

研究報告

「兄弟式国際ボタン音素変換標準規格」
 に至る長くまっすぐな道
**THE LONG AND STRAIGHT ROAD TO “BROTHER’S
 BUTTON-TO-PHONEME TRANSFER STANDARD FOR
 INTERNATIONAL LANGUAGE”**

三輪真弘（フォルマント兄）
 Masahiro MIWA
 情報科学芸術大学院大学
 Institute of Advanced Media Arts and Sciences

佐近田展康（フォルマント弟）
 Nobuyasu SAKONDA
 名古屋学芸大学
 Nagoya University of Arts and Sciences

概要

音声（人声）合成技術はこれまでテキストを読み上げる技術（Text to Speech = TTS）として進展して来た。つまり、文字テキストをを事前に打ち込みそれを発話/歌わせるシステムである。しかし TTS は、テキストを作成する行為とその読み上げという二段階のプロセスから構成される以上、そこには避け難い時間的遅延が内在しており、それは自然な発話行為とは根本的に異なるものである。それにもかかわらず、合成音声のリアルタイム・コントロールの実際的な試みは現在でもほとんど行われていない。おそらくその最大の理由は、その制御パラメータの膨大さにある。発話や歌唱を不自然ではない程度の精度で制御するには、当該言語の音素選択、音高、音量はもとより、音色をめぐる非常に多くのパラメータを同時に、しかも非常に短い時間内で行わなくてはならず、人声における音高や音色の微細でめまぐるしい変化を身体的動作によつて的確に操作することは不可能にみえるからである。また、さらに問題となるのは、たとえある個人の鍛錬によつてそれが実現できたとしても、それはあくまである個人の「技」の範疇にとどまり、その複雑さにおいて確定的に書き記すことも他人に伝えることもできない、すなわち「記述」ができないからである。これらの諸問題に挑むべく 2000 年に結成されたアーティストユニット、「フォルマント兄弟」は結成以来一貫して独自に開発した音声合成システムを人間の手で操作（演奏）し、リアルタイムで発話/歌わせる試みを続けてきた。今回は 2013 年に発表した MIDI アコーディオンを用いた「兄弟式国際ボタン音素変換標準規格」にまで至る様々な試行錯誤とそこから生まれた作品を紹介する。

The technology of speech synthesis has been developed

greatly and has become ubiquitous in our daily life. Now in Japan, even software producers are offering virtual singer programs for amusement purposes. However, they are always text-to-speech technology which requires a pre-composed text in advance. A technology which allows the creation of speaking or singing voices “directly” according to our immediate creative wishes, does not exist.

Formant Brother’s challenge for 10 years has been the production and control of artificial voices in real time. For this we developed not only an original Formant synthesis engine for vocalizations using Max/MSP, but also a standard which defined the mapping of all the phonemes in international language to specific notes for controlling the synthesizer. It is called “Brother’s Button-to-Phoneme Transfer Standard for International language (BBPTSI Standard)”. According to this standard, each phoneme is usually played as combination of MIDI-Accordion’s buttons, a combination of notes for vowels and consonants.

1. 「兄弟式日本語鍵盤音素標準規格」と「和音平均化アルゴリズム」

フォルマント兄弟（以下、F 兄弟）は、父親違いの異母兄弟として 2000 年以来、「〈声〉」をテーマに、人工音声を人間が演奏し、機械の声で歌わせる／喋らせる」試みによつて創作を続けて来た。

2003 年に大阪のブリッジというライブスペースで、その後の F 兄弟の活動がある意味で決定付けるようなパフォーマンスが行われた。それが〈兄弟 de ピザ注文〉(2003) である。

この時に人工音声を鍵盤で演奏するために兄弟が制定したのが「兄弟式日本語鍵盤音素標準規格」である。これは日本語のすべての音節を、1 オクターブ内の異なる和音によって同定するための規格である。

その後、メロディの指定にも工夫を加えた。すなわち様々な民族音楽において特徴的な微妙なコブシ回しや音程の微妙な揺れなどを表現するためには、西洋の半音単位の12音平均律では分解能が低すぎるため、「和音平均化アルゴリズム」を考案した。単旋律の音高指定において右手で和音を弾くとそれぞれの音高の平均値で単音の音高が決まるという仕掛けである。これらの規格に準拠することで、旋律の歌い回しやビブラートなども含む「歌唱そのもの」をより高い精度で譜面に書き、それを練習し、演奏できるようになったのである。その可能性を実証すべく、F兄弟は＜NEO 都々逸六篇＞(2009)を作曲した。

Brother's Keyboard-to-Phoneme Transfer Standard for Japanese language

兄弟式 日本語鍵盤 音素変換標準規格

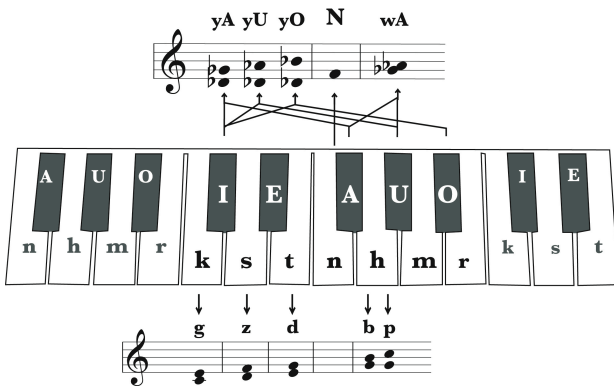


図1. 身体の規格化と文化

鍵盤キーボードを応用するのではなく、専用の特別なコントローラーを開発した方がより効率的に人工音声をコントロールできると考える人がいたとしても、兄弟は鍵盤やその後に取り組んだアコーディオンのような、既存の楽器を「流用」することを重視してきた。

その理由は、西洋音楽の五線譜が読め、それを読めば演奏できる高度な技能をもった人がすでに日本の社会に少なからずいるという事実にもとづいている。音符を見れば、10本の指が機械のように操れる驚くべき身体がそこにあるからである。それは一朝一夕に獲得できるものではなく、高価な楽器を購入して、年少の頃から長くて厳しいトレーニングを続けて初めて可能となった現実であり、同時に極東の果ての島国でありながら、高価な投資や長く苦しい訓練をしても、西洋文明のピアノを弾

けることが「価値がある」と日本の文化で見なされて来たからこそ、獲得された身体なのである。その身体を別の目的で使うという試みは、これまでの文化を別の観点で組み替える挑戦だとF兄弟は考えている。

今日の稽古は



図2. Neo 都々逸第一篇「今日の稽古は」楽譜

2. MIDI アコーディオンと「兄弟式日本語ボタン音素標準規格」

「和音平均化アルゴリズム」を用いた＜NEO 都々逸＞の作曲後、兄弟は近年、MIDI アコーディオンという楽器と向き合っている。アコーディオンにはピアノ式の鍵盤もあり、左手にたくさんのボタンがある。このボタンで日本語の音素を指定し、「和音平均化アルゴリズム」を継承した鍵盤でメロディを指定する規格を新たに制定した。それが「兄弟式日本語ボタン音素標準規格」である。

しかし、アコーディオンに注目した理由は左手ボタンによる音素指定の可能性だけではなく、何よりこの楽器には「蛇腹（ベロー）」があるという点である。MIDI アコーディオンが出力する蛇腹の情報は、単に空気圧が高いか低いかというひとつの数値変化だけなので、フットペダルでも、何かスライダーやツマミのようなものでも代用できるはずだが、しかし、蛇腹は呼吸する、声を出すという身体イメージに非常に近い行為を伴うため、結果として出て来る人工音声も「表情」のようなものが生まれ、それを極めて直感的にコントロールできるという点に大きな可能性を感じたのである。その可能性を実証すべくF兄弟は＜NEO 演歌、夢のワルツ＞(2012)を作曲した。

兄弟式日本語ボタン音素変換標準規格 BBPTSJ ver. 2.1 / 2012.07.20
 Brother's Button-to-Phoneme Transfer Standard for Japanese language

【シングルボタンでのコントロール】

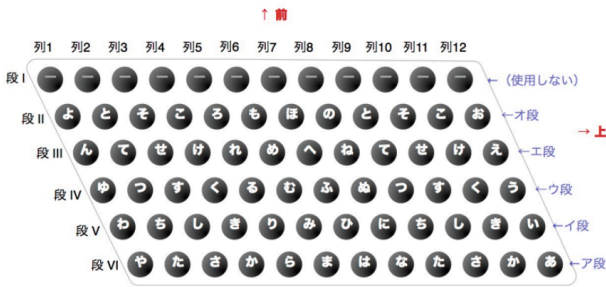


図 3. 兄弟式日本語ボタン音素変換標準規格

3. 「兄弟式日本語ボタン音素標準規格」の改訂

F 兄弟にとって仮名文字に対応する「五十音韻表」として知られる日本語に使われる音節をアコーディオンの左手ボタンにどのように配置するかは難問であった。つまり、それを操作するためにはどちらにしても「体で覚える」しかなく、ASCII (QWERTY) キーボードがそうであるように必ずしもシステムチックに (= ABC 順に) 並べる必要がないことは明白だった。しかし、ではどのような根拠に基づいて日本語音節をボタンにマッピングすれば最善なのかを判断する手がかりもみつからなかった。最終的に、数多くある左手ボタンに原則「あいうえお順」で音節をマッピングし、濁音、半濁音、二重母音などを除けば「1 ボタン = 1 音節」に対応させることにした。この原則は音節を連続して素早く呼び出す (= ボタンを押す) 際に複数ボタンのコンビネーションを連続的に弾き分けるのは困難だろうと想像したからである。

この前提に従って < NEO 演歌、夢のワルツ > をピアニスト岡野勇仁によって見事に実演してもらえた後も、やはりこのマッピングに対する「根拠のなさ」とその実演を通して明確になった広い範囲のボタンをまたぐ「演奏の難しさ」に対して F 兄弟は更なる検討を加えた末、規格の大幅な改訂を行うことにした。

それまで日本語のみを想定していた「日本語」という部分を「国際」に変更した「兄弟式国際ボタン/音素変換標準規格 BBPTSI」である。変更の主要な点は

- ボタンのマッピングを音節ベースではなく、音素ベースに変更した
- それによって「1 ボタン = 1 音節」の原則を廃し「2 ボタン = 1 音節」で発音を呼び出す
- 音素ベースにしたことで原則わずか 12 の子音ボタン間の移動だけですべての音節を呼び出す

- 母音は子音ボタンを起点とした相対的位置によって動的にマッピングされる
- すなわち原則子音ボタンを押す人差し指とそれと組み合わせられる 2 本目の指、その「手の形」が母音を決定する「口の形」に対応する
- 子音 (ボタン) 単独の発音も可能になる
- 単独子音及び、子音と母音セットの変更、拡張によって外国語発音に対応する可能性が開かれる
- 複数の母音ボタンの組み合わせによって中間母音の発音なども可能になる

などである。

兄弟式国際ボタン/音素変換標準規格 BBPTSI / 2013.05.18
 Brother's Button-to-Phoneme Transfer Standard for International language

【子音+母音の指定】—— 子音ボタンを押すとそれを起点に「動的母音エリア」が出現し母音を指定できる

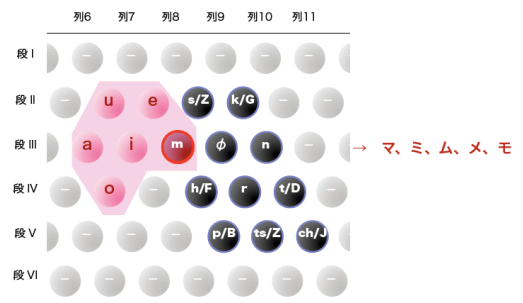


図 4. 兄弟式国際ボタン/音素変換標準規格

4. 未来に向けて

現時点で、人口音声における制御において F 兄弟は人間の声帯・口腔に依存しない人間のリアルタイム操作による発話・歌唱の実験的な可能性として考え得るほぼ唯一の方向性を示すことができたと考えており、残されているのは実証実験のみである。ただし今後は、研究協力者でもある岡野勇仁氏のようなピアニストとして高度な教育を受け、才能に恵まれた特別な人だけではなく、この試みが一般の人々にも扱えるものなのか、またそうならば、そこにどのような意義を見いだせるのかを考えていきたい。

一方、「和音平均化アルゴリズム」は単旋律における微細な音高変化を確定的に記述し再現可能にした、きわめて有効な手法だと考えているが、すでに存在する民謡などの伝統的な歌唱をこの手法に従って「楽譜化」するには多大な労力と時間を必要とすることを F 兄弟は経験している。なぜなら、目まぐるしく変化する微細な音高変化をどこまでの精度で楽譜に反映させるのか、そして何よりも和音移行時の運指の制約のもとで、それは人間に演奏可能なものなのかを事前にひとつひとつ判断していかなければならないからである。それらの課題を自動化

すべく F 兄弟は 2013 年度より、制約プログラミングのエキスパートである N T T データセキスイシステムズの山崎雅史氏の協力を得て「和音平均化旋律・運指法」の研究を開始し、その成果は近い将来報告できるはずである。

5. 参考文献

- [1] 竹川 佳成, 寺田 努, 西尾 章治郎 “鍵盤奏者のための小型鍵盤楽器を用いた文字入力インタフェースの評価” 情報処理学会研究報告. [音楽情報科学] 2006(19), pp:35-40
- [2] 竹川 佳成, 寺田 努, 塚本 昌彦, 西尾 章治郎 “鍵盤を用いた PC 用入力インタフェースの設計と実装” 情報処理学会研究報告. [音楽情報科学] 2004(41), pp:27-32

6. 著者プロフィール

フォルマント兄弟 (Formant Brothers)

フォルマント兄弟 (Formant Brothers) は、三輪真弘 (兄) と佐近田展康 (弟) という父親違いの異母兄弟によって 2000 年に結成された作曲・思索のユニット。テクノロジーと芸術の今日的問題を《声》を機軸にしながら哲学的、美学的、音楽的、技術的に探求し、21 世紀の《歌》を機械に歌わせることを目指す。

これまで発表した作品には次のものがある。トランペットの音を声に変換し歌わせる作品「La internacio」(2001、曾我部清典『トキノコダマ 2』ALM RECORDS に収録)。チェロをいじめたり笑わせたりする作品「N 市役所福祉局保育課業務日誌より」(2002)。人工音声の連弾演奏で宅配ピザを注文するパフォーマンス「兄弟 de ピザ注文」(2003)。“録楽”と名付けた音楽の在り方を考察し亡きロックスターに日本語で革命歌インターナショナルを歌わせる「フレディの墓／インターナショナル」(2009、Prix Ars Electronica 2009 入賞)、改訂版兄弟式日本語鍵盤音素変換標準規格によりマイクロトーンなコブシを歌わせる「NEO 都々逸 六編」(2009)。

また作品と一体となったテクノロジー論／芸術論の言説でも注目を集め、東京藝術大学、ロンドン・グリニッジ大学、東京大学などで講演発表・シンポジウム等を行う。