

「自分の声」と「他人の声」を識別する 音響学的特徴を解明

～思春期における自己像形成の理解に向けて～

首都大学東京大学院 人文科学研究科 萩原裕子教授の研究グループ

首都大学東京大学院 人文科学研究科 萩原裕子教授の研究グループは、「自分の声」の識別に関する音響学的特徴を明らかにするために、音声の基本周波数^{注1)}と周波数帯域を操作した結果、母音や子音の情報を形成するフォルマント^{注2)}の影響が大きいことを明らかにしました。

私たち人間の「声」は個人により違いがあり、「顔」とともに個人を同定する手がかりになると考えられます。複数の他人の声を識別するのと同様に、自分の声を他人の声から識別することもできますが、どのような音響学的特徴が「声の自己像」をなしているのかについてはあまり知られていませんでした。今回、音声の高さや構造といった特徴について調べたところ、「声」の認識全般には高さや母音が重要な役割を担っており、それらの情報が制限されている状況では、「他人の声」よりも「自分の声」の方が認識しやすいことが分かりました。この結果は、言語による精神機能の自己制御の神経基盤解明の基礎資料となるもので、健全な精神発達の基盤である「自己像」の形成過程の理解と、自己像に問題を抱える精神疾患の病態解明への貢献が期待されます。

この研究は、文部科学省科学研究費新学術領域研究（研究領域提案型）「精神機能の自己制御理解にもとづく思春期の人間形成支援学」（領域代表者 東京大学医学部教授 笠井清登）（平成23年7月～平成28年3月）の一環として行われ、本研究結果は、2013年10月11日に米国の国際科学誌「Frontiers in Psychology (Language Sciences）」（オープンアクセス誌）に掲載されました。

研究内容の詳細及び注1)、注2)解説は別紙参照

【お問合せ先】

首都大学東京大学院 人文科学研究科 言語科学教室
TEL: 042-677-2217 / FAX: 042-677-2950

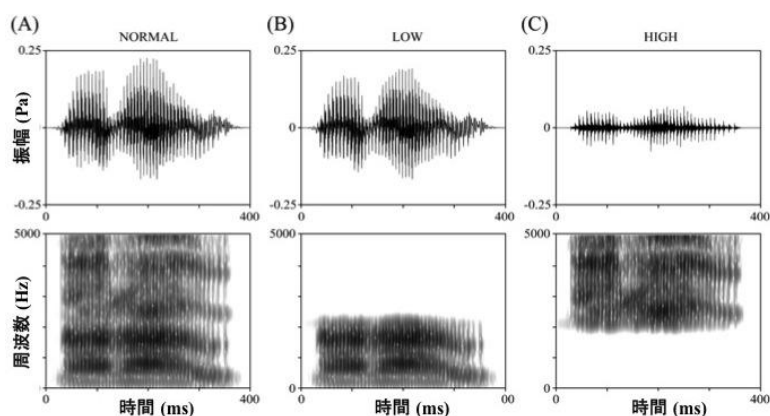
研究の背景

言語機能は私たち人間だけがもっている高度な認知能力の一つであり、精神機能の発達において重要な役割を担っています。ヒトは言語を用いて自己像や他者の存在を明晰化し、社会からの評価によって自己像を確立します。言語を介して形成された自己像は脳機能として刻まれますが、その際、意欲や内発性や性格傾向といった内的要因が、メタ認知や自己像の形成に少なからず影響を及ぼすので、その働き方には個人差が生じます。健常者では自己像と他者（社会）からの評価との間で、多少の情報の食い違いがあっても適切に修正を行い、自己像を保つことができます。しかし精神疾患患者ではこのメカニズムがうまく働いていない可能性があるようです。例えば、統合失調症患者では、「幻聴」にみられるように、自己の内言が「他人の声」で命令されているように聞こえる場合があります。自己の内言を「自分の声」として捉えられないことは、自己像に関する何らかの異常を反映していると想定されます。このことから、声をどのようにして認識しているのかを明らかにすることは、自己像の形成とその揺らぎについて理解を深めることにつながると考えられます。

研究の内容

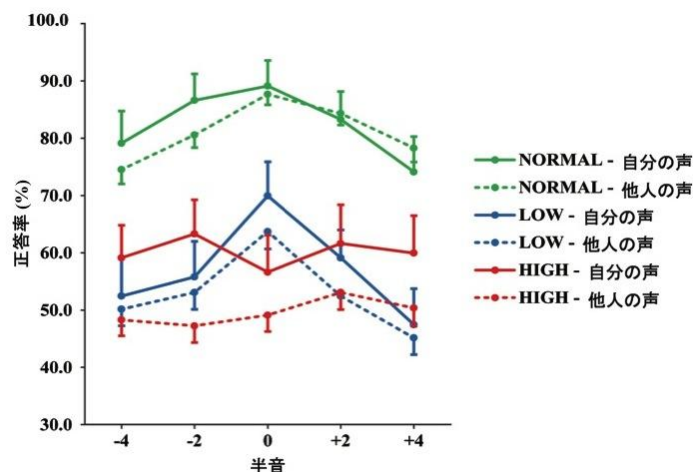
そこで、私たちは、自分の声と他人の声を識別する音響学的な特徴を探りました。声の認識とは、実際に耳から聞こえてくる聴覚情報（音声）と、その話者に対する聞き手の長期記憶（イメージ）とを照合する作業です。今回は、音声の高さ（基本周波数）と音声の構造（フォルマント周波数）に着目して行動実験を行いました。具体的には、5段階の基本周波数（元となる音声のFOを中心に半音2つずつ上下に設定）のそれぞれについて、3種類の周波数帯域（加工なしの原型（NORMAL）、第2・第3フォルマントの平均値を境界として低い周波数帯域のみ抽出（LOW）、高い周波数帯域のみ抽出（HIGH））を設定し（図1）、3モーラからなる単語の音声刺激を、合計15種類作成し、30名の成人日本人（自分及び同性の友人4人の合計5人のグループが6グループ、男女各3グループ）に、声の話者が誰であるのかを特定してもらいました。

図1



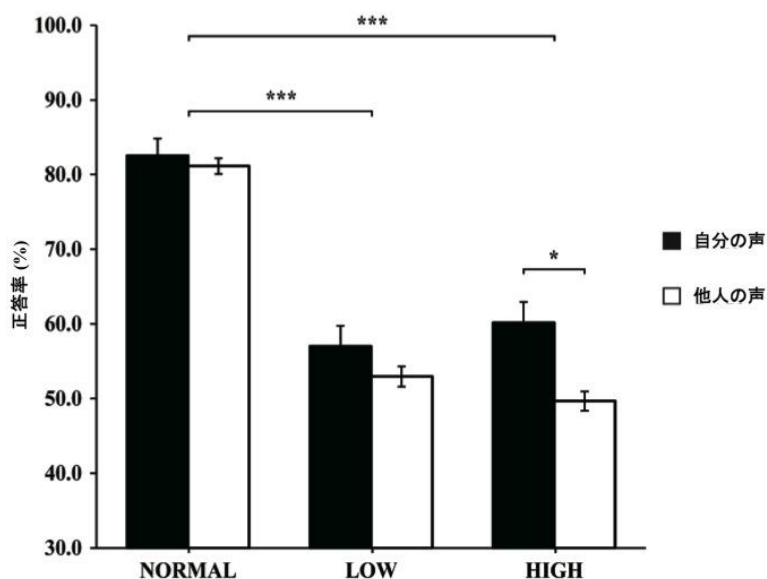
その結果、NORMAL と LOW の場合は、基本周波数が本来の FO よりも離れる（ピッチが高く/低くなる）ほど、自他ともに声の特定は難しいことが分かりました（図2）。これにより、「声」の特定全般には基本周波数及び第3フォルマント以下の周波数帯域で構成される母音が重要な役割を担っていることが明らかになりました。

図2



一方、母音の情報が少なく、主に高い周波数帯域のみという、音声情報が極端に制限されている状況 (HIGH) では、面白いことに、自分の声の方が他人の声よりも識別しやすいことが分かりました(図3)。自分の声の認識は、音声に含まれる音響情報を利用することに加えて、例えば発話に関わる運動の表象などの複数の情報を統合して行っていることが考えられます。他人の声だけではなく、自分の声を正しく認識することが、ひいては健全な自己像の形成に繋がっている可能性が示唆されました。

図3



研究の社会的意義

人間の自我機能は主に思春期に形成されますが、精神疾患の多くはこの時期に発症しています。今回の成果は、ヒトの声の特徴を「精神機能の自己制御」という観点から捉えた初めての研究で、思春期における自己の形成と発展の理解のための基礎資料となるものです。このような研究は、精神疾患、言語コミュニケーション障害、発達障害などの支援に貢献するものと期待されます。

研究成果論文 [英文] について

○論文名

Acoustic cues for the recognition of self-voice and other-voice,
Frontiers in Psychology, 4:735. doi:10.3389/fpsyg.2013.00735 (2013)

○著者名

Mingdi Xu, Fumitaka Homae, Ryu-ichiro Hashimoto and Hiroko Hagiwara
徐鳴鏞 保前文高 橋本龍一郎 萩原裕子

○論文 URL

http://www.frontiersin.org/Language_Sciences/10.3389/fpsyg.2013.00735/abstract

用語解説

注1) 基本周波数

音声においては、発声時に声帯が振動する際の基本周波数 (F0) をさし、聞こえとしては声の高さに対応する。日本語ではアクセントやイントネーションのような要素を構成する。本研究では、元となる音声に含まれる基本周波数の時間的な変化パターンは保持したまま、音程を上げる、もしくは下げるといった操作を加えている。

注2) フォルマント

声帯から口唇までを発声時に音波が通る体腔であるとした場合に、その経路の音響特性によって決まる共振周波数が音声のスペクトル上でピークとして現れる。ピークを示す周波数の低い方から順に第1フォルマント (F1)、第2フォルマント (F2)、…と呼ぶ。