

A dialectometric approach to Romanian intonation*

Paolo Roseano

UNED &
University of South Africa

Adrian Turculeț

University of Iași
Rumania

Anca-Diana Bibiri

University of Iași
Rumania

Ramon Cerdà Massó

Universitat de Barcelona
España

Ana M.^a Fernández Planas

Universitat de Barcelona
España

Wendy Elvira-García

Universitat de Barcelona
España

ONOMÁZEIN | Special Issue XI – Current Approaches to Romanian and Spanish Intonation: 105-139

DOI: 10.7764/onomazein.ne11.07

ISSN: 0718-5758



Paolo Roseano: Universidad Nacional de Educación a Distancia & University of South Africa.

| E-mail: paolo.roseano@flog.uned.es

Adrian Turculeț: Universitatea Alexandru Ioan Cuza din Iași, Rumania. | E-mail: aturcu@uaic.ro

Anca-Diana Bibiri: Department of Social Sciences and Humanities, Institute of Interdisciplinary Research, University of Iași. | E-mail: anca.bibiri@gmail.com

Ramon Cerdà Massó: Universitat de Barcelona, España. | E-mail: elmook17@gmail.com

Ana M.^a Fernández Planas: Universitat de Barcelona, España. | E-mail: anamariafernandez@ub.edu

Wendy Elvira-García: Universitat de Barcelona, España. | E-mail: wendyelvira@ub.edu

Received: March 2019

Accepted: October 2019

Abstract

This paper presents the first dialectometric analysis of intonational variation of standard spoken Romanian. The data analysed (26,680 sentences) were gathered within the AMPER-ROM project and the dialectometrical analysis has been carried out with ProDis. The results are consistent with the geolinguistic descriptions of Romanian varieties. ProDis identified two major areas for yes-no questions: a central-eastern one (Moldo-Walachian), with rising contours, and a central-western one (Transylvanian, discontinuously continuing over Bukovina to Bessarabia—the Republic of Moldova—), which presents also falling contours. The study suggests that intonational areas are somehow more conservative of and (relatively) independent from segmental dialectal areas.

Keywords: Romanian; intonation; dialectometry; falling question contours.

* Acknowledgement: This paper was financially supported by a grant from the Romanian National Authority for Scientific Research and Innovation, CNCS – UEFISCDI, PNII-RU-TE-2014-4-2298.

1. Introduction

The description of the Romanian language intonation has a long tradition that began before studies dedicated to prosody were issued in the 80's (Pike, 1948; Avram, 1972; Rankin, 1973; Chitoran et al., 1984, among others). However, the depiction of the dialectal variation of the Romanian language intonation represents a relatively new field of study whose results and outcomes are based on qualitative comparative methods (Avram, 1973; Dascălu-Jinga, 2001). The present study aims at researching this direction more deeply and proposing a first classification of the geographical varieties of the standard Romanian language in relation to the prosody characteristics, by applying certain measuring techniques of the prosodic distances of a dialectometric type.

2. State of the art

2.1. Traditional classifications of Romanian dialects

In the Romanian linguistics the term *dialect* is usually used for the four historical Romanian languages: DacoRomanian, at the north of the Danube, which became the Romanian national language, that has developed a standard variety, and the three dialects of the south of the Danube: Aromanian (Macedoromanian), Meglenoromanian and Istroromanian. The subunits of the Dacoromanian dialects are usually called subdialects and, rarely, dialects. There is no universally accepted solution for the problem of the distribution of the Dacoromanian dialects. The most important distinction is between the northern and south-western dialects (or Moldavian, including Banat, Transylvania, Crișana, Maramureș, Bukovina and Bessarabia) and the southern subdialect, comprising Wallachia, Oltenia, Dobrogea and the southern Transylvania. This distinction was largely discussed by Philippide (1927), Gheție (1964, 1975) and Vasiliu (1967). It is the oldest differentiation that can be analyzed by using written texts: mainly texts from the sixteenth century (Rosetti, 1978; Gheție & Mareș, 1974) and some texts from the fourteenth and fifteenth centuries (Gheție, 1994a, 1994b).

The theory of a bipartite distribution of the Romanian subdialects was mostly supported by language historians, while dialectologists regarded it as insufficient for the current state of these subdialects. To the three 'dialects' corresponding to the regions of Wallachia, Moldavia and Banat, as identified by Weigand (1908, 1909), an additional group of varieties covering the remaining part of Transylvania was added, two of which are more clearly distinguished, in Crișana (Petrovici, 1941; Ivănescu, 1944-1945) and Maramureș (Pop, 1950; Petrovici, 1954; Todoran, 1956). A special discussion should be dedicated to Transylvania, which has the greatest dialectal fragmentation. Even if prior to the appearance of the Romanian language geography, certain linguists (Tiktin, 1888; Gaster, 1891) also distinguished, according to the historical provinces, a Transylvanian dialect (associated to Transylvania itself, Crișana and Maramureș). The authors of the Romanian language atlases (Weigand, *ibidem*; Pușcariu, 1976 [1940]; Petrovici, *ibidem*) claimed that Transylvania itself does not have a special dialect, since its southern part is

united with Wallachia, the north-western with Moldavia, and the western part is united with Crișana and Banat. Ivănescu (*ibidem*, 1976, 1980) speaks about a ‘less individualized Transylvanian dialect’ which is, at the same time, ‘a dialect of transition’. Marin and Marinescu (1984) consider that ‘dialects of Transylvania’ (considered in a broad sense together with Crișana and Maramureș) have a ‘unit basis’ in which a very individualized area can be distinguished, represented by the subdialects spoken in Crișana, Țara Oașului and Maramureș, and a less individualized area comprising Deva, Dej, Rodna Mountains, Miercurea Ciuc, Brașov, Sibiu.

The regional variants of Romanian standard language differentiate two main types: Wallachian and Moldavian (Petrovici, 1961); Iancu (1974) added the Transylvanian variant and Turculeț (1993) a Bessarabia variant.

2.2. Previous approaches to geoprosodic variation of Romanian

Previous approaches to prosodic variation of Romance languages have often focused on yes-no questions (total interrogatives in Romance tradition) and, especially, where a falling contour occurs. Since Romanian is a final nucleus language the intonational contour is best characterized by the final part of the utterance, more specifically, by the tonal movements comprised between the last stressed syllable of the utterance until the end of the intonational group. Different linguistic traditions have called this specific part of intonational contours nucleus, nuclear configuration, terminal contour (CT) and *toneme*. In Romanian, the falling contour (CT) of yes-no questions consists of a steady tone at the last stressed syllable of the utterance, followed by a falling stressed syllable¹: H+L*L% (see figure 4). A common variant of this pattern is \downarrow H+L*L%, where the tone of the stressed vowel begins higher than the preceding syllable. This intonation pattern was first documented in Dascălu-Jinga (2001) and it was described as typical of ‘Transylvanian’ questions, nevertheless, it can also be found in some subdialects in Maramureș and Bihor.

The problem of the areas where the falling pattern of yes-no questions occurs was first addressed by Turculeț et al. (2008). They based on the auditory and acoustic analysis of the partial data in AMPER-ROM: 9 statements recorded in 4 ways: declarative—affirmative and negative—and totally interrogative—affirmative and negative—. The subjects were graduates of higher education from 10 Romanian cities: Baia Mare, Brașov, București, Chișinău, Cluj-Napoca, Craiova, Iași, Oradea, Sibiu, Timișoara. Four types of yes-no questions (affirmative and negative) were identified, their spreading areas being represented, in the work quoted above, on the map on the page 19. The falling pattern for yes-no questions occurs in all Transylvanian survey points (Timișoara, Cluj-Napoca, Sibiu, Brașov, Oradea, Baia Mare, but also

1 We called this CT *high descending*, in order to distinguish it from the CT *low descending* of the declaratives (see figures 8-11).

in Chișinău—Republic of Moldova—), but it occurs along with other patterns in all points (except for Timișoara, where the falling pattern is the only pattern used).

For a more objective auditory and acoustic analysis, the ‘distance/proximity measure’ in prosody was used between București and the other 9 points, calculating, by means of the Pearson formula, the coefficients of cross-correlation of the average melodic contours of the 9 statements uttered by the 2 subjects in Bucharest, on the one hand, and the subjects from other cities, on the other hand. The inter-correlation provides a mathematical evaluation of similarity/difference overall between the 2 curves; the maximum similarity has the index 1, that is a reduced percentage according to differences; in the case of significant differences, the ratio is negative: for example, the correlation coefficient of interrogative-negative statements in Timișoara, in relation to those in București, is -0.6.

The correlation coefficients of the melodic contours of the declaratives (affirmative and negative) do not show great differences between subjects in the same survey point or between different points; the average index of the prosodic distance between Bucharest and the other 9 points is 0.81. In the case of the yes-no questions (affirmative and negative), the indexes show differences between subjects of the same point, and especially between survey points: the average indexes correlation between Bucharest and the other points are 0.22 in affirmative yes-no questions and -0.095 for the negative ones. The highest ‘prosodic distance’ from Bucharest occurs in Timișoara, Chișinău, Cluj-Napoca, Oradea, Baia Mare and Brașov.

Statistical analysis confirms the geoproisodic data: if we represent the ‘prosodic distance’ referred above on a map, for questions, the map would be similar to the map illustrated in figure 28, which distinguishes 2 areas: a central-eastern one (Wallachia-Moldavia) and a rather discontinuous ‘Transylvanian’ one, which includes Timișoara, Chișinău, Cluj-Napoca, Oradea, Brașov and Baia Mare. Sibiu (with an average index of 40.5) would make the transition between the two areas, being closer to the centre. If we take into consideration the four ways envisaged, the average of the ‘prosodic distances’ from Bucharest would be, in increasing order: Iași, 0.70; Sibiu, 0.63; Craiova, 0.62; Brașov, 0.51; Baia Mare, 0.49; Oradea, 0.33; Chișinău, 0.31; Cluj-Napoca, 0.29; and Timișoara, 0.18. The Transylvanian area with the most significant ‘prosodic distances’ includes Timișoara, Cluj-Napoca, Chișinău and Oradea. Sibiu belongs to the central-eastern area, while Brașov and even Baia Mare are closer to the central-eastern one.

A recent study (Turculeț, 2017) establishes an accurate territorial distribution of the falling yes-no question pattern, based on the complete database of *AMPROM Part I. Regional varieties of the literary language* for 25 points of survey². In each point, 2 speakers were recorded, male and

2 There were added on the map (figure 1) the following points of survey: Arad, Satu-Mare, Târgu Mureș and Bălți.

female, all of them with higher education levels. Three respondents were from Bălți, Chișinău, three from Iași, and five from Cluj-Napoca. Overall, 55 subjects were recorded. Thus, the corpus contains 14850 utterances, representing yes-no questions (both affirmative and negative).

2.3. The AMPER Project in Romania

AMPER-ROM[ÂNIA], the part of AMPER dedicated to the Romanian language, presents some specific features that makes it different of the common Romance project. The fixed AMPER-ROM questionnaire comprises 45 questions, therefore, it omits some stress structures included in the standard questionnaire. For example, AMPER-ROM does not contain *twt* (paroxytone subject and paroxytone object) and *pwp* (proparoxytone subject and proparoxytone verb). Instead, it includes affirmative and negative sentences both in statements and questions, e.g.: *pwtā* = *Pasărea vede nevasta* ‘The bird sees the wife’, *pwtñ* = *Pasărea nu vede nevasta*, *pwti* = *Pasărea vede nevasta?* ‘Does the bird see the wife?’ and *pwtm* = *Pasărea nu vede nevasta?* ‘Doesn’t the bird see the wife?’.

So far, the AMPER database at Grenoble was sent the answers to the fixed questionnaire from 11 points of survey, representing recordings concerning both the diatopic dialectal varieties and the diatopic varieties of the literary/colloquial language, the latter recorded from subjects with higher education.

With the start of investigation for AMPER-ROM in 2007, Romanian multimedia prosodic atlas (AMPROM)³ was initiated according to the AMPER model. An innovation of this atlas is represented by a second questionnaire, which contains various statements (focusing broadly on structures containing different types of questions: partial, alternative, confirmation, etc.), with a semi-spontaneous character. AMPROM was designed in three parts: the first part is dedicated to literary Romanian language at the spoken level (standard language); the second part, to the Romanian subdialects, and the third part, to the Aromanian (Macedoromanian) and Meglenoromanian dialects. Surveys were conducted in 60 survey points (including those for AMPER-ROM) in the whole area of Romanian language. To date, data for the first part of the atlas, and partially for the rest, have been analyzed.

3. Methodology

3.1. Corpora

This research will use the responses to the fixed questionnaire AMPER-ROM that recorded subjects who speak standard Romanian language in 21 survey points from the AMP-ROM I

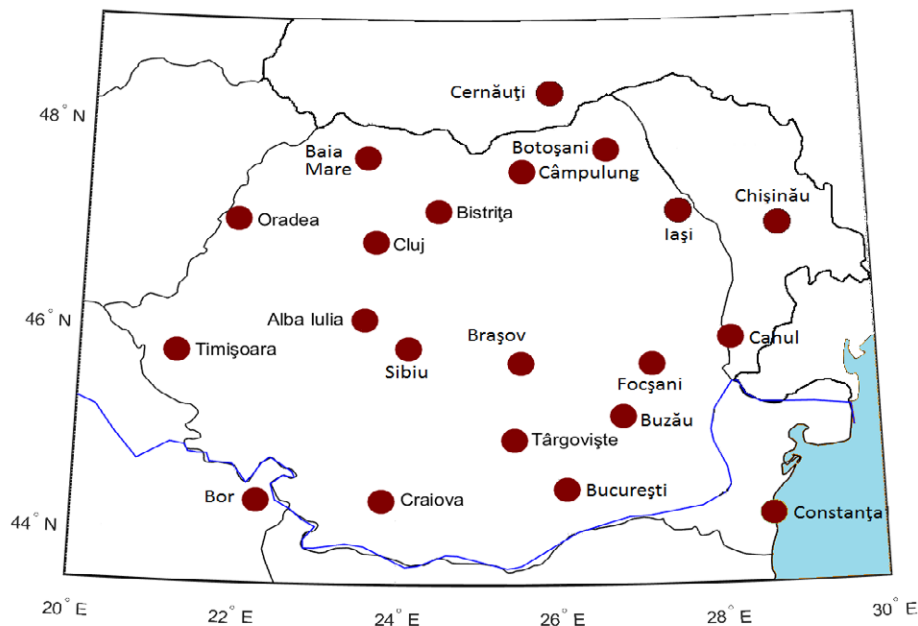
3 Turculeț & Minuț, 2007; Turculeț, 2013; Bibiri, Panaite & Turculeț, 2012.

network. Each of the 45 statements with different emphasized morphosyntactic structures was recorded in four distinct modalities: statements (affirmative and negative) and yes-no questions (affirmative and negative) (see § 2.3).

Each speaker recorded 3 repetitions of each and in each survey points two speakers were recorded (male and female). Therefore, the present study uses data/recordings from 21 survey points and a total number of 26.680 statements. The geographical distribution of the survey points is illustrated in figure 1.

FIGURE 1

The geographical distribution of the survey points



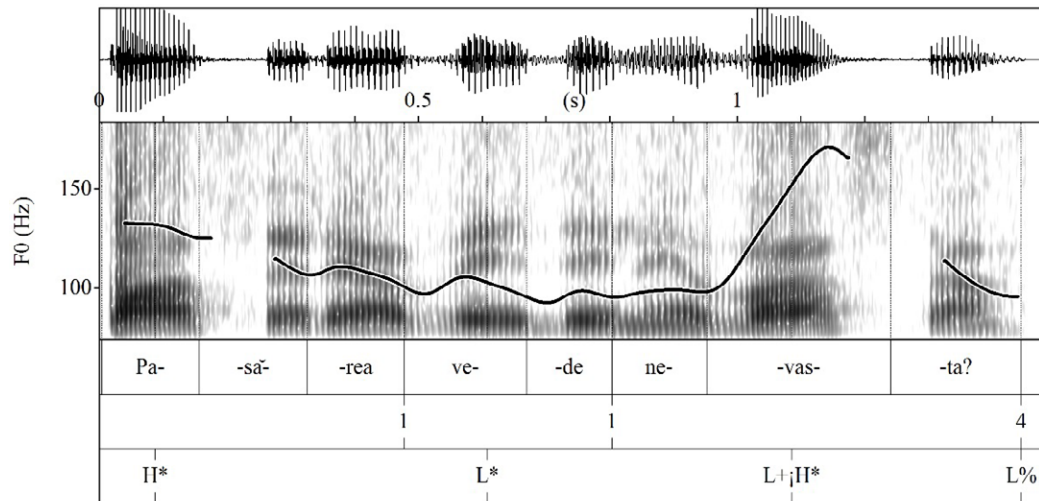
3.2. The main intonational patterns of the corpus

The corpus contains 26.680 recordings that are characterized by the heterogeneity of their intonation patterns. If we are to observe exclusively the nuclear contour or toneme, there are at least six main intonation patterns for questions (without taking into account their probable allotonic versions), some of them have rising, others falling, and some others have circumflex contours. From each category we present an example and a transcription of intonation, wherever possible, according to the first proposal of transcription of Romanian intonation, using the Ro_ToBI system (Jitcă et al., 2015).

The first intonation pattern has a nuclear configuration of the L+;H*L% type, in which Fo increases dramatically on the last stressed syllable of the utterance and decreases significantly at the end of the utterance (figure 2).

FIGURE 2

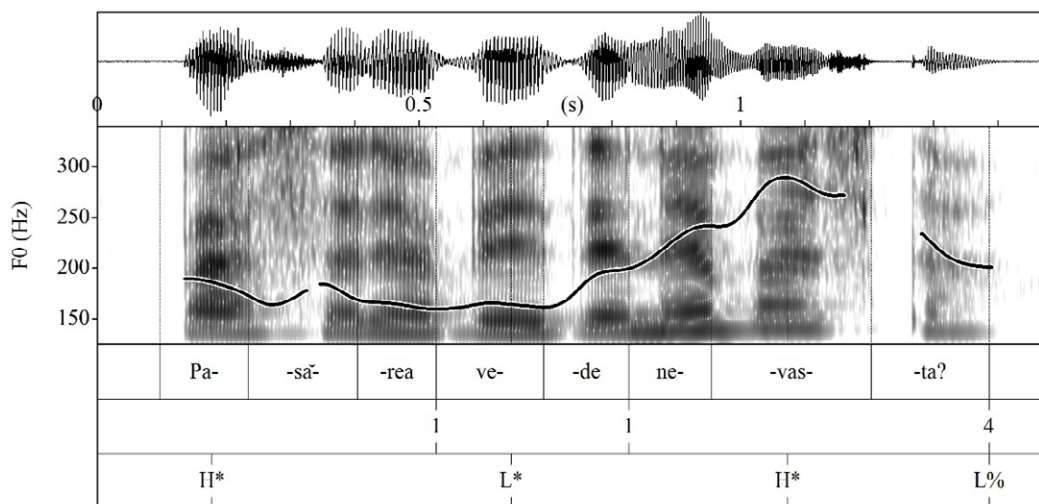
Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the absolute interrogative of the enunciation *Pasărea vede nevasta?* 'Does the bird see the wife?', uttered by a speaker from Iași



The second category of questions patterns is H* L%, which is characterized by a relatively flat pre-nucleus until the stressed syllable of the verb, where the minimum value F0 is reached, followed by an increase whose peak is on the last stressed syllable of the utterance and a decrease towards the end (figure 3).

FIGURE 3

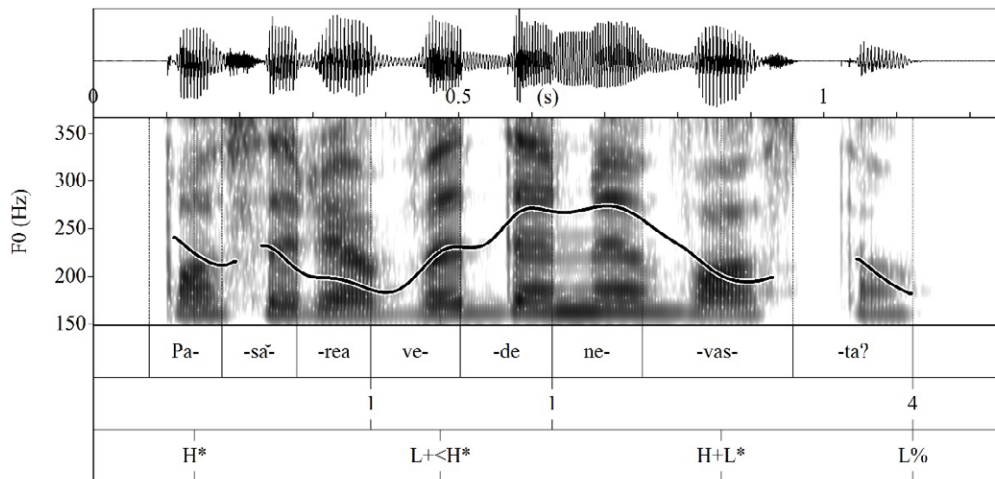
Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the yes-no question of the sentence *Pasărea vede nevasta?* 'Does the bird see the wife?', uttered by a speaker from Baia Mare



The third interrogative pattern, which is also falling (of the H+L* L% type), has a nuclear configuration where there is a fall between the last pre-tonic syllable and the stressed syllable. The final tone is low, as depicted in figure 4.

FIGURE 4

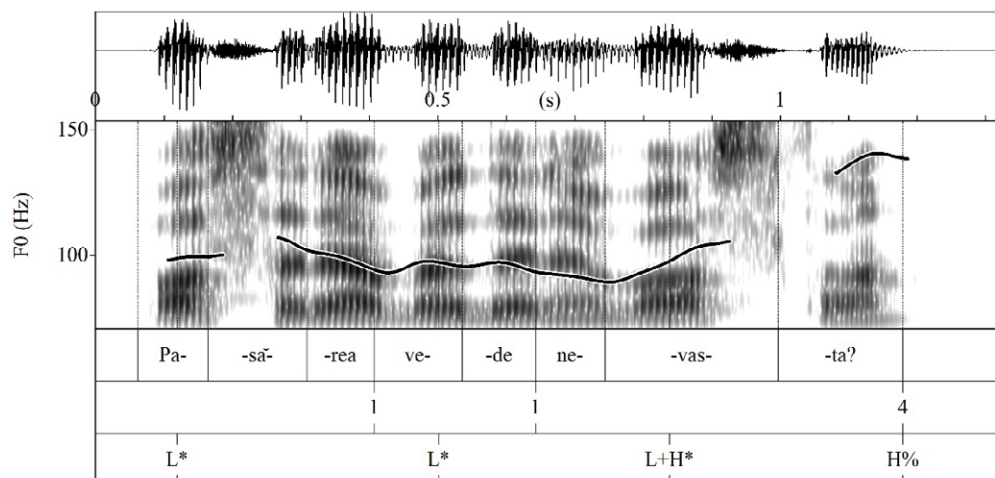
Oscillogram, spectrogram, F0 curve, orthographic transcription and prosodic annotation of the question *Pasărea vede nevasta?* 'Does the bird see the wife?', uttered by a speaker from Alba Iulia



The fourth interrogative pattern is rising (L+H* H%). It is characterized by a relatively homogeneous nucleus, followed by a rise that begins on the stressed syllable and continues throughout the utterance (figure 5).

FIGURE 5

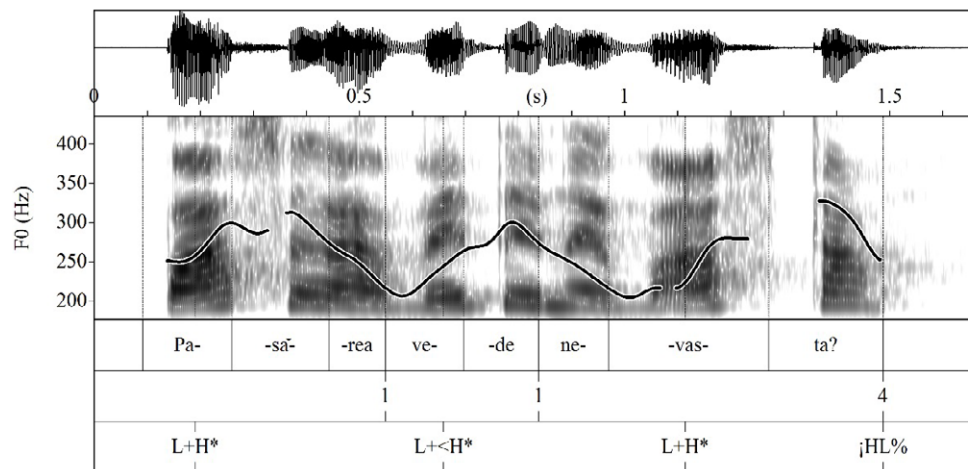
Oscillogram, spectrogram, F0 curve, orthographic transcription and prosodic annotation of the question *Pasărea vede nevasta?* 'Does the bird see the wife?', uttered by a speaker from Alba Iulia



The fifth question pattern, illustrated in figure 6, is characterized by a pre-nucleus with well-defined tonal peaks, followed by a complex tonal movement (L + H * HL%) in the final part of the utterance.

FIGURE 6

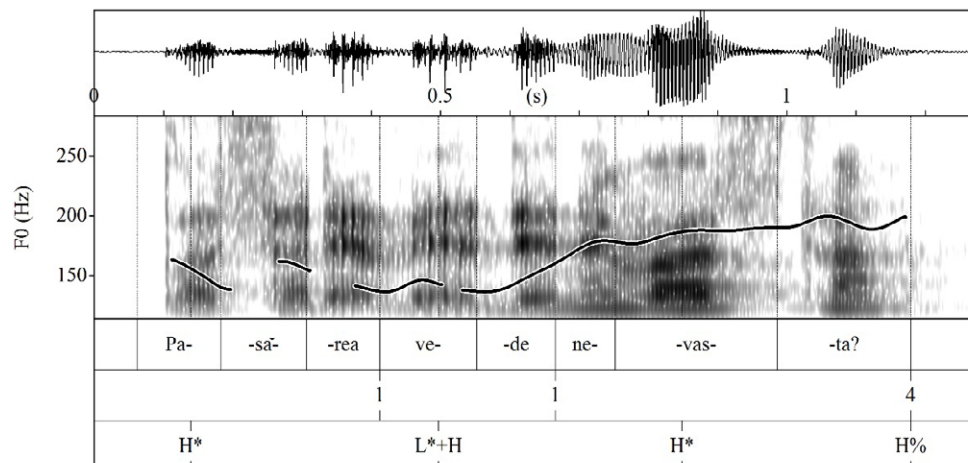
Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the question *Pasărea vede nevasta?* ‘Does the bird see the wife?’, uttered by a speaker from Brașov



The sixth interrogative pattern is the only one with a specific geographical distribution, as we shall argue later in this study. In our database, it occurs in Bor, a city of the Republic of Serbia. In fact, the same rising pattern (H* H%, figure 7) has been also documented in Serbian (Godjevac, 2005).

FIGURE 7

Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the absolute interrogative of the enunciation *Pasărea vede nevasta?* ‘Does the bird see the wife?’, uttered by a speaker from Bor



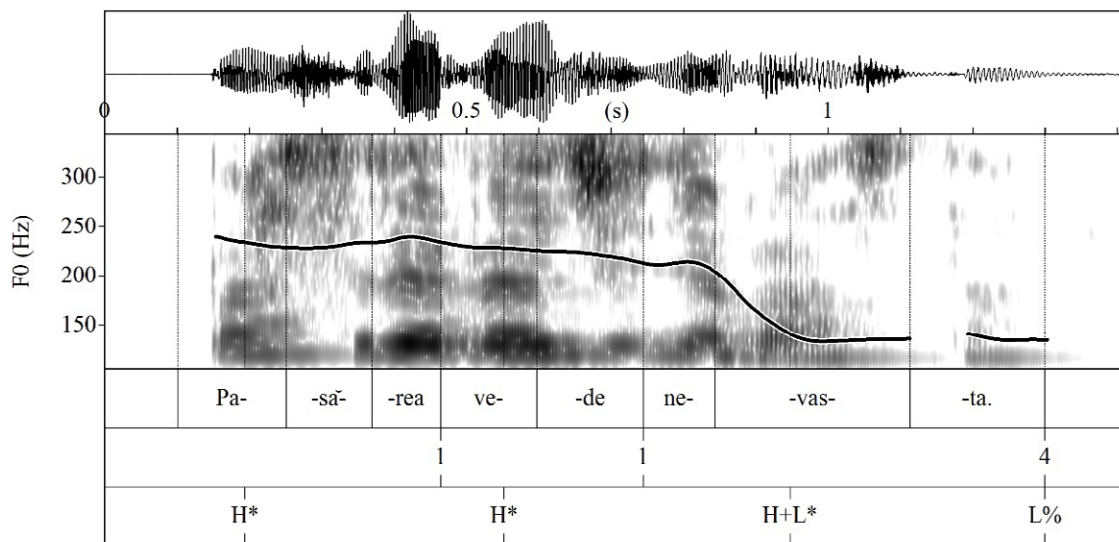
The intonation patterns presented above characterize the questions with a paroxytone final word: *nevásta*. The statements that have an oxytone final word (*căpitán*, *elegant*) have differences in questions, such as H* H% (in the Walachia-Moldavia area) and H-L* L% (in Transylvania).

In the declarative sentences, a few variant patterns can also be noticed. However, the differences among patterns are smaller, since in the case of declarative sentences all patterns are falling. The differences in declarative patterns can be found both on the nuclear accent and on the pre-tonemes.

The first declarative pattern, that we shall describe as H+L* L%, is characterized by a high plateau that lasts until the last pre-tonic syllable, then there is a sudden fall of F₀, that remains stable until the end of the utterance (figure 8).

FIGURE 8

Oscillogram, spectrogram, F₀ curve, orthographic transcription and prosodic annotation of the declarative *Pasărea vede nevasta* 'The bird sees the wife', uttered by a speaker from Bor



The second declarative pattern, exemplified in figure 9, has a typical *hat pattern* contour, characterized by an initial rising, followed by a high central plateau and the fall of the fundamental frequency F₀ that begins with the nuclear syllable (H*+L L%).

The third declarative pattern has a nuclear configuration that is similar to the previous one (H*+L L%), but differs from it due to the presence of a peak in the first stressed syllable, as stressed in figure 10.

FIGURE 9

Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the declarative *Pasărea vede nevasta* 'The bird sees the wife', uttered by a speaker from Iași

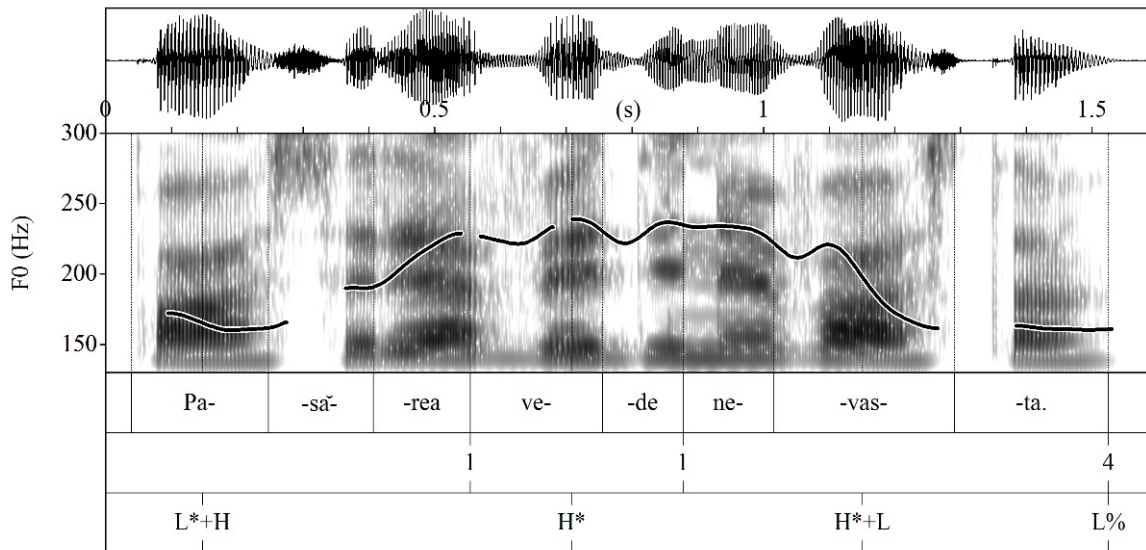
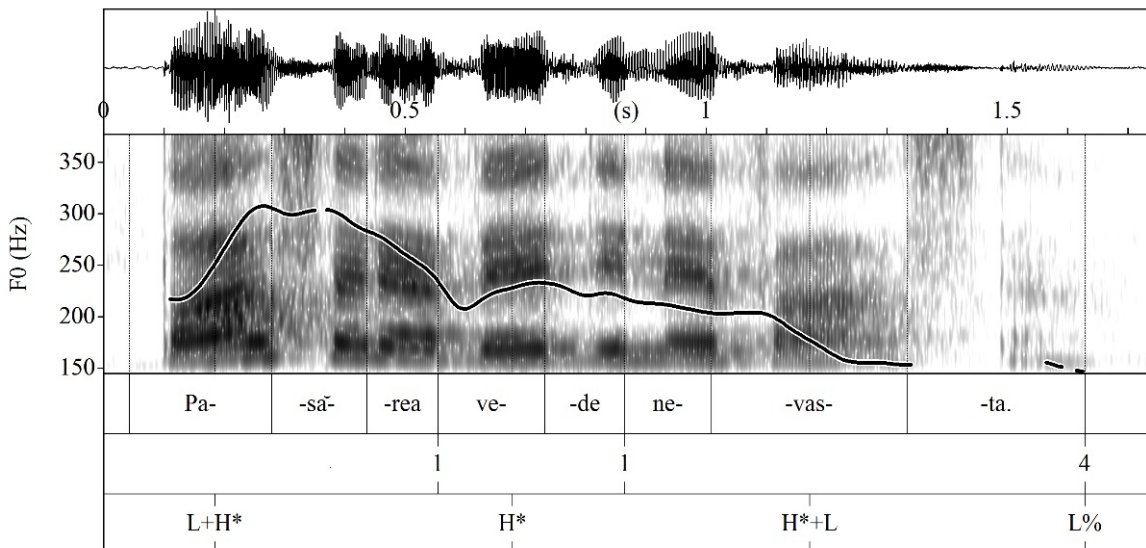


FIGURE 10

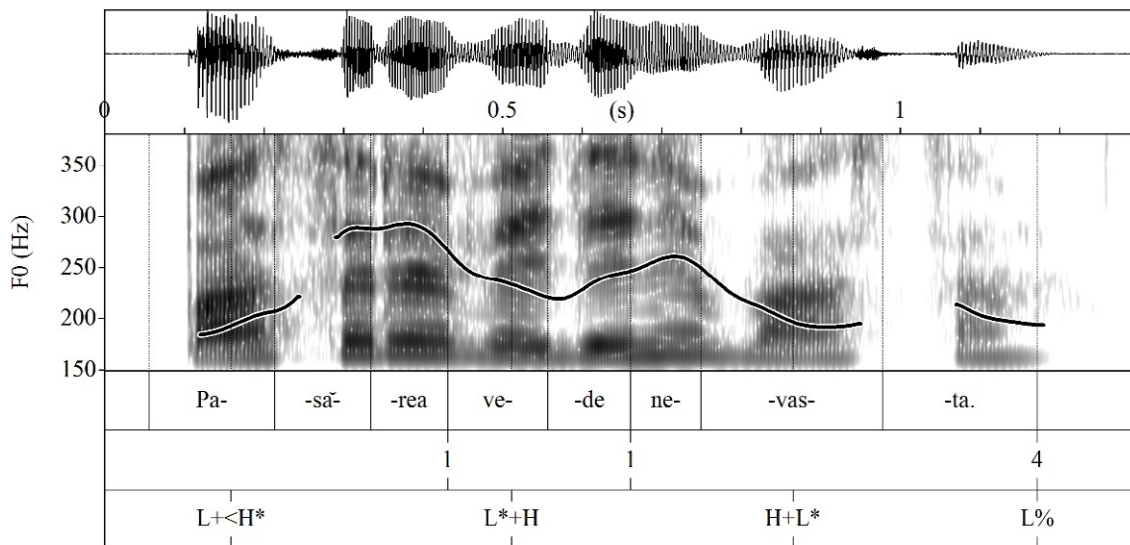
Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the declarative *Pasărea vede nevasta* 'The bird sees the wife', uttered by a speaker from Câmpulung



The fourth declarative pattern (figure 11) is the most dynamic one, as it displays two pre-nuclear peaks before the descent specific to the nuclear configuration of the H+L* L% type.

FIGURE 11

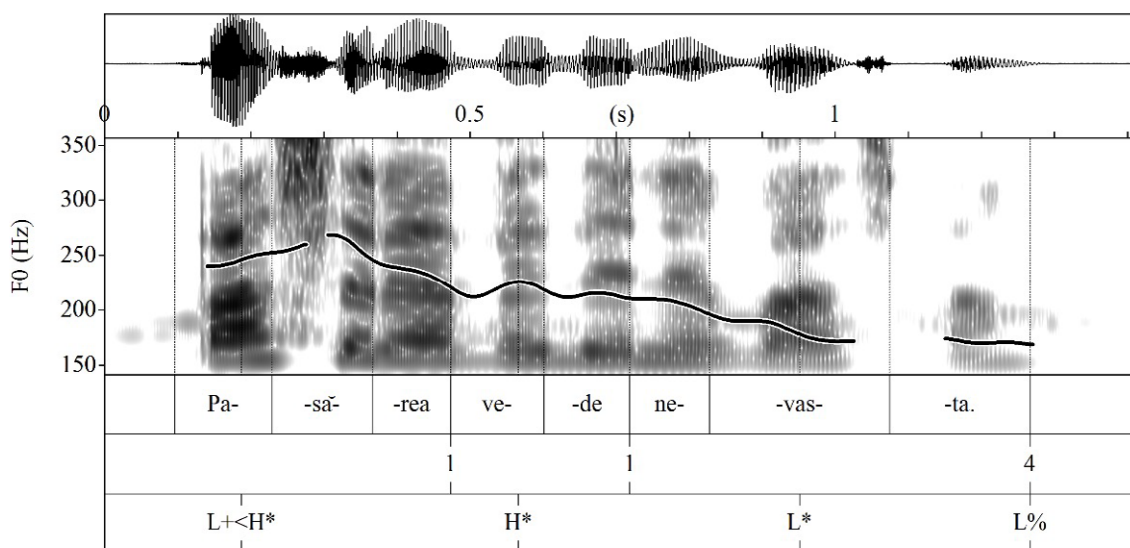
Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the declarative *Pasărea vede nevasta* 'The bird sees the wife', uttered by a speaker from Chișinău



The fifth declarative pattern (figure 12) is characterized by a tonal peak associated with the first lexical stress and a smooth falling line of the Fo fundamental frequency that is maintained until the nuclear configuration, transcribed as $L^* L\%$.

FIGURE 12

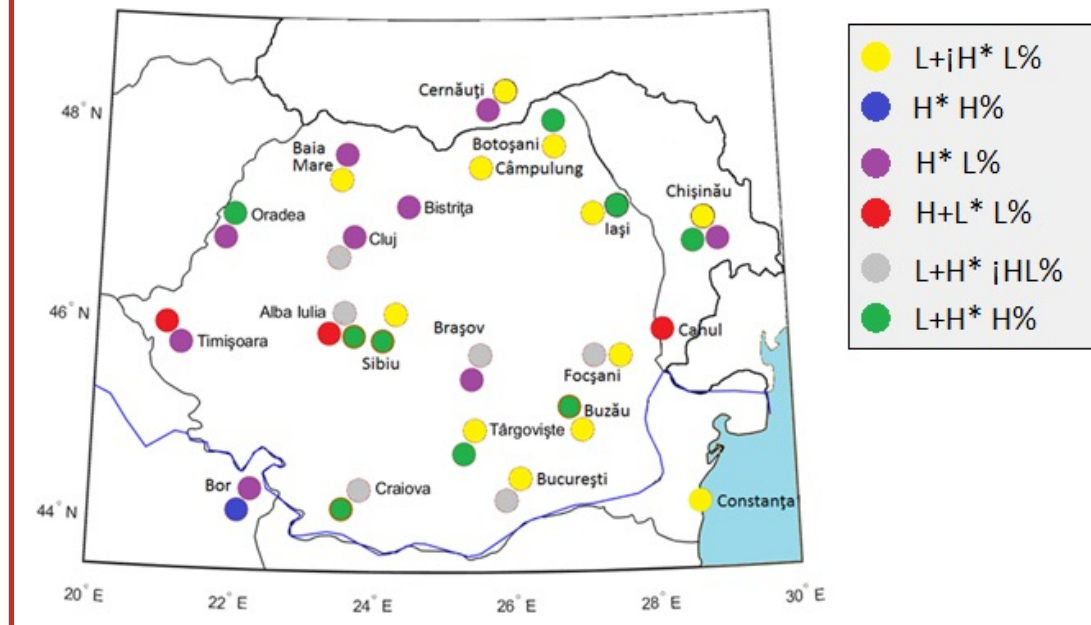
Oscillogram, spectrogram, Fo curve, orthographic transcription and prosodic annotation of the declarative *Pasărea vede nevasta* 'The bird sees the wife', uttered by a speaker from Cluj



From a dialectological point of view, the presence of such a high number of intonation patterns (both declarative and above all, interrogative) represents a challenge, since the majority of these patterns occurs with different frequencies in most of the survey points. In other words, because of the high quantity of heterogeneous data, it is very difficult for a researcher to establish which intonation pattern has the highest frequency in each point and whether there are any relatively compact geoprosodic areas. Figure 13 exemplifies the distribution of the intonational patterns for the affirmative question *Pasărea vede nevasta?* ‘Does the bird see the wife?’. As one can notice, it seems practically impossible, at a first sight, to identify any regularity as far as the geographical distribution of the intonation patterns is concerned (in figure 13 they are distinguished by different colours). The consequence of this is the impossibility to perform a ‘manual’ analysis of the territorial distribution of intonation patterns. Hence, the need of using instruments such as ProDis, whose characteristics are described in the next section.

FIGURE 13

Geographical distribution of intonation patterns documented for the affirmative interrogative *Pasărea vede nevasta?* ‘Does the bird see the wife?’



3.3. ProDis

ProDis is a tool developed by the Phonetics Laboratory of the University of Barcelona for the dialectometrization of the prosodic data in the AMPER project, namely the values of the fundamental frequency F_0 , the duration and intensity of vowel elements of the utterances included in the corpus, as described in section 2.3. In this study we used the F_0 values weighted by its duration and intensity (Elvira-García et al., 2018).

Within this study, ProDis is used for two types of analysis: a cluster analysis and a multidimensional analysis (or MDS). Both types aim at classifying the various survey points in groups, according to their similarity, so that homogeneity within each group and the differences between the groups are as high as possible. According to the statistical analyses under discussion, ProDis provides a series of graphic results, as described in section 3, where we shall explain the graphics as well as their interpretation key.

The first map is a standard deviation map that displays standard deviation for each speaker. Cool colours indicate a low standard deviation (from a linguistic point of view, speakers that tend to use the same intonation patterns). Warm colours indicate speakers whose contours display a high standard deviation, that is, from a linguistic perspective, that the speaker displays a greater variety of intonation patterns within the same type of utterance.

The second type of graphic representation is a correlation map. In this case, as described above, instead of displaying numerical values (from +1 to -1), ProDis uses a chromatic scale to represent them. Lighter colours (closer to white) indicate a high correlation between two survey points (in linguistic terms, this means that the intonation patterns are similar). Colours closer to dark red indicate a low degree of correlation (different melodic patterns).

The third type of graphic is a dendrogram, i.e. a tree diagram, drawn starting from cluster analysis, in which survey points are grouped according to their prosodic similarities. ProDis indicates the significant statistic groups by means of different colours. It should be restated here that in a dendrogram it is important to identify not only the groups themselves, but also the levels at which they are united (the more connections in the right side of the graphics, the more differences between the groups).

The fourth type of representation is an MDS map, which is a virtual space in which the points representing the survey areas are distributed according to their intonation similarities. When producing the MDS map, ProDis uses colours in order to help the reader find, on the same map, the groups identified in the dendrogram.

Last but not least, the fifth graphic result provided by ProDis is a geographical map (the Eastern Balkan area, in this specific case) in which the survey points are indicated with different colours. Each colour corresponds to a group on the dendrogram.

4. Results

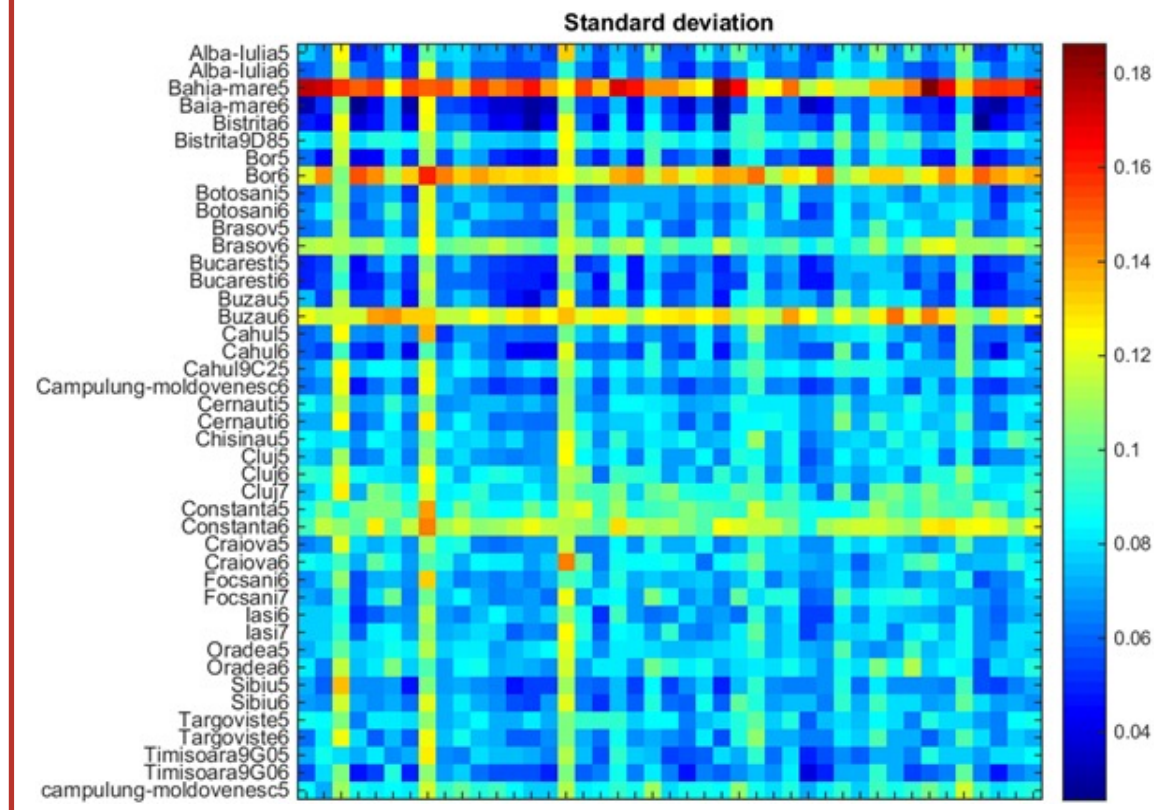
This section presents the results of the dialectometric analysis performed with ProDis for the data described in section 2.3. The section provides a description of the results of the dialectometric analysis of the investigated modalities, followed by an overall perspective upon the analysis.

4.1. Declarative and negative statements

If declarative modalities (both affirmative and negative) are analysed together (figure 14), cool colours prevail. This indicates that there is coherence in the choice of intonation patterns for declaratives (with very few exceptions) within speakers.

FIGURE 14

Standard deviation for declarative sentences



In the graphic displaying the correlation map (figure 15) warm colours prevail, indicating that the declaratives generally display high correlations between most localities. In other words, there are few dialectal variations for this modality, a result that reinforces the observations derived from the qualitative analysis provided in section 2.3.

The dendrogram in figure 16 allows us to conclude that two dialectal sub-groups can be identified: a main one that includes almost all survey points and a small group that contains only three elements. However, it is important to notice that the two groups are linked at a low level (level 1.1), a fact that indicates that the differences between them are significant (when compared, for instance, with the dendrogram in figure 12, where the groups are linked at level 2).

FIGURE 15

Correlations between investigation points for the declarative modality

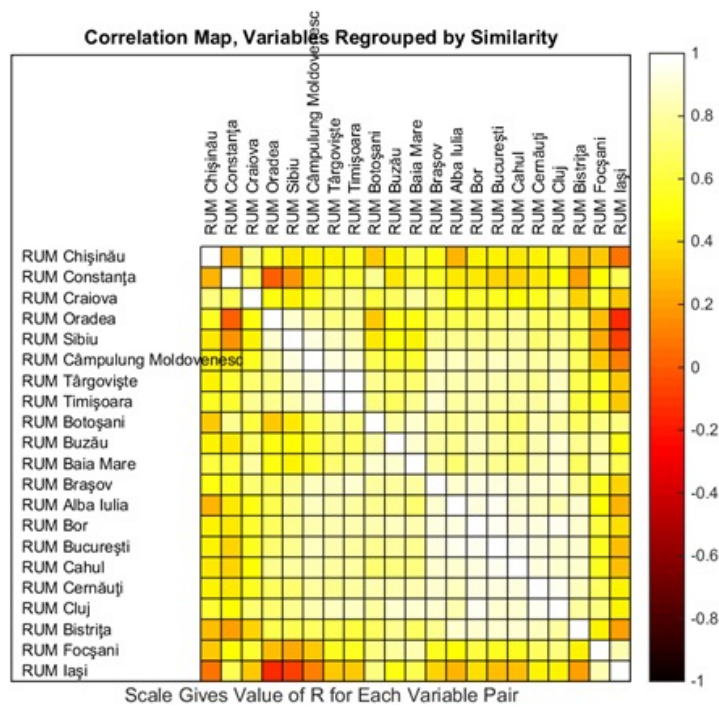
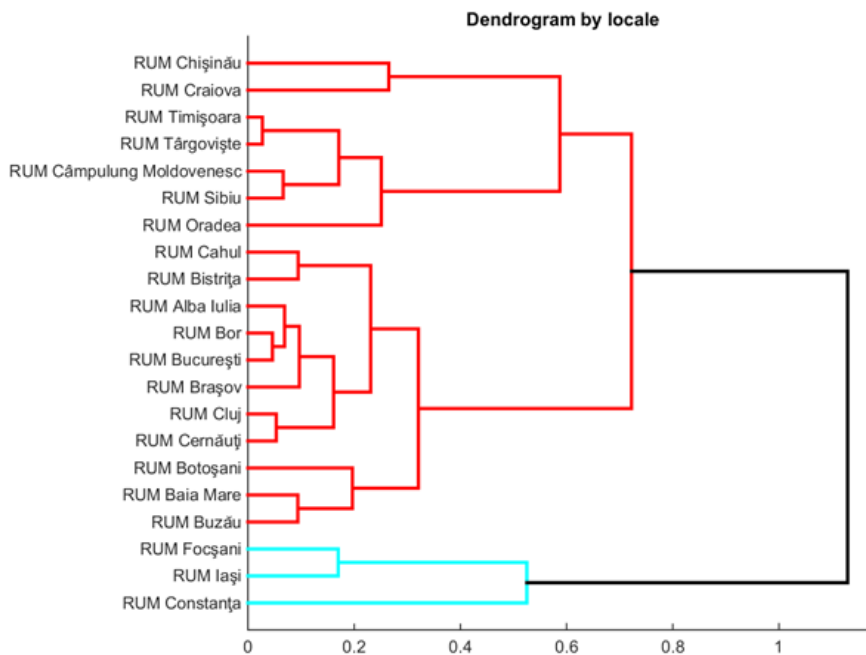


FIGURE 16

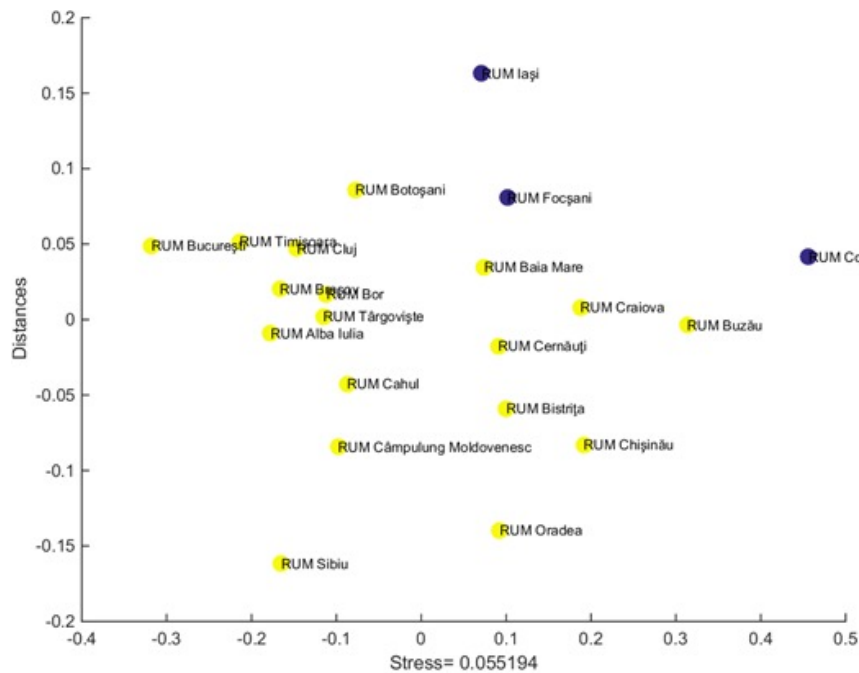
Dendrogram for the investigation points in the declarative modality



The MDS map (figure 17) can add significance to the results of the dendrogram, as it indicates that the two groups are in fact very close.

FIGURE 17

The MDS map for the investigation points in the declarative modality



Overall, according to the real geographic distribution of the points that unite the two groups (figure 18), we can observe that the three points that form the small group display a distribution that is difficult to explain from a dialectal perspective.

To sum up, we can conclude that declarative sentences display small dialectal variation (or smaller variation than questions, as described in the following sections). Moreover, the geographic distribution of the identified intonation patterns cannot clearly distinguish between the dialectal groups. This conclusion reinforces the observations upon several Romance languages in which yes-no questions usually display the highest inter-dialectal differences and, therefore, is the decisive modality as far as dialectometry is concerned (Roseano & Fernández Planas, 2019; Fernández Planas et al., 2015a, 2015b).

4.2. Affirmative yes-no questions

The analysis of the standard deviation of affirmative yes-no questions allows us to observe how they differ from the declaratives (figure 18). The standard deviation is higher in questions than in declaratives. This means that different speakers (especially the ones indicated by warm colours in the figure) use several intonation patterns to express the same function.

FIGURE 18

Map of survey points, classified according to the results of the cluster analysis of the data in the declarative modality

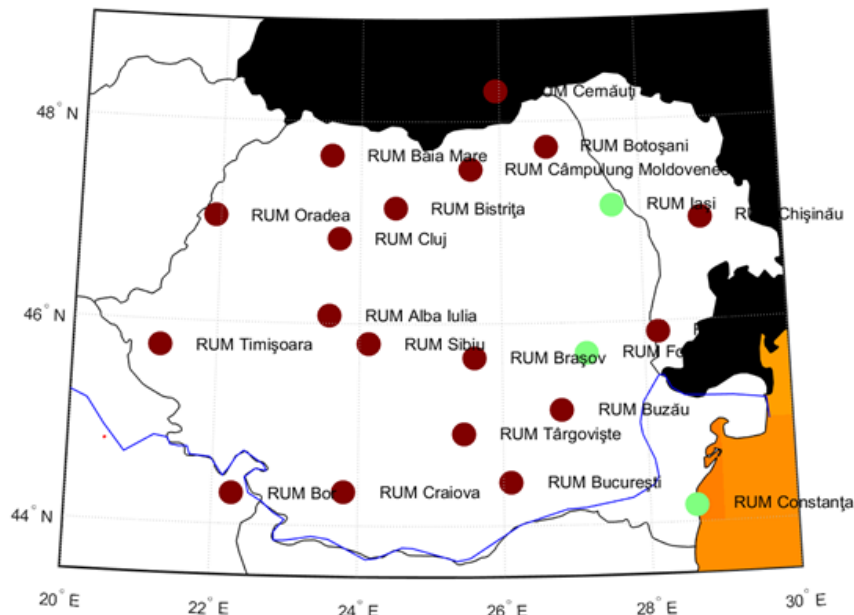
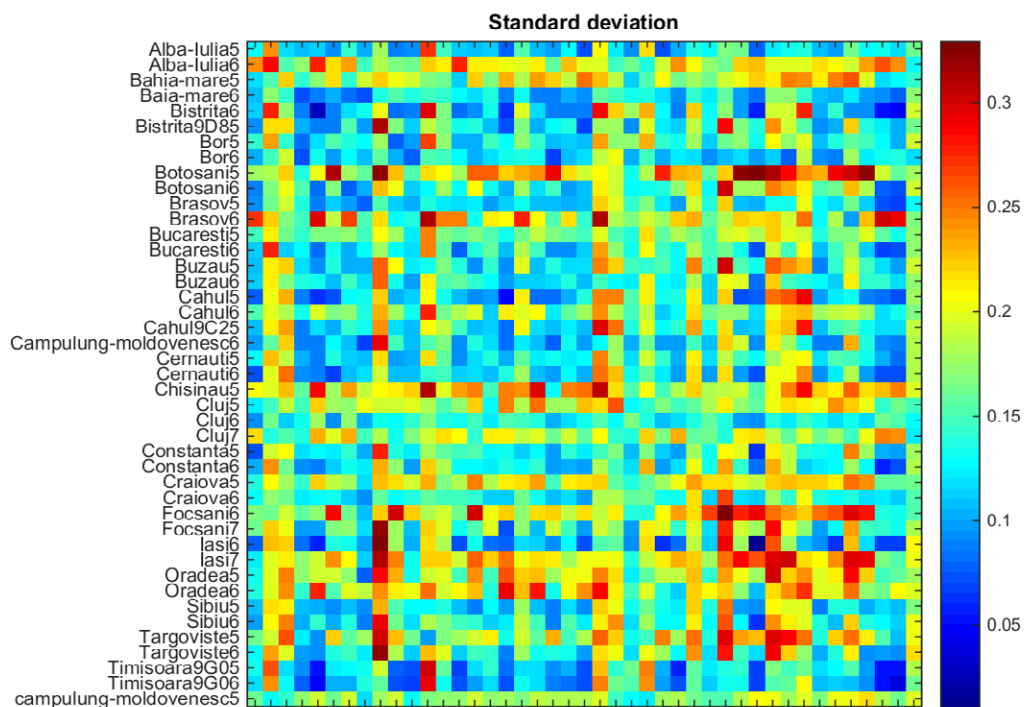


FIGURE 19

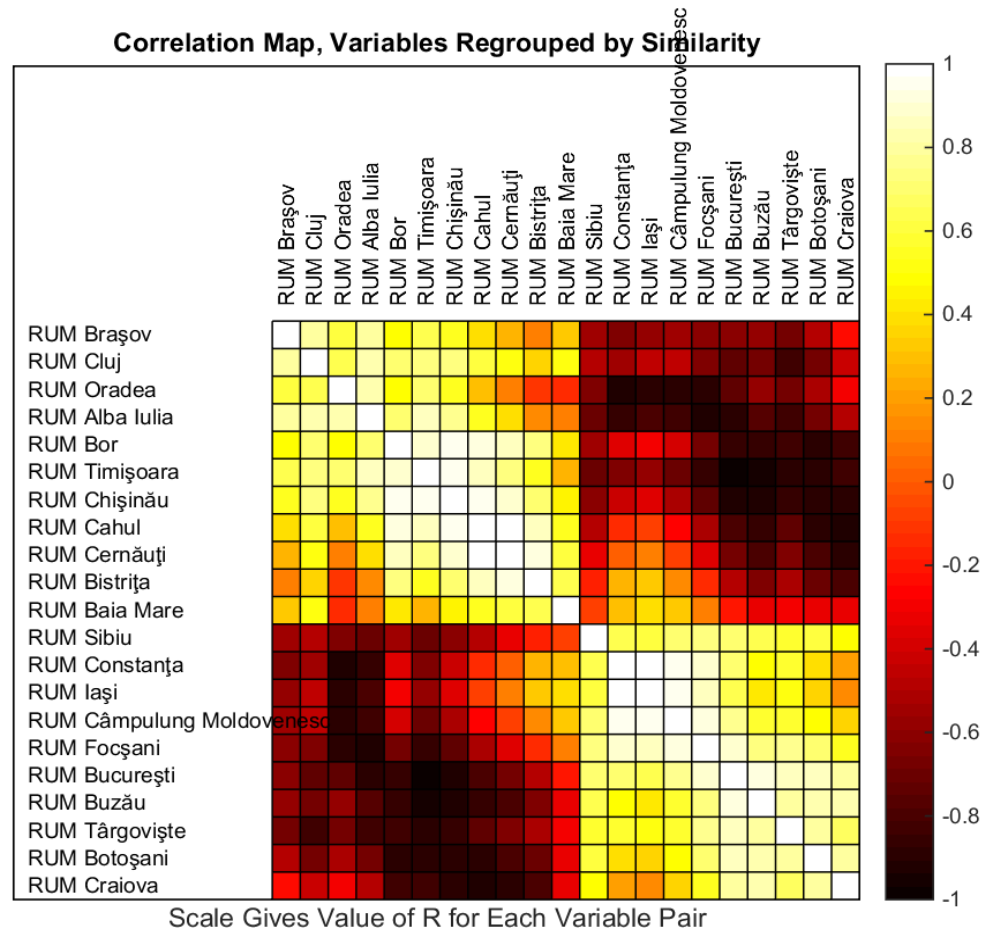
Standard deviation of total affirmative interrogatives



If we consider the correlations map (figure 20), we can see two distinct groups. Each of them has strong correlations with other members of the same cluster (light colours) and weaker correlations with other investigation points of another cluster (dark colours).

FIGURE 20

Correlations between the investigation points of the total affirmative interrogative modality



The dendrogram (figure 21) confirms the fact that the two groups are internally homogeneous while being well differentiated. In the dendrogram, both groups are linked at a low level (approximately 1 on the horizontal axis) and, at the same time, are linked between them at a much higher level (approximately 2 on the same axis).

The MDS map (figure 22) confirms the results of the dendrogram, indicating that each of the two groups of the dendrogram includes points that are grouped in two distinct areas and which can be clearly distinguished in the virtual map area.

FIGURE 21

Dendrogram for the investigation points of the total affirmative interrogative modality

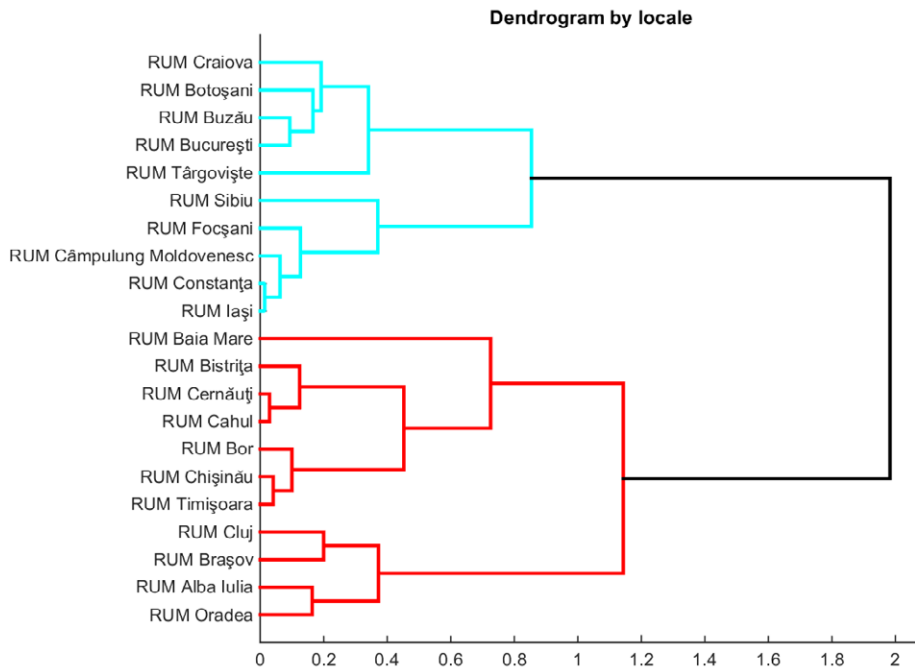
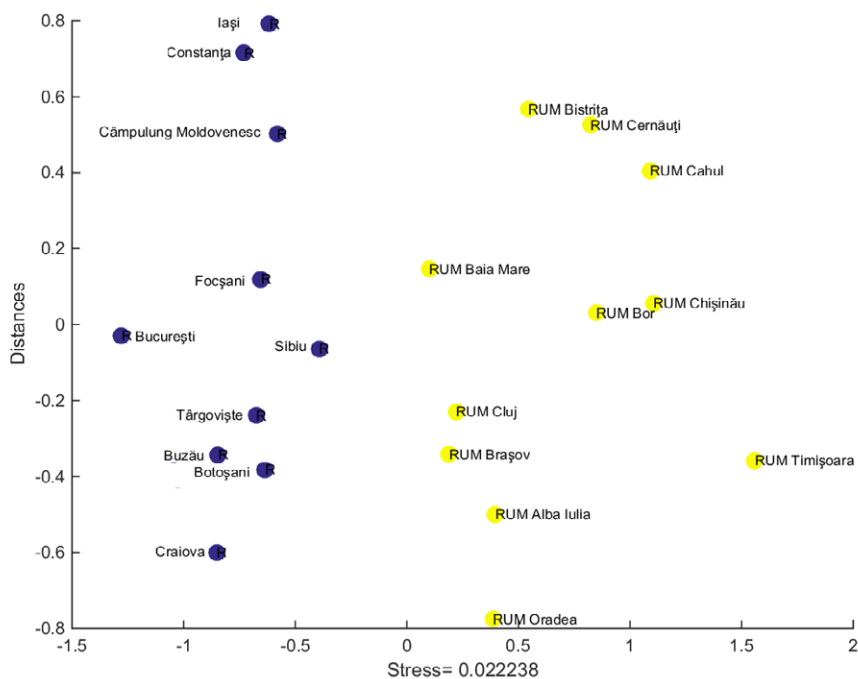


FIGURE 22

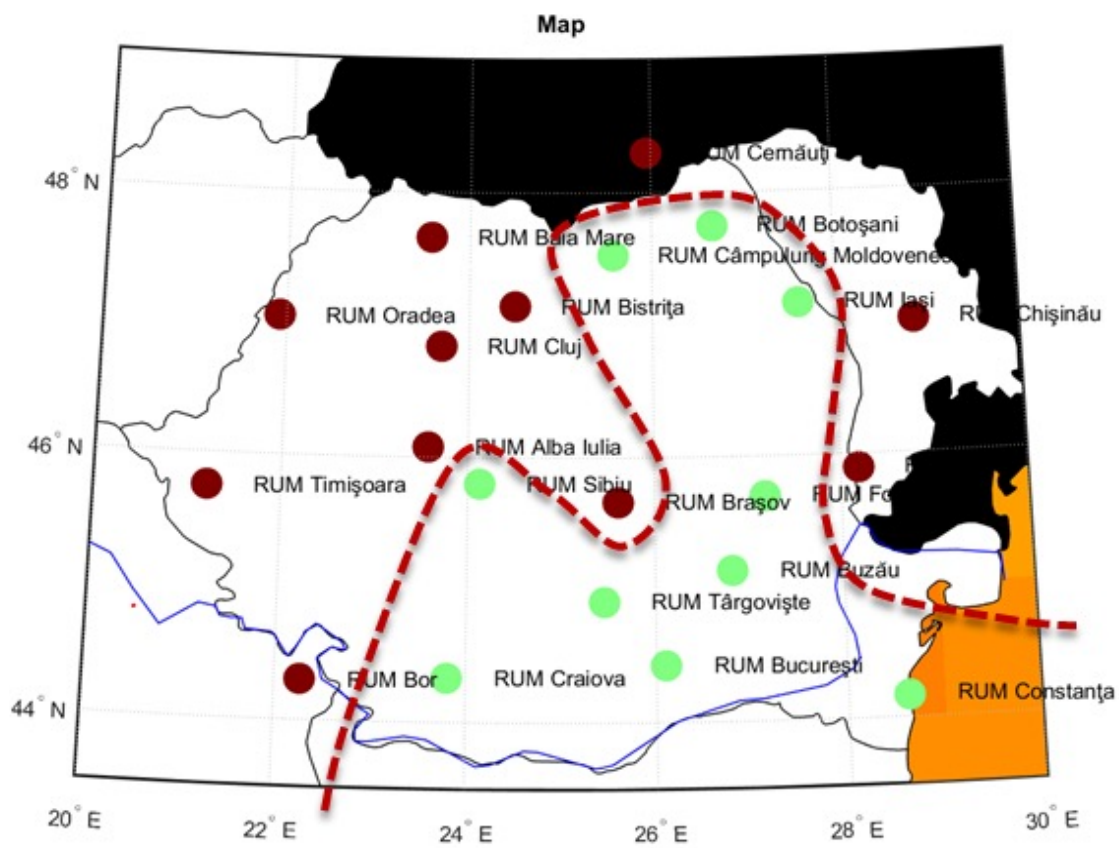
MDS map for the survey points of the affirmative yes-no questions



When the points are transferred to a real map (figure 23), we notice that the distribution is not determined by causality, since there exists some contiguity between the analysed points and, within certain limits, there also exists a territorial contiguity. The first of the two areas under discussion—depicted in green—has a triangle-like shape; the angles being Botoșani, Craiova and Constanța. The second area has the shape of a half-moon depicted in brown and surrounds the triangle representing the first area.

FIGURE 23

Map of the survey points classified according to the results of data cluster analysis of the affirmative yes-no questions



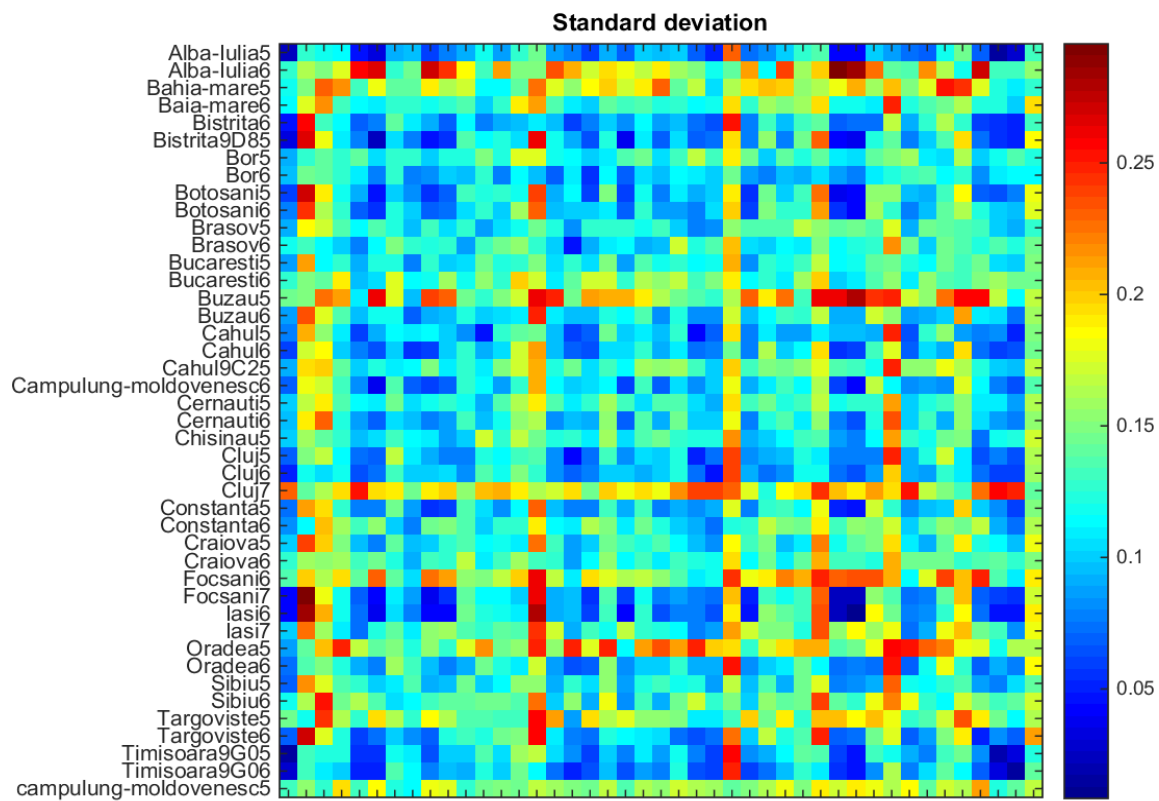
To conclude this section, two main ideas should be reiterated. On the one hand, questions, as opposed to statements, display notable variations. Moreover, these variations are not accidental. Cluster analysis and MDS allow the identification of two distinct geoprosodic groups that, on the real geographic map, have a distribution that cannot be regarded as accidental. Such a distribution shall be further argued on and interpreted, from both a historical and dialectological perspective, in the section dedicated to discussions.

4.3. Negative yes-no questions

The intonation of negative interrogatives has been described in previous studies as being very characteristic (Turculeț et al., 2008; Turculeț & Bleorțu, 2013; Jitcă et al., 2015; Turculeț, 2017). The dialectometric analysis of the data provided by this modality has led to the conclusion that, as in the case of affirmative interrogative modality, many speakers display a high standard deviation (represented by warm colours in figure 24). In other words, for the modality under analysis, there are different intonation patterns that the speakers use alternatively in their utterances.

FIGURE 24

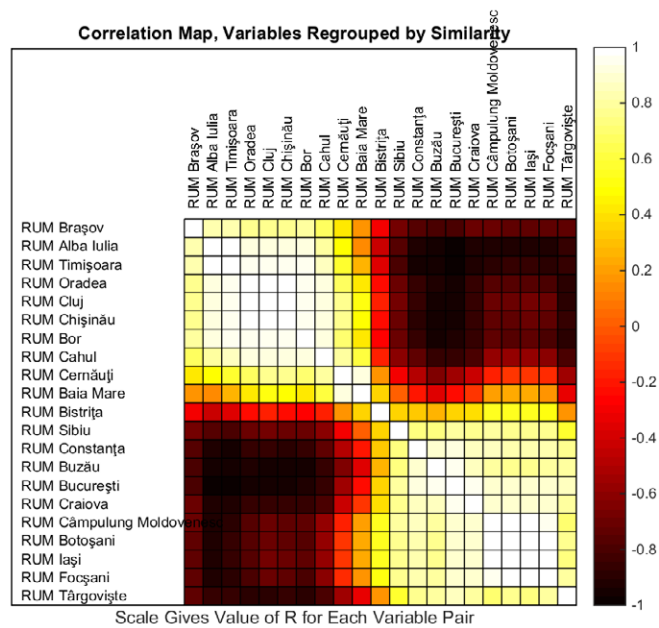
Standard deviation of the total negative interrogative modality



The analysis of the correlations (figure 25) indicates that there are two distinct groups, as it was also noted in the case of affirmative interrogatives. Once again, the points in each of the two groups have high correlations (represented by cold colours on the graphic) with the points in the same group, while the localities in the other groups have low correlations (warm colours). Furthermore, three survey points—Baia Mare, Bistrița and Cernăuți—do not seem to fit perfectly in either of the two main groups indicated in the next figure.

FIGURE 25

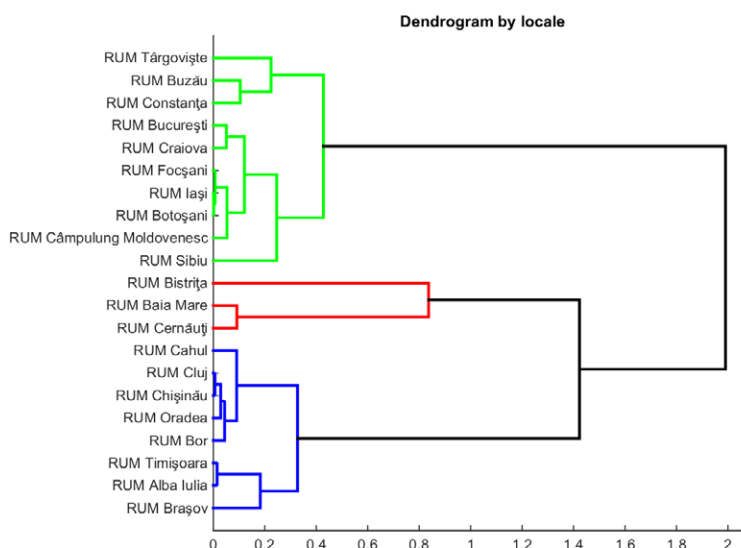
Correlations between the investigation points for the total negative interrogative modality



In the dendrogram that represents the results of the cluster analysis (figure 26) there seem to be three dialectal groups instead of two. However, one of them is very small (it includes just three investigation points: Baia Mare, Bistrița and Cernăuți) and it is linked with the other one at a very low level (1.4), an indication that the two groups are similar.

FIGURE 26

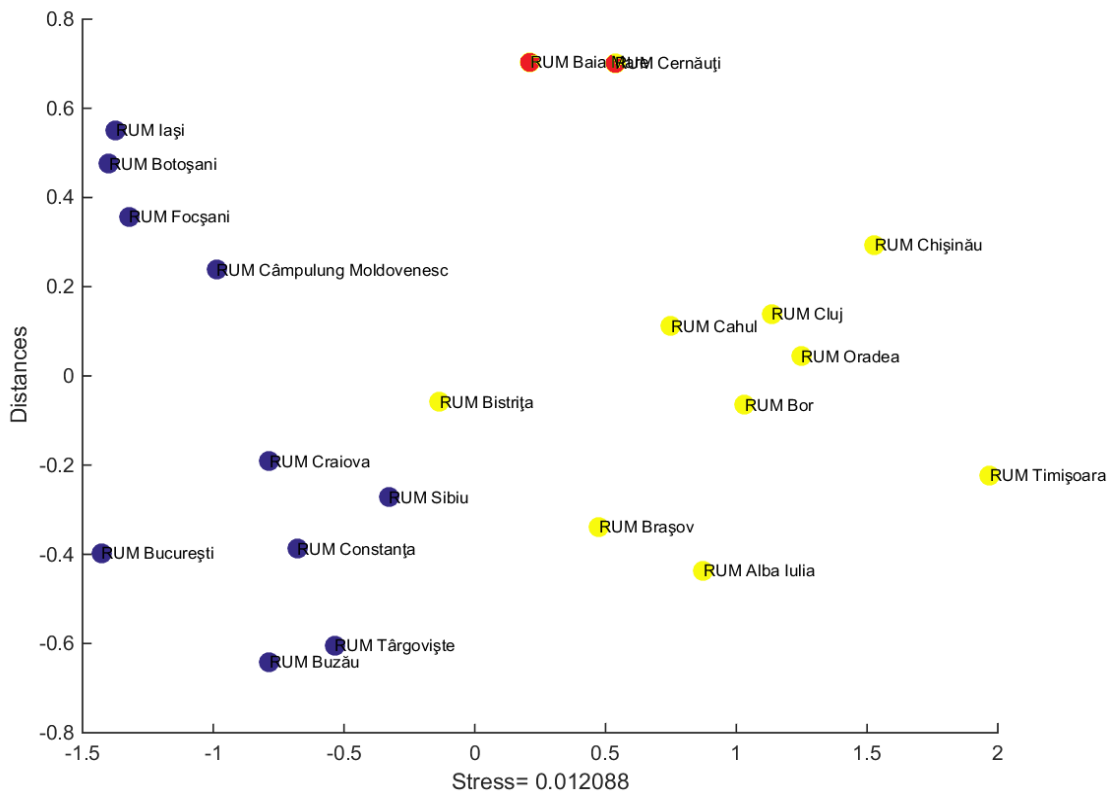
Dendrogram for the survey points for negative yes-no questions



The MDS map (figure 27) also confirms the existence of three groups, the ones appearing in the dendrogram depicted in the previous figure.

FIGURE 27

The MDS map for the investigation points for the absolute negative interrogative modality



The group appearing on the map in figure 23 has a half-moon shape. Noticing the distribution of the three groups on the real map (figure 28), we reach the conclusion that the central group that can be seen in figure 23 (in green) remains compact. This time, when divided in two areas, we can notice the existence of slightly different sub-groups, situated around the Northern part of the border area between Romania and Ukraine. The group under discussion, marked on the map with red, includes three survey points that have been mentioned a few times along this section (Baia Mare, Bistrita and Cernăuți).

4.4. Statements and yes-no questions

The following section comprises the analysis of the data of the four modalities together. Firstly, we can notice that the standard deviation remains high on the whole, mainly due to the variation observed in the analysis of interrogative utterances.

FIGURE 28

Map of the survey points classified according to the results of the cluster analysis: negative yes-no questions

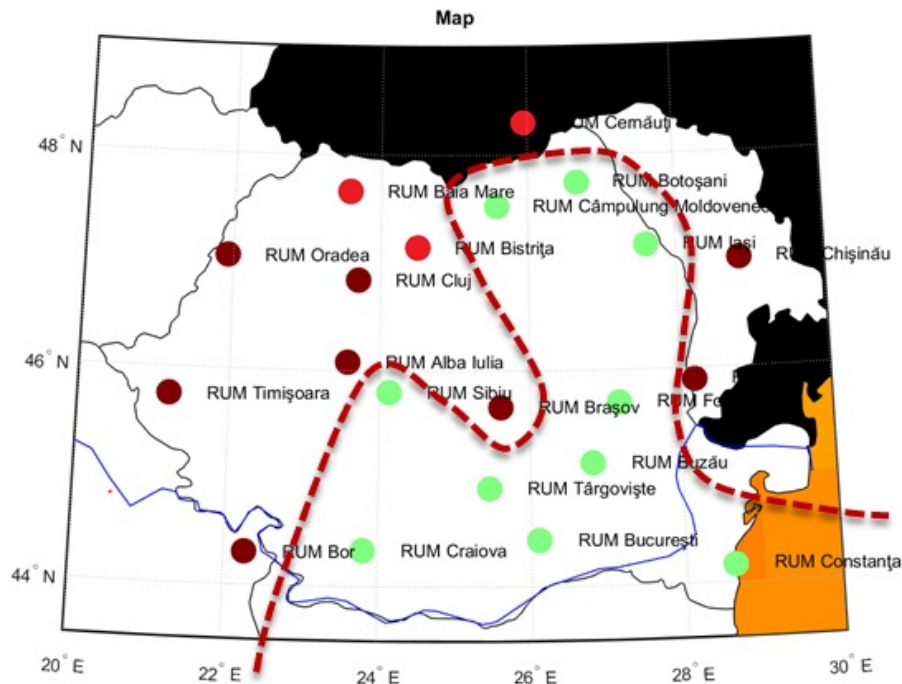
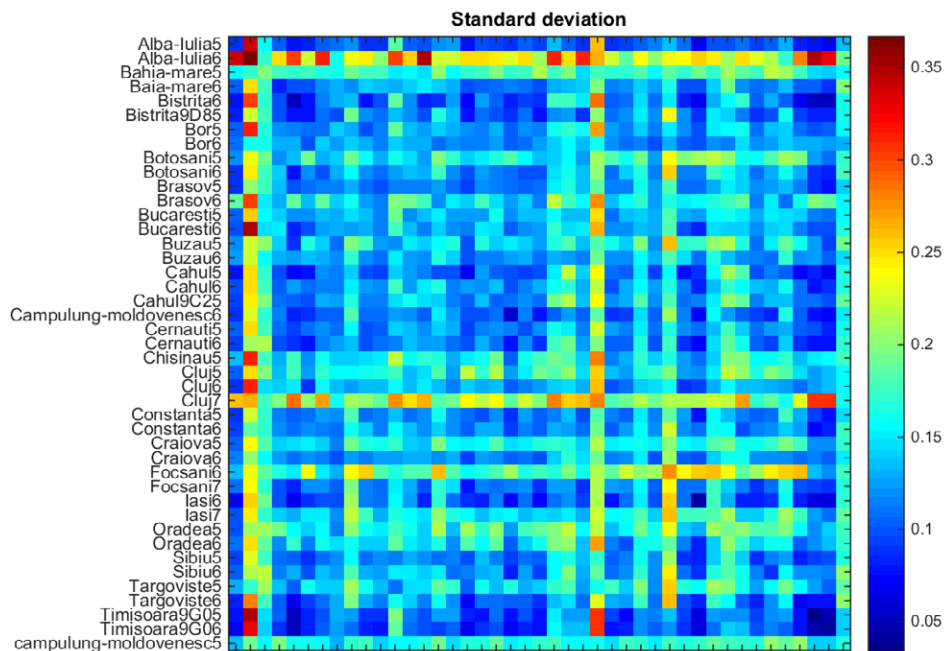


FIGURE 29

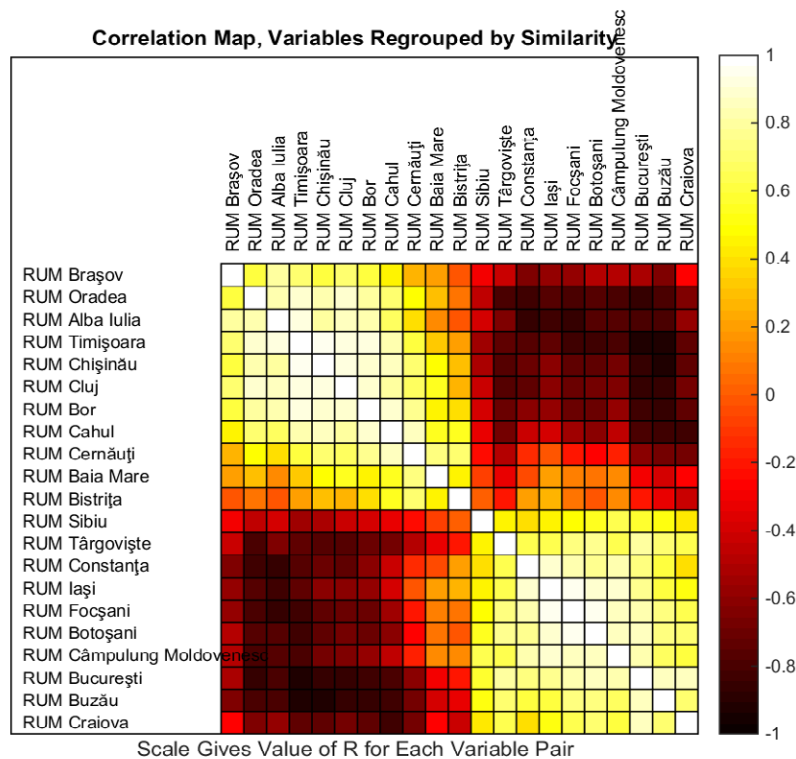
Standard deviation of the four analysed modalities



The analysis of the correlation between the survey points indicates quite obviously, once again, the existence of two groups. Therefore, the three survey points in which negative interrogatives do not fit in either of the two groups (Baia Mare, Bistrița and Cernăuți) are now included in one of the two groups.

FIGURE 30

Correlations between the survey points for the four modalities



The dendrogram (figure 31) reveals the existence of two homogenous groups (that are linked at a level of 1 or lower) that are well differentiated (as they link at the following level, namely level 2). It is thus not surprising that the MDS map (figure 32) reflects this bipartition.

If the results of the cluster analysis were transposed on a real geographic map (figure 33), we could notice the same bipartition that appeared on the map of the affirmative questions (figure 23). This indicates that the affirmative interrogative modality makes a decisive contribution to the definition of the prosodic groups.

5. Discussion

The dialectometric analysis allowed us to observe that, starting from the analysis of the four modalities (affirmative declaratives, negative declaratives, affirmative questions and nega-

FIGURE 31

Dendrogram for the survey points of the four analysed modalities

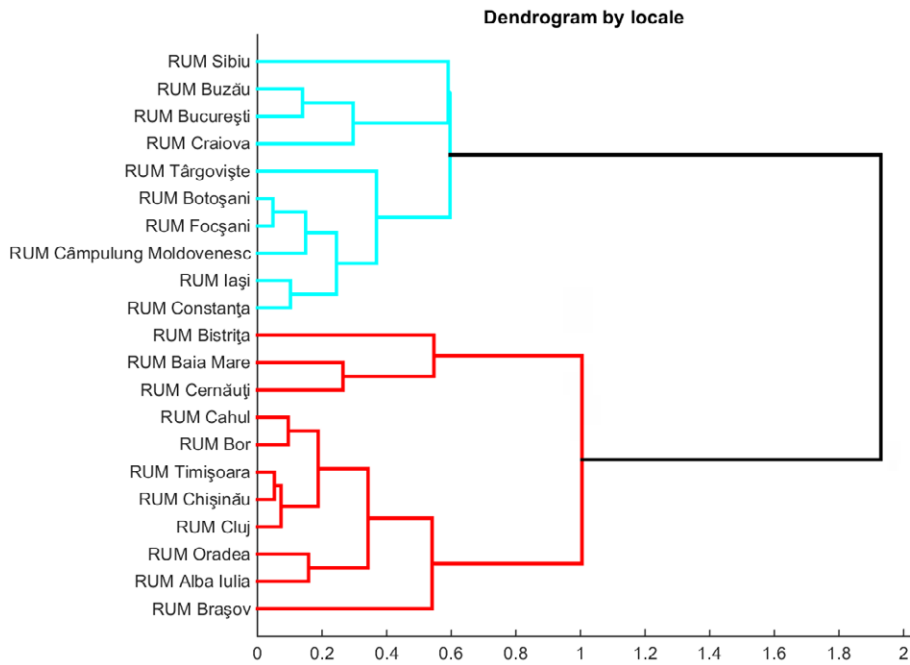


FIGURE 32

MDS map for the survey points of the four analysed modalities

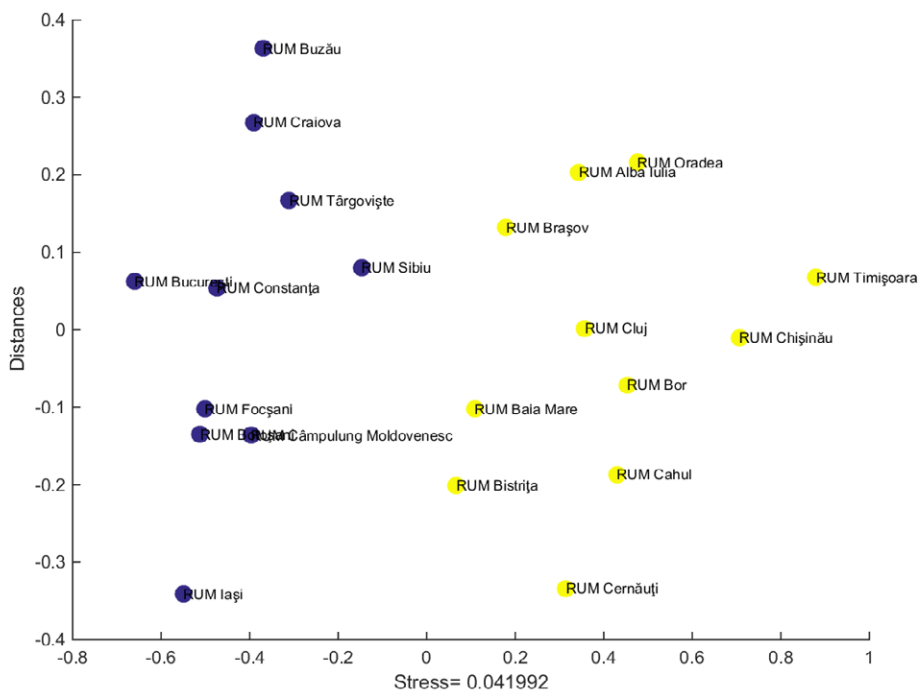
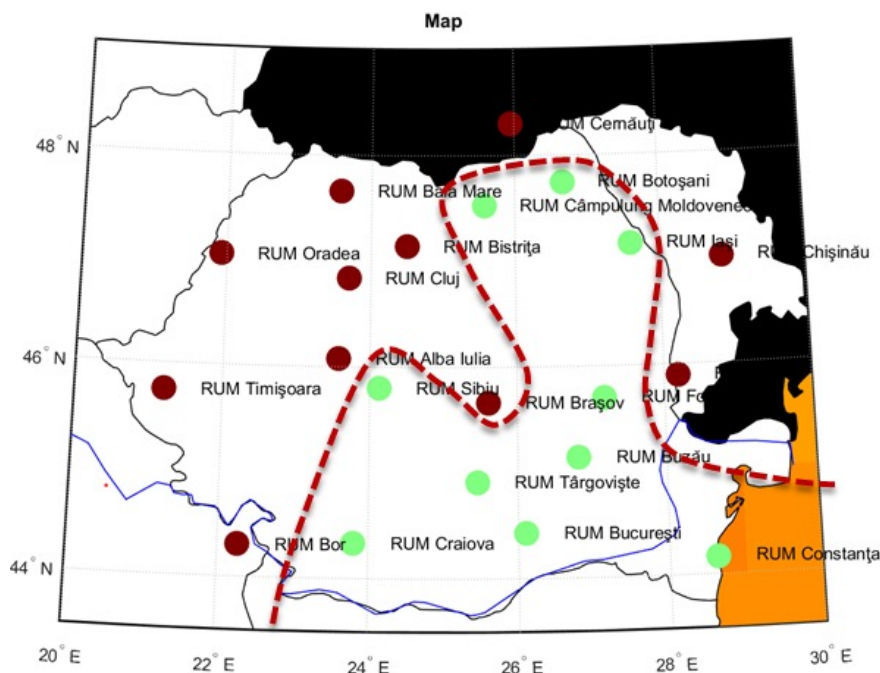


FIGURE 33

Map of the survey points classified according to the results of the cluster analysis of the data provided by the four analyzed modalities



tive questions), the geographic varieties of the spoken standard Romanian are divided in two geoprosodic areas. As previously mentioned, the variable that contributes most to the definition of such two areas is the tune of affirmative questions.

The first of such areas, that can be considered central, includes Craiova, Sibiu, Târgoviște, București, Buzău, Constanța, Focșani, Cămpulung Moldovenesc, Botoșani and Iași. In other words, we can state that we are generally speaking about the historical regions of Oltenia, Wallachia, Dobruja and Moldavia (except for Sibiu, which belongs to Transylvania from a historical perspective). This *central* area displays mainly L_iH* L% pattern for questions (which is present in almost all survey points) and L+H* H%. Moreover, the peripheral pattern (H* L%) does not occur in these survey points.

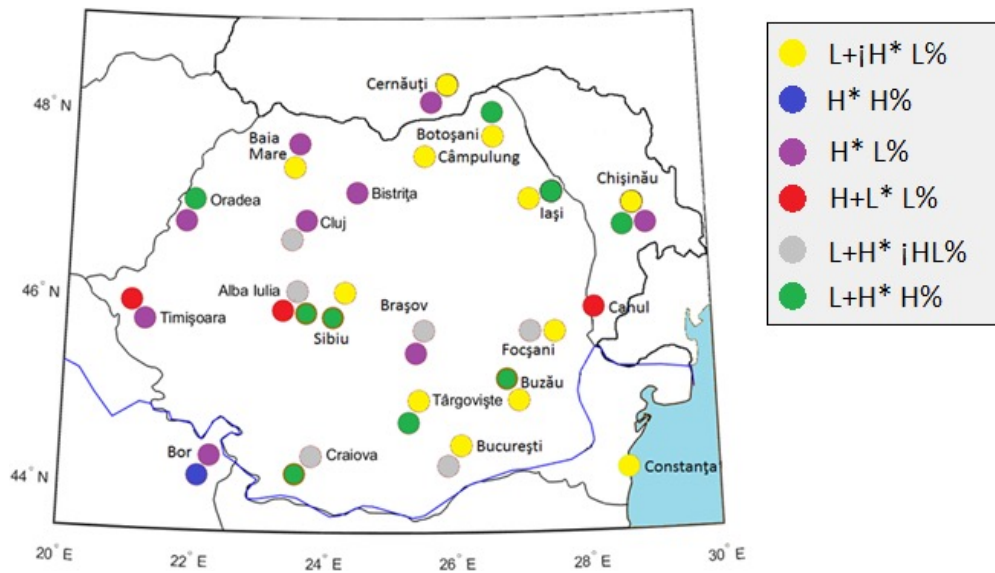
The second area, which we can regard as *peripheral*, includes the localities Bor, Timișoara, Oradea, Cluj, Alba Iulia, Bistrița, Cernăuți, Chișinău and Cahul. These localities belong to the historical regions of Transylvania, Banatului, Crișana, Maramureș, Bukovina and Bessarabia. The peripheral area is characterized by the presence of several intonation patterns, the most typical being H* L%, but also H+L* L% and H* H%.

In other words, the two areas differ insofar as the central one virtually only presents L_iH* L% and L+H* H% questions, while the peripheral area shows—in addition to these two pat-

terns—also H* L%, H+L* L% and H* H%. There is, thus, no one-to-one correspondence between an area and a single intonational pattern, but rather each area is defined by a specific combination of several patterns (see figure 34).

FIGURE 34

Overlap of the areas resulting from the cluster analysis and the geographical distribution of intonation patterns for the question *Pasărea vede nevasta?* (Does the bird see the wife?)



The results of the dialectometric analysis presented above, based on the correlations of the melodic contour computed using Pearson's formula (see § 3), are identical or very similar to those obtained by means of classic geolinguistics. This proves the accuracy of the two types of approaches, with an added value to the dialectometric analysis with ProDis, given that it provides a mathematical complex data analysis, taking into account, in addition to Fo, duration and intensity, and excludes the subjectivity of the researcher (see § 2.4; Elvira-García et al., 2018).

Assuming that the regional varieties of the Romanian standard language reflect the areas of the Daco-Romanian dialects, we can conclude that the geoprosodic areas shown on the maps in figures 28 and 34 prove the conservative nature of intonation, given that the dialectal intonation patterns remain in the current standard speech⁴.

Following the analysis of the diatopic variation of Romanian intonation, two major areas can be identified: a central-eastern one and a discontinuous central-western one, extended in the

4 For other languages, cf. Romano, 2001a, 2001b; Lai, 2002.

eastern part of the Daco-Romanian territory. Comparing this with the traditional classification of the Daco-Romanian sub-dialects in two large groups: the northern (Moldavian) and the southern (Walachian) or in five sub-dialects (see § 2.1) highlights the *independent nature* of geoprosody, the intonation areas versus the classical dialectal areas drawn by the phonetic, lexical and morphological features⁵. In no dialectological classification, the Moldavian sub-dialects are put together with Walachian ones; the Moldavian sub-dialects, those that are different, are surprisingly included in the Transylvanian area. The central-eastern area contains Wallachian sub-dialects and (most of the) Moldavian ones, separating them at the same time from the Transylvanian group. It provides, first of all, an argument supporting the existence of a special group of Transylvanian sub-dialects (Ivănescu, 1944-1945; Șandru, 1938; Țâra, 1975; Frățilă, 1984; Turculeț, 2002; Botoșineanu, 2007; Turculeț & Bleorțu, 2013).

Only assumptions can be issued regarding the age of the Walachian-Moldavian intonation area. Considering, on the one hand, the conservative character of the intonation and, on the other hand, the relative independence of the prosodic areas compared to the ‘classic’ ones, it does not seem a totally risky assumption to date this area in a period prior to the XIV-XVI centuries, when phonetic, lexical and morphological features indicating the limits between Walachian and Moldavian sub-dialects were attested (see § 2.1). Ivănescu (1980: 323-324) considers that ‘the primitive Romanian language’ presents three dialects on the Daco-Romanian territory: 1) Transylvanian, 2) Banat-Crișana-Maramureș-Ardeal area, 3) Moldavian-Walachian.

As concerning the discontinuous area of falling contour, a plausible hypothesis seems to be that ‘Transylvanian’ intonation spread to the northern part of Moldavia (later called Bukovina), to lands between Prut and Nistru⁶.

Less surprising is the fact that Brașov and Sibiu, which are part of a transition area, both belong (in a previous research, see § 2.1), or only Sibiu (after the dialectometric analysis, see figure 28 and 33), to the Transylvanian intonation area. Although the southern and especially the south-eastern part of Transylvania belong to the Walachian sub-dialect, they have some common features with the Transylvanian sub-dialects. The differences between the two analyses on Sibiu are motivated by the low percentage of the falling CT occurrences: 33% - male subject. Instead, in Brașov, the falling CT pattern appears only in the male speaker, but with 49% frequency. Low frequencies of falling CT appear in the western and northern boundaries of the area: Arad (21.3 %), Satu Mare (17.03 %), Baia Mare (21.11%), Cernăuți (27.96 %).

5 A similar discussion for Catalan in Fernández Planas et al., 2015.

6 Philippide (1927) considers that the similarity of the northern subdialects (or ‘Banato-Transcarpathian branch’) is explained by population movement on the route Banat - Transylvania (including Crișana and Maramureș) - Bucovina - Moldova - Basarabia.

6. Conclusions

This paper presents the first dialectometric analysis of intonational variation of standard spoken Romanian. The data analysed were gathered within the AMPROM (part I) project and the statistical analysis has been carried out with the recent software—ProDis—developed at the Phonetics Laboratory of the University of Barcelona. Using ProDis allows researchers to cope with the huge amount of data contained in the corpus (26.680 sentences) and with the considerable intonation variation contained in the data (this is especially true for yes-no questions with broad focus, which show considerable variation between different points of survey and, to a more limited extent, also within the same survey point; see figures 2-13).

The results of statistical dialectometric analysis are consistent with the geolinguistic descriptions of Romanian varieties. As far as the intonation patterns of yes-no questions are concerned, ProDis identified two major areas: a central-eastern one (Moldo-Walachian) with ascending final contours, and central-western one (Transylvanian, discontinuously continuing over Bukovina to Bessarabia—the Republic of Moldova—), which presents, in different proportions, falling contours. In general terms, this study suggests that intonational areas are somehow more conservative of and (relatively) independent from segmental dialectal areas. In addition, the results of this dialectometric intonational analysis show the need to take into account the geoprosody in the characterization and delimitation of the dialectal areas.

7. Acknowledgements

This research has been possible thanks to the projects FFI2015-64859-P, PNII-RU-TE-2014-4-2298 and the Juan de la Cierva program.

8. References

AVRAM, Andrei, 1972: “Intonation curves in the English of Romanians” in Dumitru CHITORAN (ed.): *Romanian–English Contrastive Analysis Project*, Bucharest: Editura Universității București, 7-42.

AVRAM, Andrei, 1973: “Particularități ale intonației interogative în graiul din Muscel”, *Fonetică și dialectologie* 8, 43-64.

BIBIRI, Anca, Oana PANAITE y Adrian TURCULEȚ, 2014: “The Romanian Multimedia Prosodic Atlas” in Fabio TOSQUES (ed.): *20 Jahre digitale Sprachgeographie*, Berlin: Humboldt-Universität zu Berlin, 155-162.

BOTOȘINEANU, Luminița, 2007: *Graiul de pe valea superioară a Someșului Mare*, Iași: Alfa.

CHITORAN, Dumitru, Hortesia PARLOG & James AUGEROT, 1984: *The Sounds of English and Romanian*, Bucharest: Editura Universității București.

DASCĂLU-JINGA, Laurenția, 2001: *Melodia vorbirii în limba română*, București: Univers Enciclopedic.

ELVIRA-GARCÍA, Wendy, Simone BALOCCO, Paolo ROSEANO & Ana M.^a FERNÁNDEZ PLANAS, 2018: “ProDis: A Dialectometric Tool for Acoustic Prosodic Data”, *Speech Communication* 97, 9-18.

FERNÁNDEZ PLANAS, Ana M.^a, Josefa DORTA, Paolo ROSEANO, Chaxiraxi DÍAZ, Wendy ELVIRA-GARCÍA, José A. MARTÍN GÓMEZ & Eugenio MARTÍNEZ CELDRÁN, 2015a: “Distancia y proximidad prosódica entre algunas variedades del español: un estudio dialectométrico a partir de datos acústicos”, *Revista de Lingüística Teórica y Aplicada* 53 (2), 13-45.

FERNÁNDEZ PLANAS, Ana M.^a, Paolo ROSEANO, Wendy ELVIRA-GARCÍA, Ramon CERDÀ-MASSÓ, Lourdes ROMERA-BARRIOS, Josefina CARRERA-SABATÉ, Dorota SZMIDT, Sabela LABRAÑA & Eugenio MARTÍNEZ CELDRÁN, 2015b: “Cap a un nou mapa dialectal del català? Consideracions a partir de dades prosòdiques tractades dialectomètricament”, *Estudios de Fonética Experimental* 24, 257-286.

FRĂȚILĂ, Vasile, 1984: *Studii de dialectologie*, Timișoara: TUT.

GASTER, M., 1891: *Chrestomație română I*, Leipzig-București: Socec & Co.

GHEȚIE, Ion, 1964: “Cu privire la repartitia graiurilor dacoromâne. Criterii de stabilire a structurii dialectale a unei limbi” in *Studii și cercetări lingvistice* 15 (3), 317-346.

GHEȚIE, Ion, 1975: *Baza dialectală a românei literare*, București: Editura Academiei Române.

GHEȚIE, Ion, 1994a: *Graiurile dacoromâne în secolele al XIII-lea – al XVI-lea, până la 1521*, București: Editura Academiei Române.

GHEȚIE, Ion, 1994b: *Introducere în dialectologia istorică românească*, București: Editura Academiei Române.

GODJEVAC, Svetlana, 2005: “Transcribing Serbo-Croatian Intonation” in Sun-Ah JUN (ed.): *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford: Oxford University Press, 146-171.

IVĂNESCU, Gheorghe, 1944-1945: “Problemele capitale ale vechii române literare” in *Buletinul Institutului de Filologie Română “Alexandru Philippide”* XI-XII, 1-412.

IVĂNESCU, Gheorghe, 1976: “Les subdivisions territoriales du roumain” in *Actes du XIIIe Congrès international de Linguistique et Philologie romanes* 2, Quebec: 167-178.

IVĂNESCU, Gheorghe, 1980: *Istoria limbii române*, Iași: Editura Junimea.

- JITCĂ, Doina, Vasile APOPEI, Otilia PĂDURARU & Samuil MARUSCAI, 2015: "Transcription of Romanian Intonation" in Sónia FROTA and Pilar PRIETO (eds.): *Intonation in Romance*, Oxford: Oxford University Press, 284-316.
- MARIN, Maria, & Bogdan MARINESCU, 1984: "Graiurile din Transilvania" in Valeriu RUSU (coord.): *Tratat de dialectologie românească*, Craiova: Scrisul Românesc, 354-390.
- PETROVICI, Emil, 1941: "Graiul românesc de pe Crișuri și Someș", *Transilvania* 72 (8), 551-558.
- PETROVICI, Emil, 1954: "Repartiția graiurilor dacoromâne pe naza Atlasului lingvistic român", *Limba română* III (5), 5-17.
- PETROVICI, Emil, 1961: "Unele tendințe fonetice ale limbii române actuale", *Cercetări de lingvistică* VI (2), 329-337.
- PHILIPPIDE, Alexandru, 1927: *Originea românilor*. Volumul al doilea. *Ce spun limbile română și albaneză*, Iași: Viața Românească.
- PIKE, Kenneth L., 1948: *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language*, Ann Arbor: University of Michigan Press.
- POP, Sever, 1950: *La Dialectologie. Aperçu historique et methods d'enquêtes linguistiques. I: Dialectologie romane*, Louvain: Imprimerie J. Duclot.
- PUȘCARIU, Sextil, 1976 [1940]: *Limba română*, vol. I: *Privire generală*, second edition, București: Minerva.
- RANKIN, L. Robert, 1973: "Romanian intonation and stress" in James E. AUGEROT, (ed.): *Romanian Grammar Pamphlets II*, Arlington, Va.: Center for Applied Linguistics, 3-28.
- ROSEANO, Paolo, & Ana M.^a FERNÁNDEZ PLANAS, 2019: "Anàlisi estadística de les distàncies entonatives: el cas del friülà", *Estudis Romànics* 41, 235-266.
- ROSETTI, Al., 1978: *Historia limbii române. I. de la origini până în secolul al XVII-lea*, second edition, București: Editura științifică și enciclopedică.
- ȘANDRU, Dumitru, 1938: "Enquêtes linguistiques du Laboratoire de phonétique expérimentale de la Faculté des Lettres de Bucarest. District de Năsăud" in *Bulletin linguistique* VI, 173-230.
- TIKTIN, Hariton, 1888: "Die rumänische Sprache" in Gustav GRÖBER (ed.): *Grundriss der romanischen Philologie* I, Strassburg: Karl J. Trübner, 438-460.
- TODORAN, Romulus, 1956: "Cu privire la repartiția graiurilor dacoromâne", *Limba română* V (2), 38-50.

TURCULEȚ, Adrian, 1993: “Variantele regionale ale românei standard”, *Fonetică și dialectologie* XII, 179-198.

TURCULEȚ, Adrian, 2002: *Graiul din zona Câmpulungului Moldovenesc. Fonetica*, Iași: Editura Universității “Alexandru Ioan Cuza”.

TURCULEȚ, Adrian, 2013: “Atlasul multimedia prozodic român (AMPROM)”, *Fonetică și dialectologie* XXXII, 157-172.

TURCULEȚ, Adrian, 2017: “Contribuții prozodice la descrierea și delimitarea varietăților regionale ale românei literare vorbite”, *Anuar de Lingvistică și Istorie Literară* LVII, 365-378.

TURCULEȚ, Adrian, & Cristina BLEORȚU, 2013: “¿Hay una frontera prosódica entre las hablas de Transilvania y las de Moldavia?”, *Estudios de Fonética Experimental* XXII, 81-127.

TURCULEȚ, Adrian, Luminița BOTOȘINEANU, Ana-Maria MINUȚ & Constantin-Ioan MLADIN, 2008: “Aspects de la variation diatopique de l'intonation au niveau de la langue roumaine standard” in Adrian TURCULEȚ, (ed.): *La variation diatopique de l'intonation dans le domaine roumain et roman*, Iași: Editura Universității Alexandru Ioan Cuza, 21-72.

TURCULEȚ, Adrian, & Ana-Maria MINUȚ, 2007: “De la AMPER la AMPROM” in Luminița HOARȚĂ CĂRĂUȘU (coord.): *Rezultate și perspective actuale ale lingvisticii românești și străine. Lucrările Sesiunii de comunicări din cadrul Zilelor Universității*, Iași: Editura Universității Alexandru Ioan Cuza, 349-361.

WEIGAND, Gustav, 1908: “Despre dialectele românești”, *I Convorbiri literare* 42 (4), 441-448.

WEIGAND, Gustav, 1909: *Linguistischer Atlas des dacorumä*, Leipzig: Sprachgebietes.