Take a breath: Respiratory sounds improve recollection in synthetic speech

This study revisits Whalen et al. (1995) by evaluating English speaking participants in a perception experiment to determine if their recollection is affected by including breath noises in sentences generated by a speech synthesis system. Whalen et al. found an improvement in recollection for sentences that were preceded by a breath noise compared to sentences without one. While Whalen and colleagues used formant synthesis to render the English sentences, we use a modern concatenative synthesis system. The present study uses inhalations of three different lengths: 0 ms (no breath noise), 300 ms (short breath noise), and 600 ms (long breath noise). Our results are consistent with Whalen and colleagues for the 600 ms condition, but not for the 300 ms condition, indicating that not all inhalations improved recollection. The present study also found a significant effect for sentence length, illustrating that shorter sentences have higher accuracy for recollection than longer sentences. Overall, the present study indicates that respiratory sounds are important to the recollection of synthesized speech and that researchers should focus on longer and more complex types of speech, such as paragraphs or dialogues, for future studies.

Raphael Werner
(work together with Susanne Fuchs, Jürgen Trouvain, Bernd Möbius)

Inhalations in speech: acoustic and physiological characteristics

This paper examines the acoustic properties of breath noises in speech pauses in relation to similar speech segments and with regard to their inhalation speed. We measured intensity, center of gravity, and formants, as well as kinematic data (via Respiratory Inductance Plethysmography) for inhalations, aspirations of stops, glottal fricatives, and schwa vowels. We find that inhalations within speech are louder than those initiating speech, share spectral properties (center of gravity) with the aspiration phase of /k/-realizations, and generally involve a more open vocal tract (higher F1) than schwa-realizations. Intensity, center of gravity, and F1 are found to be positively correlated to inhalation speed. Overall, we conclude that jaw openness and inhalation speed are major contributors to inhalation noises in speech pauses.