

平成29年度 情報工学コース卒業研究報告要旨

戸田 研究室	氏 名	岡 田 慎 太 郎
卒業研究題目	空気・体内伝導音声の混合による自己聴取音声の生成	

発話者自身が聴取する自身の声を「自己聴取音声」といい、空気を伝わる「空気伝導音声」と比べると、言語情報は同一であるものの声質が異なるものとなる。これは、発話時に、空気伝導音声のみでなく、発話者の体内を伝わって聞こえる「体内伝導音声」も同時に聴取しているためである。その結果、録音した音声を発話者自身が聴取した場合、普段知覚している声質との間に差異が生じるため、違和感を感じる事になる。自己聴取音声を収録する機器は存在しないが、仮に自己聴取音声を生成することができれば、自分の音声を含む自分の音響体験を、他人も体験できるアプリケーションの実現などが期待される。

自己聴取音声を生成する従来の試みとして、骨伝導マイクで収録された骨伝導音声と空気伝導マイクで収録された空気伝導音声を混合する手法が研究されている。しかしながら、骨伝導マイクでは、収録可能な音声の音量範囲が限られるため、その利用範囲が限定されてしまったり、収録される体内伝導音声の品質が劣化するなどの課題がある。また、そもそも自己聴取音声の生成についての研究事例が極めて少なく、十分な検討がなされていないことから、未だ不明な点は多い。

本研究では、通常音声から極めて小さなささやき声といった幅広い音声を収録可能な体内伝導マイクとして、非可聴つぶやきマイク（NAMマイク）に着目し、NAMマイクで収録された体内伝導音声と空気伝導音声の混合による自己聴取音声の生成処理（図1）について検討する。従来研究の知見に基づき、NAMマイク収録による体内伝導音声と空気伝導音声の混合比率を変化させることで様々な声質の混合音声を作成し、それらと自己聴取音との類似度について発話者自身による評価を行う。その際に、空気伝導音声波形と体内伝導音声波形に対して、平均パワー正規化処理を施すことで、正確な混合比率を算出する。

上記の生成法を用いて混合音声を生じ、20代の被験者7名（男性5名、女性2名）に対して、評価実験を実施した。実験では、日本語5母音を対象とし、各母音において、様々な混合比率により生成された混合音声と自己聴取音声との類似度を評価した。評価実験の結果、空気伝導音声に対してNAMマイク収録による体内伝導音声を混合することで、自己聴取音声との類似度を高めることができることが分かった。また、従来研究と類似した傾向として、最適な混合比率には大きな個人差があること、また、個人でも母音によって最適な混合比率が異なることが確認された（図2左）。さらに、混合処理による自己聴取音声生成処理の効果が良く確認された被験者においては、最適な混合比率が母音/a/や/i/では低く、母音/u/では高くなる傾向が見られることが分かった（図2右）。



図1：自己聴取音声生成処理

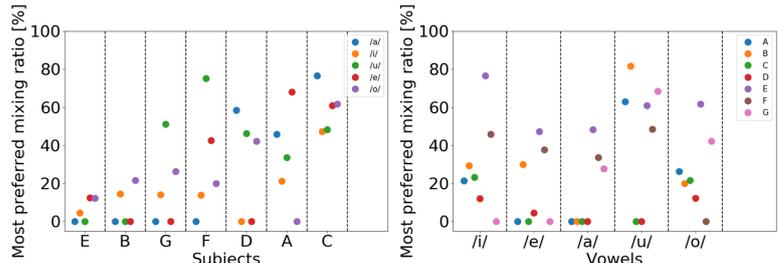


図2：各母音および各被験者における最適な混合比率
（左図：横軸を被験者とした場合、
右図：横軸を母音とした場合）