

HOW DO THEY SOUND? SUPRASEGMENTAL FEATURES AND THEIR ROLES IN SWAPPED ANIMATED MOVIE (2026)

**Naila Sakhshiya Akmalia¹, Diva Nahdah Ariela², Suci Surianti³,
Sri Melianti Hasibuan⁴, Yani lubis⁵**

naila0304242104@uinsu.ac.id¹, dnahdahariela@gmail.com², sucisurianti21@gmail.com³,
meliantihisibuan29@gmail.com⁴, yanilubis@uinsu.ac.id⁵

Universitas Islam Negeri Sumatera Utara

ABSTRACT

This study investigates the suprasegmental features found in the animated movie Swapped (2026). Suprasegmental features—including stress, intonation, pitch, rhythm, pause, and juncture—play a crucial role in conveying meaning beyond individual phonemes. The animated movie genre offers authentic spoken English that reflects natural prosodic patterns, making it a valuable source for phonological analysis. This research employs a descriptive qualitative approach. Data were collected through systematic observation and transcription of selected scenes from the movie. The findings reveal diverse patterns of stress, intonation, pitch variation, rhythmic structures, and the use of pause and juncture as discourse markers. The results are expected to contribute to the teaching and learning of English pronunciation, particularly in the context of suprasegmental phonology.

Keywords: *Suprasegmental Features, Stress, Intonation, Pitch, Rhythm, Pause And Juncture, Animated Movie.*

INTRODUCTION

Language is a fundamental means of human communication, encompassing both segmental and suprasegmental dimensions. While segmental phonology focuses on individual sounds or phonemes, suprasegmental phonology addresses features that operate above the level of individual segments, such as stress, intonation, pitch, rhythm, pause, and juncture (Roach, 2009). These features are essential in spoken language as they carry meaningful distinctions that cannot be conveyed through written text alone.

In the context of English as a Foreign Language (EFL) learning, mastery of suprasegmental features is often considered secondary to segmental accuracy; however, scholars have argued that suprasegmental errors may cause greater communication breakdowns than segmental ones (Jenkins, 2000). This underscores the importance of exposing learners to authentic spoken English that demonstrates natural prosodic patterns.

Animated movies present a rich and accessible medium for phonological analysis. Unlike live-action films, animated movies feature carefully crafted voice acting in which actors must rely entirely on vocal cues—stress, intonation, pitch variation, and rhythm—to convey character emotions and intentions. The 2026 animated movie Swapped provides a contemporary, linguistically diverse dataset suitable for examining these suprasegmental features in natural conversational contexts.

Given the limited studies focusing specifically on suprasegmental analysis within recent animated film productions, and recognizing the pedagogical potential of such analysis for EFL contexts, this study aims to describe and analyze the suprasegmental features present in the movie Swapped (2026).

LITERATURE REVIEW

1. Phonology

Phonology is the branch of linguistics concerned with the systematic organization of sounds in languages. It studies the abstract, cognitive aspects of sound patterns, including how sounds function and pattern in a specific language or languages in general (Carr, 2012). Phonology can be divided into two major areas: segmental phonology and suprasegmental phonology. Segmental phonology deals with the smallest units of sound—phonemes—while suprasegmental phonology encompasses those phonological features that extend over more than one segment, such as syllables, words, phrases, and utterances.

Phonology is not only concerned with the description of speech sounds but also with how speakers mentally organize and interpret sound systems within a language. A solid understanding of phonology helps language learners improve both pronunciation and listening comprehension (Yule, 2020; Jones, 2011).

The study of phonology is essential in language learning, as it governs the pronunciation system of a language. A comprehensive understanding of phonological principles enables learners to produce and perceive spoken language more accurately and effectively.

2. Suprasegmental Features

Suprasegmental features, also known as prosodic features, are phonological elements that apply to units larger than individual segments. They operate at the level of syllables, words, phrases, and entire utterances (Crystal, 2008). The major suprasegmental features include stress, intonation, pitch, rhythm, pause, and juncture. These features play a fundamental role in spoken communication by conveying grammatical, semantic, pragmatic, and affective information.

Suprasegmental competence is considered a crucial component of communicative competence because prosodic features influence intelligibility and listener comprehension more significantly than many segmental features (Levis, 2018; Derwing & Munro, 2015).

a. Stress

Stress refers to the relative prominence or emphasis given to a syllable within a word or to a word within an utterance (Roach, 2009). In English, stress is realized through a combination of greater loudness, higher pitch, longer duration, and clearer vowel quality. Stress can be classified into word stress and sentence stress. Word stress determines which syllable in a polysyllabic word receives primary emphasis, while sentence stress (also called phrasal or contrastive stress) highlights particular words within an utterance for semantic or pragmatic purposes.

Stress patterns in English are not entirely predictable and must often be memorized as part of a word's lexical entry. Misplacement of stress can lead to unintelligibility or misunderstanding, making it a critical feature for EFL learners to acquire.

Research has shown that incorrect stress placement can reduce intelligibility and create misunderstandings even when individual sounds are pronounced accurately (Field, 2005). Therefore, mastery of stress patterns is essential for successful oral communication in English (Gilbert, 2008).

b. Intonation

Intonation refers to the systematic variation of pitch over the course of an utterance (Cruttenden, 2001). It is realized through rising, falling, or combination tones and serves multiple functions in spoken language: signaling sentence type (statement, question, command), marking information structure (given vs. new information), expressing attitudes and emotions, and organizing discourse. The major intonation patterns in English include the falling tone (commonly used for statements and wh-questions), the rising tone

(used for yes/no questions and continuations), and the fall-rise tone (used for implications and reservations).

Intonation also functions as an indicator of speaker attitude, interpersonal meaning, and discourse organization. Learners who fail to use appropriate intonation patterns may sound unnatural or difficult to understand (Celce-Murcia et al., 2010).

c. Pitch

Pitch is the perceptual correlate of the fundamental frequency of vocal fold vibration (Ladefoged & Johnson, 2014). In suprasegmental analysis, pitch refers to the relative highness or lowness of a speaker's voice at any given point in an utterance. Pitch is closely related to intonation; however, while intonation refers to the movement of pitch across an utterance, pitch level can also function independently to signal contrast, emphasis, or emotional state. High pitch is often associated with excitement, surprise, or heightened emotion, while low pitch may signal finality, seriousness, or boredom.

Variations in pitch contribute significantly to emotional expression and pragmatic interpretation in spoken discourse. Different pitch ranges may indicate politeness, confidence, excitement, or uncertainty (Pennington, 2021).

d. Rhythm

Rhythm in spoken language refers to the regular recurrence of stressed and unstressed syllables in an utterance (Couper-Kuhlen, 1993). English is often described as a stress-timed language, meaning that stressed syllables tend to occur at roughly equal time intervals, regardless of the number of unstressed syllables between them. This contrasts with syllable-timed languages, where each syllable takes approximately equal time. Rhythm contributes to the overall flow and naturalness of spoken English and is an important aspect of English pronunciation that EFL learners must develop.

Rhythm contributes to speech fluency and naturalness. According to Underhill (2005), learners who develop awareness of English rhythmic patterns tend to achieve better pronunciation and listening performance.

e. Pause and Juncture

Pause refers to a break or cessation of sound in spoken language, while juncture refers to the phonological boundary between linguistic units (Crystal, 2008). Pauses can be silent or filled (e.g., 'uh', 'um') and serve various functions: signaling grammatical boundaries, allowing time for speech planning, marking topic shifts, or emphasizing what follows. Juncture describes how sounds at the boundary between words or morphemes are articulated, affecting whether adjacent sounds blend together or are kept distinct (e.g., 'an aim' vs. 'a name'). Together, pause and juncture structure spoken discourse and aid in listener comprehension.

Pauses and junctures assist listeners in processing information by signaling grammatical boundaries and discourse structure. Effective use of these features improves comprehensibility in spoken interaction (Tench, 2011).

3. Movie as a Learning Medium

Movies have long been recognized as valuable authentic materials for language learning. Authentic materials are those created primarily for native speakers of a language rather than for language learners, and they expose learners to natural language use in meaningful contexts (Nunan, 1999). Movies provide visual and auditory input simultaneously, which supports contextual comprehension and affective engagement.

Movies provide rich multimodal input that supports language acquisition through contextualized exposure to authentic communication. Audio-visual materials help learners develop pronunciation, listening comprehension, and pragmatic awareness simultaneously (Zhang, 2009).

In the context of phonology instruction, movies offer exposure to natural prosodic patterns including connected speech features, stress shifts, intonation variations, and natural rhythm—elements that are often difficult to simulate in controlled classroom settings. Animated movies, in particular, present speech that is carefully produced and often more clearly articulated than spontaneous conversation, while still maintaining naturalistic prosodic features. Several studies have demonstrated the effectiveness of using films and video materials in developing learners' phonological awareness and pronunciation skills (Mishan, 2005; Tamo, 2009).

Pronunciation instruction supported by authentic audiovisual materials has been found to enhance learners' awareness of stress, intonation, and rhythm in real communicative contexts (Gilakjani, 2016; Derwing, Munro, & Wiebe, 1998).

4. Synopsis of Swapped

Swapped (2026) is an animated feature film Produced by Skydance Animation. Netflix holds global distribution rights to the film, with the film premiering simultaneously (via streaming and limited theatrical release) on May 1, 2026. The story centers on Ollie, a small, otter-like forest creature (Pookoo), and Ivy, a beautiful bird (Javan), are suddenly swapped bodies after encountering a magical plant. As mortal enemies, they must work together to survive and save their colony. The film features a diverse cast of voice actors portraying characters who vary in age, personality, and communicative style, resulting in a rich variety of prosodic expressions throughout the dialogue. The plot details a deep-seated feud within a magical ecosystem called the Valley, pitting the Pookoos—small, land-dwelling mammals similar to sea otters—against the Javans, a group of giant, lush-green birds resembling a cross between an owl and a cockatoo. The central conflict begins when Ollie (voiced by Michael B. Jordan), a naive young Pookoo, teaches a young Javan bird how to crack open a hard seed shell. This kind of gesture leads to disaster, as the Javans use the trick to wipe out the Pookoos' entire food supply, sparking significant tension between the groups. In the midst of this dispute, Ollie and Ivy (voiced by Juno Temple), a free-spirited Javan bird, get into an argument in an ancient magical area of the forest that results in their instantaneous body-swapping. A new life in the wrong body gives rise to iconic scenes that combine slapstick comedy with a deep psychological message. The audience is treated to humorous and frustrating moments as Ivy, now trapped in a small mammal's body, must learn to crawl and sniff the ground, while Ollie, ensconced in a giant bird's body, struggles to overcome his fear of heights in order to learn to flap his wings and fly. One of the most emotional scenes in this Skydance Animation production occurs when Ollie has to hear firsthand how his own people, the Pookoo, hurl hateful and prejudiced remarks at the Javan bird form he inhabits. However, those differences ultimately crumble in the climactic scene, where Ollie and Ivy combine the physical agility of birds and the tactical ingenuity of mammals to work together to defeat the Firewolf, a ferocious creature that threatens the safety of the entire Valley ecosystem, while simultaneously awakening both colonies to the importance of empathy and unity.

RESEARCH METHOD

1. Research Design

This study employs a descriptive qualitative research design. Qualitative research is concerned with understanding phenomena in their natural setting and interpreting the meaning of human experiences and language (Creswell, 2014). A descriptive approach is appropriate for this study as it seeks to systematically describe and analyze suprasegmental features as they naturally occur in the dialogues of the animated movie Swapped (2026), without manipulation or experimental control.

Content analysis serves as the primary analytical framework, allowing the researcher to systematically identify, categorize, and describe the linguistic features present in the data (Fraenkel, Wallen & Hyun, 2012).

2. Data Source

The data source of this study is the animated movie *Swapped* (2026). The primary data consist of the spoken dialogues of the main characters in the movie. The dialogues were accessed through streaming platform such as Netflix and LK21. The movie has a total duration of approximately 1 hour 42 minutes. For analytical manageability and representativeness, Some scenes were selected based on purposive sampling, prioritizing scenes that demonstrate varied communicative contexts, emotional registers, and interaction types.

3. Data Collection Technique

Data were collected through the following steps:

1. Watching the animated movie *Swapped* (2026) in its entirety to gain an overall understanding of the narrative and characters.
2. Selecting scenes based on purposive sampling criteria, including variety of character interactions, emotional contexts, and discourse types.
3. Transcribing the spoken dialogues from the selected scenes, using standard orthographic transcription supplemented with prosodic notation.
4. Marking and coding the suprasegmental features (stress, intonation, pitch, rhythm, pause and juncture) in the transcripts using established phonological notation systems.
5. Conducting member checking and peer review of the transcription and coding for reliability.

4. Data Analysis Technique

Data were analyzed using Miles, Huberman, and Saldaña's (2014) interactive model of qualitative data analysis, which involves three concurrent activities: data condensation, data display, and conclusion drawing/verification.

Data condensation involved selecting, focusing, simplifying, and transforming the data. The transcribed dialogues were coded according to the five categories of suprasegmental features. Data display involved organizing the condensed data into tables and descriptive accounts for each feature category. Conclusion drawing involved interpreting the patterns and functions of the identified suprasegmental features, supported by relevant theoretical frameworks.

The validity of the analysis was ensured through triangulation of data sources and analyst triangulation, whereby the researcher's coding was cross-checked with a co-coder knowledgeable in English phonology.

RESULTS AND DISCUSSION

Analysis of Stress in *Swapped*

This section presents the findings related to stress patterns identified in the dialogues of *Swapped* (2026). The data reveal two main types of stress: word stress and sentence stress. The following table summarizes the distribution of stress types found in the analyzed data.

The description displays a tragic and tense scene inside a forest at night. A massive tree resembling an ancient throne or nest is seen caught in a raging fire, with bright orange flames glowing and thick plumes of black smoke billowing into the sky. At the bottom of the screen, the dialogue subtitle reads: "He slayed the **FIRST** **FOUR**!". And that used to analyze the contrastive stress feature, where primary vocal emphasis is placed on the words "FIRST" and "FOUR" to clarify the quantity and sequence of the targets

successfully defeated.

The most frequent type observed was contrastive stress, appearing in 1 utterances (100%). Notable examples include: “He slayed the FIRST FOUR!”. In this utterance, the primary stress falls on 'first four' particularly the words “first” and “four”, which serves to highlight important information and emphasize the intended number and sequence. This emphasis helps the listener understand that the speaker is emphasizing a particular achievement, namely the success of killing or defeating the first four targets, making this information the main focus of the conversation. This finding aligns with Roach's (2009) assertion that contrastive stress frequently signals information focus in English discourse.

Analysis of Intonation in Swapped

The description shows a close-up interaction between two furry animal characters in front of a tree house or nest. The character on the right is an otter-like creature (Pookoo) with a thick white mustache and beard. He appears to be wearing brown clothing and is holding a large wooden bowl filled with round fruits (resembling apples or onions). In front of him (on the left), the back of the head of another character wearing a blue head covering with protective goggles is visible. The subtitle below reads: “Really? Big rainbow fish?”. The illustration shows the analysis of intonation, specifically the rising tone pattern commonly used when a character expresses doubt or seeks confirmation.

The analysis of intonation in Swapped (2026) reveals three primary tonal patterns: falling tone, rising tone, and fall-rise tone. These patterns were identified across a range of utterance types including declaratives, interrogatives, imperatives, and exclamations.

The falling tone was the most prevalent pattern, consistent with its function as the default tone for completed statements and wh-questions in English (Cruttenden, 2001). Rising tones were predominantly found in yes/no questions and utterances conveying uncertainty or seeking confirmation. The fall-rise pattern appeared in contexts where characters implied reservations or qualified their assertions.

Analysis of Pitch in Swapped

The description captures a close-up facial expression of three furry creatures inside a dark cave or room. On the left, an old otter character with a white beard looks forward with a flat expression. In the center, there is a character with long, messy gray fur holding a lit candle as a light source. On the right, a large-eyed character looks on with a confused or worried facial expression. The subtitle for this scene reads: ““Eh”? What do you mean, “Eh”?”. And that used to analyze pitch variation, where the high pitch in the dialogue represents emotions of surprise, confusion, or urgency.

Pitch variation in the Swapped (2026) dialogues functions as a primary carrier of emotional and pragmatic meaning. Three pitch levels were identified in the data: high pitch, mid pitch, and low pitch. These levels co-occurred with specific communicative functions across the analyzed utterances.

High pitch was most frequently associated with expressions of surprise, excitement, and urgency. Low pitch correspondingly marked finality, resolution, and somber emotional states. Mid pitch was the most common level, appearing in neutral conversational exchanges. These findings are consistent with Ladefoged and Johnson's (2014) account of pitch as a paralinguistic cue for emotional expression.

Analysis of Rhythm in Swapped

The description features a unique and expressive purple-faced fish character. The fish has colorful, rainbow-like top fins and appears to be pouting its thick lips with sharply narrowed eyes. The background of the scene shows a landscape of rocky cliffs and a clear sky. There is a short subtitle text at the bottom that reads: “Fire Wolf!”. This photo serves as the subject for rhythm analysis, illustrating how the alternation of stressed and

unstressed syllables is spoken during an emotional, climactic, or rhetorical utterance.

The rhythmic analysis of *Swapped* (2026) focuses on the alternation of stressed and unstressed syllables and the degree to which the movie's dialogues reflect the stress-timed nature of English. Rhythmic patterns were categorized as regular, irregular, or emphatic based on the distribution of primary stresses in each utterance.

Regular stress-timed rhythm was the dominant pattern observed, reflecting the voice actors' adherence to naturalistic English prosody. Irregular rhythm was found in utterances involving hesitation, emotional disruption, or speech reflecting non-native speaker characters. Emphatic rhythm, characterized by evenly stressed syllables for rhetorical effect, appeared in climactic or commanding speech acts.

Analysis of Pause and Juncture in *Swapped*

The description presents a beautiful underwater view filled with blue and greenish light hues. In the center of the frame is a light blue aquatic or sea-fairy creature with hair/head ornaments made of rainbow-colored leaves (yellow, orange, green). In the right background, a similar but smaller character is visible. The subtitle at the bottom of the screen reads: "There was only one pod, and we said, "Boogle," before touching it.". And that used to analyze pause and juncture, where the punctuation marks indicate a silent pause that functions to separate clauses, making the information easier for the listener to process.

Pause and juncture analysis examines the distribution and functions of breaks and boundaries in the spoken dialogues of *Swapped* (2026). Two types of pauses were identified: silent pauses and filled pauses. Juncture was examined in terms of open juncture (maintained word boundaries) and close juncture (assimilation or linking).

Silent pauses predominantly occurred at syntactic boundaries—between clauses and before significant discourse shifts—serving a structural and planning function. Filled pauses ('uh', 'um') appeared in contexts of character hesitation or uncertainty. Linking junctures were common in fast, casual speech exchanges, while open junctures appeared in speech directed at conveying precise or formal information.

A representative example is found in the utterance: "There was only one pod // and we said 'Boogle' before touching it." The symbol (//) indicates a silent pause occurring between two clauses. This pause functions as a boundary marker that separates the introduction of information ("There was only one pod") from the subsequent action ("and we said 'Boogle' before touching it"). The juncture observed is an open juncture, as the speaker maintains a clear separation between the clauses, allowing the listener to process the information effectively. This finding supports Roach's (2009) view that pauses and junctures contribute significantly to speech organization, comprehension, and the signaling of meaningful discourse boundaries.

Discussion

The analysis of suprasegmental features in *Swapped* (2026) demonstrates that the movie provides a rich and varied representation of English prosodic features. Across all five categories—stress, intonation, pitch, rhythm, and pause and juncture—the data reveal systematic patterns that are both linguistically grounded and communicatively functional.

The dominance of falling intonation and regular stress-timed rhythm reflects the naturalistic American English phonological norms exemplified by the voice cast. The expressive use of pitch variation and strategic deployment of pause and juncture highlight how animated dialogue is crafted to convey nuanced emotional and pragmatic meanings through vocal performance alone.

These findings corroborate the theoretical frameworks proposed by Roach (2009), Cruttenden (2001), and Ladefoged and Johnson (2014), and support the argument that

animated movies constitute authentic and pedagogically valuable sources for suprasegmental analysis. The diversity of prosodic features documented in this study suggests that *Swapped* (2026) can serve as an effective teaching material for EFL pronunciation classes, particularly for exposing learners to natural English prosody in contextualized, affectively engaging discourse.

CONCLUSION AND SUGGESTION

Based on the results and analysis presented in the previous chapter, it can be stated that the prosody in the dialogue of the *Swapped* (2026) animated movie reflects the characteristics of natural English. Stress in words and sentences, especially contrastive stress and emotional emphasis, is used to highlight new information and to express the speaker's attitude. Descending intonation is the most common and serves as a standard model in declarative sentences, while rising and falling intonations appear to mark questions, continuations in discourse, or feelings of doubt or implication. Changes in high, medium, and low pitch respectively depict emotions such as enthusiasm, impartiality, and assertiveness. In terms of rhythm, patterns organized by stress are the most dominant, with irregular rhythms and rhythms of stress appearing in emotional or rhetorical utterances. On the other hand, silent pauses mark boundaries in grammar and discourse, filled pauses indicate uncertainty, and the use of connectors and open pauses reflects the difference between casual and formal speech styles.

Suggestion

Based on the findings and conclusions of this study, the following suggestions are proposed:

For EFL Teachers: It is recommended that teachers incorporate animated movies such as *Swapped* (2026) as authentic listening and pronunciation materials. The natural prosodic features found in these films offer learners exposure to contextualized English suprasegmental patterns that supplement textbook-based instruction.

For EFL Students: Students are encouraged to engage actively with English-language animated films, paying particular attention to prosodic features such as stress, intonation, and rhythm. Imitation and repetition exercises based on movie dialogues can be an effective strategy for developing suprasegmental competence.

For Future Researchers: It is suggested that future studies expand the scope of analysis to include a comparative study of suprasegmental features across different animated movies, genres, or character types. Additionally, studies examining EFL learners' acquisition of suprasegmental features through movie-based instruction would provide valuable pedagogical insights.

REFERENCES

- Carr, P. (2012). *English phonetics and phonology: An introduction* (2nd ed.). Wiley-Blackwell.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed.). Blackwell.
- Couper-Kuhlen, E. (1993). *English speech rhythm: Form and function*. John Benjamins.
- Cruttenden, A. (2001). *Gimson's pronunciation of English* (6th ed.). Arnold.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Ladefoged, P., & Johnson, K. (2014). *A course in phonetics* (7th ed.). Wadsworth.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

- Mishan, F. (2005). *Designing authenticity into language learning materials*. Intellect Books.
- Nunan, D. (1999). *Second language teaching and learning*. Heinle & Heinle.
- Roach, P. (2009). *English phonetics and phonology: A practical course* (4th ed.). Cambridge University Press.
- Tamo, D. (2009). The use of authentic materials in language classrooms. *Foreword*, 2(1), 64–68.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393–410.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39(3), 399–423.
- Gilbert, J. B. (2008). *Teaching pronunciation: Using the prosody pyramid*. Cambridge University Press.
- Gilakjani, A. P. (2016). English pronunciation instruction: A literature review. *International Journal of Research in English Education*, 1(1), 1–6.
- Jones, D. (2011). *Cambridge English pronouncing dictionary* (18th ed.). Cambridge University Press.
- Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.
- Pennington, M. C. (2021). *English pronunciation teaching and research: Contemporary perspectives*. Palgrave Macmillan.
- Underhill, A. (2005). *Sound foundations: Learning and teaching pronunciation* (2nd ed.). Macmillan Education.
- Yule, G. (2020). *The study of language* (7th ed.). Cambridge University Press.