Japanese vowel devoicing offers an important contribution to the understanding of sound change and the nature of phonological representations. Some studies report that Japanese vowel devoicing results in a deleted vowel [1], suggesting that it is an abrupt phenomenon. Other studies suggest that Japanese vowel devoicing is a gradual phenomenon, in the sense that temporally reduced vowels are observed before deletion takes place [2]. Whether Japanese vowel devoicing is abrupt or gradual impacts the nature of phonological representations: is the devoiced vowel deleted or does it leave a trace in the phonological representations? The goal of this study is to provide a better understanding of Japanese vowel devoicing as a sound change phenomenon and contribute towards the debate on the dynamic nature of phonological representations. It will be argued that lab speech is an important tool to understand how phonological representations are built and evolve in time. We suggest that the results presented in this paper provide evidence for the dynamic nature of phonological representations.

Typically, in Japanese, [i] and [u] are devoiced between two voiceless obstruents, as in (k[y]kaku ‘division’, k[ij]kaku ‘plan’) [3]. However, devoicing of non-high vowels, as in k[a]karu ‘take’, k[e]tae ‘digit’, k[o]tae ‘answer’, has also been observed, although at lower rates [4]. In lab conditions devoiced vowels are shorter than voiced ones [2] and voiced high vowels are shorter than voiced non-high vowels [5]. In this study we investigated the duration of devoiced vowels in comparison to their voiced counterparts, as well as the duration of voiced high and non-high vowels. Our analysis is novel in two respects: (1) the use of electroglottography (EGG) and (2) the use of a corpus of spontaneous speech for acoustic analysis (The Corpus of Spontaneous Japanese - CSJ [6]). For a full description of the methodology consult [7, 8].

With EGG, we aimed at evaluating the articulatory correlates of acoustic properties related to vowel devoicing in Japanese. An additional purpose of this experiment was to consider the gradual implementation of vowel devoicing in Japanese from an articulatory perspective. Vowel devoicing has been studied under lab conditions [2, 9]. These studies focused on the acoustic properties of vowel devoicing showing that devoiced vowels present a drop in intensity and shorter duration. In our EGG experiment we investigated the relationship between articulatory and acoustic properties of vowel devoicing in Japanese. A group of 26 native speakers of Tokyo Japanese were recorded using simultaneous audio and EGG. The subjects uttered watashi wa ____ to iimashita “I said ____”. The target word contained three to four moras and only the first vowel, which was unaccented, was analyzed. The duration of the first vowel of the target word was measured (based on the EGG), as well as the contact quotient (CQ). CQ is the duration of vocal fold contact during each glottal cycle [10]. Vowel duration of 848 vowels and CQ of 624 vowels were fitted to three different linear mixed models with predictors of voicing and vowel type and speaker as random variables. P-values were obtained through log-likelihood tests. The results showed that partially devoiced vowels are approximately (s.e.: 1.73) 17.35 ms shorter than voiced vowels (\(\chi^2(1)=98.69, \ p <0.001\)). The slope estimates for high vowels (-9.87 for [i] and -16.2 for [u]) have greater magnitude than the estimates for non-high vowels (-2.25 for [e] and -1.61 for [a]), in comparison to the baseline [a]. This result indicates that high vowels are significantly shorter than non-high vowels, confirming previous findings with articulatory correlates provided by the EGG [3, 5].

The results of a likelihood-ratio test showed that the CQ of partially voiced vowels was lowered by 2.95% (s.e.: 0.52) in comparison to voiced vowels (\(\chi^2(1)= 31.96, \ p <0.001\)). Moreover, CQ was raised approximately 0.61% (s.e.: 0.36) for the vowel [e] and approximately 1.55% (s.e.: 0.36) for the vowel [o] in comparison to the baseline [a] (\(\chi^2(2)= 7.69, \ p = .004\)). We conclude that partially devoiced vowels are produced with smaller duration for each glottal cycle and that different duration patterns for each cycle are observed depending on vowel quality and height. These results provide further evidence for the claim that Japanese vowel devoicing is a gradual articulatory phenomenon.

Regarding the acoustic analysis based on the CSJ the focus was on vowel duration. Previous studies had shown that duration is affected by vowel height: high vowels are shorter than non high vowels [5]. Thus, the purpose of this experiment was to evaluate if data from naturalistic speech, which was not collected to evaluate vowel devoicing in Japanese, provided similar results to lab speech presented in previous studies. The data from 819,439 vowels show that duration was affected by voice quality (\(\chi^2(1)=21073, \ p <0.001\)) and by vowel height (\(\chi^2(1)= 53154, \ p <0.001\)). The log transformed duration of devoiced vowels was 0.32 log ms lower in comparison to voiced vowels and the duration of high vowels was 0.29 log ms lower in comparison to non-high vowels. These results indicate that gradual temporal reduction is observed in spontaneous speech from the CSJ corpus. These findings contribute to the debate on methods for analysing speech [11]. Specifically, results from this study show that lab speech may provide results compatible to naturalistic speech.

Additionally, the findings of study foster better understanding of the nature of sound variation and change regarding Japanese vowel devoicing. The first contribution is to provide evidence that lab speech
may provide similar results to naturalistic speech. A second contribution of this paper is to show that the implementation of Japanese vowel devoicing through time is gradual, and not an abrupt phenomenon. Vowels are temporally reduced in a gradient fashion not only in their durational values but also within each glottal cycle, as demonstrated by the experiment with EGG. We therefore suggest that vowel devoicing in Japanese is dynamic in the sense that time plays an important role in its implementation, both in physical terms (duration) and articulatory ones (CQ). If vowel devoicing is dynamic and gradual how is it represented phonologically?

We claim that the gradual nature of vowel devoicing in Japanese should be part of phonological representations. Instead of being a redundant property to be discarded we suggest that the fine phonetic properties from vowel devoicing contribute towards the abstract nature of phonological representations. We assume that Exemplar Models provide the tools to design a dynamic approach to phonological representations [12, 13, 14, 15, 16]. Within this view phonological representations may be understood as dynamic and potentially changing through time. In the case under investigation here, we showed that vowel devoicing may range from a full vowel, through a devoiced vowel and eventual vowel deletion. Our proposal is based on lab speech which we suggest provide an important tool to understand how phonological representations are organized and evolve through time.

References