American Journal of Social and Humanitarian Research

Vol. 6 Issue 8 **| pp.** 1972-1978 **| ISSN:** 2690-9626
Available online @ https://globalresearchnetwork.us/index.php/ajshr

Article

The Role of Speech Signal Variability in Affecting Speech Perception Received by Iraqi EFL Learners

Hayder Sadeq Naser

1. Instructor, Mustansiriyah University, College of Arts, Translation Department
* Correspondence: hayderenglish@uomustansiriyah.edu.iq

Abstract: Speech perception plays a pivotal role in language comprehension, involving the translation of acoustic signals into meaningful linguistic units. However, speech signals exhibit inherent variability, complicating this process, especially for learners of English as a Foreign Language (EFL). Variability in speech signals arises from numerous factors such as speaker anatomy, speaking rate, environmental noise, and contextual influences. These variables significantly impact the clarity of speech perception, particularly in non-native learners. Although much has been studied regarding the variability of speech signals, its specific effect on EFL learners, especially in noisy or fast-paced speech contexts, remains under-explored. This gap calls for research to investigate how these factors influence speech perception. The study aims to evaluate how variations in speech signals particularly speed and ambient noise affect the perception of speech among Iraqi EFL learners. Findings suggest that while students could perceive part of the speech signal, none fully transcribed the original utterance correctly. The factors of noise and speaking rate significantly hindered comprehension, with male students generally outperforming female students. The research highlights the specific challenges faced by EFL learners in conditions of high variability, offering insights into the perceptual limits of speech processing in a non-native context. This study emphasizes the importance of exposing EFL learners to diverse, real-world listening environments to improve their ability to decode speech signals under variable conditions.

Keywords: speech signal, variability, speech perception, Iraqi EFL learners.

Citation: Naseer, H. S. The Role of Speech Signal Variability in Affecting Speech Perception Received by Iraqi EFL Learners. American Journal of Social and Humanitarian Research 2025, 6(8), 1972-1978

Received: 10th May 2025
Revised: 16th Jun 2025
Accepted: 24th Jul 2025
Published: 10th Aug 2025

1. Introduction

Psycholinguistics, as a field of study, is concerned with the discovery of psychological processes by which humans acquire and use language. Generally speaking, psycholinguistics addresses three major concerns: ***comprehension*** which means how people understand spoken and written language which is a broad area of investigation including how speech signals are interpreted (speech perception), how the meanings of words are determined (lexical access), how the grammatical structure of sentences is analysed to obtain larger units of meaning (sentence processing), and how longer conversations are appropriately formulated and evaluated discourse). The second component of psycholinguistics is ***speech production*** which is concerned with how people produce language, and the third component is ***acquisition*** which is defined as how people learn language [1].

According to Clark and Clark communication with language is carried out via tow primary human activities which are speaking and listening. These two activities are of particular concern to psycholinguistics since they are mental activities that hold clues concerning the nature of human mind. Thus, in speaking ideas are put into words to talk about perceptions, feelings, and intentions to other people to grasp them. In listening, the

words, are turned into ideas in order to reconstruct the perceptions, feelings, and intentions they are meant to grasp. In such a way, speaking and listening reveal what is going on inside the mind and how the mind deals with perceptions, feelings and intentions. So, speaking and listening represent the tools that people use in more global activities to convey facts ask favor, and make promises, and the role of those who listen is to receive this information. In this way, speaking and listening are very complex activities that contain complex processes to make speech signal understood and comprehended clearly and accurately [2].

Speech perception

Speech signal is a physical entity that contains vertical boundaries between consonant and vowel segments that are fairly apparent. These show the directness that defines the basis of auditory processing rather than being apparent in terms of visual processing. Another thing is that speech perception processing is sequential over time, in which these apparent segments vary widely in their duration, which makes them candidates for the initial stages. Speech perception is done via a shift from the physical-acoustic to the more perceptual auditory domain which is accomplished in the transition from the acoustic buffer to the auditory memory. Thus, phonemes are the immediate output of the signal processing and are identified based on sub-phonemic cues in the signal [3].

According to Fernandez and Cairns, the hearer's task is like a mirror of the speaker's task. The hearer's task is to use information from the acoustic signal to reconstruct the information into a phonological representation. To do so, the hearer moves to the phase of entering into the lexicon by using that phonological representation to retrieve the lexical item that matches the phonological representation. By having this done, the hearer is permitted to recover the semantic and structural details of the words in the message. This step enables the hearer to move to the next step to reconstruct the structural organization of the words to create a syntactic representation required for recovering the meaning of the sentence. Thus, speech perception is concerned with setting a final syntactic representation necessary for retrieving the meaning of the sentence [4].

Speech perception is a fundamental component of language comprehension and plays a pivotal role in both native and non-native language processing. It involves decoding a continuous stream of acoustic information into meaningful linguistic units, a task made more complex by the inherent variability of speech [5]. Variability arises from numerous factors, including speaker-specific anatomical differences, variations in speech rate, emotional intonation, and environmental noise [6]. These elements significantly impact the clarity of speech signals, especially for learners of English as a Foreign Language (EFL), who may not yet have robust phonological representations in the target language [7].

Features of the speech signal

The hearer's task is to identify the original phonemic elements since the hearer plays the role of inspector. Speech signal is characterized by having three features that speech perception must deal with. The first feature is that the speech signal is ***continuous***: no spaces are available between the vowels and the consonants, and even between words. Thus, speech perception mechanisms work on segmenting this continuous signal into discrete units like phonemes, syllables, and ultimately words. The second feature of speech signal is the ***parallel transmission*** of information about phonetic segments. Liberman, cited in Fernandez and Cairns illustrate how information about phonological units in the word ***bag*** is distributed across the word [8]. It is shown from the spectrogram waveform that the vowel has a duration of approximately 250 milliseconds, of which approximately 50 to 75 milliseconds carry information about all three phonological units. In such a way, properties of the consonant /b/ have its effect on vowel and persist through, see Figure 1.

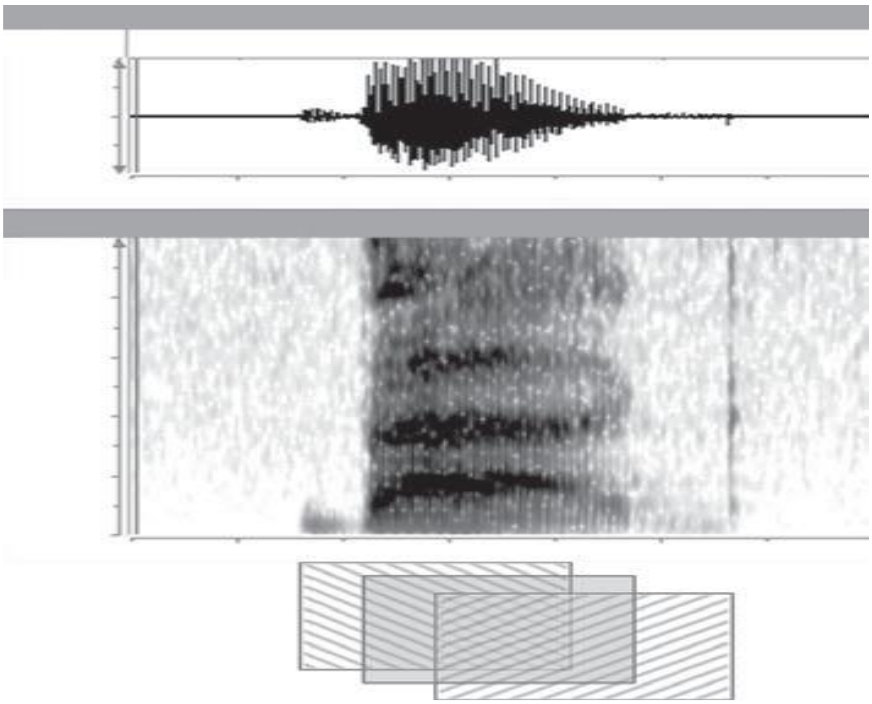

Figure 1. Factors Influencing Speech Signal Variability

The beginning of the word-final /g/. Another thing is that the properties of the final consonant /g/ began with the change of /b/ and continued through the second half of the vowel. So, the vowel /æ/ has its influence on the pronunciation of the entire word since it carries the acoustic information about both of the consonants in the word. The third feature of the speech signal is its ***variability***. Variability means that a speech signal may vary greatly each time it is produced. This variability is due to many factors, like:

1. Variability among speakers themselves because of their human anatomy.
2. Variability within speakers in terms of speaking fast, speaking low, shouting, speaking with a feeling of sadness or joy.
3. Variability caused due to ambient noise.
4. Variability caused according to the context (the articulation of a phoneme is affected by the phonemes around it.

Thus, all these variables in the speech signal affect the reception of the speech signal depending on the hearer.

2. Materials and Methods

The main purpose behind selecting variability as a corpus for this study is to check the effectiveness of these variables on the reception of a speech signal by Iraqi EFL learners. Thus, the researcher tries to examine how the existence of any of the above-mentioned variability affects the speech signal perception via listening to an American song and asking Iraqi EFL learners to write what they hear. The researcher chose ten Iraqi EFL learners (five men and five women), all of whom are third-year college students in the department of English language. Each student has a time to listen to the song, recognize the words, and then write what they hear on a sheet of paper to compare what they heard with what they wrote and see what the difference is between what they listened to and what they wrote. The song contains the utterance "***let me see you step,***" which is uttered fast with noise.

Data analysis

This section is the considered as the practical part of the research. It includes the analysis of the above-mentioned clause in which the students heard. Thus, the researcher analyzes the outcome of the ten students according to what they wrote after listening to the song to check the effect of the variables particularly ***ambient noise*** and ***speaking fast*** on the acoustic signal of speech perception.

Male student 1

After listening to the script, this student wrote "***let me see you staire***". According to this student, every word in his clause is compatible with the word in the script except the last one. The student missed only the second part of the word "***step***" since his last word is "***staire***" which means "***stare***" that contains the same two sounds of the word "***step***" and even the vowel in the song is almost the same as the same vowel in "***step***". In such a way, it can be said that the noise, as well as uttering the clause fast, affects the perceiving part of a word rather than the whole clause.

This student accurately perceived the initial four words but misinterpreted the final word "step" as "***staire***". Phonetically, both share an initial /st/ cluster and similar vowel qualities, indicating that the student recognized part of the word but failed to capture the final plosive /p/. This may be attributed to the effect of fast articulation and masking from background noise, which tends to obscure stop consonants. His grammatical structure remained intact, demonstrating awareness of syntactic expectations.

Female student 1

This female student wrote "***let see start***" as a result of what she listened to. In this case, it can be seen that the noise and speaking fast have a greater effect than the first case since this female student missed "***me***", "***you***", and also the second part of the original word "***step***" in the song. Thus, it becomes clear that the grammatical level of the first male student is better than the first female since he was able to recognize that "***let***" must be followed by an object pronoun comparing to the female student who was not able to recognize this by writing immediately the verb "***see***" after "***let***".

This response reveals multiple errors. The student omitted "me" and "you" and replaced "step" with "start." The misperception may stem from top-down processing, where lexical familiarity influences perception. "start" is a high-frequency word. The missing object pronoun "me" suggests syntactic processing difficulties under degraded input, possibly due to higher cognitive load.

Male student 2

After listening to the script, this student wrote "***la me see the sparie***". This indicates that the effective factors play their roles again to guide this student to write "**la**" instead of "***let***" and "***sparie***" instead of "***step***". Therefore, there are two mistakes committed by this student distributed in the initial position of the utterance represented by the word "***la***" and the last position in the clause represented by the word ***"sparie".*** It becomes clear that this student was not able to recognize the word "***let***" since the noise makes the acoustic signal unclear enough to enable the student to perceive what should have uttered in the song because the singer uttered it fast with noise.

The substitution of "la" for "let" reflects a significant deviation, likely due to rapid speech onset and reduced clarity of the word-initial segment. The final word "***sparie***" shares some phonetic features with "step" but introduces unfamiliar syllables, possibly due to mishearing and an attempt to phonologically 'fill in' the ambiguous signal. This participant struggled with both onset and coda recognition [9].

Female student 2

This female student wrote "***let me see is gear***". As it is noticed here, this student missed two words to write from the original utterance which are "***you***" and "***step***" meaning that the noise and speaking fast affect perceiving the last word correctly while the acoustic signal of the first three word is perceived correctly without being affected by the ambient noise as well as speaking fast. Another thing is that the student's clause includes "***verb to be,***" which is completely incompatible with the original utterance since it does not contain the verb to be.

Thus, it seems that the two factors have their strong effect on this student since she added something out of the utterance components. Thus, the presence of the verb "is," which does not exist in the target sentence, demonstrates over-reliance on grammatical prediction. This could stem from the learner's strategy to form syntactically coherent output when faced with unclear auditory input.

Male student 3

After listening to the song, this student wrote "***let me see the stray***". This shows that this student failed to recognize the last two words in the original utterance and wrote "***the***" and "***stray***" instead of "***you***" and "***step***". The last word written by this student is similar to the original one in the first part, which is "***st***", while the second part, "***ep***", is missed. Again, it can be seen that the presence of the effective variables distracts the student's attention to perceive the acoustic signal clearly due to the loudness as well as the speed of speech, which causes distraction to the listener.

Therefore, this student-maintained a near-complete structure but replaced "you" with "the" and "step" with "stray." Both substitutions suggest semantic and phonological proximity. "Stray" shares the initial /st/ with "step," indicating partial recognition. The insertion of "the" could reflect a filler strategy to complete the syntactic frame [10].

Female student 3

This student wrote, "***let me see the step***". This student missed only one word from the original utterance, which is "**you,**" and wrote "***the***" instead, although they are not compatible in pronunciation. It seems that the student has no problem perceiving the acoustic signal clearly, except for one word that occurred before the last. Thus, the ambient noise and speaking fast have their little effect on perceiving the utterance compared with other cases that miss more than one word.

In such a way, this response is mostly accurate, with the only error being the substitution of "you" with "the." The substitution may have resulted from misidentifying the pronoun under noisy conditions. The accurate perception of "step" is noteworthy, showing resilience to variability in the speech signal.

Male student 4

This student wrote, "***let me see dear***". It seems that the student has no problem with organizing the first three words of the original utterance, but he completely missed the fourth word since he did not write anything instead, while his recognition of the fifth word was "***dear***" instead of the original one, "***step***". This outcome proves the effect of the noise as well as speaking fast on perceiving the speech signal clearly. Thus, the loss of two words from the original utterance assures that the student was distracted by having an unclear acoustic signal due to the previous factors [11].

This student successfully identified the initial phrase but omitted "you" and misheard "step" as "dear." The replacement word "dear" is semantically and phonologically unrelated, suggesting that ambient noise masked key phonetic cues, prompting a substitution based on top-down inference or lexical familiarity.

Female student 4

This female student wrote "***let me see staw***". Again, this female student missed the fourth word "***you***" to be recognized from the original utterance and failed to recognize part of the fifth word from the original utterance, which is "***step***" by writing "***staw***". So, this female student tried to recognize the word "***step***" but she did not do so because of the ambient noise, as well as speaking fast, which affected her acoustic signal from being received. Thus, the transcription lacks the word "you" and misrepresents "step" as "staw." The response indicates an effort to reconstruct the acoustic signal phonetically, where /st/ was correctly perceived but followed by an incorrect vowel-consonant sequence. This suggests that partial decoding is interfered with by the effects of coarticulation and noise [12].

Male student 5

This male student wrote, "***let me see the step***". Again, it has been noticed that a complete loss of the whole fourth word, which is represented by the word "***you***" in the original utterance, while the rest part of the utterance was written correctly by the student as it was heard. In this case, the effective factors affected only one word to be perceived clearly, although the student perceived four words correctly.

This participant made only one substitution, replacing "you" with "the." Given the correct recognition of the rest of the sentence, this case illustrates how a single misperceived function word can occur even when overall comprehension remains intact. It may reflect reduced attention to function words under rapid speech

Female student 5

This female student wrote "***let see spear***". This female student had a complete loss for the original words "***me***" and "***you***" with a failure to recognize the last word of the original utterance, which is "***step***". In this case, again, a complete loss occurred as well as perceiving another word, which is "***spear***" instead of the original one, which is "***step***". Thus, ambient noise and speaking fast altered the acoustic speech signal, causing it to be perceived differently by the listener. This is done due to the way the acoustic signal is transferred. This response shows significant loss, omitting both "me" and "you," and replacing "step" with "spear." The substitution is phonetically plausible, as both share the /sp/ onset and final /r/-like rhotic quality in some accents. This suggests that acoustic similarity influenced the error, but the omission of function words points to attentional limitations in processing under noise [13].

3. Results and Discussion

Results of the study

The outcomes of the listening task reveal key insights into the challenges Iraqi EFL learners face when perceiving speech under variable conditions, such as fast rate and background noise. None of the ten participants transcribed the five-word phrase accurately in full, highlighting the consistent effect of acoustic variability on speech perception. The following outcomes are recognized:

1. Four students (three male and one female) were able to perceive four correct words exactly as the original utterance, and they failed to perceive only one word.
2. Four students (two males and two females) were able to perceive only three correct words exactly as the original utterance [14].
3. Two students (female students) were able to perceive only two correct words exactly as the original utterance, see Table 1.

Table 1. The results of the study

Number of the correct words	Number of students	Male	Female
Five	none	none	none
Four	four	three	one
Three	four	two	two
two	two	none	two

These outcomes indicate that male students generally outperformed female students, particularly in achieving four correct words. Two female students had the lowest performance, correctly identifying only two words.

Error Patterns

Most frequently misperceived word: "step" replaced with phonetically similar alternatives such as "staire, sparie, stray, spear, staw".

Function word omission: Words like "me" and "you" were often dropped, especially under pressure from rapid speech.

Substitution with grammatical structures: In several instances, students substituted missing content with functionally or syntactically fitting words.

Interpretations

The results affirm the hypothesis that speech variability, in the form of speed and background noise, poses a significant barrier to accurate auditory decoding among EFL learners. Most students were able to capture part of the sentence, suggesting some robustness in perception, yet full comprehension proved elusive due to missing phonological cues and weakened signal clarity. Several responses suggest top-down compensatory strategies where students attempted to reconstruct grammatically coherent phrases from incomplete or ambiguous input. Such strategies indicate the activation of internal linguistic expectations during real-time processing but also highlight gaps in

phonological precision under degraded listening conditions. Overall, the findings emphasize the need to expose learners to diverse and challenging listening environments in pedagogical settings to build resilience against natural speech variability [15].

4. Conclusion

In light of the study, the following conclusions are arrived:

1. The primary conclusion is that no learner could perfectly perceive the entire utterance, underscoring the significant impact of both ambient noise and speech rate on listening comprehension. These variables disrupted the acoustic clarity of key phonological elements, particularly plosive and function words, which are already perceptually less salient. Learners frequently omitted or substituted words based on incomplete auditory input, and many relied on top-down processing to reconstruct grammatically plausible phrases, sometimes successfully, sometimes not.
2. Additionally, the study found that male participants performed marginally better than female participants, with a greater number achieving four out of five correct words. This suggests potential individual differences in auditory processing or exposure to English input, though further research would be necessary to confirm any systematic gender-based patterns.

REFERENCES

[1] H. Clark and V. E. Clark, *Psychology and Language*. New York: Harcourt Brace Jovanovich, Inc., 1977.

[2] J. Field, "Promoting perception: Lexical segmentation in L2 listening," *ELT Journal*, vol. 57, no. 4, pp. 325–334, 2003.

[3] M. E. Fernandez and S. H. Cairns, *Fundamentals of Psycholinguistics*. Sussex: Blackwell, 2011.

[4] M. Garman, *Psycholinguistics*. Cambridge: Cambridge University Press, 1990.

[5] B. J. Gleason and B. N. Ratner, *Psycholinguistics*. Harcourt Publishers, 1998.

[6] K. Johnson, "Speaker normalization in speech perception," in *The Handbook of Speech Perception*, D. B. Pisoni and R. E. Remez, Eds. Blackwell, 2005, pp. 363–389.

[7] S. L. Mattys, M. H. Davis, A. R. Bradlow, and S. K. Scott, "Speech recognition in adverse conditions: A review," *Language and Cognitive Processes*, vol. 27, no. 7–8, pp. 953–978, 2012.

[8] M. R. Younus, "Speech intelligibility: a study of Iraqi EFL learners' accented English," Ph.D. dissertation, London Metropolitan University, 2020.

[9] M. N. S. Yousif, "The Correlation between Vowel Production and Vowel Perception as Elicited in the Performance of Adult Iraqi Learners of English: Acoustic Perspective."

[10] T. N. Alharbi, "Improving Intelligibility and Comprehensibility of Segmental and Suprasegmental Speech Patterns among Saudi Beginning-Level EFL Learners," Ph.D. dissertation, The Australian National University, Australia, 2021.

[11] L. H. Ibraheem, F. J. Mohammed, and M. M. Anber, "Dialectal differences in the perceptive intelligibility of Iraqi EFL learners," *An-Najah University Journal for Research-B (Humanities)*, vol. 39, no. 9, pp. None-None, 2025.

[12] W. Alsiraih, "Voice quality features in the production of pharyngeal consonants by Iraqi Arabic speakers," Ph.D. dissertation, Newcastle University, 2013.

[13] Z. M. Hassan, "An experimental study of vowel duration in Iraqi spoken Arabic," Ph.D. dissertation, University of Leeds, 1981.

[14] A. A. W. Al Abdely and Y. N. Thai, "Learning English vowels by Iraqi EFL learners: Perceived difficulty versus actual performance," *3L: Southeast Asian Journal of English Language Studies*, vol. 22, no. 1, 2016.

[15] L. F. S. Sidgi and A. J. Shaari, "The effect of automatic speech recognition Eyespeak software on Iraqi students' English pronunciation: A pilot study," *Advances in Language and Literary Studies*, vol. 8, no. 2, pp. 48-54, 2017.