

## REVIEW OF

### **An exploratory study of voicing-related differences in vowel duration as compensatory temporal adjustment in Italian and Polish**

The paper presents acoustic data on vowel duration before voiced and voiceless stops in two languages, Italian and Polish. The paper confirms the long-described voicing effect in these languages and claims that release-to-release interval remains constant regardless of the voicing effect. This latter result is argued to support the compensatory lengthening explanation of durational differences of vowels in positions before voiced and voiceless stops.

The paper is clearly written and well structured. The statistical analysis is presented transparently and I have no major reservations (for minor comments, see below). The acoustic analysis and experimental design are standard. A few shortcomings in experimental design: the procedure includes a high number of repetitions of the same word (5-6 repetitions), which is not desired, but a general practice, and a relatively small inventory of nonce words.

Results are relevant for the discussion of the voicing effect for the following reasons. Polish has been reported not to feature the voicing effect. The paper provides additional data arguing that voicing effect exists in Polish as well. The data suggest that the release-to-release interval remains constant regardless of voicing of the following stop. According to the author, this supports the compensatory lengthening approach to the voicing effect.

The main objection against the paper is that it fails to interpret its result in relation to a vast body of literature that exist on the voicing effect. It is true that the constant rate of the release-to-release interval is predicted under the compensatory lengthening approach. However, it is likely that more than one factor influences vocalic durational differences before voiced and voiceless stops. It is possible that the durational differences result from a sum of different mechanisms (each of which contributes a weighted portion).

In sum, a well-written paper with results that yield some new information for a long-standing discussion. At the current stage, I cannot recommend publication, but with a better discussion on different factors that can influence durational differences, the paper is publishable in a journal like *Glossa*.

#### **Major comment**

As already mentioned above, the paper should clearly state that the constant value of the release-to-release interval can be due to different mechanisms operating on the vowel duration differences and does not per se point to the compensatory lengthening explanation. If the author has additional evidence in favor of why compensatory

lengthening is likely the main contributing factor, this should be stated as well. Furthermore, the author should discuss the hypothesis that laryngeal features influence vowel durations in greater detail. The few works that measure vowel duration before stops other than voiced vs. voiceless (Durvasula and Luo 2014, Beguš 2017) find evidence for the effect of laryngeal features. Because the present paper measures vowel duration before voiced and voiceless stops which also have different closure durations, the laryngeal effects and the closure effects are of course conflated. This should be addressed in the paper.

### **Minor comments**

#### **Abstract**

“factors (like perceptual biases) could”

To be fair to all proposals, I think at least the laryngeal effect hypothesis should be mentioned here.

#### **186-198**

Does the author have any guesses why the results are so different for Polish? The review of the literature is sufficient, but it would be very beneficial to the reader if the author would evaluate methodology of the studies, especially since Polish has been mentioned as a case of non-existence of the voicing effect. Is experimental design equally balanced in all studies? If not, are there patterns of experimental design in studies that don't report the voicing effect?

#### **Statistical analysis.**

The analysis is generally sound and clear. Unfortunately, the manuscript does not mention whether assumptions of the linear model are checked. Sometimes modeling durations can produce patterns in the residual plot, so it would be good to inform the reader that the residual plots look without patterns.

A short justification of why Bayes factors vs. the more standard AIC criterion is used.

#### **Release to release**

Why release to release? The author mentions in the Discussion section that release-to-release measurement is somewhat arbitrary. A broader discussion on why particularly this interval would be useful. Has the author tested any other measurements? Do the result change if, say, VOT is excluded from the interval?