

文音声の基本周波数のシフトが個人 人性知覚に及ぼす影響

北村達也 肥田友紀子 川元広樹
甲南大学知能情報学部

研究の背景と目的

- ❖ 音声中的特徴量は発話のたびに変動するにもかかわらず、個人性は安定して知覚される
- ❖ 変動に対する許容範囲を知ることができれば、知覚上の話者間距離を定義可能
- ❖ 話者適応に求められる仕様も明らかに
- ❖ 本研究では、文音声の平均基本周波数のシフトに対する知覚特性を求める

先行研究

- ❖ 橋本ら (1998)
 - スペクトルと基本周波数は個人性知覚に寄与
 - 寄与度は特徴量の差異に依存
- ❖ Izumida & Kitamura (2011)
 - 2話者の識別に寄与する特徴量は相対的に決まる
- ❖ 出水田・赤木 (2012)
 - 基本周波数の変化と傾き, スペクトル傾斜の変動幅と個人性知覚の関係を調査
 - ある程度の変化は個人性に影響せず, ある点を超えると急に大きな変化を生じる

実験1

- ❖ 基本周波数のシフトに対する個人性知覚はどのようなものか？

刺激音

- ❖ 原音声：成人男性5名の文音声 (ATR音声データベースセットC)
 - 21歳から37歳
 - 未知話者
 - 標本化周波数: 20 kHz, 量子化: 16 bit
- ❖ 文
 - 文(a) 「あの坂を上れば海が見える」
 - 文(b) 「飛ぶ自由を得ることは人類の夢だった」
- ❖ 基本周波数のシフト
 - STRAIGHT
 - 十二平均律に基づき-6半音から+6半音まで2半音きざみ

-6半音から+6半音のシフト

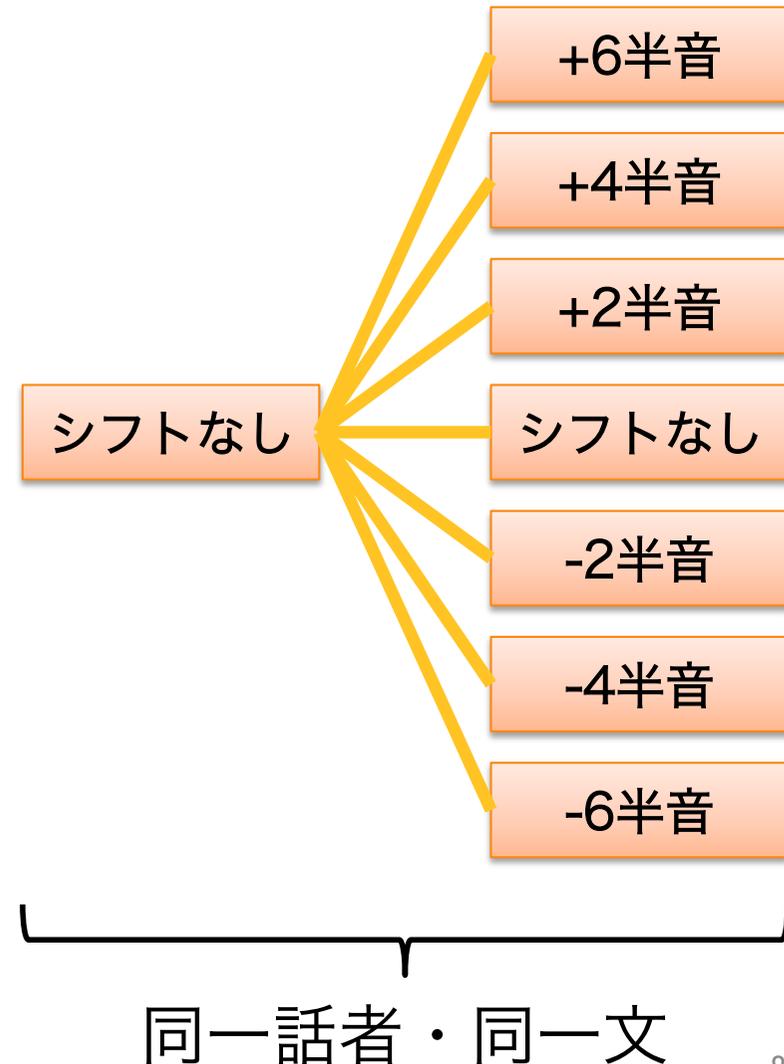
shift [semitone]	frequency [Hz]
+6	169.7
+4	151.2
+2	134.7
0	120.0
-2	106.9
-4	95.2
-6	84.9

刺激音の例

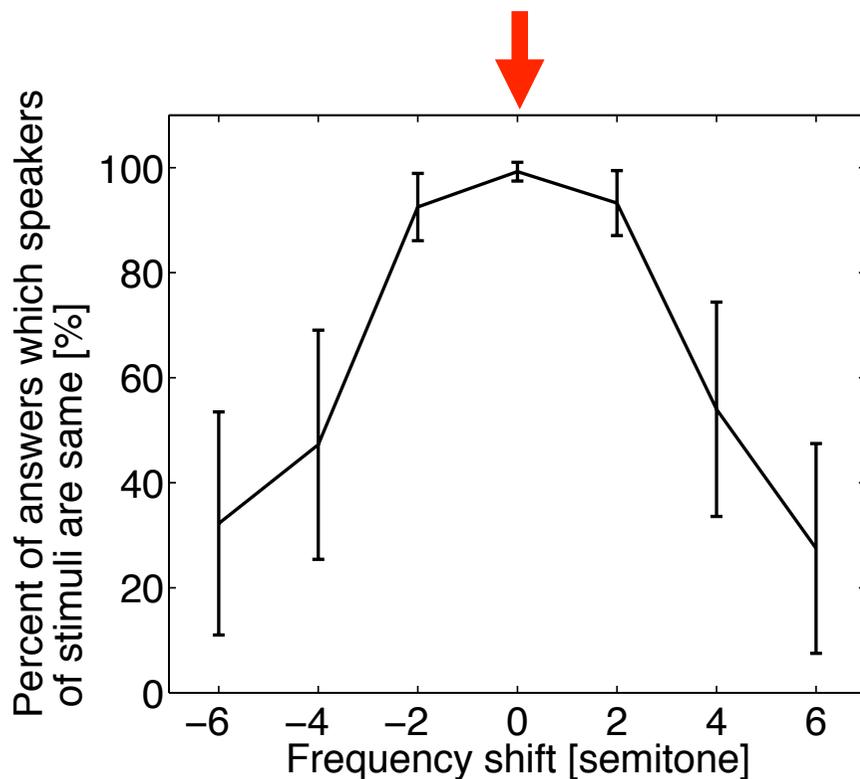
- ❖ +6半音 
- ❖ +4半音 
- ❖ +2半音 
- ❖ シフトなし 
- ❖ -2半音 
- ❖ -4半音 
- ❖ -6半音 

実験協力者・実験方法

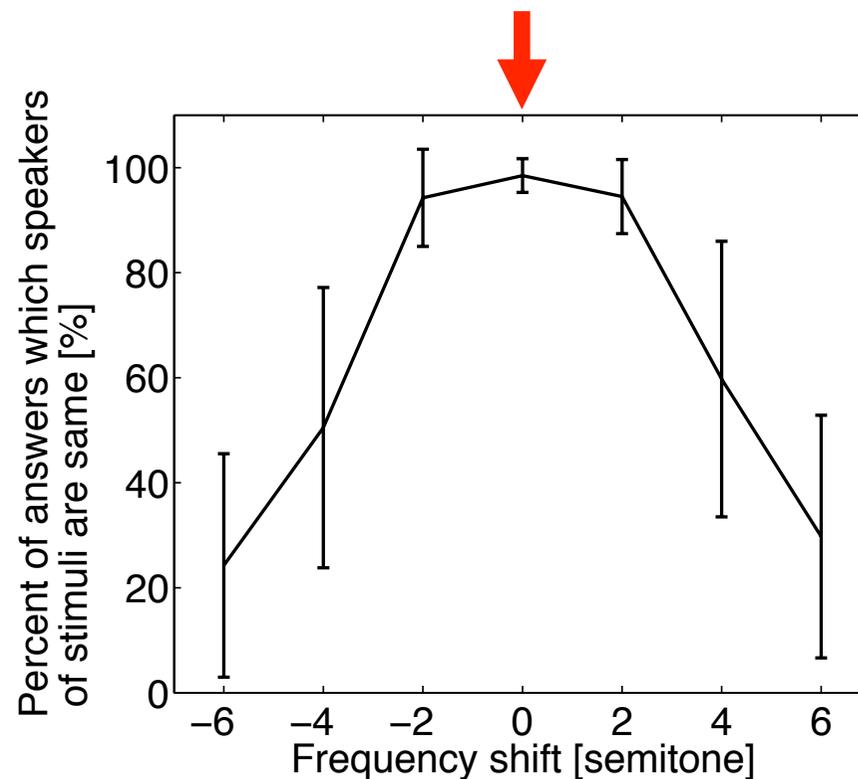
- ❖ 実験協力者：男性16名, 女性4名の計20名
- ❖ 実験方法
 - 同一話者・同一文の比較
 - 実験協力者は同一話者か否かを判定
 - Praatを用いて実施



結果(1/2)



文(a)



文(b)

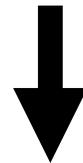
オリジナルの基本周波数を持つ文音声の刺激音 (frequency shift=0 ST) との比較において同一話者と回答された割合 (%)

結果(2/2)

- ❖ ± 2 半音のシフトは個人性知覚への影響が小さい：90 %以上
- ❖ それ以上のシフトは個人性知覚への影響が大きい
 - ± 4 半音のシフト：約50 %
 - ± 6 半音のシフト：約30 %
- ❖ 下降と上昇の影響はほぼ対称
- ❖ ある種、カテゴリーカルな知覚特性

実験2

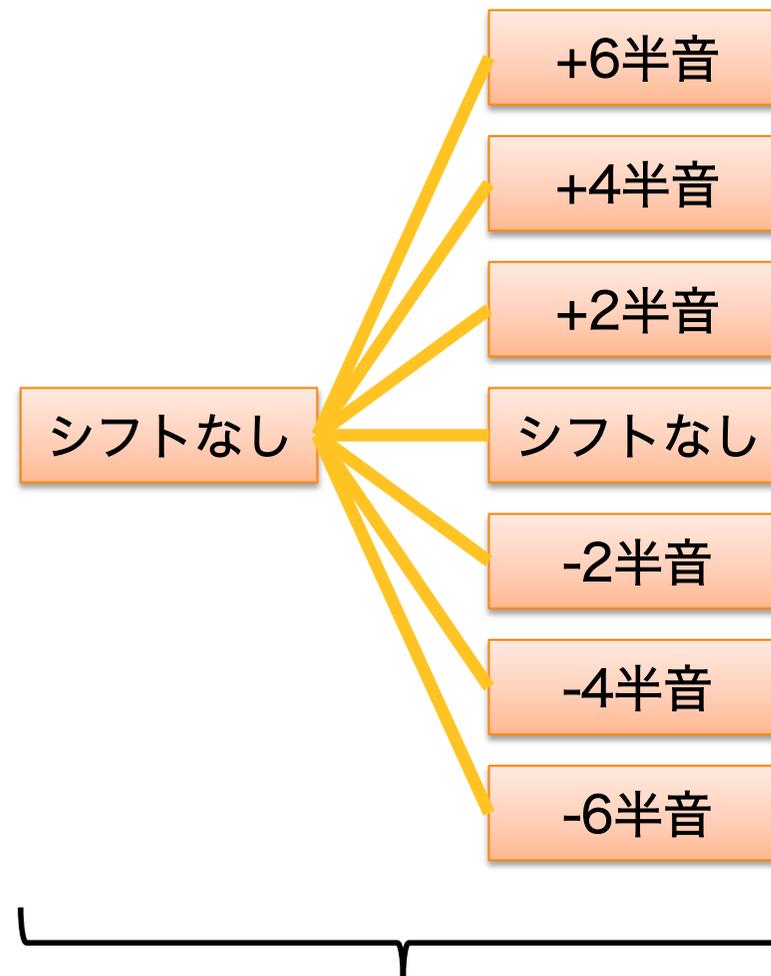
実験協力者は ± 2 半音の基本周波数シフトを検出できなかったのではないか？



刺激音のピッチを比較する実験を実施

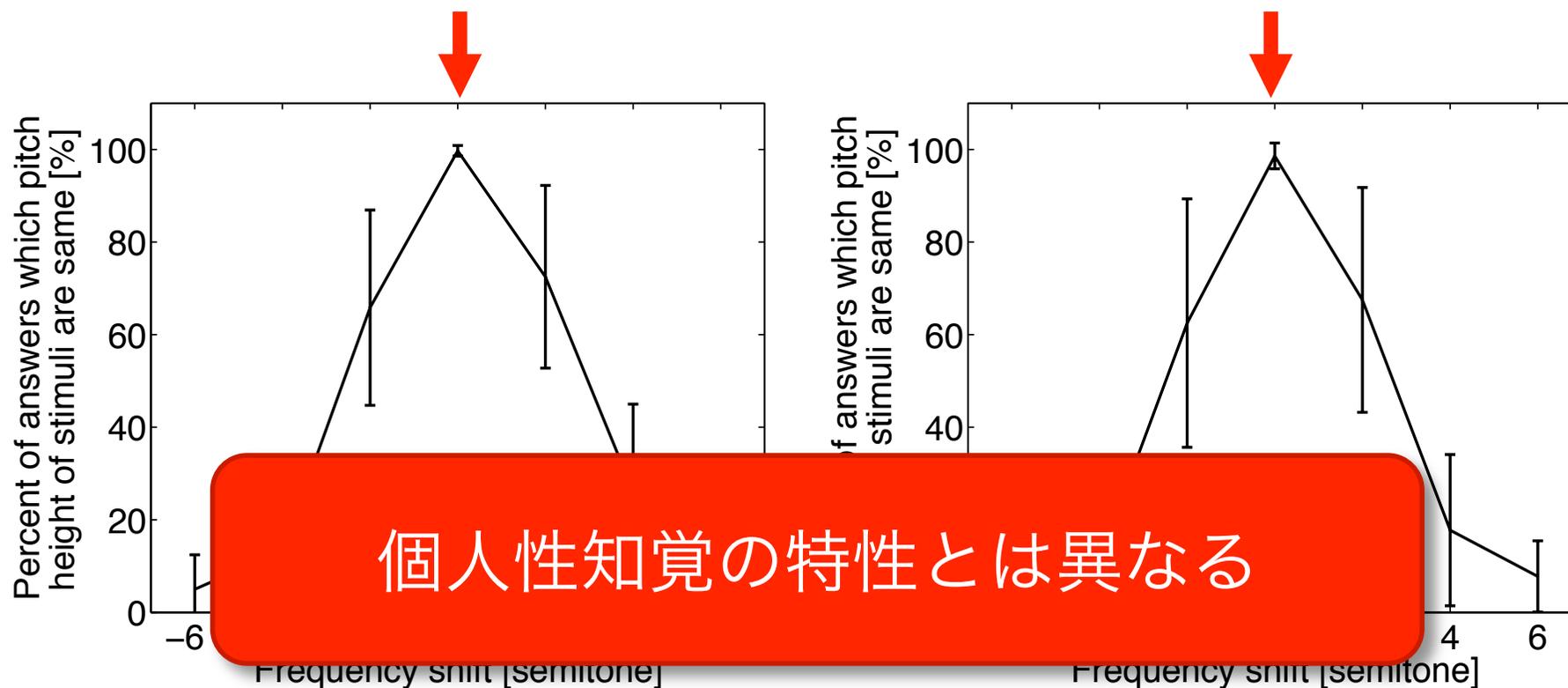
実験条件

- ❖ 刺激音：実験1と同じ
- ❖ 実験協力者：男性13名,
女性5名の計18名
- ❖ 実験方法
 - 実験協力者は声の高さが
同一か否かを判定



同一話者・同一文

結果



文(a)

文(b)

オリジナルの基本周波数を持つ文音声の刺激音 (frequency shift=0 ST) との比較において声の高さが同じと回答された割合 (%)

平均基本周波数の変動範囲

- ❖ 橋本 (1973)
 - 単語の母音重心点の基本周波数は非常に安定
 - 標準偏差は4.7 %
 - ❖ 本橋 (2002)
 - 6話者による『羅生門』の朗読
 - 25 Hz~40 Hz程度の変動
 - ❖ 水木 (1984)
 - 感情により1オクターブ程度変動
 - ❖ 宮武・匂坂 (1990)
 - 種々の発話様式の実験
 - 母音中心の基本周波数は-38.7 Hz~+32.7 Hz
- } ±2半音を
超える

日常生活の中で基本周波数が±2半音以上変動する
場面は存在する

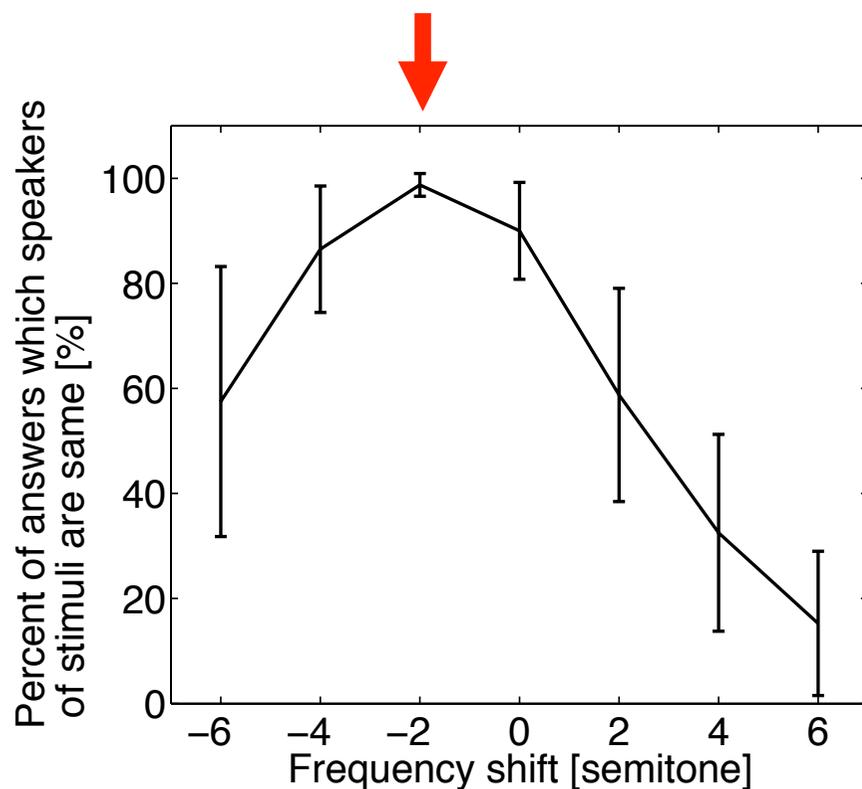
考察

- ❖ 個人性知覚において、基本周波数の一定範囲内の変動は許容され、それ以上は話者の特徴として認識される。
- ❖ ある種、カテゴリーカルな知覚特性
- ❖ ピッチ知覚特性とは異なる
 - 個人性知覚が平均基本周波数のみでは決まらない

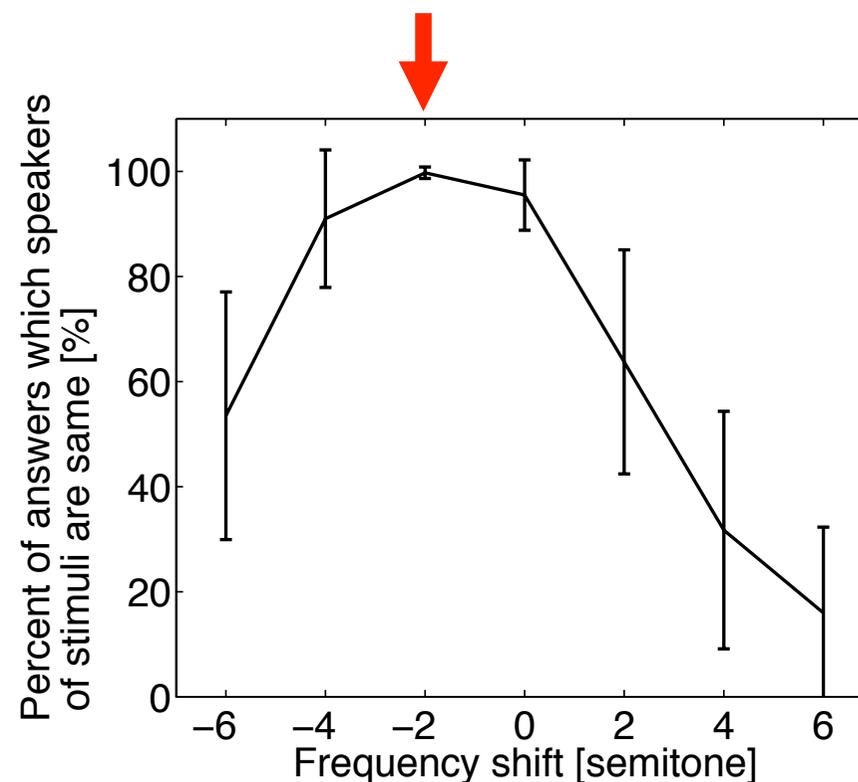
まとめと補足

- ❖ 文音声の基本周波数シフトによる個人性知覚への影響
 - 多少の変動は許容する
 - ± 4 半音以上で急激に影響が大きくなる
 - ピッチ知覚特性とは異なる
- ❖ 北村, 川元 (2013年7月EA・H研)
 - 基準の基本周波数を変えても同じ傾向
 - 比較する2文が異なっても似た傾向

基準の基本周波数を変えた結果



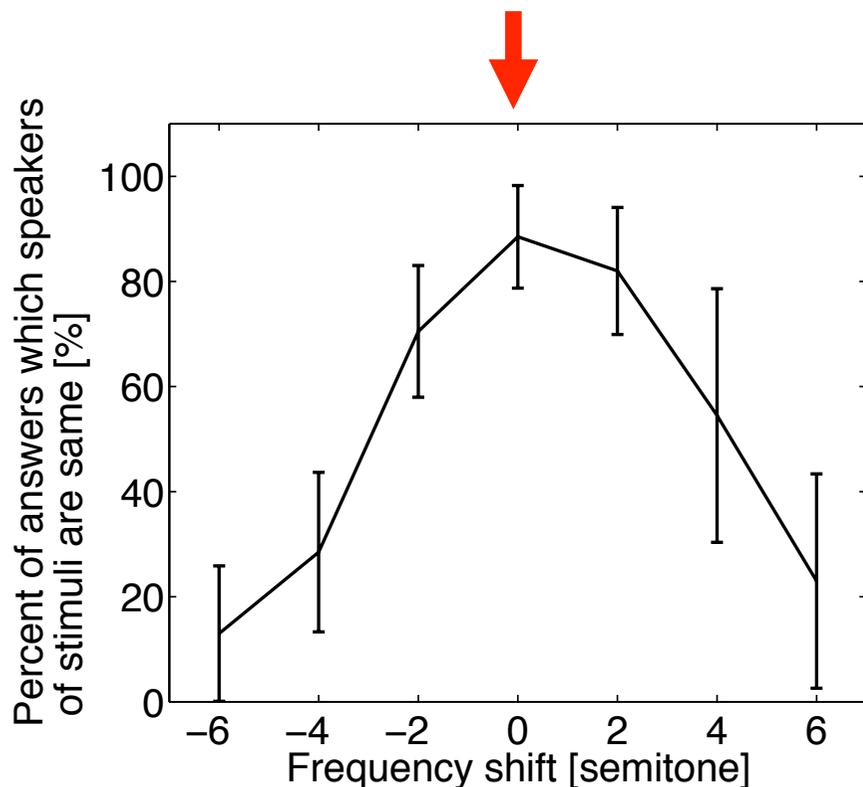
文(a)



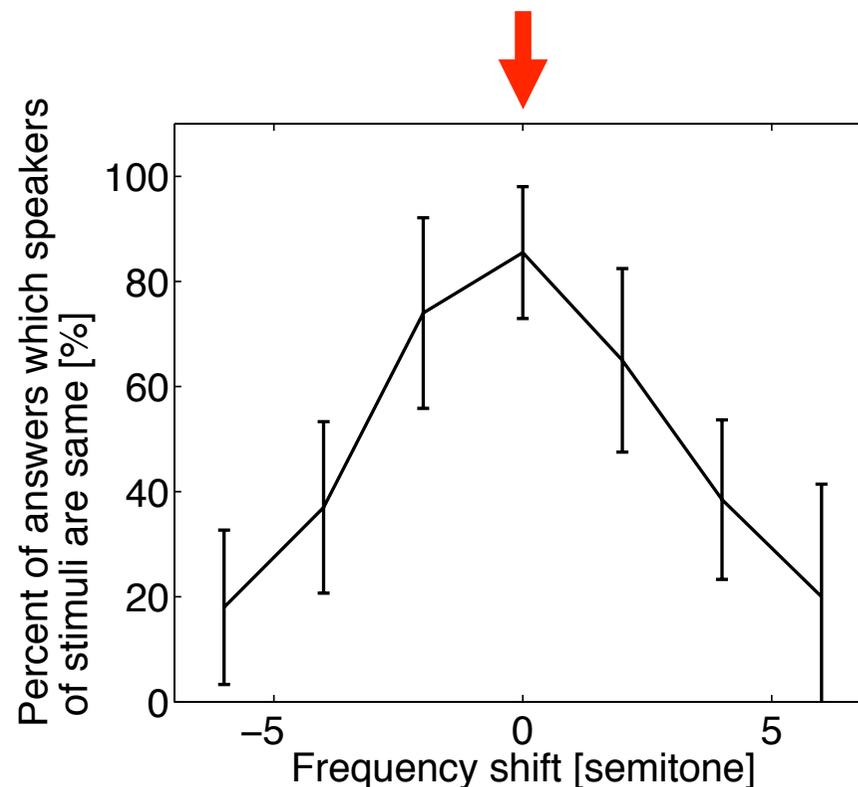
文(b)

-2半音の基本周波数を持つ文音声の刺激音 (frequency shift=-2 ST) との比較において同一話者と回答された割合 (%)

異なる2文で行った結果



基準：文(a)， 比較対象：文(b)



基準：文(b)， 比較対象：文(a)

図：異なる2文の刺激音の比較において同一話者と回答された割合 (%).
基準となる文音声はオリジナルの基本周波数をもつ.