

## 今後の展望

## 音声研究 来たるべきもの

The Shape of Speech Research to Come

森 大毅

## 1. 音声研究の実り、そして収穫の後

大量の計算資源とデータ、そして深層学習技術の急速な進展によって、音声認識や音声合成などの音声情報処理技術の性能は人間の能力に迫りつつある<sup>(1)</sup>。これらの多くは、そう遠くない将来、半ば解決されたものとして研究の対象でなくなる。一方で、ドラえもんのように言葉を使う機械が今後10~20年で実現すると信じている人は多くないだろう。

では、それらの間に横たわるものは何か？ 難しい問いだが、この問いへの答えを探し求めることこそが、今後の音声研究の進むべき道を照らすことではないだろうか。

## 2. 音声研究の潮流：実験室から実環境へ

ノイズで動的な対象である音声を科学的方法で扱うことを可能にした点で、実験音声学がこれまで果たした役割は大きい。他方、統制された環境で収録した音声は、日常の話し言葉とは異なる。ここ20年の音声学・音声情報処理の一大目標は、話し言葉（≠テキストの読み上げ）を扱えるようにすることであった。

実験室データから実環境データへの流れは、今後加速していくものと思われる。例えば、代表的な自発音声データである「日本語話し言葉コーパス」に収録されている音声の多くは講演やタスク指向対話であり、実はそれほど自発性が高くない。これに対し、国立国語研究所により開発が進められている「日本語日常会話コーパス」は、日々の活動において自然に生じた会話を収録しようとするものである。これこそが我々が普段接しているはずの音声なのだが、音声情報処理の対象としては異次元の難しさである。（e.g. 音響的条件、同時発話、発声様式、パラ言語、笑い／叫び…）

今後の音声研究は、これを汚いデータだと言って見な

いふりをするのか、それとも新たな目標に奮い立つのか？

## 3. 音声研究の新しいパラダイム

音声認識は、音声信号から話者のメッセージを推定する問題として定式化できる。音声分析にせよ音声合成にせよ、問題が一旦定式化されれば、後はその評価関数の最大化を目標にすればよかった—そう、これまでは。今後の音声研究には、そのような定式化が難しいがゆえに避けてきた問題に向き合う気概が必要である。一例として、音声からの感情の理解を挙げる。未知の発話が「怒り」「喜び」「悲しみ」…のどれなのかを推定することを、解くべき問題と定義したとしよう。このように定式化して性能を競えば、感情認識の研究は一見進展したように見える。しかし、これで機械が感情を理解したことになるのだろうか？ さもなくば、感情を理解するという問題は、一体どのように設定すればよいのだろうか？

感情に限らず、音声研究の新しいパラダイムを切り開くことは、人間をより深く理解しようとすることにつながる。音声研究には工学のほか、言語学・心理学・社会科学・医学などの側面があるが、これらに共通するのは、私たち自身＝人間に対する探究心である。音声研究は、私たち自身を知るための長い旅なのだ。

## 文 献

- (1) 中川聖一, “音声処理技術がヒトの能力を超える日,” 信学技報, SP2015-74, pp. 25-30, Dec. 2015.

(平成29年5月17日受付 平成29年6月9日最終受付)



もり ひろき  
森 大毅 (正員)

平5東北大・工・通信卒。平10同大学院博士後期課程了。博士(工学)。現在、宇都宮大学大学院(電気電子工学科)准教授。音声対話、音声合成、福祉情報工学の研究に従事。著書「音声は何を伝えているか—感情・パラ言語情報・個人性の音声科学」(共著・コロナ社)など。

森 大毅 正員 宇都宮大学大学院  
Hiroki MORI, Member (Academic Association, Utsunomiya University, Utsunomiya-shi, 321-8585 Japan).  
電子情報通信学会誌 Vol.100 No.10 p.1069 2017年10月  
©電子情報通信学会 2017