

Improvements for a German Vowel Trainer CAPT Tool

Patrick Carroll, Jürgen Trouvain, Frank Zimmerer
Saarland University, Saarbrücken, Germany
(pcarroll|zimmerer|trouvain@coli.uni-saarland.de)

Yves Laprie, Odile Mella, Dominique Fohr
LORIA, Nancy, France
(odile.mella|dominique.fohr|yves.laprie@loria.fr)

1 Introduction

The proposed CAPT tool functions as a *listening and response* exercise in which L2 learners hear native German productions of vowels in the context of a short sentence, and then record themselves producing the sentence. After a recording is completed learners are presented with visual feedback showing acoustic and durational information of their vowels alongside target values and examples from native speakers. The following proposed improvements will be discussed:

- **Selecting a corpus of suitable training sentences**
- **Improving intelligibility of the visual interface for non-phoneticians**
- **Providing additional training or instruction for uses of the tool**

2 Training Sentence Selection

Sentences for the training corpus will be selected based on the following criteria:

Pedagogic Value: A sentence of high pedagogic value should impress upon the user the concept that non-standard vowel production is not merely a cosmetic part of an L2 speakers' accent, but does in fact impact their ability to communicate [1, 2].

/l/ - /i:/	1) Ich bitte dich um einen Gefallen.	2) Ich biete die beste Qualität für Ihr Geld.
/e:/ - /ɛ:/	1) Er betet vor dem Essen.	2) Wie man sich bettet , so liegt man.
/a/ - /a:/	1) Meine Schwester lebt im Staat New York.	2) Die Stadt New York ist für seine Wolkenkratzer bekannt.
/u/ - /u:/	1) Er bucht eine Flugreise nach Italien.	2) Wir segeln in einer Bucht in Italien.

Ease of Automatic Recognition: The live feedback aspect of the vowel trainer tool depends upon vowel boundaries being recognized by a forced alignment algorithm. Typically the algorithm struggles with boundary detection when the surrounding segments are acoustically similar (e.g. other vowels, nasals, liquids). Therefore, want to select sentences where the vowels in focus are surrounded by segments which mark an abrupt acoustic change (e.g. silence, stops, fricatives) [3].

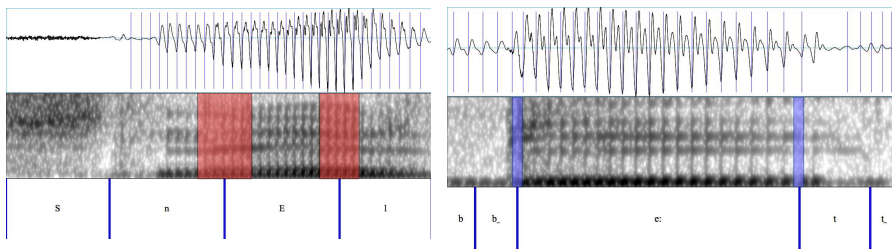


Figure 1: Ambiguous boundaries between vowel and nasal / liquid marked in red.

Figure 2: Clear boundaries between vowel and stops marked in blue.

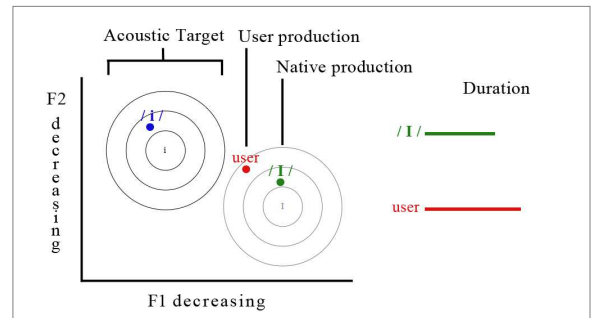


Figure 3: Original Prototype Visual Feedback.

3 Visual Interface Improvements

The current prototype of the CAPT tool [4], depicts a vowel's acoustic quality in a 2-D vowel space, and duration as a type of bar graph (figure 3). This paradigm should be intuitive to phoneticians but may not be for the average user. The following proposals offer alternative ways of visualizing vowels:

Proposal 1: Represent the principal difference between two vowel categories on a horizontal axis. Secondary differences can be seen on the vertical axis, and tertiary differences can be seen in the diameter of the point plotted. This representation encourages the user to focus on correcting major pronunciation errors first, and then fine tune.

Proposal 2: Represent acoustic quality as a shade of color by mapping the first three formant values to the RGB values of an object. Duration can be displayed as the diameter of the object. This strategy has the advantage of allowing for 4 different channels of information to be visualized simultaneously.

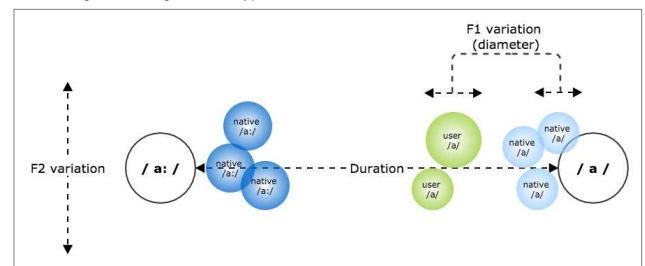


Figure 4: Vowel Axis Representation.

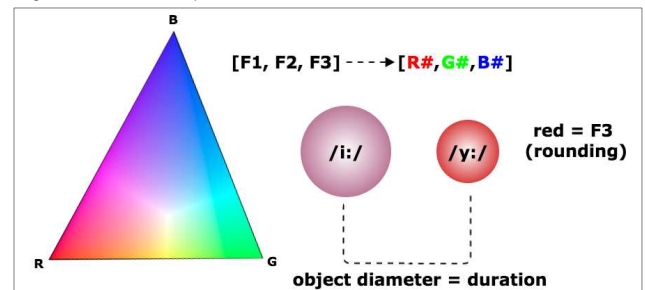


Figure 5: Color Representation.

4 Additional Training and Instruction

Before users can effectively use the tool to adjust their own vowel productions, we hypothesize that they will need some initial training in German vowel category perception. We believe that this training can be achieved in two ways:

- **Listening Exercises:** The user hears 2 sentences with different target vowels, and must identify which vowel was produced.
- **Meta-linguistic Instruction:** Provide short explanations alongside the training sentences which describe important aspects of German phonology. For example the grapheme conventions for representing long and short vowels.

Selected References

[1] Koekkoek, B. J. (1964). German vowel length and the student of american-english background. Monatshefte, pp. 267-272. [2] Zimmerer, F., & Trouvain, J. (2015). Perception of French speakers' German vowels. Interspeech 2015, pp. 1720-1724. [3] Cosi, P., Falavigna, D., & Omologo, M. (1991). A preliminary statistical evaluation of manual and automatic segmentation discrepancies. In EuroSpeech. [4] Carroll, P., Trouvain, J., & Zimmerer, F. (2015). A visual feedback tool for German vowel production. Proc. ESSV, pp. 150-157.