Problems of Speech Perception Experienced by the EFL Learners

Adi Sutrisno

Department of Languages and Literature, Universitas Gadjah Mada, Yogyakarta, Indonesia

Abstract—**Most native speakers of English modify complicated sequences in connected speech in order to simplify the articulation process. This habit has created problems in the part of EFL learners as they are not trained enough to extract word sequences from the running speech. As a consequence, misperception occurs, which often leads to the failure in the listening comprehension. This article is intended to prove that such a problem is not solely triggered by the native speaker factor. Other factors might play pivotal roles, too and need to be carefully examined. In light of this interest, a secondary research aimed at revealing factors causing speech perception problems was carried out. The research result shows that problems of speech perceptions experienced by the EFL learners were mostly caused by their phonetic knowledge and phonological competence, besides lexical and syntax knowledge.**

Index Terms—**connected speech, speech perception, misperception, listening comprehension**

I. INTRODUCTION

Problems of speech perception experienced by the EFL learners listening to native speakers of English are real and need to be investigated thoroughly. Though the common reason has been justified in many research studies, that the prime cause of the failure in the listening process is the failure to extract word sequences spoken by native speakers of English, it is important to examine further if there is any other aspect of language which might play significant role too in such a situation.

To approach problems of speech perception, it deems necessary to consider common views shared by Segui, Freuenfelder, and Halle (2001) and Collins (2008) that there has often been a 'clash' between what is conveyed by native speakers of English and what is perceived by the EFL learners while listening. This discrepancy can be the point of departure to find out the problem of speech perception. Therefore, to discuss problems of speech perception experienced by EFL learners, one needs to talk about what is going on in the listening process and what factors cause misperception in the part of the EFL learners in such a process.

Listening to an utterance is essentially a matter of data processing and interpretation of acoustic signals carried out by the listener once he receives the information in a communication event. The process begins with the reception of acoustic signals followed by the identification of speech sounds in the form of both of segmental and supra-segmental (Kent and Read, 2002, p.12). During speech sound identification, the synchronization of background knowledge, knowledge of the context of the situation, and knowledge of the language is automatically activated in the listener's cognitive domain and proceeds in such a way, synthesizing speech sounds and then interpreting them into specific meanings according to the message contained in the acoustic signal received (Bejar, 2000, p.4).

However, the process is not that simple. When the acoustic signals are received in the sensory register, they are not automatically processed in the brain. It depends on whether the listener can transform the acoustic signals into visual representation or not; and whether the visual representation has the pictorial quality or not. If the transformation of acoustic signals into visual representation succeeds, and thus produce a good pictorial quality, the identification of speech sound will be easily executed. Conversely, if the transformation experiences difficulty, the process of visualizing sound will suffer. In this situation, the incoming acoustic waves will remain stationary in the station buffer (Rozenzweig and Leiman, 1982; cited in Sprinthall and Sprinthall, 1994, p.287). Common people term this as "buffering" or "loading".

In the EFL situation, "buffering" is most likely affected by the difficulty of extracting word sequence from the running speech uttered by the native speaker of English. It explains Bond's (2006) statement, reprimanding that "in everyday conversation, speakers employ various reductions and simplifications of their utterances, so that what they say departs in significant ways from the clarity norms found in normal speech of laboratory recordings" (p. 290). Bonds (2006) elaborated further, that to restore the halted process, a listener must use up all the skills he has so as to facilitate his sense of perception to decode the meaning intended by speaker. Therefore, the interlocutor must be skillful in assimilating the acoustic signals and the discovery of the intended meaning" if he wants to be able to perceive the utterance (Sebastian-Gales, 2006). Implied in such a statement is the fact that various factors do play role in the process of listening experienced by EFL learners. So, what factors are involved in the problem of speech perception and what is the most pivotal one directly related to it?

Research studies have proven that problems of speech perceptions experienced by EFL learners can be caused by

many factors, such as the exposure to the language being studied (Brown, 1990), short term memory (Jacquemot and Scott, 2006; Ohata, 2006), language knowledge, situation knowledge, background knowledge (Bejar, 2000), multiple intelligences (Gardner, 1983), meta-cognitive ability (Pierce, 2003), cognitive ability (Teasdale and Barnard, 1993), and others. However, the immediate factor which is directly associated with problems of perception is the ability to perceive speech sound (Goh, 2000; Bond, 2006; Sebastian-Gales, 2006; Ingram, 2007; Vandegrift, 2007).

Consistent with what has been shared by Goh (2000); Bond (2006); Sebastian-Gales (2006), Ingram (2007), and Vandegriftn (2007). this article contends that in the heart of perception problems, phonetic and phonological competence plays a significant role in the identification of speech sound by the EFL learners.

II.   THE ROLE OF PHONETIC AND PHONOLOGICAL COMPETENCE IN THE IDENTIFICATION OF SPEECH SOUND

To have a better picture of the role of phonetic and phonological competence in the identification of speech sound, following is presented a detailed explanation about what is actually going on in the brain when listening is in progress.

Listening is a mental process that takes place in the human brain in response to the coming stimulus in the form of an acoustic signal (Kent and Read, 2002) that carries messages brought into the temporal lobe to the cortex. Process of listening is not as simple as "plugging ear" then all the information conveyed by the interlocutors fills short-term memory slot of the listener and then suddenly understands the messages received. According to Hewlet and Beck (2006, p.17) to be able to decode meaning conveyed by the interlocutors, one must exert all his ability to capture the acoustic signals with the aid of the language knowledge, knowledge of the context, and the ability to anticipate the direction of the subject. Thus, according to Hewlett and Beck (2006, p.17) the listening process takes place inferentially by a combination of bottom-up and top-down, in that bottom-up deals with a person's knowledge of the language; while top-down is related to the context of the knowledge and direction of anticipation made by the listener.

According to Kent and Read (2002, p.12) in the process of listening there are three domains interconnected with one another, namely the realm of physiology (the physiologic arena), acoustic (the acoustic arena), and perception (the perceptual arena). The realm of physiology is related to the articulation of sound, which is how the sound is made by the interlocutors. The realm of acoustic covers aspects of physics of sound, such as how many frequencies vowels have; how many milliseconds the sound of a spoken phrase is strung (connected); and how the sounds of certain syllables-syllable are connected (viewed with the help of the spectrogram); and the realm of perception is related to the perception of speech sounds that determines whether the process of understanding the spoken text takes place correctly or creates misperceptions which resulted in the failure of decoding the meaning of the message (Kent and Read, 2002).

To that end, in order to decode the meaning, one must know in detail the relationship between the components above, or more specifically, about what happened when communication is taking place, starting from how the message is delivered by the speakers and how the message is digested and understood by those who listen. Knowledge of such a detailed picture is needed by the EFL learners because if there is difficulty in understanding the process of information, the difficulty can be traced back to the part where the difficulty arises and what causes the trouble. The following is the theory of speech sound transfer process from the point of view of linguistics that is described by Kent and Read (2002, p.13).
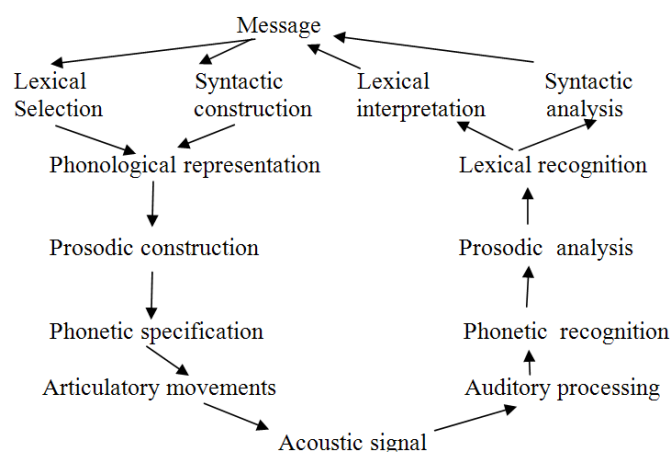
In the chart, Kent and Read (2002) described that the communication process begins with the selection of words in accordance with the messages to convey by the speaker; at the same time he must determine the appropriate sentence structure that supports the meaning of the message contained in the words he has chosen. The words selected require utterances which match the messages, associated with the sentence structure by considering aspects of prosody, and then sound specification is formed and ready to be articulated to become a message in the form of acoustic waves. Once the acoustic signal reaches the ears of those who listen, the process of this listening comprehension commences.

The process is initiated with phonetic recognition and at the same time the analysis of prosodic aspects to have the

clear picture of what words uttered by the speaker (lexical recognition). The intended meaning of the words that have been detected is sought by considering the sentence structure captured. If the words used in the speaker's sentence structure are the same as the ones captured by the listeners, it can be said that the information is accurately detected as conveyed by the speaker. In this case, there occurs accurate perception to the speech sounds conveyed the speaker. By contrast, if there is a difference of meaning or there is a difference in the message delivered by the speaker, it can be said there occurs misperceptions of speech sounds (Kent and Read, 2002).

Misperception of speech sound may occur in syllables, words, or the relationship between words in a sentence. Such misperceptions commonly occur in an oral communication using a foreign language as the language of communication. This is due to the influence of the pattern of speech sound articulation, prosodic features, foreign language syntactic constructions of the target language that are not the same as the source language (Kent and Read, 2002).

What has been outlined by Kent and Read (2002) and Hewlett and Beck (2006) indicates that phonetics and phonological competence do play a crucial role in the speech perception during the listening process. It is imperative, then, that these competences be acquired by all the EFL listeners in order that they can listen better to the English speech sound.

III. LITERATURE REVIEW

Not much has been known about the influence of phonetic and phonological competence to the process of speech perception. To see it more closely, the following is presented three research studies underlying the importance of phonetic and phonological competence to the identification of speech sound. The first part deals with the problems of speech perception experienced by native speakers of English (NS) understanding other native speakers of English (NS). The second part deals with the problems of speech perception experienced by non-native speakers of English (NNS) understanding native speakers of English (NS).

Native speakers understanding Native speakers

Speech perception problems experienced by the same L1 users, namely native speakers of English understanding other native speakers English was reported by Bond (2006). In his longitudinal study, entitled "Slips of the Ear: Errors in the perception of casual conversation", Bond (2006) asserted that

In everyday conversation, speakers employ various deductions and simplifications of their utterances, so that what they say departs in significant ways from the clarity norms found in formal speech or laboratory recordings. Both listeners and speakers are sometimes engaged in other tasks while carrying on a conversation, distracted, or preoccupied with their own ideas, so listeners vary in the amount of attention they pay to speech. Not surprisingly, sometimes listeners fail to understand what a speaker has said. Instead, a listener perceives, clearly and distinctly, something that does not correspond to the speaker's intended utterance. (Bond, p. 290).

Problems of speech perception experienced by native speakers of English were reported to occur when listeners failed to perceive unstressed vowels. These vowels are normally modified in order for the speakers to simplify the articulation process. As a result, the quality of vowels in unstressed syllables may be misperceived or may even be perceptually lost or added (Bond, 2006, p.293). Look at the example below:

(1) Grammar Workshop   ⟶   Grandma workshop
(2) Attacks in the ear   ⟶   a tax on the ear
(3) Dec writer   ⟶   decorator
(4) I teach speech science   ⟶   I teach a speech signs

The failure to perceive (1) is mainly because the listener did not pay attention to the focus of the conversation so that, instead of perceiving /ˈgræməˈwɜːkʃɒp/ <grammar workshop>, he caught /ˈgrænmɑːˈwɜːkʃɒp/ <grandma workshop>. Indeed, the phonetic distance is not far, but the lexical meaning changes as it formed another 'word sequence' which is not intended by the speaker in the running speech. Problem of perception as occurred in (2) is caused by vowel quality misperception which normally occurs in function words. The misperception of function word affects the perception of other surrounding words which then leads to global error in the sentence heard. Meanwhile, problem of perception in (3) indicates that the listener failed to detect a word boundary and altered the phonological shape of the word by reporting a spurious medial syllable; and the failure to perceive (4) shows the inability of the listener to identify the speech sound in that a syllabic nasal loses its syllabicity.

Problems of speech perception do not only occur in the failure of perceiving certain vowels having week stresses but also certain vowels having strong stresses. The following is an example of misperception in vowel having strong stress:.

(5) It's like a math problem
(6) It's like a mouth problem.

Error in which a stressed vowel is replaced by a very different vowel such as /ɪts ˈlaɪk ə ˌmæθ ˈprɒbləm/ to /ɪts ˈlaɪk ə maʊθ ˈprɒbləmʾ/ is highly unusual, in that "the misperceived vowel is not in a phonetic environment which affects vowel quality and, second, in that the phonetic distance between the target and the misperception is considerable" (Bond, 2006, p. 292). Stressed vowel should have been clear to the ear of the listener and should not create problems to the listener as what Pisoni (1981; cited in Bond 2006) stated that stressed syllable provide "an island of reliability", that is, reliable phonetic information which listeners use to interpret the stream of speech.

In a nutshell, the perception of speech sounds does not always run smoothly. Listeners often mistakenly perceive the

information conveyed by the interlocutor. Only after they hear further information to other information, they realized they have experienced a misperception on the previous information. In the concluding remark Bond (2006) summarized that there are six factors causing the failure of the perception of speech sounds, among other things:(1) the phonetic knowledge, (2) knowledge of phonology, (3) knowledge of vocabulary, (4) knowledge of structure, and (5) knowledge of semantic and pragmatic. These five factors either individually or jointly can cause failure of speech perception.

However, the samples studied by Bond (2006) were native English speakers. The difficulties experienced by native and non-native speakers will not be the same. According to Bond (2006) the misperception experienced by native English speakers is typical; namely when the subjects hear the information, they are also at the same time being preoccupied with other things that fulfill their thoughts.

Failure of perception experienced by non-native speakers of English may also be caused by the same thing as experienced by native English speakers, but open the possibility of other causes such as limited English ability, the difference of the characteristics between the source language and the target language in terms of speech patterns and the inability to identify the speech sound.

Non-native speakers understanding native speakers' utterances.

Research conducted by Goh (2002) about difficulty of understanding speech sound experienced by non-English speakers in China shows the important things that can be learned by anyone who is interested in the problems of difficulty of listening to the speech sound of English as a foreign language. Goh (2002) suggests the importance of inferential abilities as the catalyst for bottom-up and top-down process of listening to English as a foreign language. Goh (2000, 2002) reported the difficulties faced by Chinese students in understanding spoken English texts.

In a research report entitled "A cognitive perspective on language learners' listening comprehension problems", Goh (2002, p. 45) explains the difficulties faced by Chinese students at three stages of cognitive processes, which include the stages of perception, parsing, and utilization. Perception is the stage where the listeners perceive information by paying attention to the sounds they listen to (bottom-up process); parsing is the stage of input processing in the form of sound to look for meaning in short term memory. While utilization is the stage of synchronization of background information with the meaning of the sounds heard from the memory he had in long term memory networks (top-down).

Goh's research results (2000, pp 185-206) showed that difficulties of listening to the speech sound of English as a foreign language experienced by the students occurred in the three stages mentioned. In the early stage, the problem encountered is mainly phonological in nature, such as (1) the difficulty to shape or recognize the word which was already known before; (2) difficulty to recognize the following parts when recognition process is in progress; (3) difficulty to capture segments heard; (4) difficulty to catch the beginning of the text; and (5) difficulty to concentrate on running speech.

Difficulties in the mid stage or parsing includes (1) difficulty remembering what they listen to; (2) the difficulty of constructing meaning from the utterances that are heard; and (3) difficulty catching up parts because of the problems in the previous stage. Lastly, difficulties in synchronizing stage background knowledge which include (1) the difficulty of processing messages despite the familiarity of the words heard; and (2) difficulty grasping the basic idea of the message received. Thus the influential factors in the process of listening by Goh (2000, p. 201) is no longer purely limited to language skills, but leads to the ability to process the existing language in the cognitive domain.

What is this research about?

What has been outlined Bond (2006) and by Goh (2000, 2002) indicates that speech perception problem is experienced by both native speakers and non-native speakers of English, with somewhat different causes. Native speakers' failure to understand other native speakers is mainly because of their mind being, at the same time, preoccupied with other things that fulfill their thoughts; while the speech perception problem experienced by the non-native speakers of English is to a certain degree caused by the same problem as experienced by native speakers of English; but mostly, it is related to the mastery level of the language learned and the influence of their mother tongue. Interestingly, however, both Bond (2006) and Goh (2000, 2002) mention phonetics and phonology as the key issue of the speech perception problems.

In response to the Bond (2006) and Goh's (2002) findings, it is crucial to investigate further what aspect of phonetics and phonology causing problems of speech perception in the listening process. This research was intended to do so, namely examining problems of listening to English speech sound seen from micro level perspective, by researching the ability of the subjects to detect the element of intonation phrase, among other things proclitics, enclitics, interval clitics and tonics. This research at the same time was also meant to answer the speculation that problems of listening experienced by native non-native speakers of English is merely a matter of vocabulary problems. The speculation is based on the common believes that (1) vocabulary is the heart of language (Lewis, 1993, p. 89); (2) vocabulary is a prerequisite of success in the process of listening to the speech of foreign language (Segalowitz and Segalowitz, 1993); (3) vocabulary is fundamental in the process of foreign language information (Hayashi, 199, p. 151); (4) vocabulary is necessary for a verbal communication because without vocabulary nothing can be conveyed (Wilkins, 1972, pp. 111-112).

IV. METHODS

The present research used 3 research instruments including: (1) TOEFL, (2) Listening test in two versions, and (3)

Speech Perception Test. The first instrument, namely TOEFL was used to discriminate language proficiency levels of the subjects (n=30) so that all the selected subjects could represent the expected range of score from the lowest (31) up to the highest (68). The second instrument, listening test with two versions was meant to find out the different result between listening to a narration with its transcript on it and to a narration without any transcript on the text. This test was intended to find out whether the problem of listening was barely due to the lack of vocabulary or was closely related to the insufficient knowledge of phonology among the subjects. The third instrument, namely speech perception, was adopted to investigate the role of speech perception in the process of listening comprehension. For this purpose, the subjects were assigned to listen to eight sentences spoken by a native speaker of English to test subjects' ability to detect the 'tonicity' and 'tonality' represented by six elements of clitics and tonics and two mixed elements of clitics and tonics, forming connected speech. While listening, they were required to jot down every string of word they heard. All the sentences they wrote were analyzed to find out some possible slip of the ears on the 'onset', 'middle tonic', and 'tonic' (tonic elements), and 'proclitic', 'interval clitic', and 'enclitic' (clitic elements). The misperceptions occurred reflect problems faced by the subjects in comprehending the intonation phrase they listened to. To make it easier to locate the misperceived part, a model of speech perception proposed by Cauldwell (2002) accompanied with spectogram was employed. After misperception was identified a comparison of phonetic/phonological explanation of both languages was made to explain the cause of misperception, and a summary was drawn to justify the role of speech perception ability in listening comprehension.

V. RESULTS AND DISCUSSION

The following is the result of the listening test from the second instrument, namely listening test in two versions: version A and version B with a total number of 30 subjects participating in the research. The subjects being studied represent all students having the range of score of 31-68 on the TOEFL Listening section, as grouped by the TOEFL test, the first instrument of the research.

Paired Samples Statistics

TABLE 2
PAIRED T TEST OF THE LISTENING TESTS

	Score of *Listening*	Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Version B	55,7333	30	7,66961	1,40027
	Version A	49,9667	30	9,12550	1,66608

The data in the above table shows that the average score of listening without transcripts of 30 respondents was lower, ie 49.9667 than listening scores accompanied with transcripts, which reached 55.7333, or there is a difference of 5.7666. The significant difference between the two groups is illustrated in the following table

TABLE 3
THE SIGNIFICANCE OF THE PAIRED T TEST

		t	df	Sig. (2-tailed)
Pair 1	Version A - Version B	9,766	29	,000

The data shown in table 3 above indicates that the difference of the scores achieved in the test accompanied by transcripts and the one without transcripts   was significant. It is shown by the Sig. (2 - tailed) = 0.000 or <0.05.

With reference to the evidence of test data that showed a negative trend in the test version A in comparison with the results of the listening test version B, it can be concluded that the difficulty faced by the subject of research is not solely caused by the limited vocabulary of the subject but is related to phonological abilities of the subjects as well.

Another result, which was based on the data obtained from the third instrument (n = 30) showed the occurrence of misperceptions on the onset by 40%, middle tonic 36.66%, 76.66% tonic, proclitic 80%, interval clitic 73.33% and enclitic 76.66%. These results prove that 'tonicity' or the stress pattern in the English language is the cause of the difficulties in the process of understanding English speech sound. Other evidence in this study is that the 'tonality', which is a combination between the elements that make up the connected speech in a speech stream is also a difficulty in the process of listening to the English speech sound. Data show that in tonality made up of elements of onset-interval-middle tonic, the misperception reached 96.66% while tonality made up of middle tonic-interval-tonic reached 90%. This figure reflects the difficulties faced by the subject in listening to connected speech.

The results also show that the tonicity and tonality spoken with the speed of 160 words per minute have triggered the speech compression, speech sound changes as a result of assimilation and elision of certain consonants or vowels. The changes of the sound quality are difficult to predict because the production of words, rhythm and stress, and all things related to the phonological aspects are of speaker controlled. The listeners can only receive and understand the speech that he listens to without having the control of it.

Problems of perceiving English speech sound, thus, is not solely caused by the limited vocabulary of a person but are also related to the ability to perceive speech sound phonologically. The results of the listening test version A and B shows that the subject of research consistently shows higher results on test A rather than test B. The implication of these findings is that one who does not have the ability to perceive speech phonologically will have difficulty understanding

English speech sound.

The ability to perceive speech sound becomes an important key in the process of understanding English speech sound because this ability is closely related to the decoding process of acoustic signal in the event of an oral communication. The process is complicated because it involves elements of sound, changes in the quality of certain vowels and consonants as a result of the assimilation of speech caused by weak and strong stress, speech rhythm and tempo of speech compression. When the decoding process takes place smoothly then speech perception was running correctly. Conversely when the process of decoding speech sound doesn't run as it should, misperceptions occurred.

From the statistical data showing the speech perception problems, there is evidence that the problem of inaccurate perception of speech sounds is not only found in the elements having weak stress (clitic) such as proclitic, interval, enclitic but also the elements having strong stress (tonic) such as onset, middle-tonic and tonic. The data show a correlation between the number of misperception on the elements of clitic and the adjacent tonic

That is, when there is a misperception on a clitic, the chance of the occurrence of misperceptions on the nearest tonic is high. This event occurs because the dynamics of clitic and tonic is volatile following the tonicity and tonality on the sentence heard.

The significance of these findings is that the process of assimilation of sounds on the intonation phrases that are connected to one another potentially trigger a misperception. Nevertheless, the event of this sound misperceptions remains difficult to predict as how the speech sound was articulated is entirely under the authority of the speaker. Theoretically, content words get stronger stress than the function word; syllable stress tends to fall in a regular tempo; and the vowels are articulated as they are depicted in the theory of phonetics and phonology. But the results of this study prove that the speakers are not always bound by such rules or conventions. It is not uncommon that the speaker put a stress on the words which are, in theory, not commonly given; speaker put stress on the syllables in irregular rhythms; and speakers pronounce vowel sounds and consonant without following the pattern as described by the experts of phonetics and phonology. These events reinforce the view of Brazil (1996) that the speech sound in the event of mutual communication tends to be speaker-controlled, purpose-driven, interactive, cooperative, context-referenced, and context-changing. As a result, the difficulty of perceiving speech sounds accurately occur in all elements, namely elements having weak stress and elements having strong stress.

The quality of the speech sound which tends to be unclear as affected by the existence of prominent and non-prominent stressed elements as seen in this study is caused by the rapid change of sound on each element. Changes in the quality of the sound and shifting stress on the elements of the intonation phrase has caused 'deafness' or speech illusion, namely the events up unreal irregularities in the speech sound in an inter language communication as a result of changes in the quality of the sound and shifting location of syllable stress.

This event is a manifestation of the inability of the research subjects to identify certain sounds that came out in an unbroken series or connected speech affected by speech compression and the articulation process of the speech sounds such as assimilation, elision, liaison, soothing, deletion, and lagging. This proves that the ability to perceive speech sound is an important part of the linguistic mechanism in the process of understanding the spoken text in English, and proves that differences in language characteristic that include differences in the characteristics of Indonesian as phonetic language and English as non-phonetic language, stress patterns, syllable and word stress, and the differences in tonicity and tonality effect   the ability of the Indonesians to perceive English speech sounds.

VI. CONCLUSION

The evidence reviewed here leads to the conclusions the perception of speech sounds does not always run smoothly. Listeners often mistakenly perceive the information conveyed by the interlocutor. In the case of English native speaker listening to other English native speakers (NS – NS), the misperception occur particularly when the listeners' mind are engaged in other tasks while carrying on a conversation, distracted, or preoccupied with their own ideas, so listeners vary in the amount of attention they pay to speech. Not surprisingly, sometimes listeners fail to understand what a speaker has said. Instead, a listener perceives, clearly and distinctly, something that does not correspond to the speaker's intended utterance.

In the case of non-native speakers of English listening to native speakers of English (NNS – NS) Problems of speech perception faced by the EFL learners to a certain extent are caused by the native speaker factors. Yet, it is naïve to blame native speakers as the factors causing misperception in the part of the EFL learners. Factors other than native speakers have been sought and evidence has brought to the conclusion that phonetic and phonological competence does play a significant role in the problems of speech perception experienced by EFL learners.

In response to the speculation about the vocabulary as the sole factor of the emergence of speech perception problems this research refuted the speculation. Knowledge of Vocabulary proved to be important, but when knowledge of phonetics and phonology does not in support to it, the knowledge of vocabulary alone is simply inadequate to facilitate comprehension. These findings confirm the difficulty of speech perception as the problem related to the phonetic and phonological knowledge besides vocabulary knowledge of the listener.

REFERENCES

[1] Bejar, I. et al. (2000). TOEFL 2000 "Listening Framework: A Working Paper". Monograph Series. *Educational Testing Service*, New Jersey.
[2] Bond, Z.S. (2006). "Slips of the Ear". In Pisoni and Remez (Eds). *The Handbook of Speech Perception* (290 – 310) Oxford: Blackwell Publishing.
[3] Brown, G. (1990). Listening to Spoken English. New York: Longman.
[4] Gardner, H. (1983). Frames of Mind: The Theory of Multiple Intelligences. New York: Basic Book.
[5] Goh, C.C.M., (2000). "A Cognitive Perspective on Language Learners' Listening Comprehension Problems". *System*, 30, 185 – 206.
[6] Goh, C.C.M. (2002). "Learners' Self Reports on Comprehension and Learning Strategies for Listening". *Asian Journal of English Language Teaching*, 12, 45 – 68.
[7] Hayashi, T. (1991). "Interactive Processing Words in Connected Speech in L1 and L2". *International Review of Applied Linguistics,* 29(2), 151–169.
[8] Hwelett, N. dan Beck, J. (2006). An Introduction to the Science of Phonetics. New Jersey: Lawrence Erlbaum Associates.
[9] Ingram, J.C.L. (2007). Neurolinguistics: An Introduction to Spoken Language Processing and Its Disorders. New York: Cambridge University Press.
[10] Jacquemot, C. & Scott, S. K. (2006). What is the Relationship Between Phonological Short-term Memory and Speech Processing? *Trends Cogn. Sci.* 10, 480-486.
[11] Kent R and Read C. (2002). Acoustic analysis of Speech, second edition, San Diego: CA Singular Publishing Group.
[12] Lewis, M. (1993). The Lexical Approach. Hove, UK: Language Teaching Publications.
[13] Ohata, K. (2006). "Auditory Short-Term Memory in L2 Listening Comprehension Processes", *Journal of Language and Learning.* 5 (1), 21–27.
[14] Pierce, W. (2003). Metacognition: Study Strategies, Monitoring and Motivation. Retrieved from https://www.academic.pgcc.edu/wpierce/MCCCTR/index.html (accessed 13/6/2014).
[15] Sebastián-Galés, N. (2006). "Cross Language Speech Perception". In Pisoni and Remez (Eds). *The Handbook of Speech Perception.* Oxford: Blackwell Publishing.
[16] Segalowitz, N and Segalowitz, S. (1993). "Skilled Performance, Practice, and the Differentiation of Speed-up from Automatization Effects: Evidence from Second Language Word Recognition". *Applied Linguistics*. 14, 369–381.
[17] Sprinthall, R.C. and Sprinthall, N.A. (1994). Educational Psychology: A Developmental Approach. New York: McGraw-Hill, Inc.
[18] Sutrisno, A., (2015). Kesulitan Mempersepsi Bunyi Ujaran Bahasa Inggris sebagai Bahasa Asing di Indonesia dan Penyebab-penyebabnya [Problems of Perceiving the Speech Sound of English as a Foreign Language in Indonesia and Their causes]. Unpublished Disertation, Program Pascasarjana Fakultas Ilmu Budaya, Universitas Gadjah Mada, Yogyakarta, Indonesia.
[19] Teasdale, J, D and Barnard P.S. (1993). Affect, Cognition and Change: Re-modeling Depressive Thought. Hove, UK: Lawrence Erlbaum Association Ltd.
[20] Vandegrift, L. (2007). "Recent Developments in Second and Foreign Language Listening Comprehension Research". *Language Teaching,* 40, 191–210.
[21] Wilkins, D.A. (1972). Linguistics in Language Teaching. London: Arnold.

Adi Sutrisno is a lecturer at Universitas Gajah Mada of Indonesia. His research interests include speech perception, language testing, and language teaching. Email: adisutrisno@ugm.ac.id.