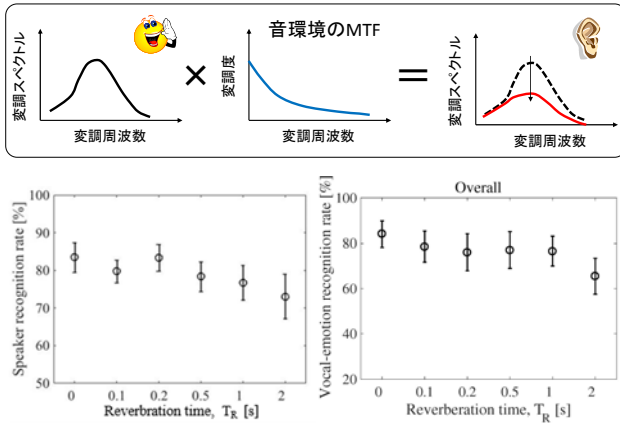


振幅変調の概念に基づいた聴知覚における質感認識メカニズムの理解



研究代表者 鷓木 祐史（北陸先端科学技術大学院大学情報科学研究科・教授）



図：音源・音環境・聴知覚の流れからみた質感認識（上）と音源の非言語情報知覚（左下：話者性，右下：感情）に対する残響環境の影響

○研究の背景と目的

質感認識は、脳による物体の本性の解釈であるといわれています。聴覚においてもヒトは、音を聞いただけで音を発した物体の性質（例えば、生物／無生物、個体差など）を瞬時に見極めることが可能です。さらに、聴覚では、音源となる物体の性質だけでなく、音がどのような環境で発せられたかという環境の性質も、同じ音から同時に認識することができます。例えば、リビングで聞いた声と風呂場で聞いた声の質感は異なりますが、声を発したヒトが誰であるのか、どういった感情で話したのかといった音源の性質はきちんと見極めることができます。

本研究の目的は、音環境の質感認識にターゲットを絞り、「振幅変調」の概念に基づいて、その物体の発する音の質感だけでなく、それが耳に到達するまでに通った伝送経路の質感も認識するメカニズム（図1上）を理解することです。

○これまでに得られた成果

大きく分けて、次の二つの成果を得ました。

(1) 音源の変調周波数と言語・非言語情報の知覚

まず、聴覚フィルタバンク、振幅包絡線検波、変調フィルタバンクで構成される雑音駆動型音声分析合成系を構築しました。次に、この雑音駆動型音声分析合成系を利用して包絡線情報の帯域制限を施した4モーラ単語（親密度別単語音声了解試験用データベース）に対する単語了解度試験を行いました。その結果、言語情報として正しく単語を認識するに

は変調周波数 4 Hz 以下の情報が必要であることが分かりました。次に、同様の操作から、話者 10 名 10 文章（ATR 音声データベース C セット，男性 5 名，女性 10 名）に対する話者認識実験を行いました。その結果、変調周波数 16 Hz 以下の情報が必要であることが分かりました。最後に、5つの感情（平静，喜び，悲しみ，抑えた怒り，荒げた怒り）をもつ日本語感情音声から作成した雑音駆動音声に関して感情認識実験を行いました。その結果、雑音駆動声の感情を認識できることが分かりました。また、認識率が感情によって差があることも分かりました。

(2) 残響特性が非言語情報の認識に与える影響

室の変調伝達特性（残響特性）と音源の変調成分の関係ならびにそれらに関する聴知覚の検討を行いました。実験では、残響環境における非言語情報の知覚を想定し、Schroederの室内インパルス応答（残響時間 0~2.0 s）と(1)と同様の音声刺激を利用しました。ここでは、静音／残響環境下で観測された音声信号を雑音駆動型声分析合成系で得られた振幅包絡線情報に関し、非言語情報（話者性と感情知覚）の知覚が、室の音響特性（残響の変調伝達関数）によってどのような影響を受けるか調査しました。その結果、音源の話者性（図1左下）や感情（図1右下）のような非言語情報の知覚は残響の影響を受けにくいことが分かりました。

これらの結果から、我々は音源の質感と音環境の質感を容易に切り分けて認識していることが示唆されます。現在、これらの結果に基づき、様々な音環境（例えば、雑音環境や雑音残響環境など）での音源の質感の影響（分離知覚）を調査しています。

○関連する研究発表論文

1. Zhu Z., Nishino, Y., Miyauchi, R. and Unoki, M.: Study on linguistic information and speaker individuality contained in temporal envelope of speech, *Acoustical Science and Technology*, 37(5), 258-261, Sept. 2016.

学会発表など

1. Zhu Z., Miyauchi, R., Araki, Y., and Unoki, M., "Feasibility of vocal emotion conversion on modulation spectrogram for simulated cochlear implants," *Proc. EUSIPCO2017*, pp. 1884-1888, Kos Island, Greece, Aug. 2017.