Voice Onset Time Enhanced User System: a web graphic interface for the analysis of plosives' release phases

2017 set the beginning of the celebrations for the 50th anniversary since the first definition of Voice Onset Time (VOT) as the temporal measure between the burst of a plosive and the beginning of glottal pulsing [1]. Abramson and Whalen [2], anticipating a dedicated special issue of the *Journal of Phonetics*, recently wrote a retrospective commentary on the sub-segmental feature, pointing out future disciplinary challenges with attention to automatic measurement of VOTs. In particular, the authors describe the potentialities of AUTOVOT, a discriminative large-margin learning algorithm [3] later integrated in a software [4] based on declarative programming. AUTOVOT has been mainly used for the optimization of the annotation of large corpora of spontaneous speech. The real-time studies on the Glaswegian English led by Jane Stuart-Smith (e.g. [5]) tested the software through the segmentation of 9898 VOTs taken from recordings made in the 1970s and the 2000s. The type of speech material did not affect the functionality of the algorithm; moreover, 62,6% of the total number of VOTs was correctly recognized after the training phase, and an additional 15,8% was easily adjusted by a manual annotator. AUTOVOT was also employed in studies on mid-term, individual, spontaneous VOT shifts [6] and research projects on phonemic accommodation [7], leaving aside its various implementations with corpora of read speech. We argue that the scientific community has just begun to scratch the surface of the benefits provided by automatic speech annotation. For example, sociophoneticians have to master a wide variety of competences, both humanistic and scientific. The lack of expertise in one of its essential components results in being detrimental to the field itself. Following recent similar proposals [8], we present VOTEUS, an open-access framework for semiautomatic VOT annotation of speech corpora. The tool, which is written in Python and integrates AUTOVOT, presents a user-friendly web based interface to automatically initialize and manually refine VOT annotations. Since VOTEUS aims at cutting to the bare minimum the required scientific skills, this can be done without relying on any form of programming. At the same time, given its modularity, it allows researchers to integrate their own algorithm to compare or complement different annotation methods. VOTEUS will be easily deployable to permit batch processing of large corpora and training of new models. To maintain interoperability with existing software the user will be able to import and export WAV files and PRAAT textgrids for further analysis.

[1] LISKER, L., ABRAMSON, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. In *Word*, 20:3, 527-565.

[2] ABRAMSON, A. S., WHALEN, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. In *Journal of Phonetics*, 63, 75-86.

[3] SONDEREGGER, M., KESHET, J. (2012). Automatic measurement of voice onset time using discriminative structured predictions. In *The Journal of the Acoustical Society of America*, 132(6), 3965-3979.

[4] KESHET, J., SONDEREGGER, M., KNOWLES, T. (2014). AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction. https://github.com/mlml/autovot/

[5] STUART-SMITH, J., SONDEREGGER, M., RATHCKE, T., MACDONALD, R. (2015). The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. In *Laboratory Phonology*, 6, 505–549.

[6] SONDEREGGER, M., BANE, M., GRAFF, P. (2017). The medium-term dynamics of accents on reality television. In *Language* 93(3), 598–640.

[7] CHEN D., KOZBUR, D., YU, A. (ongoing project). Pandering vs. Persuasion? Phonemic accommodation in the U.S. Supreme Court.

[8] VAN DE VELDE, H., HEERINGA, W. (2017). Visible Vowels. https://fryske-akademy.nl/fa-apps/VisVow/