THE IMPROVEMENT OF MISPRONUNCIATION ENCOUNTERED BY MOST YOUNG ENGLISH LEARNERS

Dianatalia Agustine Widhiastuti[1)], Indah Arvianti[2)], Eko Heriyanto[3)]
Faculty of Language and Culture, University of AKI
Email: agustined8@gmail.com

Abstract

Mispronunciation is a critical issue encountered by most young English learners in a second language learning that causes low spoken English proficiency, which plays important roles today. The objectives of this study were to identify the segmental changes resulted in the articulation of English mispronounced words produced by most young learners, to describe the different elements between Indonesian and English sound systems, and to explain the phonological interference of mother tongue in English mispronunciation. This study uses descriptive qualitative research design performed on ten primary school students selected by using purposive sampling in Semarang. Data were collected by using simak method and analyzed by using padan fonetik artikulatoris method.

The data suggest that the majority of the participants had difficulty in articulating the words with absent phonemic segments /ɪ, æ, əː, ɔ, ʌ, eɪ, ʊə, ɪə, aɪə, ʃ, θ, ð, ʒ, voiced v/, silent letters <b, d, e, t, w, gh, and r in some cases>, spelling patterns <th, ow, ph, ie, ueue, ough>, and tricky words including several English borrowings that have inconsistent spelling pronunciation. The results showed that the interference of L1 sounds and pronunciation rules occurred through substitution of closest native sounds for absent phonemes, insertion of native sounds, and negative transfer of L1 spelling pronunciation. It can be concluded that most of the young learners mispronounced on words with absent phonemic segments, silent letters, spelling patterns of many letters combination, and several English borrowings.

Keywords: *mispronunciation, phonetics, young learners, interference*

1. Introduction

The success of learning global English is important to improve the low English proficiency in Indonesia. Although the language has neither official status nor crucial roles in national life, it is the first foreign language widely taught in schools. English is learned as an optional subject content or a compulsory subject in many Indonesian schools with national curriculum, and is learned as a second language mostly spoken as the primary medium of instruction in many excellent national schools with international-adopted curricula, where the dominant students came from upper class or elite families. In the same way, Dardjowidjoyo (2000) in his previous research indicated that "English is the second language of the educated

urban elite and is also the first foreign language taught in schools, but often with only limited success" (as cited in Kirkpatrick, 2010: 10). In a social context, most Indonesian people whose second languages are regional languages such as Javanese, Sundanese, and Balinese, speak and learn English as a foreign language. Few people or families whose parents are couples of a mixed marriage, English is spoken as a second language. The practical use of this global language in both educational and social context demonstrates that English is mostly learned as foreign language with limited proficiency.

Mispronunciation is a common issue encountered by English learners, especially the young primary school-aged children with low speaking skill. Incorrect pronunciation may cause the utterance unclear and unintelligible for listeners. The speakers sometimes feel less confident at school or even being downgraded by other English speakers in social interaction. Thus, mispronunciation causes the low spoken English proficiency.

Pronunciation problems are mainly caused by the absence of nativesoun ds in foreign sound system and first language interference. Most second language learners have to manage the difference between the target language and their native sound systems. In his book, Ramelan (2003) exemplified the different elements causing difficulty in learning the second language such as individual sounds, phonetic features of similar sounds, and distribution of equivalent sounds. In learning foreign language pronunciation, most of the learners tend to substitute the closest native sounds for the absent English phonemes, hence the mispronunciation.

Considering the important role of spoken English proficiency, which is globally in high demand, mispronunciation is a critical issue that should be managed. As the potential major cause of low English speaking skill, pronunciation problems often occur in the beginning process of foreign or second language learning. Therefore, the mispronunciation produced by young learners, who often encounter problems in learning English, is the main interest of this study. This thesis entitled "The Improvement of Mispronunciation Encountered by Most Young English Learners".

Based on the background of the study above, the problems of this research are formulated as the following questions.

1. What segmental changes are identified in English mispronunciation encountered by primary or elementary students in Diana private course?

2. What are the differences between Indonesian and English sound systems causing pronunciation problems?
3. How does the phonological interference of first language or mother tongue occur in second language pronunciation learning?

The purpose of this study is to examine the pronunciation problems encountered by the young second language learners which are mainly caused by different phonetic features of English sound system and phonological inference in language transfer. The objectives of this study are :

1. To identify the segmental changes in the articulation of English mispronounced words produced by primary or elementary students in Diana private course.
2. To examine the different elements between Indonesian and English sound systems causing pronunciation problems.
3. To explain the occurrence of first language phonological interference in the second language pronunciation.

2. Review of Related Literature

2.1. Phonetic

Phonetics deals with speech in its purely physical aspects, which are the way sounds are articulated by the speaker, the acoustic properties of sound waves, and the effects that these have on the ear of the hearer (Kreidler, 2004). The phonetics of a language concerns the concrete characteristics (articulatory, acoustic, auditory) of the sounds used in language (Cruttenden, 2014). The study is divided into two major components namely segmentals and suprasegmental phonetics. Segmental phonetics is concerned with "segments" of speech that divides concrete utterances into individual speech sounds, while suprasegmental phonetics is concerned with the larger units of connected speech such as syllables, words, phrases and texts (Skandera & Burleigh, 2005; Vrabel, 2009). Gut (2009) differentiated three areas of phonetics, which are articulatory phonetics, acoustic phonetics and auditory phonetics.

2.1.1. Speech Sounds

The sounds are the smallest and indivisible segments singled out in the flow of speech as separate discrete elements, which are produced by the vocal organs including the mouth and the respiratory organs (Ramelan, 2003; Vrabel, 2009). In other words, phones are the phonetic units

or sound-type produced in actual speech (in the mouth) and written in square brackets (Yule, 2010).

Vowels are sonorant speech sounds that are produced by using the front, centre and back of the tongue articulated between palatal and velar (Ashby, 2011). Articulatory properties of vowels are represented using three-way labels, such as tongue height (high–mid–low), frontness–backness or tongue position (front–central–back), and lip shape (rounded–neutral-spread). For example, the vowel /iː/ in 'cheese' is high, front and spread (Przedlacka, 2018). General British (GB) English has 20 vowels consisting of thirteen long phonemes and seven short phonemes (Cruttenden, 2014). Vrabel (2009) described the vowels into five long, seven short monophthongs, and the rest are referred to long sounds with different a glide to [i], [u], and [ə].

Consonants mostly involve an obstruction to the airflow through a contact of articulators along the entire length of the vocal tract (Przedlacka, 2018). Rogerson-Revell (2018) explained that English has 24 common consonants, which can be described in term of the way consonant sounds are produced;

(a) Place of articulation (where the sounds are produced in the vocal tract)

(b) Manner of articulation (how they are produced)

(c) voicing (whether or not there is vibration of the vocal cords/folds)

Figure 1.

English Consonants Chart

MANNER OF ARTICULATION	POINT OF ARTICULATION															
	Bilabial		Labio Dental		Dental		Alveolar		Palato - alveolar		Palatal		Velar		Glottal	
	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd
Plosive ←	p	b					t	d	tʃ	dʒ			k	g	ʔ	
Fricative			f	v	θ	ð	s	z	ʃ	ʒ					h	
Affricate																
Nasal		m						n						ŋ		
Lateral								l						(l)		
Rolled								/r/								
Flapped								ɾ								
Semi - vowel		w										y		(w)		

Note. All types of consonants are categorized by place or point of articulation (horizontal axis), manner of articulation (vertical axis) and voicing, i.e. the pairs of phonemes in a single cell

represent a voicing contrast (i.e. a minimal pair of voiced/voiceless phonemes). Symbols between () are secondary articulation and // are dialectical sounds.

Source: Ramelan, 2003, p. 108

2.1.2. Speech Organs

Ramelan (2003) described the subdivision of speech organs based on the function into (a) initiator, (b) phonator, and (c) articulators. The main initiator refers to the lungs that set air into motion to produce speech sounds (pulmonic air), while the phonator refers to the vocal cords in the larynx to produce voiced sounds. The articulators are the speech organs that are used to obstruct the out-going air in the production of speech sounds.

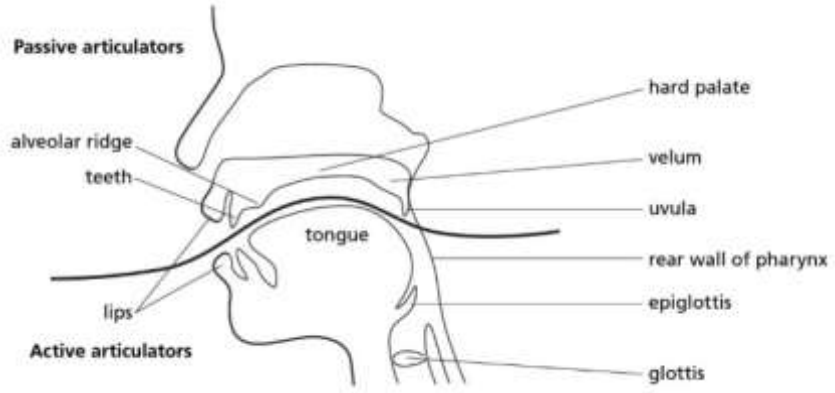
2.1.3. Articulatory Phonetics

Articulatory phonetics is defined as a subdiscipline dealing with the production, perception, identification, and categorization of speech sounds (Bickford & Floyd, 2006). Thus, theories of speech production or how speech sounds are made or articulated can be learned in this Study (Yule, 2010; Ashby, 2011). Articulators, places of articulation, and manners of articulation play important roles in the subdiscipline.

The speech organs or articulators are divided into two types (Ashby, 2011; Przedlacka, 2018). The first type is active articulators such as tongue tip, tongue body, and lower lip, which movable in speech production. The second is passive articulators such as teeth, alveolar ridge, hard palate, and the wall of the pharynx, that are fixed or stationary.

Figure 2.

Midsagittal Vocal Tract Drawing of Active and Passive Articulators



Source : Ashby, 2011, p. 33

A place of articulation is any passive point that is approached or contacted by its relevant active organ (Ashby, 2011). Eleven names of articulatory place such as bilabial, labiodental, dental, alveolar, postalveolar, retroflex, palatal, velar, uvular, pharyngeal, and glottal are identified along the horizontal axis of IPA chart.

The manner of articulation describes the degree of impedance of the airstream and the type of closure that produces that impedance (Bickford & Floyd, 2006). Yule (2010) investigated that some sounds of the same category in terms of where they are articulated can differ in their manner of articulation, that is, in the way they are pronounced. For example, [t] and [s] are both voiceless alveolar sounds. However, in terms of how they are articulated, the [t] sound is one of a set of sounds called stops and the [s] sound is one of a set called fricatives. The degree of narrowing or 'stricture' of the airflow caused by movement of the articulators has three possible levels; (a) complete closure, (b) partial closure and (c) approximation (Rogerson-Revell, 2018).

(a) Complete closure – a complete 'stoppage' of the air flow, producing plosives, e.g. /p/ in 'pan' bilabial plosive, /k/ in 'kit' velar plosive

(b) Partial closure – a narrowing of the air stream

(i) Causing friction between articulators, producing fricatives, e.g. /f/ in 'fun' labiodental fricatives, /θ/ in 'think' dental fricatives

(ii) Causing air to 'go round' the closure, producing laterals and nasals, e.g. /l/ in 'lid' alveolar lateral, /m/ in 'men' bilabial nasal

(c) Approximation – proximity of the articulators without actual contact, producing approximants, e.g. /r/ in 'run' alveolar approximant, /w/ in 'wet' bilabial approximant, /y/ in 'yes' palatal approximant.

2.2. Phonology

Phonology has several definitions in terms of the function of phonemes in a language. Dardjowidjojo (2009) stated that the study handles how words are phonologically formed and the arrangement of sounds to form a syllable and a word. According to Yule (2010), phonology is essentially the description of the systems and patterns of speech sounds concerning with the abstract or mental aspect of the sounds that allows us to distinguish meaning in the actual physical sounds we say and hear. Phonology concerns how sounds function in a systemic way in a particular language (Cruttenden, 2014). Therefore, this study investigates the phonetic

phenomena from the point of view of their use, which is how speakers systematically use a selection of units (phonemes) to express meaning.

2.2.1. Phonemes

A phoneme is a distinctive speech sound which can differentiate meanings of the word. A phoneme is defined as the mental representation of a specific speech sound (Gut, 2009). Howe (2003) emphasized that phonemes are by themselves meaningless but acquire meaning in combination. As the smallest phonological units, the phonemes or sound segments can create a linguistic difference in meaning, such as the vowel phonemes /e/ and /ɪ/ distinguishing the words 'pen' and 'pin', and the consonant phonemes /t/ and /d/ distinguishing the word 'lit' from 'lid' (Rogerson-Revell, 2018).

2.2.2. Segmental Phonology

The composition, combinations, and the function of speech sounds can be studied in segmental phonology. Segmentals refer to sound units arranged in a sequential order, for example, the utterance 'good heavens' /gud_hɛvəns/ has nine segmentals (Ramelan, 2003). Each segment has phonological features regarding the phonetic properties. The features are psychological entities defined in terms of articulatory realization which provides the link between cognitive representation of speech and its physical manifestation (Hall, 207). Segmental phonology focuses on speech sounds (segments), which refer to their internal composition and external interactions (Howe, 2003). Unlike phonetics, segmental phonology is not interested in the production, the physical properties, or the perception of these sounds, but in the function and possible combinations of sounds within the sound system (Skandera & Burleigh, 2005).

2.3. Second Language Learning

Language learning is viewed from the perspective of the student or learner rather than the teacher, and as the background for language teaching and language use that sets their potentials and constraints. Language learning stresses the contribution of the learner as a group representative and as an individual to the process of acquiring a second or foreign language, whether in a naturalistic or classroom setting (Pennington & Rogerson-Revell, 2019).

Then, the learning processes of Second Language (L2) differ from the First Language (L1) acquisition fundamentally. While the L1 acquisition is inextricably interwoven with the development of muscular control and cognitive abilities, most L2 learners have full command

over their speech organs and fully developed cognitive competence. They have already acquired patterns of muscular activities and have formed the corresponding mental representations for the production of speech in their L1 (Gut, 2009).

English as a foreign or second language is learned later at any ranges of age after the native or L1 acquisition. Nowadays, English is learned mostly at young ages because of the influence of the global importance of English skills. In learning a target language, pronunciation is one of important basic skill to master spoken English. However, English learners, especially the young learners, often encounter many difficulties in obtaining language proficiency. Mispronunciation, which is caused by phonological problems, is the main cause of low spoken language performance that has to be improved.

2.3.1. Second Language Learners

Younger learners, particularly children aged six to twelve, are better than adults in learning a second language in terms of psychological and social factors. Steinberg & Sciarini's (2006) revealed that children have better intellectual processing, memory ability, and motor skills to produce good pronunciation. The child who is often exposed to second language speech and remembers what he or she has heard will be able to analyze and discover the generalization or rule that underlies that speech.

2.3.2. Pronunciation

Pronunciation involves both phonetics and phonemics, which is a traditional approach to phonology that analyses the stream of speech into a sequence of contrastive segments (Cruttenden, 2014). Language teachers prefer the term 'pronunciation' referring to an area of proficiency in language learning or a type of skill in spoken language performance, rather than phonology.

English pronunciation is tricky because of the lack of consistency in this language. Modern English has been formed from and influenced by different languages namely Latin, Greek, French, and German to its evolution from Old and Middle English, so that how words are spelt and pronounced often seem inconsistent (Farlex International, 2017).

2.3.3. Mispronunciation

Mispronunciation mau hinder communication. Incorrect pronunciation can result in the unclear and less intelligible speech so that the listeners may misunderstand the utterance of the

speaker. Inaccurate pronunciation of individual vowels or consonants can sometimes cause real problems in communication. Segmental mispronunciation may interfere with understanding and communicative purpose to a greater or lesser degree (Pennington and Rogerson-Revell, 2019).

Differences of L1–L2 sound systems are two main problems causing mispronunciation.

a. Differences of L1–L2 Sound Systems

Phonemes of a target language or a second language (L2) may be different from the native or first language (L1) ones, which often interfere on L2 pronunciation. Pennington and Rogerson-Revell (2019) explained that the sound system of a language consists of its individual phonemes, the distinctive consonant and vowel sounds of the language, and their contextual variants (or allophones), the specific pronunciations of the phonemes in different contexts. Speech sounds, the meaningful components of the phonological system, are one of the sound system aspects which are likely to differ for L1 and L2 (Saville-Troike, 2006).

- Mispronounced phonemic segments

Mispronounced phonemic segments were produced through a substitution of equivalent or nearest available L1 phoneme for L2 phoneme, which is absent and difficult to pronounce. Consequently, the learners tend to replace the difficult L2 phoneme with the available L1 phoneme which has the closest phonetic properties. For example, the fricative /v/ in the word ’van’ is often mispronounced as voiceless /f/ because the two sound have similar articulatory features except the voicing.

- Silent letters

Farlex International (2017) stated that the presence of silent letters is a result of assimilation of various English spelling and pronunciation since the language has evolved from several different sources (such as Latin, Greek, French, German, Old English). Silent letters help distinguish homophonous words and indicate the meaning or origin of a word. Pronunciation of English words is influenced by silent vowels or consonants, such as silent E, U, B, C, D, G, H, K, L, M, N, P, T, W, TH, and GH.

- Tricky words

Tricky words are also called common exception words, which are the words that do not follow the common phonetic spelling rules. Young learners always have difficulty to recognize and articulate tricky words because they are not decodable using tye normal or regular rules and

letter-sound in phonics. The irregular letter combinations in a vowel or consonant sounds blend in a tricky word can be difficult to pronounce correctly (Farlex International, 2017).

b. L1 Phonological Interference

The prior knowledge of the native language commonly influences target language learning. Saville-Troike (2006) divided cross-linguistic influence or transfer of prior knowledge from L1 to L2 into two major types;

(1) positive transfer, when an L1 structure or rule is used in an L2 utterance and that use is appropriate or "correct" in the L2;

(2) negative transfer (or interference), when an L1 structure or rule is used in an L2 utterance and that use is inappropriate and considered an "error."

Transfer from L1 to L2 phonology may occur in all levels of interlanguage; vocabulary, pronunciation, grammar, and all other aspects of language structure and use (Gut, 2009). Language learners, whose first L1 requires different phonological representations of phonological units and rules from the L2, might, at least at the beginning of language learning, inappropriately use these units and rules in the L2.

Negative transfer from L1 to L2 often cause mispronunciation. Hayes (2009) explained negative transfer as a phenomenon of pronunciations in a second language in ways attributable to the phonology of the first language. Any structure in L2 which has a form not occurring in L1, where there is partial overlap but not equivalence in form, meaning, and/or distribution, are most likely to cause interference (Saville-Troike, 2006). The interference may decrease as the learner develops familiarity with the L2 (Yule, 2010).

3. Method of Research

3.1. Type of Research

The research design used in the present study is descriptive qualitative research. Qualitative research is an activity or process of "understanding" the reality of phenomena with a natural setting, focusing on descriptive data provided by triangulation for analysis, thus generating holistic understanding based on the participant's perspective according to the context (Muhammad, 2011). The scope of the research is English as foreign or second language learning in a naturalistic context, where most young learners use the target language as a primary medium

of instruction at school or a means of communication in daily life. The study focuses on English pronunciation learning to obtain the obvious understanding about mispronunciation, the cause, and the improvement.

3.2. Data Source

Data obtained in this descriptive study are in the form of words or orthography, and phonetic and phonemic transcriptions taken from the speech sounds of the mispronounced words spoken by young English learners in Diana private course. Primary data are gathered from participational through conversation in one-to-one tutoring located in each participant's house. Secondary data are retrieved from an online dictionary, internet articles, and ebooks. The substantive data, which are the transcript of speech sounds of the mispronounced words, are used to identify the different elements and interference causing the mispronunciation. The phonetic transcription of English vocabulary including the correct pronunciation of the mispronounced words is retrieved from online Lexico Oxford Dictionary.

This study has three limitations, which are the scope of data analysis, variables and the sample size. First, this research analyzed the articulatory features of lexical pronunciation on segmental level. Second, this study involved two variables, which are age and learning setting. This research used young learners as subject and outside classroom or informal setting. Third, the sample consists of ten participants to obtain detail information and obvious understanding of the topic.

3.2.1. Population

Population refers to all members of the community (Buchstaller & Khattab, 2013). The population in this study is all young English learners, which are dominantly elementary school students in Semarang.

3.2.2. Sample

A sample is a group of the population who becomes the target participants for data collection. In linguistic research, a representative sample is the subgroup of people that reflects the population as a whole in terms of their social and linguistic characteristics (Buchstaller & Khattab, 2013). This study sample is a group of ten young English learners aged between seven and eleven years old in Diana private course. The participants are the students from several excellent private and national elementary schools as listed in the table below.

Table 1.

Participants of the Study

No	Names of Participant	Grades	Names of Elementary School
1	Kenston	5	Maria Regina
2	Darren	3	Marsudirini
3	Olsen	2	Daniel Creative School
4	Kayla	2	Daniel Creative School
5	Matthew	2	Daniel Creative School
6	Azka	5	Nasima
7	Sarah	5	PL Bernardus
8	Sofi	3	PL Bernardus
9	Cinta	5	Kutowinangun 12
10	Zelly	3	Kutowinangun 9

Referring to the objectives of the research, the sample was selected by using purposive sampling based on two variables, which are age and learning setting. In purposive sampling, the sample units are chosen because the informants or participants have particular features or characteristics which will enable detailed exploration and understanding of the central themes and puzzles which the researcher wishes to study (Ritchie & Lewis, 2003). The sample in this study was chosen based on the writer's judgement about the most informative data potentially provided by the sample units. Young learners were preferred because primary school-age children, especially bilinguals or multilinguals, often encounter pronunciation problems at the beginning of the foreign language learning process. Thus, the sample will provide detailed information and obvious understanding of mispronunciation.

3.3. Method of Data Collection

Simak and cakap are two methods generally used for data collection (Sudaryanto, 1993). The research data in this study are collected by using simak and cakap methods. The first method is applied with a basic technique sadap, followed by two further techniques, simak libat cakap and catat. The method is conducted through a participational observation. The second method is

applied with basic technique pancing, followed by cakap semuka and catat techniques. The method is conducted through a semi-structured face-to-face interview.

Muhammad (2011) explained that simak method is equal to observing method in anthropology or social researches, and employed by observing the language use. The method uses a basic technique sadap, which is conducted by tapping the participants' speech. In the participational observation, the writer applies sadap technique by tapping the informants' speech and simak libat cakap by participating in the conversation. Cakap method is equal to interview method, where the interviewer directs the dialogue to obtain all the required information muhammad, 2011). In the semi-structured interview, the writer uses pancing technique by persuading the informants with spontaneous questions, and cakap semuka technique through face-to-face discussion on the problems encountered in English pronunciation learning.

3.4. Method of Data Analysis

The research data are analyzed by using padan fonetik artikulatoris method with a basic technique named teknik pilah unsur penentu and further technique named teknik hubung-banding. The method uses speech articulators as key factors to determine the research object, which is the speech sounds of mispronounced words. In the basic technique, the consonantal and vowel segments of the mispronounced words are described based on the articulatory features. In the further technique, the writer attempts to explain that the mispronounced segments representing the distinctive phonemes which may result in different meanings of the word and cause interpretation. By comparing and contrasting the phonetic properties and phonological features of the distinctive phonemes, the problems contributing to the mispronunciation can be managed.

According to Muhammad (2011), padan method is a way of analyzing data to answer the research problems using nonlinguistic aspects as determiners or key factors such as reference, articulatory organs, foreign language, orthography, speech partners, and social aspect. The method employs teknik pilah unsur penentu as a basic technique and teknik hubung-banding as the further technique (Sudaryanto, 1993). The basic technique in padan fonetik artikulatoris method uses a set of speech articulators as the key factor for analysis, whereas the further technique is conducted by identifying the domain's constraints and the relevant key factor.

3.5. Method of Data Representation

Data can be represented by using formal and informal methods. Formal method is a method of data representation by using marks and symbols, whereas informal method is a method that uses words to describes the research focus according to the domain's constraints and the correlation of the features (Muhammad, 2011). Sudaryanto (1993) mentioned marks such as plus (+), minus (-), asterisk (*), The arrow (→), round brackets (()), curly brackets ({}), and square brackets ([]), are commonly used in formal method. Common symbols used in formal method are letters in abbreviation and acronyms.

The data in this study will be represented by using both formal and informal methods. In formal method, the writer uses several marks such as square brackets ([]) to enclose the phones, phonetic transcriptions, and phonological features, double slashes (//) to enclose the phonemes, angle brackets (<>) to enclose the spelling, period (.) to separate the syllables in the phonetic transcriptions, and plus-minus (±) to mark the presence or absence of binary phonological features. Since the writer uses British English, the data in this study will be represented in British transcription.

4. Result and Discussion

Data showed segmental changes in 24 mispronounced the words consisting of 11 words with mispronounced phonemic segments, 6 words with silent letters, and 7 tricky words. The 11 words such as sick, she, van, three, the, usually, owl, circle, sphere, fiery, and crayon had vowel or consonant changes. The participants also said the silent letters which should be unpronounced. The tricky words are mostly borrowings which have inconsistent spelling pronunciation and do not follow the common phonetic spelling rules.

Table 2.

Research Data of Mispronunciation

No	Words	Correct pronunciation	Mispronunciation	Number of Participants who mispronounced
1	circle	[ˈsəː.kəl]	[ˈsɪr.kəl] [ˈsiːr.kəl]	2 4
2	crayon	[ˈkreɪ.ɔn], [ˈkreɪ.ən]	[ˈkraɪ.ɔn]	3
3	eight	[eɪt]	[eg]	6
4	fiery	[ˈfaɪə.ri]	[ˈfiː.ri]	5
5	gauge	[ɡeɪdʒ]	[gaʊdʒ]	4

			[gaʊg] [gɔːg]	3 1
6	listen	[ˈlɪs.ən]	[ˈlɪs.tən]	7
7	mosque	[mɔsk]	[mɔskyuː] [mɔːskuiː]	4 3
8	owl	[aʊl]	[əʊl] [ɔwəl]	4 3
9	queue	[kyuː]	[kyuː.iː] [kyuː.iː.uː]	3 4
10	she	[ʃiː]	[siː]	3
11	sick	[sɪk]	[siːk]	5
12	sphere	[sfiər], [sfɪr]	[spɪr] [sfɛr] [spɛr]	2 1 3
13	subtle	[ˈsʌt.əl]	[ˈsʌb.təl] [ˈsuːb.təl] [ˈsʌbt.liː] [ˈsuːbt.liː]	4 1 1 2
14	the	[ðə]	[də]	3
15	three	[θriː]	[triː]	6
16	through	[θruː]	[θruːg] [θrɔːg] [trɔːg]	2 2 2
17	tongue	[tʌŋ]	[tɒŋ] [tɔːŋ.yuː] [tɔːŋuːe]	2 1 3
18	tortoise	[ˈtɔː.təs]	[ˈtɔː.tɔɪs]	6
19	usually	[ˈyuː.ʒʊ.ə.li]	[ˈyuː.ʃʊ.ə.li] [ˈyuː.sʊ.ə.li]	3 3
20	van	[væn]	[fæn] [fɛn]	3 3
21	vegetable	[ˈvɛdʒ.tə.b(ə)l]	[ˈvɛ.dʒə.teɪ.bəl] [ˈvɛ.gə.teɪ.bəl] [ˈfiː.giː.teɪ.bəl] [ˈfɛ.gə.tɛ.bəl]	1 1 1 5
22	vehicle	[ˈviː.ə.kəl]	[ˈviː.haɪ.kəl] [ˈfe.hiː.kəl]	4 3
23	write	[raɪt]	[wraɪt] [wriːt]	2 2
24	Wednes-day	[ˈwɛnz. deɪ]	[ˈwɛt.nəs.deɪ] [ˈwɛt.nɛs.deɪ]	5 3

Difference between English and Indonesian phonemic inventories result in unavailability or an absence of L2 phonemes, that becomes the main problem in the articulation of the absent

phonemes. English has 44 phonemes consisting of 20 vowels and 24 consonants, whereas Indonesian has 32 phonemes consisting of 23 consonants, 6 monophthongs, and 3 diphthongs. The twenty-three consonants acknowledged by Indonesian today are composed of eighteen native consonant phonemes are /p, b, t, d, k, g, s, h, c, j, m, n, ɲ, ŋ, I, r, w, y/, four loan consonants /f, ʃ, z, x/, and a glottal /ʔ/ (Lapoliwa, 1981). However, the absent English phonemes such as /ɪ, æ, əː, ɔ, ʌ, eɪ, aʊ, ʊə, ɪə, aɪə, ʃ, θ, ð, ʒ, voiced v/ cause mispronunciation.

The absence of many English phonemes in Indonesian sound system contributes to either vowel or consonant changes in the mispronounced words. Based on the difference between L1 and L2 phonemic inventories, English has more phonemes that are not available in Indonesian phonology. The absent English vowels such as /ɪ, æ, əː, ɔ, ʌ, eɪ, ʊə, ɪə, aɪə/ have distinct vocal tract configurations, whereas the consonants like /ʃ, θ, ð, ʒ, voiced v/ have different places and manners of articulation. Consequently, most of the participants had difficulty in the articulation and tended to replace the sounds.

The absent English silent letters are often incorrectly pronounced. Many participants incorrectly pronounced [w] of silent W in the word ‘write’, [t] of silent D in the word ‘Wednesday’, [g] of silent GH in the word ‘eight’, [b] of silent B in the word ‘subtle’, [t] of silent T in the word ‘listen’, and [ə] in the second syllable of the word ‘Wednesday’ and ‘vegetable’. The results are consistent with the previous finding reported by Aswad et al. (2020) about pronunciation errors caused by English common silent consonant letters such as silent W in the initial position of the word, silent D, and silent T in the medial position of the word.

To improve the mispronunciation, the participant must increase the knowledge of pronunciation rule on English silent letters. W is silent when the letter comes before R. D is silent when the letter appears after N. Silent GH usually appears after spellings <ou>, <au>, <ai>, <ei>, and letter I like in the words ‘bought’, ‘caught’, ‘straight’, ‘eight’, and ‘high’. B is silent when the letter proceeds T and comes after M. T is silent when the letter comes after S and is followed by endings <-le> and <-en> like in the words ‘castle’ and ‘listen’.

The inconsistent English pronunciation rules on spelling patterns may cause the mispronunciation. The pronunciation of many English spelling patterns such as <ow, ir, ie, ay, ph, th> differ from Indonesian’s pronunciation, and the patterns like <ueue, ough> are not available in native sound system. English digraph <ow> may represents either [əʊ] like in the

word ‘low’ or [aʊ] like in the word ‘cow’. The digraph <ow> in the word ‘owl’ represents the diphthong [aʊ]. The digraph <ir> in the word ‘circle’ represents [əː]. The spelling <ie> in the middle of the word may represent [iː] like in the word ‘believe’ and the triphthong [aɪə] like in the word ‘fiery’. The spelling <ay> always represents [eɪ]. The digraph <ph> represents [f], whereas <th> represents [θ] sound. The tetragraph <ueue> in the word ‘queue’ represents [yuː]. The tetragraph <ough> may represent four sounds, which are [əʊ] like in the word ‘thorough’, [aʊ] like in the word ‘drought’, [ɔː] like in the word ‘fought’, and [uː] like in the word ‘through’.

Interference or negative transfer of L1 takes place through the substitution of native sound with closest articulatory features for the absent L2 phoneme. The consonantal and vowel segments of the mispronounced words have phonetic properties and phonological features that differs from the correct English segments. Since the prior knowledge of L1 of most young English learners influences L2 pronunciation learning, most of the participants, who mispronounced the words, inappropriately use the native pronunciation rules and change the absent English phonemes with other available sounds that are mostly similar in articulatory features. This is supported by Arvianti’s (2012) about the interference of mother tongue resulting in vowel change.

The [ɪ] and the mispronounced [iː] segments in the word ‘sick’ are two distinctive phonemes. The /ɪ/ is [-tense], while /iː/ is [+tense], which is produced with a greater tension of the tongue muscle than the lax vowel /ɪ/. The /ɪ/ is an unrounded half-close or mid-high front vowel, while /iː/ is an unrounded high or close front vowel. Both vowels are produced by spreading the lips and raising the front of the tongue. The only difference between the two vowels is tongue height. Since /iː/ has higher tongue position and greater tension than /ɪ/, the mispronunciation can be improved by lowering the tongue from the close or high to the half-close or mid-high position and opening the jaws a bit wider.

As reported and supported by Mulya and Mujiyanto’s (2018), Ambalegin & Hulu (2019), and Rahman et al. (2020), the evidence in this study points to the absence of [æ] sounds in Indonesian sound system contributed to mispronunciation. The [æ] and the mispronounced [ɛ] segments in the word ‘van’ are two distinctive phonemes that differ in tongue height and jaws opening. The /æ/ is [+low], an unrounded low or open front vowel, whereas /ɛ/ is [-low], an unrounded mid-high or half-close front vowel. To improve the mispronunciation, the articulation

of mid-high /ɛ/ should be moved into low /æ/ by spreading the lips, widen the jaws opening, and retracting the tongue front until the tip of the tongue touches the bottom front teeth.

The [aʊ], the mispronounced [əʊ] and [ɔ] segments in the word 'owl' are three distinctive phonemes with different articulatory features. /aʊ/ is a front-back upgliding or closing full diphthong, /əʊ/ is a central-back half diphthong, and /ɔ/ is an open or low back vowel. In articulating /aʊ/, the front part of the back (central part) of the tongue is slightly raised to the first element /a/ position and then moves to the direction of /ʊ/. To improve the mispronounced /əʊ/, the central part of the tongue should be moved from the mid /ə/ to the low /a/ element of the diphthong. To improve the mispronounced /ɔ/, the tongue should be moved or glided to /ʊ/ with lips rounding.

The absent English sound [əː], which is often mispronounced into [ɪ] or [iː], has distinctive phonetic features. In previous finding, the substitution of [iː] for [əː] also took place in the English mispronunciation produced by Buginese and Makassarese students (Utami et al., 2007). The [əː], the mispronounced [ɪ] and [iː] segments in the word 'circle' are three different phonemes with different tongue positions. /əː/ is a half-open or mid-low central vowel that is produced by raising the central part of the tongue to the half-open position with neutral lips. To improve the mispronounced [ɪr], the tongue should be moved or retracted from high-close or mid-high /ɪ/ to the half-open or mid-low /əː/ and the /r/ is unpronounced. To improve the mispronounced [iːr], the tongue should be retracted from high /iː/ to the mid-low /əː/ position.

Regarding the pronunciation or the presence of [r], letter R in English can be either pronounced or silent in some cases. When R appears in letter combinations such as <br, gr, phr, the> in the front and medial position of a word, [r] is always pronounced. When R appears as a single sound in the front of a word, the [r] is also pronounced. However, R as a single sound appearing in the medial and the end of a word is pronounced or silent.

The [eɪ] and the mispronounced [aɪ] segments in the word 'crayon' are two distinctive phonemes with different phonetic properties. Both are front closing half diphthongs that differ in tongue height of the first element. The /eɪ/ is produced by putting the tongue on the front /e/, which is a little lower than half-close position, then moving in the direction of /ɪ/, which is a little above the half-close. The mispronunciation can be improved by moving the tongue from the low or open front /a/ up to the mid-high or half-close /e/, then gliding to /ɪ/.

The [ʃ] and the mispronounced [s] segments in the word ‘she’ are two distinctive phonemes. /ʃ/ is a rounded voiceless palato-alveolar fricative and /s/ is an unrounded voiceless blade-alveolar fricative, where both phonemes are produced by raising the tongue blade to the teeth ridge and the outgoing air escapes with the frictional sound. The differences between the two phonemes are the lips shapes and the points of articulation. /ʃ/ is articulated by putting the tongue blade at the point between the teeth ridge and the hard palate with rounded lips, whereas /s/ is articulated by putting the tip and the blade of the tongue very close to the teeth ridge with spread lips. Therefore, the mispronunciation can be improved by rounding the lips and moving the tongue blade further backwards to a point between teeth ridge and hard palate.

The absence of [ð] sound in Indonesian consonantal inventory causes the substitution of the native [d] for [ð] sound. This result is confirmed by Bui (2016) and Ambalegin and Arianto (2018), who reported the negative transfer of Javanese sound that caused President Joko Widodo to mispronounce the sound.

The [ð] and the mispronounced [d] segments are articulated through different places and manners of articulation. /ð/ is a voiced dental fricative involving construction on the narrowing between the tongue front and the back part of the upper teeth (post-dental), whereas /d/ is a voiced alveolar plosive involving plosive airstream release from a complete closure between the tongue tip and teeth ridge. Thus, the mispronunciation can be improved by moving the tongue front forwards to the back part of the upper teeth.

Since [θ] sound does not exist in the native sound system, most young Indonesian students tend to replace the English consonant with either alveolar [t] or interdental [t] sound. This finding is consistent with the previous result showing the substitution of [t] for the absent [θ] in the native sound system (Kosasih, 2017; Utami et al., 2017). This result is also supported by Bui (2016), who reported the EFL learners' tendency of replacing /θ/ with Vietnamese /t/ in the initial, medial and final positions of the words.

The [θ] and the mispronounced [t] segments in the words ‘three’ and ‘through’ are two distinctive phonemes with different phonetic properties. The fricative /θ/ is [+continuant] [+distributed], which is produced impeding but not completely blocking so that the primary constriction in the vocal tracts does not block the continuous airflow. The obstruction in articulating the fricative extends over a considerable area along the middle-line of the oral tract. Whereas the plosive /t/ is [-continuant] [-distributed], which is produced by completely blocking

the airflow through the centre of the vocal tract with a smaller area of contact. Phonetically, /θ/ is a voiceless dental fricative that is produced by putting the tongue tip very close to the upper teeth forming a narrow passage through which the airstream escapes with audible friction. Whereas /t/ is voiceless alveolar plosive that is produced by putting the tongue tip close to the teeth ridge (alveolum) so that the airstream escapes with a plosive sound. The mispronunciation, thus, can be improved by moving the tongue tip from the alveolum to the upper teeth.

The substitution of [f] for [v] because both [f] and [v] are voiceless in Indonesian sound ssystem However, the [v] and the mispronounced [f] segments in the words 'van', 'vehicle', and 'vegetable' are two distinctive phonemes with a slight difference. Although both phonemes are labiodental fricatives, /v/ is voiced or [+voice] and /f/ is voiceless or [-voice]. Therefore, the mispronunciation can be improved by vibrating the vocal cords.

The absence of English consonant [ʒ] in the native consonantal inventory caused the young learners to replace the sound with either [ʃ] and [s] sound. This result confirms previous findings in the literature. The substitution of [ʃ] for [ʒ] in the medial position of the word is consistent with the result reported by Rahman et al. (2020) and Kosasih (2017). The absence of both [ʃ] and [ʒ] in Indonesian sound system caused the tendency to replace [ʒ] in the middle of the word with the native [s] sound. This is also supported by previous findings (Hentasmaka, 2015; Utami et al., 2017; Komariah, 2018).

The [ʒ], the mispronounced [ʃ] and [s] in the word 'usually' are three distinctive phonemes with different articulatory features. /ʒ/, a voiced blade-alveolar fricative, is the counterpart of /ʃ/, a voiceless palato-alveolar fricative. In producing both /ʒ/ and /ʃ/, the tongue blade is raised towards a point between teeth ridge and hard palate, the main body of the tongue is raised towards the hard palate, and the tongue tip is retracted. In producing /s/, the tip and the blade of the tongue almost touch the teeth ridge. To articulate voiced /ʒ/ correctly, the first pronunciation can be improved by vibrating the vocal cords and raising slightly the tongue blade while articulating voiceless /ʃ/. Whereas the second pronunciation can be improved by retracting the tip and the blade of the tongue from the teeth ridge toward the hard palate and vibrating the vocal cords.

The three of seven tricky words, namely 'mosque', 'queue', and 'gauge' are borrowings from French that have spelling pronunciations follow the former rules, not the common English

spelling patterns. Thus, the participants must enhance the knowledge of English borrowings and avoid the incorrect syllable segmentation to pronounce the words correctly. The four tricky words such as 'vehicle', 'tortoise', 'tongue', and 'through' do not follow the common phonetic spelling rules spelling and are not decodable using the regular rules so that most of the participants mispronounced the words. To improve the mispronunciation, the participants must enhance the vocabulary on tricky words.

5. Conclusion and Suggestion

5.1. Conclusion

Mispronunciation is often caused by two main problems, namely different elements between L1 and L2, and phonological interference of the mother tongue. The different elements mostly include distinctive phonemes of the two phonemic inventories and different spelling pronunciation rules of L1-L2 sound systems. The majority of the participants demonstrated some degrees of difficulty in the articulation because of the absence of English phonemes, silent letters, and spelling patterns in Indonesian sound system. Absent phonemes such as /ɪ, æ, əː, ɔ, ʌ, eɪ, ʊə, ɪə, aɪə, ʃ, θ, ð, ʒ, voiced v/, absent silent letters <b, d, e, t, w, gh, r in some cases>, unavailable spelling patterns such as <th, ow, ph, ie, ueue, ough>, and tricky words including English borrowings such as 'mosque', 'queue', and 'gauge' with inconsistent spelling pronunciation caused the mispronunciation.

The different elements may trigger the occurrence of L1 phonological interference, where the prior knowledge of L1 phonology transfers negatively into L2 so that the young learners tend to substitute the closest sounds for the absent phonemes, add the sounds for silent letters, and apply L1 pronunciation rules. Also, the lack of knowledge of English tricky words and low phonemic awareness contributed to L2 pronunciation through negative transfer of L1 sounds. Therefore, the phonological interference of L1 occurs when the mispronounced phonemic segments change both the phonetic form and the meaning of the words.

The mispronunciation can be improved by developing the familiarity with the L2 phonology. The young English learners must increase the phonemic awareness, enhance the English vocabulary, and do more practice on the correct pronunciation in verbal communication.

5.2. Suggestion

1. For English Teachers

- To perform a proficient model of pronunciation.
- To teach the students to be excellent users of English.
- To apply more effective methods and techniques namely spelling pronunciation, homophones, homographs, and tongue twisters can be applied through quizzes or games to improve young learners' pronunciation.

2. For English Learners
   - To be more aware of the significant roles of pronunciation.
   - To practise conversational skills everyday.
   - To have an enthusiasm for learning English.
3. For those who are interested in researching English mispronunciation
   - To investigate the broader ranges of pronunciation features, such as suprasegmental features, phonological rules and processes.
   - To involve social and psychological variables such as age, learning setting, motivation, and attitude.
   - To provide the valid and authentic data such as video or voice records for data collection.
   - To use pronunciation assessment systems or technology-based devices to obtain more accurate data.

6. References

Ambalegin & Arianto, T. 2018. *English Vowels and Consonants Mispronunciation of the Seventh President of Republic of Indonesia in His Official English Speeches and Its Influencing Factors.* Language Literacy: Journal of Linguistics, Literature and Language Teaching, 2 (2), 111-125. http://doi.org/10.30743/ll.v2i2.678.

Ambalegin & Hulu, F. 2019. *EFL Learners' Phonological Interference of English Articulation.* Journal BASIS, 6(2), 145-154. https://doi.org/10.33884/basisupb.v6i2.1415.

Arvianti, I. 2012. *Karonesse Interference in Indonesian as the Reflection of Karonesse Maintenance.* In A. Subyanto, Mualimin, & Prihatini (Eds.), Proceedings of International Seminar "Language Maintenance and Shift II" (pp. 58-62). Diponegoro University. http://www.eprints.undip.ac.id/54074/1/Proceedings_International_Seminar__July_5-6_2012_no_scure_-_Barans_Irawan_Palangan.pdf.

Ashby, P. 2011. *Understanding phonetics*. Bookpoint Ltd.

Aswad, M., Nurchalis, N. F., Yassi, A. H., Nasmilah, Pammu, A., & Arbain. 2020. *Common Silent Consonant Letters Pronounced Incorrectly by Freshmen of English Education Program.* Asian EFL Journal, 27 (1), 145-158. https://www.asian-efl-journal.com/wp-content/uploads/AEFL-Volume-27-Issue-3.1-June-2020.pdf.

Bickford, A. C. & Floyd, R. 2006. *Articulatory Phonetics*. SIL International.

Buchstaller, I. & Khattab, G. 2013. *Population Samples*. In R. J. Podesva & D. Sharma (Eds.), Research methods in linguistics (pp. 74-95). Cambridge University Press.

Bui, T. S. 2016. *Pronunciations of Consonants /ð/ and /θ/ by Adult Vietnamese EFL Learners*. Indonesian Journal of Applied Linguistics, 6(1), 125-134. http://dx.doi.org/10.17509/ijal.v6i1.2744.

Cruttenden, A. 2018. *Gimson's Pronunciation of English (8th ed.).* Routledge.

Dardjowidjojo, S. 2009. *English Phonetics & Phonology for Indonesians.* Yayasan Pustaka Obor Indonesia.

Farlex International. 2017. *The Farlex Grammar Book: Complete English Spelling and Pronunciation Rules (Vol. 3).* Farlex International.

Gut, U. 2009. *Textbooks in English Language and Linguistics (TELL) (Vol. 1)*. Peter Lang GmbH.

Hall, T. A. 2007. *Segmental Features*. In P. De Lucy (Ed.), The Cambridge handbook of phonology (pp. 311-334). Cambridge University Press.

Hayes, B. 2009. *Introductory Phonology*. Blackwell.

Hentasmaka, D. 2015. *Analyzing the Pattern of L1 Sounds on L2 Sounds Produced by Javanese Students of STKIP PGRI Jombang*. Indonesian EFL Journal 1(1), 20-30. http://ejournal.kopertais4.or.id/mataraman/index.php/efi/article/view/190.

Howe, D. 2003. *UCALGARY.CA: Segmental Phonology [PDF]*. University of Calgary. http://www.ai.mit.edu/projects/dm/featgeom/howe-segphon-book.pdf.

Kirkpatrick, A. 2010. *The Routledge Handbook of World Englishes*. Routledge.

Komariah, A. 2018. *Problems in Pronouncing the English Sounds Faced by the Students of SMPN 2 Halong, Banjar*. Journal of English Language and Pedagogy 1(2), 1-10. http://dx.doi.org/10.36597/jelp.v1i2.4127.

Kosasih, M. M. 2017. *Native Language Interference in Learning English Pronunciation: A Case Study at a Private University in West Java, Indonesia.* International Journal of Education and Research, 5(2), 135-150. https://www.ijern.com/journal/2017/February-2017/11.pdf.

Kreidlerr, C. W. 2004. *The Pronunciation of English: A Course Book (2nd ed.)*. Blackwell.

Muhammad. 2011. *Metode Penelitian Bahasa*. Ar-Ruzz Media.

Mulya, D. & Mujiyanto, J. 2018. *The Influence of Serawai Melayunese Dialect towards Students' English Pronunciation.* English Education Journal, 8 (3), 290-300. http://journal.unnes.ac.id/sju/index.php/eej.

Pennington, M. C. & Rogerson-Revell, P. 2019. *English Pronunciation Teaching and Research.* Springer Nature Ltd.

Przedlacka, J. 2018. *An Overview of Phonetics for Language Teachers*. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), The Routledge handbook of contemporary English pronunciation (pp. 39-57). Routledge.

Rahman, A., Asmaradhani, A. I., & Sutarto. 2020. *Phonological Investigation into Students' Mispronunciation of English Words Containing Phonemes Absent in Indonesian.* International Journal of Scientific & Technology Research, 9(4), 1620-1625. https://www.ijstr.org/final-print/apr2020/Phonological-Investigation-Into-Students-Mispronunciation-Of-English-Words-Containing-Phonemes-Absent-In-Indonesian.pdf.

Ramelan. 2003. *English Phonetics*. UPT UNNES Press.

Ritchie, J., & Lewis, J. 2003. *Qualitative Research Practice*. SAGE Publications Ltd.

Rogerson-Revell, P. 2018. *English Vowels and Consonants*. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), The Routledge handbook of contemporary English pronunciation (pp. 93-121). Routledge.

Saville-Troike, M. 2006. *Introducing Second Language Acquisition.* Cambridge University Press.

Skandera, P. & Burleigh, P. 2005. *A Manual of English Phonetics and Phonology.* Narr Francke Attempto Verlag GmbH + Co.KG.

Steinberg, D. D. & Sciarini, N. V. 2006. *An Introduction to Psycholinguistics (2nd ed.)*. Pearson.

Sudaryanto. 1993. *Metode dan Aneka Teknik Analisis Bahasa: Pengantar Penelitian Wahana Kebudayaan Secara Linguistik.* Duta Wacana University Press.

Utami, D. H., Wello, B., & Atmowardoyo, H. 2017. *The Phonological Interference of Students' First Language in Pronouncing English Sounds (A Case Study on Buginese and Makassarese Students).* ELT Worldwide 4 (2), 205-212. https://doi.org/10.26858/eltww.v4i2.4414.

Vrabel, T. T. 2009. *UDC 811.111'342: Lectures in Theoretical Phonetics of the English Language and Method-Guides for Seminars [PDF]*. PoliPrint. http://www.kmf.uz.ua/hun114/images/konyvek/vrabel_tamas_lectures_in_theoretical_phonetics_of_the_english_language.pdf.

Yule, G. 2010. *The Study of Language (4th ed.)*. Cambridge University Press.