



Perception of an Existing and Non-existing L2 English Phoneme behind Noise by Japanese Native Speakers

Mako Ishida^{1,2}, Takayuki Arai¹

¹ Sophia University, Japan,

² Research Fellow of the Japan Society for the Promotion of Science

mako.ishida@sophia.ac.jp, arai@sophia.ac.jp

Abstract

This study investigates how similarly a person hears an existing and non-existing speech sound behind noise in L2, as compared to L1 reported in Mattys, Barkan, and Samuel (2014). Participants were Japanese native speakers who spoke English as a second language. They listened to English words and non-words in which a phoneme was covered by noise (added) or replaced by noise (replaced). The target phoneme was either a liquid or a nasal. In experiment, participants listened to a pair of a word with noise (added or replaced) and a word without noise (normal) in a row, and evaluated the similarity of the two by using an 8-point scale. The results suggested that L2 listeners perceived the added and replaced sound significantly differently. L2 listeners found the added sound (a phoneme + noise) more similar to a normal sound than the replaced sound (noise only), as was also reported in L1 listeners. At the same time, they also perceived the illusory sound of a missing phoneme in the replaced condition. A missing nasal was significantly more restored than a missing liquid. There was no lexical effect in perceptual restoration of phonemes among L2 listeners, although it was reported among L1 listeners.

Index Terms: phonemic restoration, perceptual restoration, speech perception, second language listening

1. Introduction

Phonemic restoration is an auditory illusion in which listeners hear a physically non-existing speech as if the original speech is intact [1, 2]. When a phoneme is deleted from acoustic waves and replaced by noise, listeners often hear an illusory sound of the missing phoneme [3, 4]. It is reported that phonemic restoration is observed when the replacing and replaced sounds have similar acoustic characteristics [5, 6]. For example, a fricative is well-restored when replaced by white noise, while a vowel is well-restored when replaced by pure tone. It is also reported that a liquid and a nasal are in the medium range of restorability when replaced by white noise. It is considered that listeners perceptually restore a gap in speech when masking sound has similar acoustic characteristics as original speech.

The restoration of missing phonemes also relates to lexical knowledge. Listeners restore a missing phoneme in a word better than that in a non-word [5, 6, 7]. For example, a deleted phoneme /r/ in a word “accelerate” is restored better than that in a non-word “vabbellerate”. It seems that listeners perceive speech by predicting a particular phonemic sequence or a

possible lexical item. Perceptual restoration seems to take place under certain acoustic and lexical conditions.

On the other hand, listeners also perceptually differentiate an existing and non-existing speech sound behind noise [5, 6, 7], despite the fact that a non-existing speech sound is often perceptually restored. When listeners listen to a word in which a phoneme is covered by noise (phoneme + noise; added) or replaced by noise (noise only; replaced), listeners perceive these sounds differently. When listeners listen to a word with noise (added or replaced) and a word without noise (normal) in a row, and evaluate the similarity of the two by using an 8-point scale (8: very similar, 1: not similar), they respond that the added sound is more similar to a normal sound than the replaced sound [6, 7]. It seems that the presence of acoustic waves behind noise is perceptually differentiated.

The earlier studies suggested that a missing part of speech can be restored under certain conditions. At the same time, a missing phoneme is perceived differently from a physically existing phoneme. These studies so far were mostly conducted in L1, and we do not know what would happen in L2. The present study investigates the perception of an existing and non-existing L2 English phoneme behind noise by Japanese native speakers, in search of perceptual accuracy and phonemic restoration in L2.

2. Experiment

Japanese native speakers listened to English stimuli, in which a phoneme is covered by noise or replaced by noise. This study examined the perceptual sensitivity to an existing and non-existing English phoneme behind noise, and the perceptual restoration of a non-existing L2 English phoneme behind noise in words and non-words. This experiment examined the perceptual sensitivity to acoustic details, and explored the affective factors for the perceptual restoration of L2 phonemes: 1. Noise (added vs. replaced), 2. Phonemes (liquid vs. nasal), and 3. Lexical factors (word vs. non-word).

2.1. Materials

There were 90 pairs of a word and a non-word, adopted from Mattys, Barkan, and Samuel (2014) [7]. All pairs had the same stress patterns, and each item consisted of four to five syllables. The first 60 pairs were test stimuli, and the last 30 pairs were fillers. The test stimuli pairs had the same last two syllables, with the target phoneme of a liquid (/l/ or /r/) or a nasal (/m/ or /n/). It was located in the first phoneme of the last syllable. The target phoneme was either covered by signal-correlated noise at 0 dB SNR (added), or replaced by signal-correlated noise (replaced), as shown in Figure 1. The filler

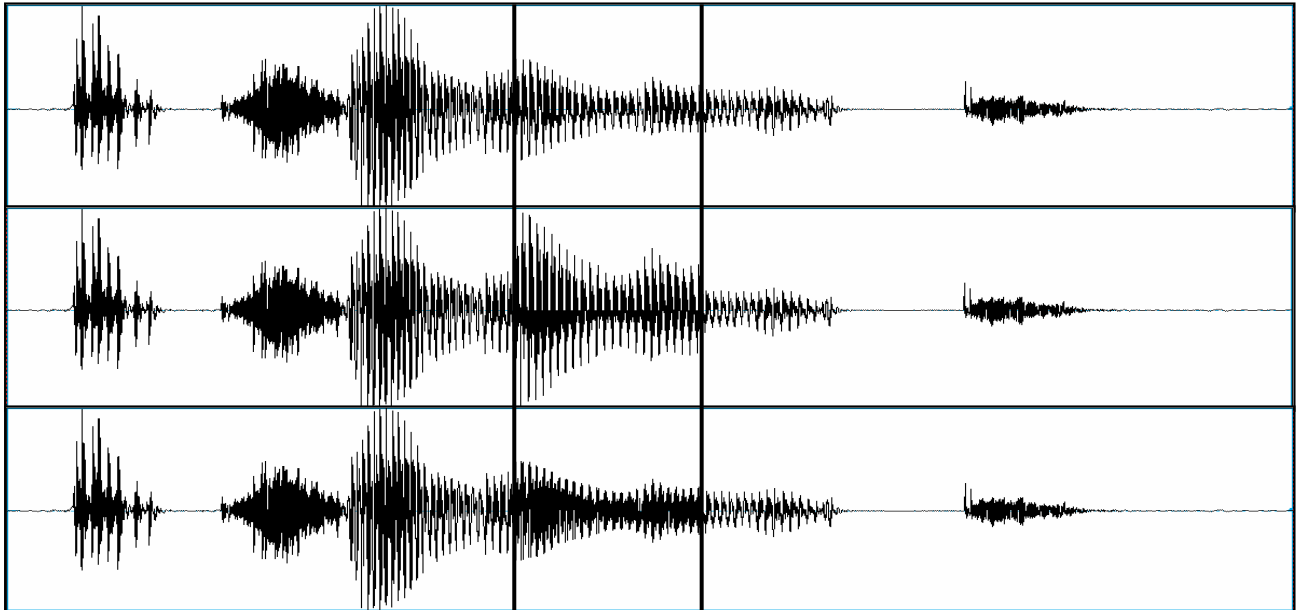


Figure 1: Acoustic waves of “accelerate” with the target phoneme /r/. The top panel shows the original speech, and the middle panel shows accelerate (added), and the bottom panel shows accelerate (replaced).

stimuli, on the other hand, were the pairs that had the same first syllable, with the target phoneme within the first syllable or at the beginning of the second syllable. The filler stimuli were prepared to make sure that listeners listen to the whole parts of the stimuli. The target phoneme was either covered by signal-correlated noise or replaced by signal-correlated noise as mentioned above. All the stimuli were spoken by a female native speaker of Standard Southern British English (see details in [7]), and the stimuli pairs were re-organized into a set of 90 words (60 test words + 30 fillers) and a set of 90 non-words (60 test non-words + 30 fillers).

2.2. Participants

Twenty seven Japanese native speakers (17 female, 10 male) participated in this study. They spoke English as a second language. Their age ranged from 20-60 years old. They had no reported hearing or speech impairments.

2.3. Procedures

There were two sessions in this experiment: (1) a word session and (2) a non-word session. Half of the participants started with the word session followed by the non-word session, and the other half started with the non-word session followed by the word session, for the counterbalancing purpose. Each session had 180 trials, half of which consisted of stimuli with an added phoneme, and the other half with a replaced phoneme. All items were presented in a random order. Each session lasted about 15 minutes, and the total duration of the experiment was about 40 minutes.

Participants wore headphones (SONY MDR-CD900ST), and sat in front of a computer in a sound treated room. They listened to an English word (or non-word) with noise (added/replaced) and an English word (or non-word) without noise (normal) in a row, and evaluated the similarity of the two by using an 8-point scale (8: very similar, 1: not similar). Listeners responded with 8 if they clearly perceived the

original speech behind noise, while they responded with 1 if they did not perceive the original speech behind noise. The audibility of the original speech was evaluated from 1 to 8 in an ascending order.

3. Results

Table 1 shows the mean similarity scores of words and non-words with an added or replaced phoneme (liquid or nasal) on an 8-point scale (8: very similar, 1: not similar), as compared to their normal counterparts. The words and non-words with an added phoneme yielded the higher similarity score than those with a replaced phoneme. In addition, the words and non-words with a nasal yielded the higher similarity score than those with a liquid.

The size of restoration of a missing phoneme is indicated with scores in the replaced condition (no original sound existing behind noise). When the difference between the added and replaced score is small, it means, these two different sounds were equivalently perceived. It can be understood that listeners perceived the existing and non-existing sound similarly, as if a missing phoneme was there in the replaced condition. Listeners seem to have restored the deleted phoneme behind noise.

Table 1. Mean similarity score on an 8-point scale.

	words		non-words	
	liquid	nasal	liquid	nasal
added	6.13	6.32	6.14	6.41
replaced	5.32	5.78	5.22	5.78

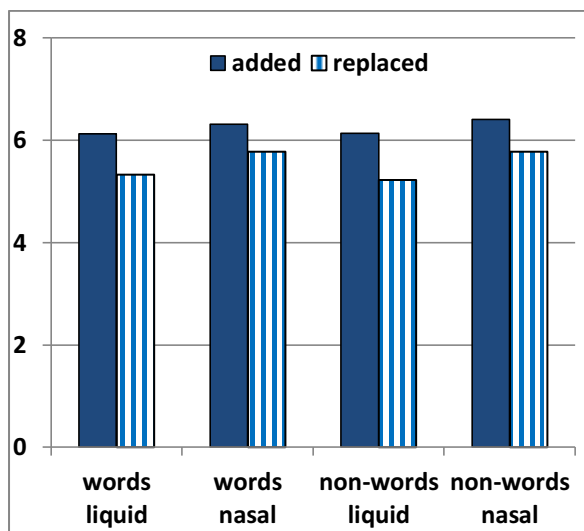


Figure 2: The mean similarity score of an added and replaced phoneme in words and non-words.

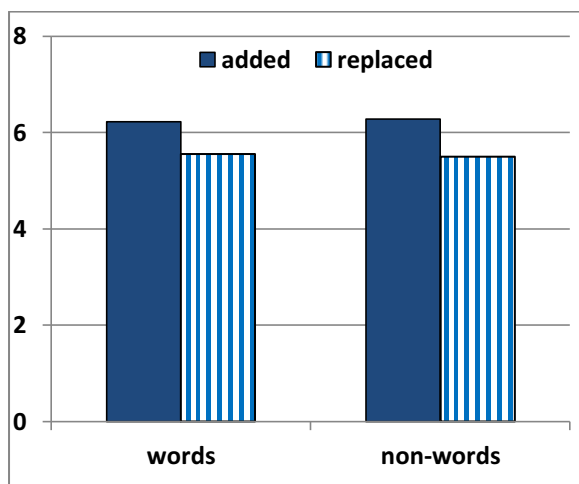


Figure 3: The mean similarity score of an added and replaced phoneme in words and non-words.

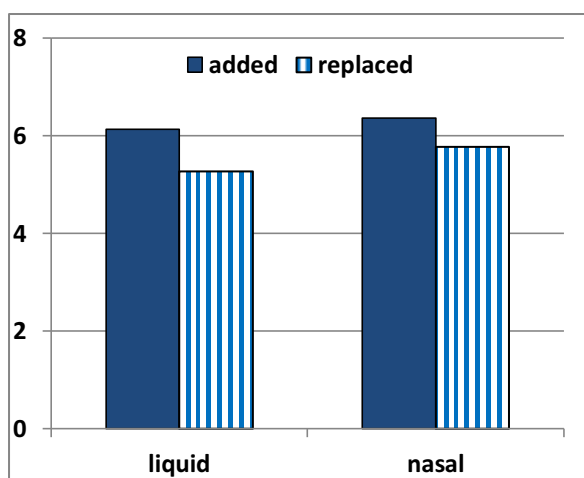


Figure 4: The mean similarity score of an added and replaced liquid and nasal.

Figure 2 shows the mean similarity score of the added and replaced liquid and nasal in words and non-words. While the added sounds were almost equivalently evaluated, the replaced sounds were evaluated differently depending on the phonemes. This suggests that the restorability of a missing phoneme might depend on the acoustic characteristics of the original phoneme.

Figure 3 shows the mean similarity score of the added and replaced sounds in words and non-words. It seems that there was no big difference between words and non-words. The mean score of words (added) was 6.22 and the mean score of words (replaced) was 5.55. On the other hand, the mean score of non-words (added) was 6.27, and the mean score of non-words (replaced) was 5.50.

Figure 4 shows the mean similarity score of the added and replaced liquid and nasal. There seems to be a slight difference between the liquid and nasal. The mean score of a liquid (added) was 6.12, while the mean score of a liquid (replaced) was 5.27. On the other hand, the mean score of a nasal (added) was 6.36, while the mean score of a nasal (replaced) was 5.78.

An ANOVA was performed with two noise conditions (added vs. replaced), two phonemes (liquid vs. nasal), and two lexical conditions (words vs. non-words) as within-subject factors. The results suggested that a noise-added phoneme was perceived significantly more similar to a normal sound than a noise-replaced phoneme, $F(1, 26) = 35.999, p < .001$. Listeners seem to have perceived the difference between the presence and absence of a L2 phoneme behind noise. In addition, a nasal in noise was perceived significantly more similar to a normal phoneme than a liquid in noise, $F(1, 26) = 33.237, p < .001$, which suggests the involvement of phonemic-acoustic details in perceptual restoration. On the other hand, there was no significant difference between words and non-words, $F(1, 26) = .001, p = .997$. It seems that lexical factors did not affect the perception of existing and non-existing phonemes behind noise among L2 listeners. The interaction between phoneme and noise was, as expected, significant, $F(1, 26) = 7.509, p = .011$. However, there was no significant interaction between lexical and phonemic factors, $F(1, 26) = 1.806, p = .191$, and no significant interaction between lexical and noise factors, $F(1, 26) = 1.528, p = .227$. Lexical factors did not affect the perceptual sensitivity to the presence and absence of phonemes behind noise nor perceptual restoration among L2 listeners. There was also no significant three-way interaction among lexical, phonemic, and noise factors, $F(1, 26) = .024, p = .878$.

Additionally, a paired T-test was performed with the restoration score of a liquid (replaced) and a nasal (replaced), in order to see phonemic effects in perceptual restoration. The results suggested that a nasal ($M = 5.27, SD = 1.11$) is significantly more restored than a liquid ($M = 5.78, SD = 0.96$) by Japanese native speakers who speak English as a second language, $t(26) = 4.834, p < .001$.

4. Discussion

The results suggested that there are common, slightly common, and very different characteristics between speech perception of L1 and L2 listeners. What was common was that listeners were able to differentiate acoustic details of noise (added or replaced). When L2 listeners listened to a word with an added phoneme and a word with a replaced phoneme, they found the word with an added phoneme more similar to a normal sound than the word with the replaced phoneme. Their perceptual system seems to be sensitive to the presence and

absence of speech signals behind noise, which was also observed among L1 listeners [7]; L1 listeners differentiated the added and replaced phoneme even when engaging in cognitive tasks (i.e., finding a specific visual pattern in a presented picture). Regardless of the size of attentional resources available for auditory processing, the acoustic details seem to be naturally detected in L1 and L2.

What was slightly common between L1 and L2 listeners was the role of phonemes in perceptual sensitivity. Samuel (1981, 1996) reported that a liquid and a nasal are in the medium range of restorability among L1 listeners when replaced by noise. The results of L2 listeners in this study also showed that a liquid and a nasal are in the medium range of restorability with the mean score of 5.32 and 5.22 for a liquid in words and non-words, and with the mean score of 5.78 and 5.78 for a nasal in words and non-words. We also analyzed the data by conducting a T-test and found that a nasal was significantly more restored than a liquid by L2 listeners. One of the possible explanations for this is that perceptual restoration is supported by listeners' L1 linguistic background. While nasal phonemes, /m/ and /n/, exist in Japanese, liquid phonemes, /l/ and /r/, do not exist in Japanese. When the phoneme co-exists in L1 and L2, the listeners' perception system would find it easy to restore the distorted portion of speech in L2. However, when the phoneme does not exist in L1 but exists in L2, the listeners' perception system might need to work harder to restore the distorted portion of speech in L2. The presence of particular phonemes in a particular language might regulate the size of perceptual restoration.

Another possible explanation for a nasal was significantly more restored than a liquid, is that this is how our perception system works. Our recent data with L1 English listeners suggested that L1 listeners might also restore a missing nasal more than a missing liquid. The size of restorability of each phoneme might be shared equally by native and non-native speakers, and the acoustic-phonemic details are processed similarly regardless of listeners' native language. This needs to be further investigated by examining the illusory perception of L1 and L2 phonemes among native and non-native listeners. At this moment, we can speculate that there might be two steps in perceptual restoration of phonemes; phonemic-acoustic processing, and lexical processing. The phonemic-acoustic processing might be shared among L1 and L2 listeners' perception system, but the lexical processing might differ depending on the listeners' language proficiency.

Finally, what was different between L1 and L2 listeners is lexical advantage in perceptual restoration. While lexical contexts helped the perceptual restoration of L1 English listeners [7], it did not help the perceptual restoration of L2 English listeners. That is, L2 English listeners seem to have perceived the missing phoneme in words and non-words equivalently. As Japanese native speakers in this study (L1 Japanese, L2 English) did not have the vocabulary size of English native speakers, there seemed to be no lexical advantage over the restoration processes in English. It is possible that the perceptual restoration of distorted speech depends on the listeners' vocabulary size. In the future study, it is worth examining the relationship between the size of receptive or productive vocabulary in a particular language and perceptual restoration, by analyzing the data based on the vocabulary size of L2 speakers.

5. Conclusions

This study explored the perceptual restoration of a missing phoneme in L2, as compared to L1. The results suggested that L2 listeners were perceptually sensitive to the presence and absence of acoustic signals behind noise, although the absence of acoustic signals is usually perceptually restored, and listeners hear the illusory sound of non-existing speech. The acoustic details seem to be perceptually detectable regardless of language proficiency. At the same time, phonemic restoration was also related to listeners' linguistic background. When phonemes co-existed in L1 and L2, these phonemes were restored relatively easily. However, when the phonemes did not exist in L1 but in L2, the size of restoration in L2 was relatively small. It seems that L1 phonemic inventory regulates the perceptual restoration in L2. In parallel, there is also a possibility that perceptual restoration takes place in two steps: acoustic-phonemic processing, and lexical processing. The acoustic-phonemic processing can be equally shared among L1 and L2 listeners, but lexical processing might differ depending on the listeners' language proficiency. Finally, this study suggested that there was no lexical advantage in perceptual restoration of L2 listeners. The size of restoration for words and non-words were equivalent for L2 listeners, while L1 listeners restored the distorted words more than distorted non-words. The lexical context helped the perceptual restoration of L1 listeners, but not of L2 listeners. In the future studies, the relationship between perceptual restoration and listeners' language proficiency (i.e., vocabulary size, listening and speaking proficiency) needs to be further investigated, to understand the critical factors for perceptual restoration in L1 and L2. Findings of these studies would contribute to understand the development of listening proficiency in L1 and L2, and auditory mechanisms in real-life situations.

6. Acknowledgements

The authors would like to thank Dr. Sven Mattys, who kindly offered the stimuli for the experiment, and Dr. Keiichi Yasu, who kindly created the 8-button response board for the study. We also would like to thank Dr. Arthur Samuel who kindly gave us comments on the paper. We also would like to thank participants for this experiment, and the members of Arai Lab of Sophia University and the members of Kashino Lab of NTT Communication and Science Laboratories. This research was supported by JSPS KAKENHI Grant # 15J00033.

7. References

- [1] Warren, R. M. (1970). Perceptual restoration of missing speech sounds, *Science*, 167, 392-393.
- [2] Warren, R. M., & Warren, R. P. (1970). Auditory illusions and confusions. *Scientific American*, 223 (12), 30-36.
- [3] Warren, R. M., & Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, 9, 358-362.
- [4] Kashino, M. (2006). Phonemic restoration: The brain creates missing speech sounds. *Acoustical Science and Technology*, 27 (6), 318-21.
- [5] Samuel, A.G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, 110, 474-494.
- [6] Samuel, A.G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, 125, 28-51.
- [7] Mattys, S.L., Barden, K., & Samuel, A.G. (2014). Extrinsic cognitive load impairs low-level speech perception. *Psychonomic Bulletin and Review*, 21(3), 748-754.