

## 発話観測システム NDI Wave のセンサ装着が発話に及ぼす影響\*

○北村達也, 波多野博顕 (甲南大), 能田由紀子 (ATR-P)

南都智紀 (森ノ宮病院), 吐師道子 (県広大), 西谷実 (アドバンスシステムズ)

### 1 はじめに

Northern Digital Inc. の Wave speech research system (以下 Wave と略) は, 舌や口唇などに貼り付けた複数のセンサの動きを追跡することによって調音運動を観測するシステムである. センサのサイズは  $3 \times 3 \times 2$  mm で, ツイストペアワイヤ (1本のワイヤの直径は 0.4 mm) が接続されている. このセンサは小型であるとはいえ, 調音運動に少なからず影響を及ぼす. 従って, Wave を研究に利用するにあたっては, センサとワイヤによって調音運動がどの程度阻害されるかを把握しておく必要がある. そこで, 本研究ではセンサ装着時の音声および調音運動を継続的に計測し, その変化を明らかにする.

### 2 方法

#### 2.1 被験者

著者のうちの 1 名 (成人男性) を被験者とした.

#### 2.2 センサの装着

Wave では, リファレンス用の 6D と一般用の 5D のセンサを使用する. 6D センサはナジオン (nasion, 鼻根点) に, 5D センサは図 1 に示す正中面上の 7 カ所に貼り付けた. 舌尖を用いる子音の調音に配慮し, T1 は舌尖から約 10 mm の位置に貼り付けた. センサの接着には医療用接着剤アロンアルファ A (東亜合成) を用いた.

#### 2.3 データ収集

調音運動の計測に先立ち, 口蓋用センサで口蓋形状を計測した.

発話資料は昔話「桃太郎」の冒頭部分である. センサ装着完了の直後 (1 回目), 10 分後 (2 回目), 20 分後 (3 回目) にデータを収集した. 調音運動の標本化周波数は 400 Hz である. Berry[1] による Wave の精度評価の結果を参考に, フィールドジェネレータから正中面までの距離を 20 cm 以内に保つようにした. 被験者にはデータ収集の

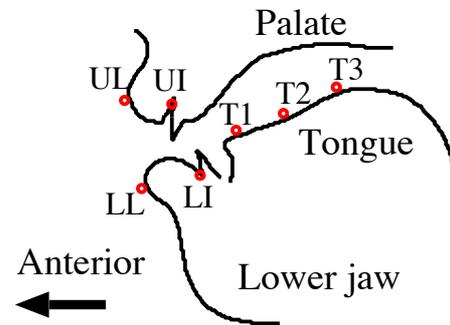


Fig. 1 Schematic illustration of midsagittal seven-sensor placement constellation. (UL: upper lip, UI: upper incisor, LL: lower lip, LI: lower incisor, Tn: tongue no. n)

合間に実験者と会話させて, 可能な限りセンサを装着した状態での発話に慣れてもらうようにした.

調音運動と同時にマイクロフォン (Sony ECM-77B) と USB オーディオ (M-Audio Fast Track) を用いて音声も収録した. この他, センサ装着せずに発話した音声も録音した. 標本化周波数 22.05 kHz, 量子化 16 bit である.

### 3 結果

#### 3.1 音声への影響

まず, 音声を聴取し違和感のある部分を選び出した. 1 回目の音声は口の中に異物があるような印象が感じられた. 特に, (1) 「し」, (2) 「じ」, (3) 「した」, (4) ラ行の音節に違和感が感じられた. 例えば, 「むかし」, 「おじいさん」, 「きました」, 「しばかりに」, の下線部である. これらはいずれも調音時に舌背や舌尖を口蓋や歯茎に接触・接近させるものである. Wave による計測では舌表面にセンサを装着させるため, これらの調音への影響を完全に排除するのは困難である. 2 回目の音声では 1 回目で感じられた違和感は減少し, 3 回目の音声ではほとんど問題がな

\*Effects of attaching sensors of NDI Wave speech research system on articulation. by KITAMURA, Tatsuya, HATANNO, Hiroaki (Konan Univ.) NOTA, Yukiko (ATR-Promotions), NANTO, Tomoki (Morinomiya Hospital), HASHI, Michiko (Pref. Univ. of Hiroshima), NISHITANI, Minoru (Advanced Systems)

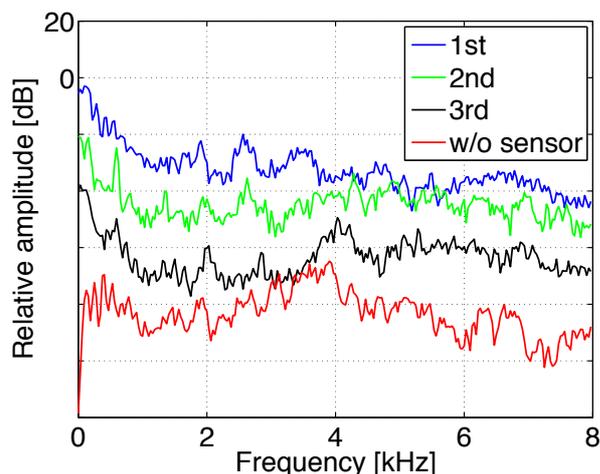


Fig. 2 Spectra of /shi/ in /mukashi/.

いといえるまでに減少した。

次に、上記の音声区間の対数スペクトルを比較した。「むかし」の/shi/部の対数スペクトルを図2に示す。なお、この区間の母音は無声化されている。音声データを標準化周波数 16 kHz にダウンサンプリングした後、フレーム長 32 ms、フレーム周期 8 ms、窓関数として Hanning 窓を用いて対数スペクトルを求めた。そして、周波数軸上で平均した。なお、各対数スペクトルには 20 dB ずつバイアスを加えている。

図2によると、1回目の音声の対数スペクトルはセンサ未装着時のものと大きく異なっている。その後、回数を重ねるに従って約 4 kHz 以上の振幅が増大し、本来の/shi/の対数スペクトルに近づいている。しかし、3回目の音声においてもセンサ未装着時の音声の対数スペクトルとの間に差異がある。

### 3.2 調音運動への影響

「むかし」の/shi/部(上の対数スペクトルを求めた音声区間に対応)のセンサ位置を図3に示す。3回の発話において LL, LI, T1(図1参照)の位置に差異が見られる。T1を例にすれば、1回目の発話ではほとんど動いていないが、3回目では 10 mm 以上動いていた。これは、センサ装着から時間が経過するに従って下顎や舌の動作範囲が広がったことを示している。

T1は下から上へ移動し、/shi/区間の最後で最も口蓋に接近する。最終的な T1 と口蓋の距離は3回目の発話が最も小さい。図2において3回目の音声の対数スペクトルがセンサ未装着時のものに近いことに対応していると考えられる。

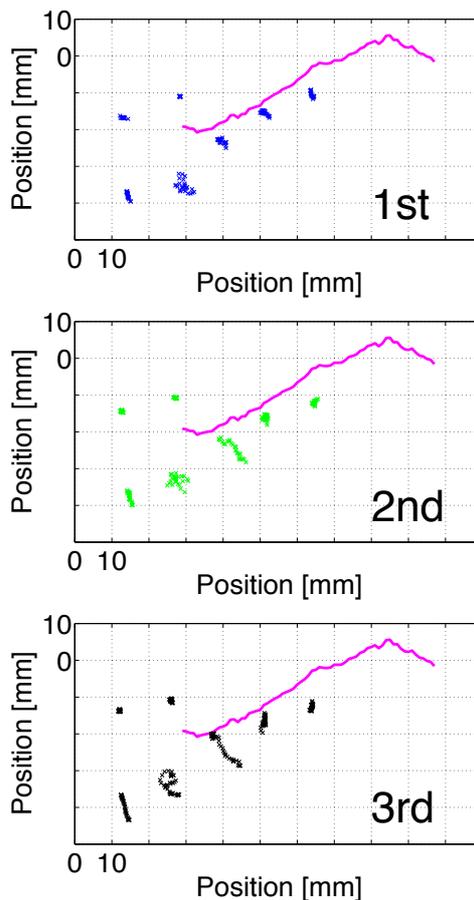


Fig. 3 Sensor position during production of /shi/ in /mukashi/. The magenta line indicates the palate.

## 4 おわりに

Wave を用いた計測では舌表面にセンサを装着させるため、普段と異なる調音になることは避けられない。その影響は特に舌と口蓋や歯茎が接触する調音で大きい。その影響を抑えるためには、被験者にセンサを装着した状態での発話を繰り返させ、普段通り発話してもセンサが外れない、もしくは外れても簡単に再装着できることを納得させることが重要である。

Wave のセンサに接続されているワイヤは柔軟性に欠ける上、十分に細いとはいえない。現在、これを柔軟で細いワイヤに交換する検討を進めており、実現すれば発話に及ぼす影響が大きく軽減できると考えられる。

**謝辞** 本研究は平成 24 年度科学研究費 (21300071, 24652085) にて実施された。

### 参考文献

[1] Berry, *JSLHR*, 54, 1295–1301 (2011).