditional schemes and test the usefulness of Legendre coefficients in intonology. The following step could turn to the listener and examine the distinctiveness of contour subtypes in perception experiments. A potential systematic relationship between the perceived perceptual differences of F0 contours and the values of their Legendre coefficients would provide further evidence for the relevance of this method.

Acknowledgements

This work was supported by Czech Science Foundation (GAČR), project GA21-14758S.

REFERENCES

Andruski, J. E., & Costello, J. (2004). Using polynomial equations to model pitch contour shape in lexical tones: An example from Green Mong. *Journal of the International Phonetic Association*, *34*(2), 125–140.

Auguie, B. (2017). *gridExtra: Miscellaneous Functions for 'Grid' Graphics*. https://CRAN.R-project.org/package=gridExtra

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.2.07). https://www.praat.org/

Bořil, T., & Skarnitzl, R. (2016). Tools rPraat and mPraat. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech, and Dialogue* (pp. 367–374). Springer International Publishing.
Daneš, F. (1957). *Intonace a věta ve spisovné češtině*. Nakladatelství ČSAV.
de Ruiter, L. E. (2011). Polynomial Modeling of Child and Adult Intonation in German Spontaneous Speech. *Language & Speech*, *54*(2), 199–223.
Fujisaki, H. (1983). Dynamic Characteristics of Voice Fundamental Frequency in Speech and Singing. In P. F. MacNeilage (Ed.), *The Production of Speech* (pp. 39–55). Springer.
Grabe, E., Kochanski, G., & Coleman, J. (2007). Connecting Intonation Labels to Mathematical Descriptions of Fundamental Frequency. *Language & Speech*, *50*(3), 281–310.
Hermes, D. J. (2006). Stylization of Pitch Contours. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schließer (Eds.), *Methods in Empirical Prosody Research*. DE GRUYTER.
Hirst, D., Di Cristo, A., & Espesser, R. (2000). Levels of Representation and Levels of Analysis for the Description of Intonation Systems. In M. Horne (Ed.), *Prosody: Theory and Experiment* (Vol. 14, pp. 51–87). Springer Netherlands.
Matura, M., & Jůzová, M. (2018). Correction of Formal Prosodic Structures in Czech Corpora Using Legendre Polynomials. In A. Karpov, O. Jokisch, & R. Potapova (Eds.), *Speech and Computer* (pp. 387–397). Springer International Publishing.
Palková, Z. (1994). *Fonetika a fonologie češtiny: S obecným úvodem do problematiky oboru* (1. vydání). Karolinum.
R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/
’t Hart, J., Collier, R., & Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge University Press.
Taylor, P. (1994). The rise/fall/connection model of intonation. *Speech Communication*, *15*(1–2), 169–186.
Volín, J., & Bořil, T. (2014). General and speaker-specific properties of F0 contours in short utterances. *Acta Universitatis Carolinae – Philologica*, 9–19.
Volín, J., Tykalová, T., & Bořil, T. (2017). Stability of Prosodic Characteristics Across Age and Gender Groups. *Interspeech 2017*, 3902–3906.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686–1691.

Michaela Svatošová
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
E-mail: michaela.svatosova@atarien.com

Jan Volín
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
E-mail: jan.volin@ff.cuni.cz