Differences in prominence perception by German natives and Italian L2 learners.
Acoustic aspects and their implications for SLA

Vincenzo Damiazzi (Milano)

Abstract

This study employs a perceptual and an acoustic analysis of a corpus of spontaneous utterances taken from German podcasts on sustainability. The objective is to identify and examine the differences in the perception of prominence between two groups of participants: a group of adult Italian learners of German as a Foreign Language (GFL) and a group of German native speakers. The results of the analysis lend support to the hypothesis that the differences in prominence perception can be attributed to perception patterns based on duration for Italian native speakers and on pitch variations for German native speakers. In light of the evidence presented, some considerations are made regarding the potential applications of acoustic analysis in GFL courses. The potential applications of this approach are investigated, based on a simplified visualisation of prosody, by means of F0 curves enriched with information on the periodic energy of the acoustic signal.

1   Introduction

The acquisition of correct intonation is a vital component in SLA (Second Language Acquisition), as it plays a crucial role for effective communication and for the overall comprehension of a foreign language. Despite the inclusion of pronunciation training in many language textbooks, it is frequently overlooked in practice. The advent of the so-called prosodic turn (*prosodische Wende*) saw a shift in focus towards the accentual and intonational aspects of the foreign language at all stages of acquisition. Subsequently, the original Common European Framework of Reference for Languages (CEFR) published in 2001 provided a detailed description of the phonological aspects to be used in SLA. However, the 2001 CEFR phonology scale appeared to imply that advancement in proficiency was analogous to the attainment of a native-like accent. This perspective failed to take into account a crucial element of language proficiency, namely intelligibility. Indeed, this is a fundamental aspect of communication, and arguably more important than achieving a native-like accent. This approach based on nativeness was subsequently challenged in the companion volume to the CEFR, which states that "the focus on accent and on accuracy instead of on intelligibility has been detrimental to the

development of the teaching of pronunciation" (European Council 2020: 133). Consequently, the focus should shift towards enhancing phonological control.[1]

Learning correct pronunciation is crucial and, in fact, the lack of adequate prosodic training is the primary cause of the so-called "foreign accent" as highlighted by Cruz Ferreira (1989).[2] However, there is another aspect that needs to be taken into account, namely the correct perception of L2 sounds and prosody.[3] While it is true that learners can produce the sounds of a foreign language correctly by chance, it is not possible to perceive correctly at random. It is therefore equally important to include perception training as part of the curriculum.

Difficulties in perceiving the intonation features of a foreign language are closely linked to the presence of a foreign accent and a lack of phonological control (cf. de Bot/Mailfert 1982). Our ears function like filters, and the so-called phonological filter, as described by Trubetzkoy in 1939, prevents the accurate perception and production of L2 sounds. This filter is formed by the phonological system of the native language, which influences how learners perceive and categorise sounds in a new language. For instance, Italian speakers learning German might struggle with the phonetic differences in vowel length and pitch that are crucial in German, leading them not just to a foreign accent but also to difficulties in the perception of these differences.

One specific instance of prosodic differences in the Italian-German language pair relates to the perception of prominence, that is the correct recognition of salient syllables within a word (lexical level) or salient words within an utterance (syntactic level). At the syntactic level, decoding prominence through the correct perception of prominence is crucial in order to understand the meaning of utterances and for the so called *Aufmerksamkeitssteuerung* ('attention control'). This term refers to the function of the accent of redirecting attention towards the most relevant element within the utterance from a communicative point of view, i. e. the word that carries a pragmatic-communicative value. Correct prominence perception and production is therefore fundamental in order to also communicate efficiently in German.

These circumstances underscore our intention to explore the differences in prosodic perception by native and non-native speakers. Through the observation and analysis of the differences in prominence perception between Italian learners of German and German native speakers, we

---

[1] The concept of phonological control refers to control of both individual sounds and prosodic features, and also takes into account the extent of influences from other languages spoken by the learners. When articulating the sounds, the degree of familiarity with the sounds of the target language and the precision with which they are articulated are taken into account. With regard to the control of prosodic features, the ability to use prosody to convey meaning is explicitly emphasised, in particular the precision with which pitch accents, intonation and rhythm are used, as well as the ability to adapt prosodic features according to the communicative intentions or the type of message the speaker wants to convey.

[2] This phenomenon is evident across various language pairs, including Italian as L1 and German as L2 or L3, affecting all phonetic levels (cf. Missaglia 2018).

[3] In this contribution we use *prosody* as a synonym for intonation in its broad definition. The term refers to all phenomena that operate at the suprasegmental level, i. e. above the level of segments, including rhythm, accent, stress, intonation contours and, more generally, variations in pitch, loudness and duration that affect the lexical and syntactic levels.

intend to identify diverging perceptual patterns. The acoustic data will inform future research and yield implications for GFL phonetic courses that focus on German prosody.

2   Background

The perceptual difficulties experienced by Italian learners of German can be attributed to two main factors. The first factor is the mismatch between phonological and phonetic perception, that is between the acoustic and auditory level. In fact, while the correlation between articulation and acoustics is direct (a higher rate of vocal fold vibration leads to a higher pitch, which can be measured acoustically in an increase of F0 or fundamental frequency), perception is often not correlated with acoustic measurements. Acoustic measurements record what has been realised through the articulatory system, but that is not sufficient to take into account how the perception of specific sounds or intonation patterns may differ between speakers of different languages (cf. Missaglia 2012, 2021, 2023; Blühdorn 2013).[4]

The second factor is related to how a language's rhythm influences phonetic and prosodic perception and, consequently, also phonetic and prosodic production. While rejecting the theoretical assumption of an acoustically measurable isochrony as postulated by Abercrombie (1964), it is possible to observe specific differences in rhythm, syllable structure, intersegmental processes, prominence and accenting and deaccenting patterns between the so-called syllable-timed Italian language and the stress-timed German language. One of the differences at the suprasegmental level that must be considered (also for GFL courses) concerns the acoustic correlates of accent realisation. Whereas in Italian accents are mainly realised through changes in duration, namely through vowel lengthening, German is characterised by pitch accents (*Tonhöhenakzent*) that are realised through changes in pitch and intensity. Another difference at the suprasegmental level concerns stressed and unstressed syllables. In German acoustic prominence is generally placed on stressed syllables while unstressed syllables are reduced and acoustically less prominent. In Italian syllables are less complex and there is far less difference between stressed and unstressed syllables and therefore also in their degree of prominence.[5]

Consequently, in order to correctly produce the foreign sounds and prosody, it is necessary to be able to perceive them correctly. However, textbooks often present only a series of prototypical prosodic patterns not considering the fact that in authentic speech prosody is extremely varied and complex (cf. Baumann/Niebuhr/Schroeter 2016; Damiazzi 2022). Prosodic variation in authentic speech presents a significant challenge for correctly acquiring prosody. Therefore, a shift in focus is required and the communicative relevance of prosody must take centre stage (cf. Atoye 2005; Alter et al. 2001). To address this, we propose to use podcasts, in order to enrich prosodic training in German as a foreign language (GFL). These resources provide

---

[4] An example of this mismatch between Italian learners of German and German natives lies in the different perception of long and short vowels. Italian learners have great difficulties in identifying and discriminating between the 15 German vowels in stressed position, which are all phonemes in German but not in Italian (that only has 7 vowel phonemes).

[5] For an overview of the rhythmic features of stress-timed and syllable-times languages cf. Missaglia 1999: 47; Pettorino/Pellegrino 2016: 13–28; Bertinetto 2021.

authentic examples of prosody in use, reflecting the natural variability and complexity of spoken language.

2.1   Using (semi-) spontaneous speech and podcasts

The selection of spontaneous or semi-spontaneous material[6] in the context of SLA is not a straightforward process because there is a need to find a common ground between the need for easy and ready-to-use renderings of day-to-day spoken interactions and the varied and complex realisations found in real conversation. Granted that learners have individual needs and that it is virtually impossible to train for the complete variation found within spontaneous speech in a foreign natural language, the advantages of using spontaneous and semi-spontaneous speech are substantial.

One of the primary benefits is authenticity. Spontaneous speech captures the genuine variability and complexity of language, including natural variations in prosody, intonation, and rhythm. This prosodic richness is often lost in controlled, laboratory settings. Labov (1972) and Milroy (1987) emphasise that the natural ebb and flow of spontaneous speech provide a true representation of how language is used in everyday interactions, making it invaluable for acquiring the nuanced features of prosody.

Another critical advantage is ecological validity. Natural materials, such as spontaneous speech, reflect real-life language use more accurately than scripted or laboratory-generated speech. This authenticity means that the findings from studies using spontaneous speech are more likely to be applicable to real-world contexts (cf. Bucholtz/Hall 2005). In contrast, laboratory settings can create artificial environments that do not fully capture the dynamics of everyday communication, thus leading to conclusions that may lack external validity.

The use of spontaneous speech also results in the collection of data which is less prone to the observer's paradox. This paradox occurs when participants alter their behaviour because they are aware that they are being studied. Spontaneous speech, collected in natural settings, mitigates this issue, providing a more genuine insight into how people speak (cf. Johnstone/Andrus 2024). This authenticity is crucial for researchers aiming to understand the true nature of language use, free from the distortions introduced by artificial observation conditions.

Additionally, spontaneous and semi-spontaneous speech capture the inherent complexity and variability of natural language. Laboratory settings sometimes simplify language to control variables, but this simplification can strip away the rich, multifaceted nature of real language use. Couper-Kuhlen/Selting (2018) argue that the complexity found in spontaneous speech includes a range of linguistic features – such as hesitations, repairs, and overlaps – that are crucial for a full understanding of language dynamics. These features are often omitted or underrepresented in laboratory studies, leading to an incomplete picture of language use.

---

[6] The distinction between spontaneous and semi-spontaneous language or utterances is about the level of planning, preparation, and naturalness involved in producing them. Spontaneous language is produced naturally without prior planning. Semi-spontaneous language involves some level of planning or preparation and is often based on prior thought or structure. In podcasts it is possible to find examples of both of speech types (cf. Labov 1972; Tannen 1989; Chafe 1994).

Practical application is another significant advantage of using (semi-) spontaneous materials. Insights gained from natural language data are more readily applicable to language learning and teaching. Understanding language in use can inform more effective teaching methods and materials. When learners are exposed to the variability and complexity of spontaneous speech, they are better prepared for real-world communication, as opposed to the more sanitised and simplified language often found in textbooks (cf. Ellis 1994; Long/Doughty 2011).

Podcasts represent an exemplary instance of material that incorporates elements of both spontaneous and semi-spontaneous speech. In terms of the degree of spontaneity, they occupy a position that is midway between the poles of unedited and unplanned speech and prepared and scripted speech. As a consequence, podcasts can be a highly beneficial tool in the context of SLA, as they facilitate the utilisation of natural language, without the inherent challenges associated with the high variability of spontaneous speech structures and the dysfluencies that they are characterised by.

Moreover, podcasts are a media genre intended to resonate with the subjective perspectives of the listeners and to tie them in with the narrator's perspectives, thus contributing to the degree of naturalness of the proposed material (cf. Kalch/Schlütz 2022) and therefore coming closer to the reality of the learners.

2.2   Aims of the research

The pilot study described here constitutes the first phase of a larger study aimed at analysing the perceptual patterns of non-native speakers and at assisting Italian learners of German as a foreign language (GFL) in accurately perceiving and interpreting authentic German utterances. In light of the aforementioned background, three aims are identified for further investigation in this contribution:

a) to investigate prominence perception by Italian GFL learners and German native speakers;
b) to analyse language-specific perceptual patterns and investigate the link between the phonetic (acoustic) nature of prominence and its perception;
c) to identify a way to visualise the cues that steer prominence perception in the two groups and to explore how this visualisation could be used in SLA.

These aims underscore our intention to explore how prosody is perceived by native and non-native speakers, to analyse, on an acoustic basis, cases of perceptual differences between the two groups and to reflect on the didactic significance of using a simplified visualisation of these differences to train learners in the correct perception of prominence in German.

3   Methodology

In line with the aims of this research, a corpus of target sentences was compiled and utilised for both a perception experiment and an acoustic analysis. The methodology employed in the compilation of the corpus, the design of the perceptive experiment and the selection of correlates for the acoustic analysis will now be outlined.[7]

---

[7] For further details on the perception experiment and acoustic analysis cf. Damiazzi (in press).

3.1   Corpus

The corpus consists of 20 utterance phrases (UPs) taken from 5 German-language podcasts on the macro topic of sustainability[8], namely *ZEIT für [KLIMA]* ("Kann die Kultur das Klima retten?", broadcasted on 17.11.2021; "Wie treiben wir die Energiewende richtig voran?", broadcasted on 31.03.2022), *MDR-Investigativ* (2023a: "Kohleabbau – ist Mühlrose das neue Lützerath?", broadcasted on 08.03.2023; 2023b: "Grüner Wasserstoff aus Afrika – Energie der Zukunft?", broadcasted on 13.01.2023) and *Kemferts Klima-Podcast* ("Wie viele Autos darf es in Zukunft geben?", broadcasted on 23.02.2023). The topics covered are the future of car driving, green hydrogen resources in Africa, the coal industry in Germany, the role of culture in climate change and strategies to promote the energy transition. The 20 UPs have a duration between 2 and 4 seconds and are divided into 10 pairs, each pair having a specific communicative intention. The UPs are realised by 8 speakers.

3.2   Perception experiment

The perception experiment aims to examine the differences in prominence perception between Italian GFL learners and native speakers of German. The experiment was therefore conducted with two groups. The first group comprised 37 Italian-speaking learners in their first year of a bachelor's degree in foreign languages at the Università Cattolica in Milan. All learners were beginners in German.

The second group consisted of 11 native German-speaking lecturers from the Università Cattolica and the University of Cologne. Data collection was carried out using a questionnaire hosted on Google Forms. Each utterance phrase (UP) was presented both as an audio file and in text format. Participants were asked three questions for each UP. For this contribution the focus will be on question a).

a)   Which word carries the main accent in this sentence?
b)   What communicative intent do you attribute to this sentence?
c)   Which element was most helpful in defining the communicative intent of the sentence?

The perception experiment with the Italian-speaking GFL learners was conducted in a supervised setting. The questionnaire was administered on-site at the university under the supervision of an instructor, ensuring that each utterance was heard no more than three times (one for each of the questions in the questionnaire). In contrast, the native speakers of German completed the questionnaire online, without supervision.

3.3   Acoustic analyses

After the perception experiment, we intended to explore the correspondence between prominence identification and acoustic correlates in the speech signal. Thus, two acoustic analyses were carried out: one was performed using Praat (cf. Boersma/Weenink 2025) and the other one with the ProPer (PROsodic analysis with PERiodic energy) package (cf. Albert et al. 2024; Albert/Cangemi/Grice 2018; Albert 2022).

---

[8] For the complete list of UPs in the corpus see Appendix.

First the utterances were analysed with Praat (figure 1): for all UPs the tones and degree of prominence were annotated on a pre-phonological level according to the guidelines of the German prosody annotation system DIMA (cf. Kügler/Baumann/Röhr 2022). Secondly, a phonological analysis of the intonation pattern was carried out in order to identify accents in nuclear and pre-nuclear position as well as their functions within the utterance (i. e. focus, emphatic or contrastive accents).
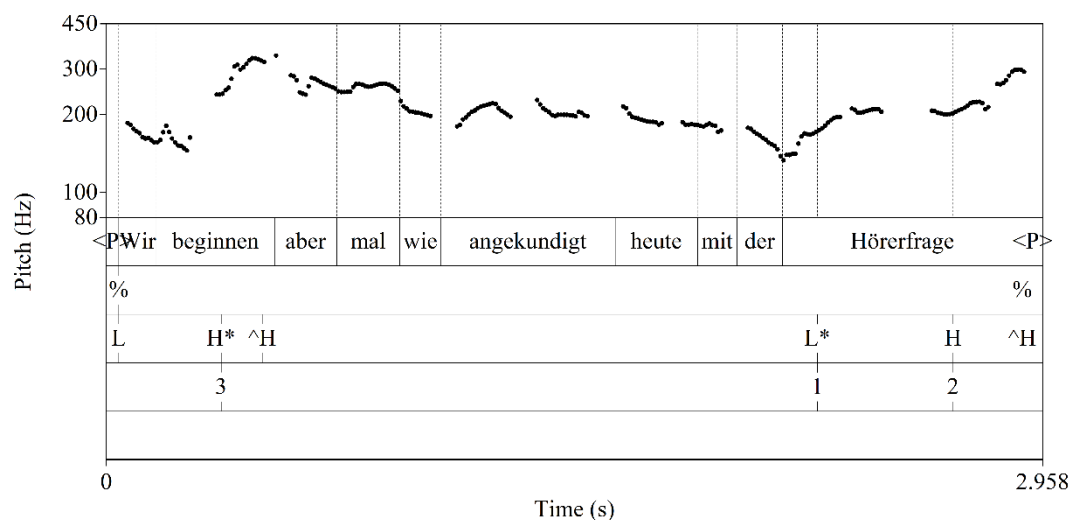

Figure 1: Analysis of UP1 with PRAAT and DIMA annotation

The second analysis was carried out with ProPer, a script package developed at the University of Cologne for use with the R software (R Core Team 2023). The aim of the ProPer analysis is to identify the correlation between the perception of prominence and a series of novel correlates, which are multidimensional in nature. Specifically, ProPer allows users to visualise so-called periograms, i. e. F0 curves enriched with information regarding periodic energy and periodic energy mass. Periodic energy is a "measurement of the acoustic power of periodic components in the signal" (Albert 2022: 55), whereas periodic energy mass measures the strength of each syllable and is "the integral of duration and power, which is the area under the periodic energy curve" (Albert 2022: 146). Therefore, the correlate of mass is multidimensional in nature because it "accounts for duration and power together in a single variable that attempts to capture the overall prosodic strength" (Albert 2022: 149).

Another correlate measured by ProPer is the ΔF0 (Delta F0). Although not multidimensional, it measures the variation of F0 between one syllable and the preceding syllable, allowing for the identification of salient F0-movements (both rising and falling).

The advantage provided by ProPer in addition to Praat is the possibility of visualising multidimensional correlates of the acoustic signal. In the context of the perception of German by Italian speakers, this implies the introduction and subsequent analysis of an additional prosodic correlate (in addition to F0, duration and intensity), namely mass, on which differences in perception between native and non-native speakers could be based.[9] Moreover, while Praat permits the

---

[9] For a more in-depth overview of the use of ProPer for prosodic research cf. Sbranna et al. 2023.

one-dimensional visualisation of the F0 curve, ProPer enables a multi-dimensional visualisation of the acoustic signal through periograms. This allows for the representation of not only the F0 pattern (represented by the curve progression), but also of duration and intensity (through the mass correlate, which is represented as the thickness of the curve).
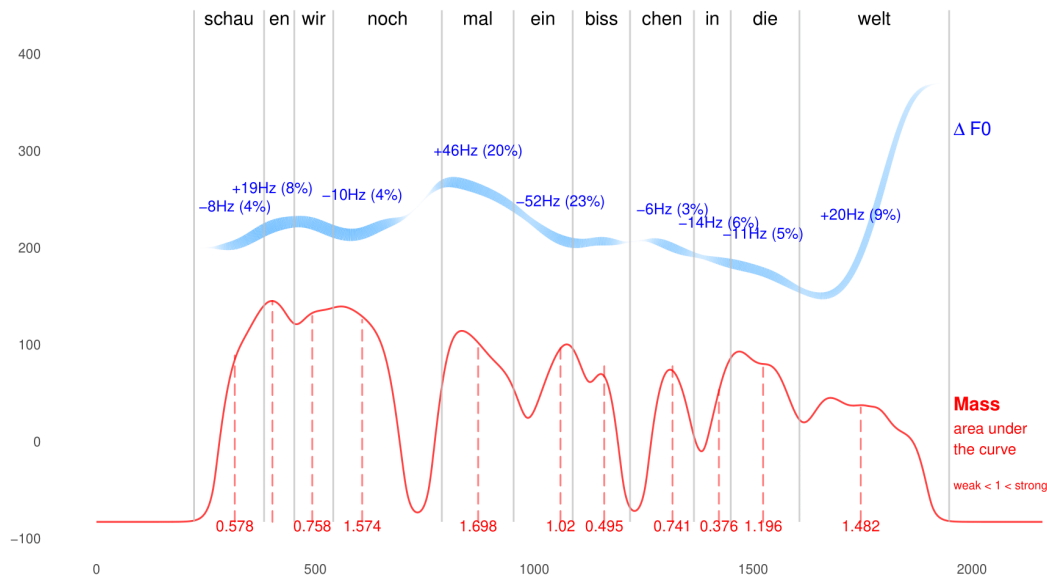

Figure 2: Visualisation of UP6 through ProPer

Figure 2 shows an example of a UP as analysed with ProPer. At the top, the utterance is segmented in syllables. The blue line is the periogram, an F0 curve enriched with information on periodic energy (the line is thicker when the periodic energy is greater). Above the blue line the positive and negative ΔF0 values (in Hz and in percentages) are indicated and they show the change in F0 between syllables by calculating the difference from the previous syllable. The red line represents the periodic energy, whereas the area under the red curve shows the periodic energy mass. The bigger the area under the curve the more salient the prosodic strength[10] of the corresponding syllable. Values above 1 indicate high prosodic strength and values under 1 weak prosodic strength.

4   Results

UP	Accent – Group 1	Accent – Group 2	F0 max.	Mass max.	ΔF0 max.
UP1	beginnen (57)	beginnen (60)	be**ginnen**	mal	be**ginnen**
UP2	zugelassen (90)	zugelassen (90)	**zu**gelassen	zuge**las**sen	**zu**gelassen
UP3	große (86)	richtig (64)	also	schon	**rich**tig
UP4	Die₁ (62)	Heimat (82)	**Hei**mat	**Hei**mat	**Hei**mat
UP5	Ihnen (49)	Klick (80)	wann	Ih**nen**	wann

---

[10] The term *prosodic strength* is defined as a summation of duration and power in a single variable (cf. Albert 2023: 149).

UP	Accent – Group 1	Accent – Group 2	F0 max.	Mass max.	ΔF0 max.
UP6	Welt (96)	Welt (60)	Welt	noch**mal**	noch**mal**
UP7	Netzwerks (22) Netzwerk (20,5)	Netzwerk (40)	ist	sein	**Netz**werk
UP8	sein (55)	sein (50)	sein	sein	sein
UP9	egal (81,5)	egal (60)	egal	egal	egal
UP10	drin (76)	drin (60)	drin	drin	drin
UP11	viele (77)	viele (70)	**vie**le	haben (ham)	**vie**le
UP12	Teilmenge (57)	Teilmenge (40)	sie**ben**	Teil**men**ge	sie**ben**
UP13	Gesetze (94)	Gesetze (90)	Ge**set**ze/ schreibt	wer	schreibt
UP14	sprechen (58)	sprechen (50)	spre**chen**	dar**über**	be**vor**
UP15	haben (70)	haben (100)	ha**ben**	ha**ben**	ha**ben**
UP16	Sie (31) schätzen (29) Zusammenarbeit (24)	Zusammenarbeit (50)	wie	Sie	wie
UP17	Daten (96)	Daten (100)	Da**ten**	Be**reich**	Da**ten**
UP18	schon (90)	schon (100)	schon	ge**nom**men	schon
UP19	Kohle (81)	Wasser (46) Kohle (27) Zukunft (27)	Was**ser**	Was**ser**	Was**ser**
UP20	Straße (98)	Straße (100)	**Stra**ße	**kle**ben	**Stra**ße

Table 1: Combined results of the perceptive and acoustic analyses

Table 1 shows the combined results of the perception experiment and the acoustic analysis. For each UP the word and the syllable(s) (**in bold**) are presented where the maximum value of either F0, mass or ΔF0 are placed within the utterance. For ΔF0, only the positive slopes were considered, i. e. sharp F0 rises respective to the preceding syllable. As far as the perception experiment is concerned, Table 1 also shows the main accent in the UP as indicated by the absolute or relative majority of group 1 (Italian-speaking GFL learners) and group 2 (native speakers of German). The percentage of test subjects that make up the absolute or relative majority is also indicated in brackets next to each word.

The two groups indicated the same word as the main accent in the vast majority of utterances, namely in 70% of the UPs (14 out of 20). In 4 UPs (3, 4, 7 and 19) the choice of the Italian learners of German did not have any correspondence with any of the analysed acoustic correlates. In turn, for the native speakers this was the case in 2 of the UPs (5 and 16). As stated in 2.2, one of the aims of this research was to analyse the cases in which the perceptions of the

two groups diverge. Only these UPs are considered since in cases where the choice of the main accent is consistent between the two groups, it can be assumed that the indication of the main accent depends not only on the correct perception of prominence, but also on other incidental factors such as information structure, focus accent position[11], syntactic structure, etc. Cases of perception mismatch engender greater consideration as they highlight instances of different perceptual preferences that may be an interesting stimulus to enrich prosodic and perception training in SLA.

A diverging perception of prominence can be observed in 6 UPs (3, 4, 5, 7, 16 and 19). In these cases, the two groups associate prominence with different acoustic cues, specifically with a high value of mass for Italian GFL learners and with a high value of ΔF0 or a salient movement of F0 for native speakers of German. By basing their perception on a high mass value, Italian learners are in fact selecting two variables, i. e. two suprasegmental correlates, as the basis of their perception: duration and intensity. This is analogous to the behaviour observed in their native language. On the other hand, native German speakers select only one variable as the basis for their perception of prominence, namely F0 peaks. UP3 (figure 3) and UP19 (figure 4) are two examples of these perception patterns.
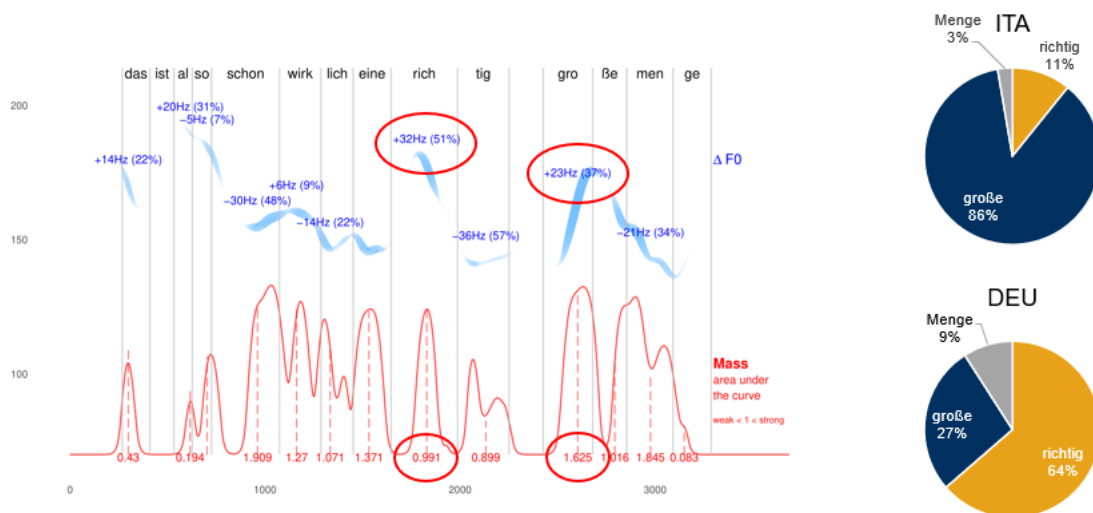

Figure 3: Acoustic and perceptive analysis of UP3

In UP3 (figure 3) it can be observed that even in the case of similar ΔF0 values (and therefore similar rising movements of F0) on the stressed syllable of the words indicated as the main accent in the utterance by the two groups, the Italian learners associate their perception of prominence to the syllable with the highest mass value (*große*). The choice of the German native speakers seems to be cued solely by pitch (on both syllables in *richtig* the mass value is below 1 which signals a weaker prosodic strength).

---

[11] Focus accent is defined as "the semantic-pragmatically most relevant actually phonetically highlighted accent of the intonation phrase, which indicates the semantic-pragmatic focus of the utterance and highlights it from the background" (Selting et al. 2009: 371).
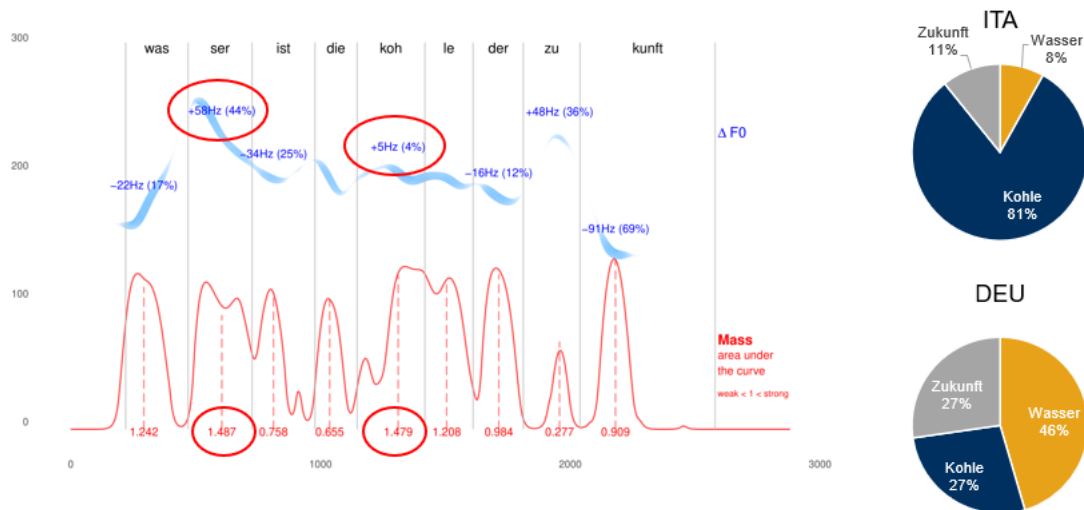
Figure 4: Acoustic and perceptive analysis of UP19

In UP19 (figure 4) it can be evidenced that the words indicated as main accent by the two groups have similar mass values. In this case, the perception of prominence in German native speakers is cued by the presence of a syllable with a high ΔF0 value (*Wasser*) whereas Italian-speaking GFL learners seem to disregard salient F0 movements and indicate a word with a mass value above 1 but a rather flat intonation pattern.
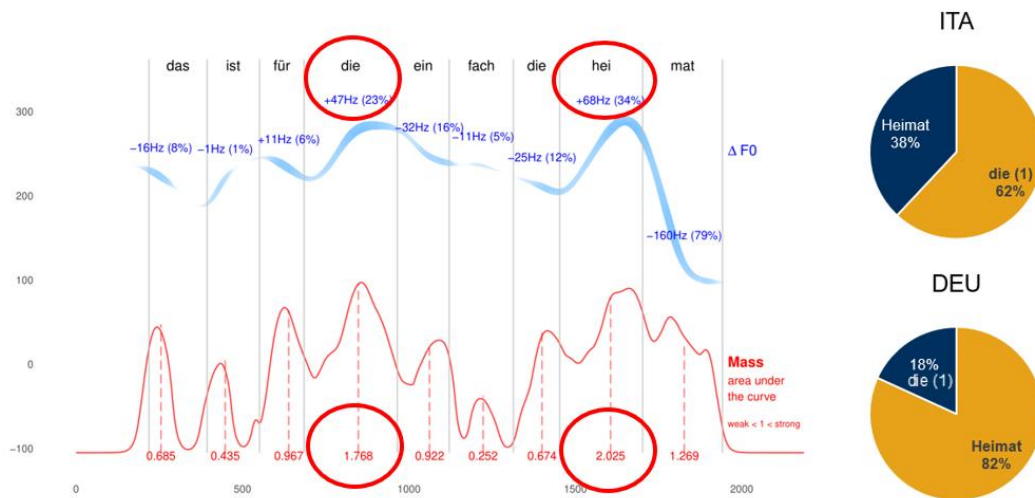

Figure 5: Acoustic and perceptive analysis of UP4

Additionally, UP4 (Figure 5) presents an illustrative example wherein the stressed syllable of the word indicated as the main accent by the majority of German native speakers is associated with a high mass value (above 2 in **Hei**mat) – a value that is higher than that observed on the word indicated by Italian learners (the monosyllabic *die*). Yet, the key aspect in the perception of prominence of German natives is the presence of rise-fall F0 pattern, i. e. a rising movement (the highest rise respective to the preceding syllable) followed by a sharp fall. Italian learners do not ascribe as much prominence to this rise-fall nuclear contour as the German natives do.

The utterances presented here represent individual cases of the different relationship between the auditory and acoustic levels that governs the perception of prominence in the two groups. The examples provided represent an initial effort to examine the perceptual differences between native and non-native speakers and they support the theses about different perception parameters and different prosodic prominence distribution between syllable-timed Italian and stress-timed German (see par. 2). Indeed, they confirm the evidence that, in cases of divergence of perception from German native speakers, Italian learners rely on mass (i. e. duration and power) as they do in their mother tongue.[12] Moreover, they do not rely on F0 peaks which, in turn, is the correlate that German native speakers rely on.

5   Teaching-related aspects

In light of the experimental data gathered through the combined perception experiment and acoustic analysis, a reflection can be made on the utilisation of acoustic measurements and the integration of relevant acoustic parameters and simplified visualisations of prosody into GFL classroom practice for the correct perception of prominence in German. Indeed, if through the auditory channel, Italian-speaking learners are unable to select the F0 peak variable in order to perceive prominence in German, it may be useful to also use the visual channel to train correct perception.

To address this, any approach should aim at developing the learners' prosodic awareness, which is essential when confronted with authentic speech. Teachers need to underline that what might initially seem like "erroneous" perception in the acquisition of a foreign language is actually a natural part of the learning process. The perception differences between natives and non-natives are not faults but innate processes that, with targeted instruction and practice, can be corrected (cf. Flege 1995; Munro/Derwing 2008). With proper guidance, learners can retrain their auditory perception to better align it with that of German native speakers. To facilitate this, we advocate for a prosodic training that makes use of periograms.[13] These visualisations aids can help making abstract prosodic features more concrete. In other words, they make visible through the eyes what the ears fail to perceive. By integrating visualisations of prosody into the learning process, learners can be made aware of the differences in the perception of prominence in German and all the while enhance the overall L2 communicative competence (cf. Lewis 1999; Trofimovich/Baker 2006).

5.1   Suggestions for perception-based prosodic training

A teaching method based on periograms is presented here. This teaching method is designed for learners with different levels of proficiency in German, starting from a lower intermediate level (A2). The utterances are presented in isolation to focus on the prosodic structures and learners are shown a simplified periogram (see figure 6), i. e. a periogram devoid of superfluous

---

[12] Evidence for the Italian-German pair already indicates that perception patterns are transferred from the L1 to the L2/L3 (cf. Rabanus 2001; Niemann et al. 2011; Avesani et al. 2015; Missaglia 2019).

[13] For more details on different methods to visualise prosody and their use in SLA/TLA cf. Damiazzi 2020 and Niebuhrr/Fischer/Schümchen 2017.

information for learners such as ΔF0 or mass values. Learners are only presented with the periogram and the periodic energy curve. These two elements and their correlation are made evident in the following steps for phonetic training (table 2).
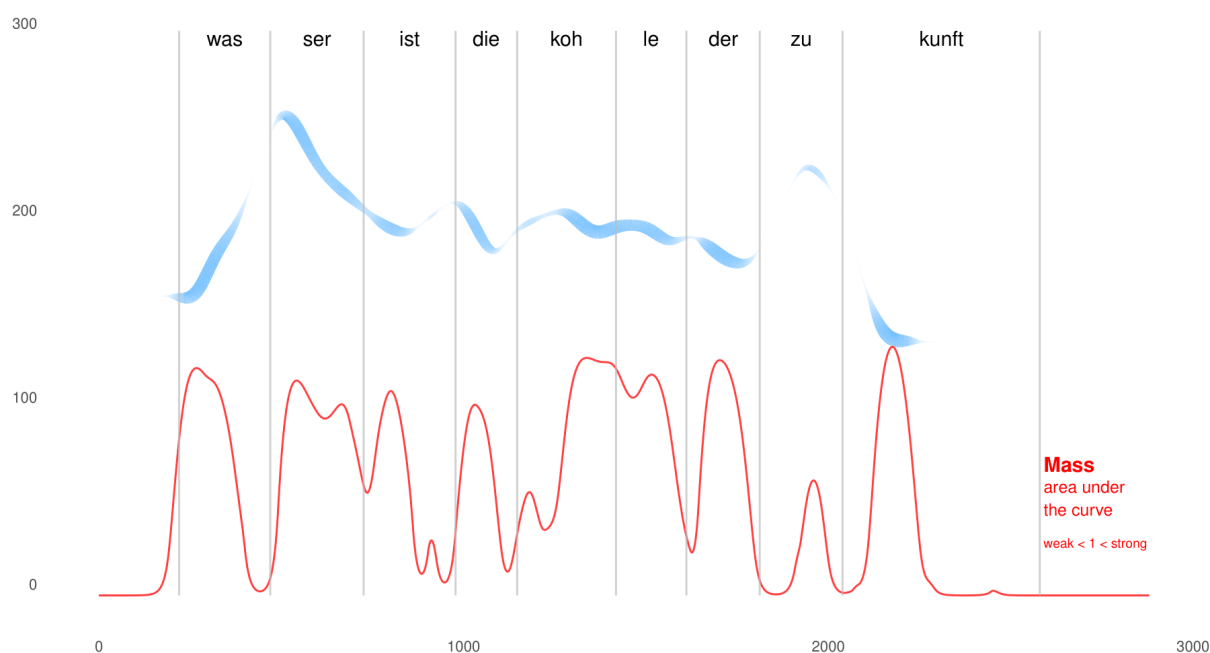

Figure 6: Simplified visualisation of UP19 through periograms and periodic energy

The process shown here was already tested with a group of first-year university students attending an introductory course on German phonetics and phonology at the Università Cattolica in Milan. Therefore, the choice of the main accent highlighted in the steps of the phonetic training below reflects not only the results of the perceptive experiment but also the answers given in a real teaching setting by Italian GFL learners.

Steps of an L2 phonetic training session based on periograms	Examples
1. Have Students (Sts) listen to the utterance and locate the word with the most important/salient stress;	Wasser ist die **Kohle** der Zukunft
2. Ask Sts to locate the indicated word on the simplified periogram (figure 5) and describe the features of the light blue curve in that position;	On *Kohle* we have a flat curve, with a slight raising movement. The curve starts off in a faded colour and then the colour intensifies.
3. The teacher helps Sts elicit the correlation between the hues of the curve and the underlying red curve;	More area means more energy in the realisation of that sound. Where there is more energy the curve has a solid colour, where the energy is lower the colour is muted;

Steps of an L2 phonetic training session based on periograms	Examples
4. Teacher explains and highlights the importance of *stark steigend/fallend* ('high rising/falling') F0 patterns in the perception of prominence by German listeners. Then ask Sts to locate these movements on the curve;	**Wasser** ist die Kohle der **Zukunft**
5. Tell Sts that those parts are more prominent for a German ear – this is important to understand the meaning of the utterance;	The foci of the utterance are *Wasser* and *Zukunft* → Water (and not something else) is the energy source of the future;
6. Have Sts listen again to the utterance and ask them if they can perceive it differently, then ask Sts to repeat the utterance.	Check for correct realisation of accent patterns in German.
7. In pairs, Sts are given a prompt and they produce a short conversation in which they have to insert a sentence with the same communicative aim and similar prosodic pattern as the example utterance.	A: Wie fandest du den Film? B: Ich fand ihn super! Und du? A: Ich auch. **De Niro** ist der beste **Schauspieler.**

Table 2: Steps of the phonetic training with examples

6   Conclusions

The research reported in this paper aimed to explore the perception of prominence in authentic speech by native and non-native speakers of German. Comparing prominence perception by Italian GFL learners and German native speakers, some differences were observed which confirm the theses that in languages that exhibit isochrony at the syllable level, such as Italian, perception is based on duration and intensity, whereas in languages that exhibit isochrony based on the stress level, such as German, perception is based on pitch variations. The acoustic analysis of the speech signal, and specifically the measurement of novel acoustic correlates based on periodic energy measurements within the presented corpus, suggest that while German native speakers consequently base their perception of prominence on a single variable (namely F0) Italian native speakers base their perception on two variables (namely duration and intensity, presented in periograms as the multidimensional correlate of mass). At the same time, the use of prosodic correlates based on periodic energy proved to be an innovative and effective way to describe the divergent perceptual habits of native and non-native speakers and has opened the way for new research approaches to perception awareness in the foreign language based on multidimensional correlates.

In addition, the analysis of periodic energy distribution allows for a more immediate and intuitive visualisation of prosody when compared to the separate visualisations offered by PRAAT (oscillogram and F0 curve). Instead of relying solely on pitch movements, it is also possible to simultaneously visualise prosodic patterns in relation to the energy that is used in the production of speech and thus also taking into account the different degrees of prominence in authentic utterances. This also has implications for SLA, since simplified visualisations of prosody, such as periograms, allow learners to capture visually how prominence is perceived by natives while

comparing it with their perception. This visual approach facilitates the training of both perception and production of intonation patterns, making abstract prosodic features intelligible to learners.

Finally, we stress the need to take into account the perceptual differences between natives and non-natives in future prosodic training proposals and, among future developments, we aim to expand the corpus to analyse more utterances from diverse podcasts to provide a robust foundation for a teaching proposal based on perception awareness.

References

Abercrombie, David (1964): "Syllable quantity and enclitics in English". In: Abercrombie, David et al. (eds.): *In Honour of Daniel Jones*. London, Longmans: 216–222.

Albert, Aviad/Cangemi, Francesco/Grice, Martine (2018): "Using periodic energy to enrich acoustic representations of pitch in speech: A demonstration". *Proceedings of Speech Prosody 2016*: 804–808.

Albert, Aviad et al. (2024): *ProPer: PROsodic analysis with PERiodic energy*. doi.org/10.17605/OSF.IO/28EA5.

Albert, Aviad (2023): *A Model of Sonority Based on Pitch Intelligibility*. Berlin: Language Science Press.

Alter, Kai et al. (2001): „Kontrastprosodie in Sprachproduktion und -perzeption". *Linguistische Arbeitsberichte* 77: 59–79.

Atoye, Raphael O. (2005): "Non-Native Perception and Interpretation of English Intonation". *Nordic Journal of African Studies* 14/1: 26–42.

Avesani, Cinzia et al. (2015): "Prosody and information status in Italian and German L2 intonation". In: Chini, Marina (ed.): *Il parlato in [italiano] L2: Aspetti pragmatici e prosodici*. Milano, Franco Angeli: 93–116.

Baumann, Stefan/Niebuhr, Oliver/Schroeter, Bastian (2016): "Acoustic cues to perceived prominence levels – evidence from German spontaneous speech". *Proceedings of Speech Prosody 2016*: 711–715.

Bertinetto, Pier Marco (2021): "Rhythm in the Romance Languages". In: Loporcaro, Michele (ed.): *Oxford Research Encyclopedia of Linguistics.* Oxford: Oxford University Press.

Blühdorn, Hardarik (2013): „Intonation im Deutschen – nur eine Frage des schönen Klangs?". *Pandaemonium* 16/22: 242–278.

Boersma, Paul/Weenink, David (2025): *Praat: doing phonetics by computer*. Version 6.4.26. praat.org/. [18.01.2025].

Bucholtz, Mary/Hall, Kira (2005): "Identity and interaction: A sociocultural linguistic approach". *Discourse Studies* 7/4–5: 585–614.

Chafe, Wallace (1994): *Discourse, Consciousness, and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

Couper-Kuhlen, Elizabeth/Margret, Selting (2018): *Interactional Linguistics. Studying Language in Social Interaction*. Cambridge: Cambridge University Press.

Cruz Ferreira, Madalena (1989): "A Test for Non Native Comprehension of Intonation in English". *International Review of Applied Linguistics* 17/1: 23–39.

de Bot, Kees/Mailfert, Kate (1982): "The Teaching of Intonation: Fundamental Research and Classroom Applications". *TESOL Quarterly* 16/1: 71–77.

Damiazzi, Vincenzo (2000): "Tecniche di visualizzazione prosodica per l'acquisizione dell'intonazione tedesca". *Nuova Secondaria* 2: 83–88.

Damiazzi, Vincenzo (2022): *Domande retoriche nel Bundestag. Un'analisi prosodica*. Berlin, etc.: Lang.

Damiazzi, Vincenzo (in press): "Perzeptive Aspekte der Prominenz in der Fremdsprache Deutsch". In: Ricci Garotti, Federica et al. (eds.): *DaF heute zwischen Anwendung und Empirie. Tendenzen und Perspektiven*. Wien, Praesens Verlag.

Ellis, Rod (1994): *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

European Council (2020): *Common European Framework of Reference for Languages. Companion Volume*. Bruxelles: European Council.

Flege, James Emil (1995): "Second language speech learning: Theory, findings and problems". In: Strange, Winifred (ed.): *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore, York Press: 233–277.

Johnstone, Barbara/Andrus Jennifer (2024): *Discourse Analysis*. Hoboken: Wiley-Blackwell.

Kalch, Anja/Schlütz, Daniela (2022): „Nachhaltigkeit auf die Ohren?! Klima und Umweltschutz Podcasts und ihre Wirkung". In: Katzenberger, Vera/Keil, Jana/Wild, Michael (eds.): *Podcasts. Perspektiven und Potenziale eines digitalen Mediums*. Wiesbaden, Springer Fachmedien: 417–442.

Kemferts Klima-Podcast (2023): "Wie viele Autos darf es in Zukunft geben?" mdr.de/nachrichten/podcast/kemfert-klima/audio-autos-studie-klimaziel-erreichen-100.html [18.01.1025].

Kügler, Frank/Baumann, Stefan/Röhr, Christine T. (2022): „Deutsche Intonation, Modellierung und Annotation (DIMA) – Richtlinien zur prosodischen Annotation des Deutschen". In: Schwarze, Cordula/Grawunder, Sven (eds.): *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion: Konzepte, Probleme, Lösungen*. Tübingen, Narr: 23–54. dima.uni-koeln.de/ [18.01.2025].

Labov, William (1972): *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lewis, John M. (1999): "Intonation in Theory and Practice, Revisited". *TESOL Quarterly* 33/1: 37–63.

Long, Michael H./Doughty, Catherine J. (eds.) (2011): *The Handbook of Language Teaching*. Hoboken: Wiley-Blackwell.

MDR Investigativ (2023a): "Kohleabbau – ist Mühlrose das neue Lützerath?" mdr.de/nachrichten/podcast/mdr-investigativ/audio-podcast-investigativ-muehlrose-kohleabbau-100.html [18.01.2025].

MDR Investigativ (2023b): "Grüner Wasserstoff aus Afrika – Energie der Zukunft?" mdr.de/nachrichten/podcast/mdr-investigativ/audio-podcast-gruener-wasserstoff-investigativ-recherche100.html [18.01.2025].

Milroy, Lesley (1987): *Observing and Analysing Natural Language*. Hoboken: Wiley & Sons.

Missaglia, Federica (1999): *Phonetische Aspekte des Erwerbs von Deutsch als Fremdsprache durch italienische Muttersprachler*. Frankfurt a. M.: Hector.

Missaglia, Federica (2012): *Deutsche Phonetik und Phonologie für Italiener*. Milano: Vita&Pensiero.

Missaglia, Federica (2018): „Deutsch nach Englisch in Italien: Welches Sieb für den Phonetikerwerb?". In: Hepp, Marianne/Nied Curcio, Martina (eds.): *Educazione plurilingue. Ricerca, didattica e politiche linguistiche*. Roma, Istituto Italiano di Studi Germanici: 181–198.

Missaglia, Federica (2019): „Deutsche Intonation: Formen und Funktionen". In: Di Meola, Claudio et al. (eds.): *Germanistische Linguistik und DaF-Didaktik*. Berlin, Frank&Timme: 47–63.

Missaglia, Federica (2021): „Kontrastive Phonetik und Prosodie: Deutsch vs. Italienisch". In: Brambilla, Marina/Crestani, Valentina/Calpestrati, Nicolò (eds.): *Deutsch im Vergleich. Theorie, Praxis, Didaktik*. Berlin etc., Lang: 207–225.

Missaglia, Federica (2023): „Deutsche Phonetik und Prosodie im DaF/DaZ-Unterricht für italienische Lernende". *Kritik* 1/23: 31–55.

Munro, Murray J./Derwing, Tracey M. (2008): "Segmental acquisition in adult ESL learners: A longitudinal study of vowel production". *Language Learning* 58/3: 479–502.

Niebuhr, Oliver/Fischer, Kerstin/Schümchen, Nathalie (2017): "Comparing visualization techniques for learning second language prosody: First results". *International Journal of Learner Corpus Research* 3/2: 250–277.

Niemann, Henrik et al. (2011): "Tones and gestures: The case of Italian and German". *Proceedings of ICPhS XVII*: 1486–1489. internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Niemann/Niemann.pdf [14.01.2025].

Pettorino, Massimo/Pellegrino, Elisa (2016): "%V and VtoV: an acoustic-perceptual approach to the rhythmic classification of languages". In: Bardel, Camilla/De Meo, Anna (eds.): *Parlare le lingue romanze*. Napoli, Il torcoliere: 13–28.

R Core Team (2023): *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria. R-project.org/ [18.01.2025].

Rabanus, Stefan (2001): *Intonatorische Verfahren im Deutschen und Italienischen*. Tübingen: Niemeyer.

Sbranna, Simona et al. (2023): "Prosodic marking of information status in Italian". *Journal of Phonetics* 97: Article 101212.

Selting, Margret et al. (2009): „Gesprächsanalytisches transkriptionssystem 2 (GAT 2)". *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion* 10: 353–402.

Tannen, Deborah (1989): *Talking Voices. Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Trofimovich, Pavel/Baker, Wendy (2006): "Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech". *Studies in second language acquisition* 28/1: 1–30.

Trubetzkoy, Nikolaï Sergeyevich (1939): *Grundzüge der Phonologie*. Prague: Cercle Linguistique de Copenhague et Ministère de l'Instruction Publique de la République Tchéco-Slovaque.

Zeit für [Klima] (2021): "Kann die Kultur das Klima retten?" zeitfuerx.de/klima/wie-passen-kultur-und-klimaschutz-zusammen/ [18.01.2025].

Zeit für [Klima] (2022): "Wie treiben wir die Energiewende richtig voran". zeitfuerx. de/klima/wie-treiben-wir-die-energiewende-richtig-voran/ [18.01.2025].

Appendix

UP-No.	UP-Text
UP1	Wir beginnen aber mal, wie angekündigt, heute mit der Hörer-Frage
UP2	In Deutschland sind knapp 50 Millionen Pkw derzeit zugelassen
UP3	Das ist also schon wirklich eine richtig große Menge
UP4	Das ist für die_1 einfach die_2 Heimat
UP5	Wann hat's bei Ihnen Klick fürs Klima gemacht?
UP6	Schauen wir nochmal ein bisschen in die Welt
UP7	Die Aufgabe des Netzwerks ist es erstmal, kein Netzwerk zu sein
UP8	Soll das die berühmte Renaissance der Atomkraft sein?
UP9	Den Leuten vor Ort ist es nicht egal, was mit dem Klima passiert
UP10	Ich glaube das ist noch ein bisschen Spielraum drin
UP11	Niedersachsen und Schleswig-Holstein haben viele große Windparks
UP12	Die G7 ist eine Teilmenge der G20
UP13	Wer schreibt Gesetze?
UP14	Bevor wir darüber sprechen
UP15	Wir haben überhaupt diese Alternative
UP16	Wie schätzen Sie diese Zusammenarbeit ein?
UP17	Der zweite Bereich, da geht's um Daten
UP18	Im Grunde genommen kann man das schon
UP19	Wasser ist die Kohle der Zukunft
UP20	Deswegen kleben sich junge Menschen auf die Straße