ПЛОВДИВСКИ УНИВЕРСИТЕТ „ПАИСИЙ ХИЛЕНДАРСКИ“ – БЪЛГАРИЯ
НАУЧНИ ТРУДОВЕ, ТОМ 60, КН. 1, СБ. А, 2022 – ФИЛОЛОГИЯ,
PAISII HILENDARSKI UNIVERSITY OF PLOVDIV – BULGARIA
RESEARCH PAPERS, VOL. 60, BOOK 1, PART A, 2022 – LANGUAGES AND LITERATURE

A METHODOLOGICAL REVIEW OF PROSODY RESEARCH

Ralitsa P. Kuncheva
Paisii Hilendarski University of Plovdiv

In recent years, there has been an increasing interest in the study of prosody and intonation. This is due to the fact that these two aspects of speech are known to play a significant role in communication. However, despite the importance of prosody and intonation, there is still lack of agreement on the best way to study them. This brief review aims to present a critical analysis of some prosody and intonation research methods. The advantages and disadvantages of different research methodologies, approaches, and data analysis methods that have been used in previous research will be discussed. The difficulties and challenges associated with prosody research will also be highlighted.

Key words: research methods, prosody, intonation

Introduction

At every level of linguistic structure, from the word level up to the discourse context, prosody communicates information about the linguistic context of an utterance by means of its acoustic correlates. Prosodic variation due to the speaker, language variety, speech style, and other situational context factors must also be taken into account when interpreting prosodic cues in terms of the lexical, syntactic, and discourse information they encode.

Intonation and rhythm, the musical aspects of speech, are frequently used to define prosody (Wennerstrom 2001). Rhythm and intonation are suprasegmental characteristics of speech for they define patterns that are mostly independent of the segmental composition, in other words the consonant and vowel phonemes, of a given word or phrase. The duration and relative timing of phones, syllables, and other speech units, as well as the auditory characteristics of pitch and loudness, are all examples of suprasegmental features. These auditory attributes in turn depend on the acoustic signal's time-varying characteristics, such as the fundamental frequency (F0), amplitude, and the length of the acoustic intervals corresponding to phones and syllables.

One of the biggest issues with prosody, as opposed to linguistic segments, is what Xu refers to as *the lack of reference problem*. By reference, he means "a pivot that serves as both a starting point of inquest and a point that one can comfortably fall back on" (Xu 2011). For instance, word identity is used as a reference in segmental research since it may be consciously accessed regardless of whether the language being studied has a writing system or even if the human informant is literate. This provides us with grounds to confidently investigate the phonetic characteristics that set one word apart from another. However, the discovery of phonological awareness has shown that awareness of the segments needs to be either introduced or enhanced by means of literacy education (Bentin et al., 1992).

Except for punctuation, which has at best uncertain meaning, very little of prosody's functionality is orthographically reflected in written language. As a result, the starting point for prosody investigations is inherently ill-defined and arbitrary, and it can be equally challenging to determine with certainty what to compare an observation to after it has been made. Our ability to distinguish between pitches is not nearly as good as is commonly believed, particularly when it comes to melodic events in prosody (Dankovicova et al. 2007).

Descriptive methods

Early prosody research relied heavily on descriptive methods, which can be categorized as analysis via introspective transcription. The researcher's intuition and nonexperimental observation were used to propose symbolic representations of prosodic events in this method.

Prosody has been studied systematically as early as Walker (1787), who developed a tone marking scheme for English intonation that is not all that dissimilar from the IPA annotations for lexical tones. This approach has been continued in the modern British intonation tradition, which includes works by Palmer (1922), Kingdon (1958), O'Connor and Arnold (1961), Halliday (1967), Crystal (1969), Cruttenden (1997), and Wells (2006). In this tradition, intonation is represented through a transcription system made up of contours (by curved lines, sometimes with arrow heads to mark the direction of pitch movements) and prominences (often represented by the size of subsequent dots corresponding to the stressed syllables).

American transcription systems that prioritize tonal levels over tonal contours run parallel to this tradition. Rush (1827) is the earliest known proponent of this approach, and later works include Pike (1945), Trager and Smith (1951), and Hockett (1958). Bolinger (1986, 1989) suggested a

variation on this approach that uses a transcription system that depicts pitch contours via word-art-like text arrangements.

The transcription system known as ToBI for TOnes and Break Indices, which is still frequently used, is the most recent significant advancement in prosodic analysis by transcription (Silverman et al. 1992). The system is based on the boundary representations presented by Price et al. (1991) and the pitch accent representations proposed by Pierrehumbert (1980).

However, because the transcription systems themselves were not created using empirical data, the aforementioned issue with the absence of reference is not fully resolved. Analyses by transcription are frequently utilized in empirical investigations in more modern research, where the transcriptions are used as measurements and are processed by means of statistical analysis (e.g., Grice et al. 2009; Mady & Kleber 2010).

Acoustic analysis

Acoustic analysis is perhaps the most straightforward approach to studying prosody and intonation. This methodology involves recording speech samples and then analyzing the acoustic properties of the signal using specialized software. This approach has the advantage of being relatively quick and easy to implement, but it can be limited in its ability to provide insights into the underlying linguistic processes involved in producing prosodic effects.

Acoustic analysis provides objective and quantitative data about speech sounds, which can be useful in investigating various aspects of prosody. For example, formant analysis can be used to measure vowel quality, pitch analysis can be used to measure intonation patterns, and intensity analysis can be used to measure stress levels. Acoustic data can also be combined with other types of data, such as linguistic data, to provide a more complete picture of prosodic phenomena.

Different tools can be utilized in acoustic analysis. As H.G. Tillmann points out "Before the mainstream of speech science, after the invention of a new kind of visible speech in the 1940s, moved on to an extensive acoustic analysis of speech as well as to the early experiments concerning auditory speech perception, there already existed a solid foundation of phonetic knowledge about the acoustic nature of speech sounds" (Tillmann 2005). The description of speech sounds goes back to the work of Willis (1830), but the main technological development that made it possible to analyze and visualize speech signals was the development of the sound spectrograph in 1945. Thanks to later advancements in digital signal processing, most notably the discrete

Fourier transform, all acoustic analyses can now be completed on a basic microcomputer. A spectrogram is a visual representation of the spectrum of frequencies in a sound. It is usually a graph with time presented on the x-axis, frequency on the y-axis, and amplitude (loudness) represented by different colors or shades. Spectrograms can be used to identify phonetic sounds, to analyze the sounds of different animals, and to study how the human vocal tract produces speech sounds.

Fourier transform spectrograms and wavelet transform spectrograms are the two main types of spectrograms. Fourier transform spectrograms are based on the Fourier transform, which decomposes a signal into its constituent frequencies. Wavelet transform spectrograms, on the other hand, are based on the wavelet transform, which decomposes a signal into both frequency and time components.

Both types of spectrograms have their advantages and disadvantages. Fourier transform spectrograms are typically faster to compute, but they can be less accurate than wavelet transform spectrograms. Conversely, Wavelet transform spectrograms can be more accurate, but they can be more time-consuming to compute.

Linear prediction is another technique used in acoustic analysis of prosodic variables. It is a mathematical technique that can be used to estimate the future values of a signal based on its past values. Linear prediction can be used to estimate the future values of speech signals, such as formants, pitch, and intensity and this makes it a useful tool in predictive analyses of prosody. However, the accuracy of linear prediction depends on how well the model represents the actual signal.

One difficulty that can arise with acoustic analysis is that the interpretation of acoustic measures is often subjective. For instance, two different researchers might look at the same data and come to different conclusions about what it means. Additionally, acoustic analysis can be time-consuming and expensive, particularly if recordings need to be made in a controlled laboratory setting.

Perceptual analysis

Perceptual analysis is another common research method used to study prosody and intonation. This approach typically involves having listeners judge various aspects of speech, such as how pleasant or natural it sounds or what emotion or attitude the speaker is conveying. Perceptual measures can provide valuable information about how people perceive prosodic features such as pitch, loudness, and duration as well as the linguistic and extralinguistic meaning they may convey.

One difficulty that can arise with perceptual analysis is that listeners' judgments are often subjective and may be influenced by factors such as their own native language or dialect. Furthermore, perceptual analysis can be time-consuming and costly, especially if recordings need to be produced in a laboratory.

Experimental phonetics

In recent years, experimental phonetics has become a very popular research method used to study prosody and intonation. This approach typically involves manipulating one or more independent variables (such as pitch or loudness) while holding other variables constant, and then observing the effect on a dependent variable (such as listener perception). Experimental methods can provide valuable insights into causal relationships between various prosodic features and listener responses.

One difficulty that can arise with experimental methods is that it can be difficult to control all of the variables in a real-world setting. Moreover, experimental methods can also be time-consuming and expensive when performed in a laboratory setting that requires specialized equipment and software.

Computer-assisted analysis

During experimental research of prosody, specialised computer software is often utilized in order to perform acoustic analysis of speech and its prosodic variables. Praat is by far one of the most popular software "for doing phonetics by computer" out of the many speech analysis programs that are now available (Boersma and Weenink 2007). However, a wide variety of signal analysis software is available with features for speedy and precise extraction of frequency, pitch contours, intensity levels, as well as on-screen presentation of speech sound waves and spectrograms. Some notable examples of such software are: Sygnalize©, WinCECIL© (Computerized Extraction of Components of Intonation in Language), MSL (Micro Speech Lab), Visi-Pitch, Speech Analyzer©, and Emu.

Making comparisons between different pieces of software is difficult because pitch-tracking systems are always being improved and updated. The author may already be using an outdated version of the software when making the comparison (Urbani 2011). Ladefoged (2003: 17) nonetheless outlined four characteristics that a program should have in order to be an effective recording system: a good frequency response, a good signal-to-noise ratio, reliability and user-friendliness, and the ability to utilize and

modify the recordings for a long time. A computer program is dependable and effective if it possesses these attributes.

Difficulties in conducting prosody research

The main difficulty in conducting prosody research is the lack of a standardized methodology. There is no single accepted approach to measuring or analyzing prosodic features, which makes it difficult to compare results across studies. This lack of standardization can be traced back to the early days of prosody research, when different schools of thought developed their own methods and terminology. This lack of agreement makes it difficult to design studies that can be replicated and compared.

In addition, prosody is a complex phenomenon that involves many different factors, both linguistic and paralinguistic. This makes it difficult to isolate the specific effects of each factor in an experiment. For example, when studying the effect of intonation on listeners' perceptions, it is important to control for other variables such as loudness, pitch range, and speaking rate. If these variables are not controlled for, it is impossible to say with certainty that any observed effects are due to intonation alone.

Another challenge in conducting prosody research is that many of the relevant features are subtle and difficult to measure accurately. For example, small variations in pitch can have a big impact on listener perceptions but are very hard to quantify precisely. This can make it difficult to obtain reliable results from experiments or corpora-based studies.

Challenges in analyzing prosodic data

Researchers of prosody often face challenges with respect to data collection. They have to decide on matters such as what data to collect, for example read versus spentaneous speech, monologue versus dialogue, and what speakers to choose, for instance male or female voices, which are different in quality and pitch range. Another important consideration is how to collect the empirical data, in a laboratory setting or in a normal social setting such as having a conversation over a cup of coffee.

Once data has been collected, there are also challenges associated with analyzing it effectively. Prosodic features are often highly variable within and across speakers, making it difficult to identify patterns or trends. In addition, acoustic measurements can be affected by factors such as microphone quality and recording environment, which can introduce errors or bias into the data.

These challenges make it difficult to draw clear conclusions from prosody research studies. However, despite these difficulties, there has been significant progress made in our understanding of prosodic phenomena over the past few decades thanks to advances in methodology and technology.

Conclusion

The study of prosody and intonation has traditionally been approached from a variety of different methodological perspectives, each with its own strengths and weaknesses. Nevertheless, modern advances of technology constantly provide new and upgraded instruments for a more precise recording and analysis of the various prosodic variables. Hopefully, this will make for more streamlined and precise research of prosody.

REFERENCES

Bentin 1992: Bentin S., Ram F. and Leonard K. *Chapter 11 Phonological Awareness, Reading, and Reading Acquisition: A Survey and Appraisal of Current Knowledge.* In: editors. Advances in Psychology. Volume 94: North-Holland, 1992. 193-210.

Boersma 2010: Boersma, P. and D. Weenink. *Praat: a System for Doing Phonetics by Computer.* Tech. Report 132, Inst. Phonetic Sciences, Univ. Amsterdam, 2001. Available at http://www. fon.hum.uva.nl/praat

Bolinger 1986: Bolinger D. *Intonation and its parts: melody in spoken English*. Palo Alto: Stanford University Press; 1986.

Cruttenden 1997: Cruttenden A. *Intonation (second edition)*. Cambridge University Press; 1997.

Crystal 1969: Crystal D. *Prosodic Systems and Intonation in English.* London: Cambridge University Press; 1969.

Dankovicova 2007: Dankovicova J., House J., Crooks A. and Jones K. *The Relationship between Musical Skills, Music Training, and Intonation Analysis Skills.* Language & Speech 2007; 50:177-225.

Grice 2009: Grice M., Baumann S. and Jagdfeld N. *Tonal association and derived nuclear accents - The case of downstepping contours in German.* Lingua 2009; 119:881-905.

Halliday 1967: Halliday MAK. *Intonation and Grammar in British English.* The Hague: Mouton; 1967.

Hockett 1958: Hockett C.F. *A course in modern linguistics*. New York: MacMillan; 1958.

Kingdon 1958: Kingdon R. *The groundwork of English intonation*. London: Longman; 1958.

Ladefoged 2003: Ladefoged, P. *Phonetics Data Analysis. An Introduction to Fieldwork and Instrumental Techniques.* Oxford: Blackwell, 2003.

Mady 2010: Mady K. and Kleber F. *Variation of pitch accent patterns in Hungarian.* In: Proceedings of Speech Prosody 2010, Chicago, 2010.

O'Connor 1961: O'Connor J.D. and Arnold G.F. *Intonation of Colloquial English*. London: Longmans; 1961.

Palmer 1922: Palmer H. E. *English Intonation, with Systematic Exercises.* Cambridge: Heffer; 1922.

Pierrehumbert 1980: Pierrehumbert J. *The Phonology and Phonetics of English Intonation [Ph.D. dissertation]*. MIT, Cambridge, MA. [Published in 1987 by Indiana University Linguistics Club, Bloomington]; 1980.

Pike 1945: Pike K.L. *The Intonation of American English*. Ann Arbor: University of Michigan Press; 1945.

Price 1991: Price P.I., Ostendorf M., Shattuck-Hufnagel S. and Fong C. *The use of prosody in syntactic disambiguation.* Journal of the Acoustical Society of America. 1991; 90:2956-2970.

Rush 1879: Rush, J. (1827). *The philosophy of the human voice.* 7th ed. Philadelphia: J. B. Ltppincott & Co., 1879.

Silverman 1992: Silverman K., Beckman M., Pitrelli J., Ostendorf M., Wightman C., Price P., Pierrehumbert J. and Hirschberg J. *ToBI: A standard for labeling English prosody*. In: Proceedings of The 1992 International Conference on Spoken Language Processing, Banff, 1992. 867-870.

Tillmann 2005: Tillmann, H.G. *Experimental and Instrumental Phonetics: History,* in Encyclopedia of Language & Linguistics (Second Edition), Elsevier Pergamon, 2005, pp. 374-389.

Trager 1951: Trager, G.L.; Smith, H.L. *An outline of English structure.* Battenberg Press, Norman, 1951.

Urbani 2011: Urbani M., *Instruments and methods for the analysis of prosody*. Linguistics Applied, vol. 4 (2011), pp. 104-115.

Walker 1787: Walker J. *The melody of speaking delineated* (printed for the author, London 1787; reprinted by The Scolar Press, Menston. English Linguistics 1500–1800: No. 218, 1970), 1787.

Wells 2006: Wells J.C. *English intonation: an introduction.* Cambridge: Cambridge University Press; 2006.

Wennerstrom 2001: Wennerstrom, A. K. *The music of everyday speech: Prosody and discourse analysis.* Oxford: Oxford University Press, 2001.

Willis 1830: Willis, R. *On vowel sounds, and on reed organ pipes.* Transactions of the Cambridge Philosophical Society III, 1830, 231-276.

Xu 2011: Xu, Yi. Speech prosody: A methodological review. Journal of Speech Sciences, 2011, 85-115.