

解説

—音声認識技術への応用(1)—

● 音声認識の導入による鉄道の車両入換作業の改善†



大石 勝^{††} 田辺 茂人^{††}

1. はしがき

国鉄における音声利用技術の応用は、プッシュホンによる座席予約システムの音声応答などの適用例を中心に進められてきた。近年、音声認識装置の急速な進歩により、音声入力を有効に適用しようという面からの検討が進められて、実用化されようとしている。ここで紹介するシステムも、その一つである。

このシステムでは無線と音声入力方式を用いることにより、入換作業の進路設定を作業指揮者(操車担当)が直接に、構内の任意の地点より行えるシステムである。

以下、音声入力方式を中心に本システム(RFRC)の概要を紹介する。

2. 入換作業方式^{1), 2)}

入換作業は、駅またはヤードにおいて到着した列車を分解・組成して、目的とする列車に仕立てる一連の作業である。これは操車担当を中心とするチームでなされる。主に以下の事柄よりなっている。

- (1) 入換計画を立てる。
- (2) 進路の設定を行う。
- (3) 入換機関車の運転を行う。
- (4) 車両の解放・連結を行う。

2.1 従来の入換作業方式

操車担当は入換計画を立て、これをもとに作業前に関係者と打合せを行う。転搬担当は操車担当の指示に基づいて、ポイント(転搬機)を扱うことにより所定進路を設定する。進路が設定されたら、引上線から車両を仕訳線に持込むか、仕訳線から車両を引上線に引上げるために、入換機関車を運転して引上線から仕訳線、仕訳線から引上線へと移動する。これは操車担当の手ばたの誘導に基づいて、機関士が入換機関車を操

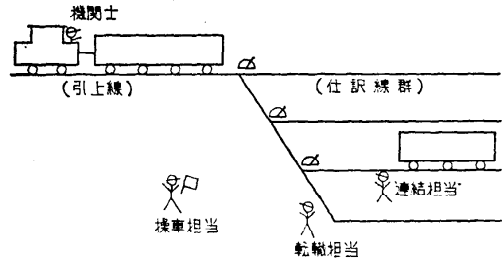


図-1 従来の入換

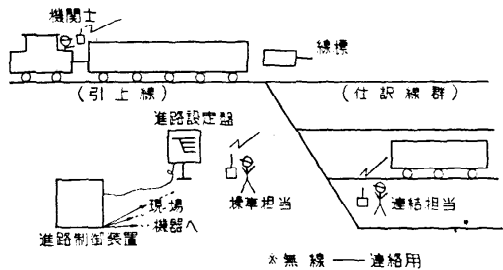


図-2 FRC方式による入換

縦することによりなされる。なお構内無線が導入されている所では、誘導を無線で行う。解放車両の連結器の解放、突放車両の停止連結等は連結担当により行われる。図-1 がその概念図である。

入換作業の改良の一環として、操車担当が直接に進路を設定する方式は、FRC (Freight-car Route Control) という名で、すでに東福山駅などに導入されている。この場合は構内の特定点に、進路設定盤を設けている。これにより転搬担当への連絡ミスの心配がなくなり、また省力化につながることになる。図-2 がその概念図である。

ところで無線方式は、構内の特定点で進路を設定することになるので、操車担当の機動性を制約する。そこで無線を利用することにより、構内の任意の地点で進路設定操作が可能な RFRC (Remote FRC) が開発されつつある。無線は連絡用としてすでに使用

† Improvement of the Shunting Using a Voice Data Input Device in J. N. R. by Masaru OOISHI and Shigeto TANABE (Railway Technical Research Institute, J. N. R.).

†† 日本国鉄道技術研究所自動制御研究室

しているものを利用すればよい。無線機での進路設定操作は、押釦入力方式と音声入力方式が考えられるが、操作性と無線機の軽量化を考慮して、音声入力方式が採用されている。

2.2 RFRC による入換作業方式

RFRC は 54 年度から開発が進められて、55 年度末より実用化試験を行っているものである。主な特徴としては以下の事項である。

- (1) 無線を利用することにより、操車担当が構内の任意の地点より進路設定可能で、作業能率がよい。
- (2) 進路設定操作に音声入力方式を用いて、操作性がよい。
- (3) 操作結果などを音声で応答アナウンスする。
- (4) 進路設定制御には従来の継電方式ではなく、マイクロ・コンピュータをベースとした方式を用いている。

RFRC において、入換作業を行う線群での進路設定を対象として、作業中は他線群への通路となる進路は固定する。それゆえ柳線のように、仕訳線尻が終端となっているところに適用されるのが望ましい。引上線が複数ある場合も、特定の一線に進路を固定する。また当該区域内で入換機関車は一機とする。概略作業手順は以下のごとくである。

- (1) 操車担当は作業計画を立てる。
- (2) 関係者で作業手順の打合せを行う。
- (3) 機関士は入機を待機させる。
- (4) 操車担当は作業域の当該進路を固定する（無線—音声入力）。
- (5) 操車担当は進路設定を行う（一々）。
- (6) 関係者は線標で構成進路を確認する。
- (7) 関係者でその進路対応の入換作業（突放、解放、連結等）を行う。
- (8) 一連の作業が終るまで(5)～(7)を繰返す。
- (9) 操車担当は作業域の当該進路の固定を解く（無線—音声入力）。

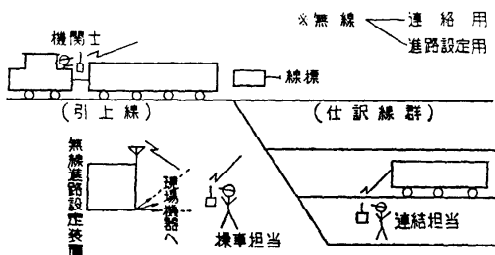


図-3 RFRC 方式による入換

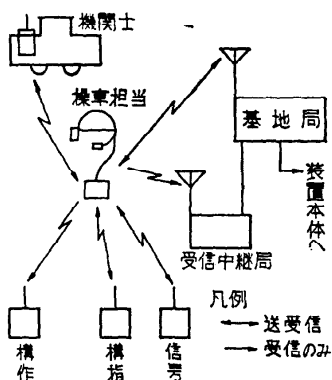


図-4 無線の使い方

図-3 は RFRC の概念図である。

進路設定用の無線は、入換作業時連絡用として、すでに使用されている無線をそのまま用いる。操車担当の携帯する無線機は、スイッチで連絡用と進路設定用に分けて利用できる。無線波は 1 波なので、進路設定用の送話は、トーン信号と共に送信される。これにより連絡用と区別する。使い方の概略は図-4 である。

主な機能は以下のようである。

- (1) 基地局は特定のトーン信号受信時のみ受信態勢となる。
- (2) 1 波プレス・トーク方式である。
- (3) 2 以上の受信局があれば、受信レベルの高い方を選択受信できる。
- (4) 基地局でトーン信号受信中は送信モードとはならない。

3. 音声入出力方式

RFRC における進路設定操作は、音声入力でなされる。処理した結果等は、音声出力される。いずれも無線で交信される。操作に用いる語は、あらかじめ使用者自らの発声入力に登録しておかなければならない（特定話者）。作業時の入力音声は、その登録パターンとの類似で評価され、認識される。

3.1 操作語および応答アナウンス

基本的な操作は以下のものがある。

- (1) 作業開始操作
- (2) 進路設定操作
- (3) 進路取消操作
- (4) 進路確認操作
- (5) 作業終了操作

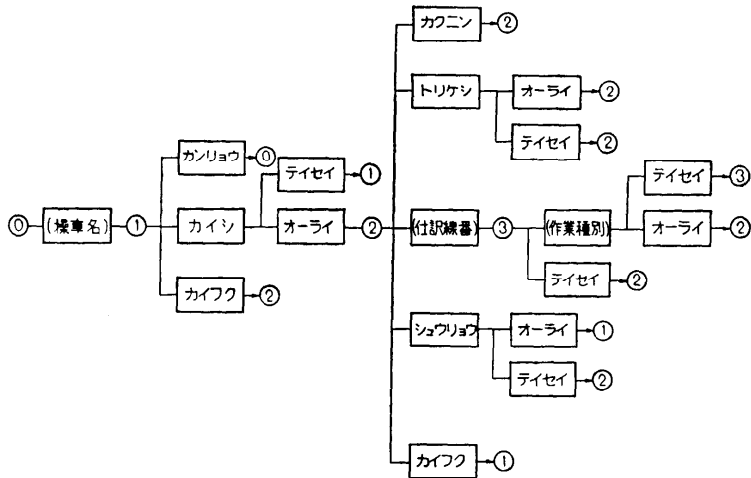


図-5 操作語のシンタックス

表-1 入換進路設定操作語

種類	操作語	記事
操車名	(扱者名そのもので)	初期設定後4名分まで選択可能
操作完了	カンリョウ	最初指定した扱者対応の操作の完了
作業の開始終了	カイシ, シュウリョウ	
仕訳線番	イチバン, フタババン, サンバン, ヨンバン, ゴバン, ロクバン, ナナバン, ハチバン	仕訳線が1番から8番までの場合
作業種別	レンビキ, ヒキアゲ, トップウ, シンニューウ, ホリユウシンニューウ, ホリユウヒキアゲ	入換のやり方
進路取消	トリケン	
進路確認	カクニン	
確認応答	オーライ, テイセイ	acknowledgement
制御語	カイフク	シンタックスの移行

表-2 入換進路設定操作の手順例

音声入力	装置よりの応答	記事
○ ○	○○さん 準備オーライです	操車名 (扱者) ひとつの時間帯の作業 作業 ひとつの 作業 作業返し のし
カイシ	作業開始です	
オーライ	番線どうぞ	
××バン	××バン 作業種別をどうぞ	
△ △	△△ です	
オーライ	××バン △△ 了解です	
：	：	
シュウリョウ	：	
オーライ	作業終了です	
：	作業終了しました	
カンリョウ	終わりました	次の時間帯の作業 操車担当の交替 (扱者)

これらの操作は音声で行われる。そのための操作語を表-1に示す。これらの語は一樣に操作されるのではなく、前後の順序関係（シンタックス）がはっきり決まっている。たとえば、仕訳線番の入力の後に入力されるのは作業種別である。このシンタックスを利用して、入力される語の認識率をよくしている。図-5がそのシンタックスである。

操作語が入力されるのに対応して応答アナウンスがある。これは入力された操作語を確認すると共に、次の操作の案内を行うためである。

実際の入換作業は、これらの進路設定操作を行いながら進める。その操作手順例を表-2に示す。

4. システムの目標と室内実験

4.1 音声入力操作に対するシステムの目標

進路設定については保安上十分な考慮がなされなければいけない。進路設定操作については、入力音声の認識という確率的な要素が入るが、設定された進路は線標で確認の上作業を行うので、保安上の問題はない。しかし入力音声の認識を誤って、別の進路が設定されると、現場扱者の信頼感を落す結果となる。また入力音声がよく判別できないとして、聞き直されることは多少あってもよいと考える。そこでシステム上の目標として以下の2項が設定された。

(1) 操作者は20回に1回位聞き直されてもよいだろう。

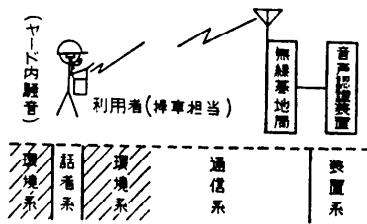


図-6 音声認識システムモデル

(2) 誤認識に基づいて現場機器の誤動作が起るとは、現用の信号機器のレベル (10^{-6} ~ 10^{-9} のエラー率と想定される) 以下でなければならない。

なお、音声認識装置の認識に対する目標としては、人間の会話程度のレベルとした。すなわち、20回に1回位の聞き直しと、1,000回に3回位の聞き間違いは起こり得ると考えた。

4.2 室内実験

前節での目標が達成可能かどうか検討するために、図-6のようなモデルを想定して、室内実験が行われた。すなわち、語が発声されて、それが当該語だと認識されるまでを以下の系に分けて、それぞれの系で何らかの要素に影響されつつ認識されるとした。これらの系での要素とその特徴を検討した。

(1) 装置系パラメータ

認識装置の性能そのものと直結したものと考えられる。装置がキャッチした音が入力音声であるかどうか判断するものと、登録語の中に該当のものがあるかどうか判断するものである。代表的なものとして以下のものがある。

(a) 発声音内最大休止

操作語を発声した実績から 400 ms の時間を設定すれば十分である (一番長い無音区間がある「トッポウ」でも普通の発声でだいじょうぶである)。操作語の入力間隔は、この値以上にしなければならないが実用上問題は無い。

(b) 発声音の最小時間と最大時間

最小時間はポップ性ノイズ* 排除に関連する。

(c) 発声音の始端と終端のレベル

認識装置への入力レベルの変動と密接に関連する。一番難しい面を含んでいる。

(d) 登録語からの抽出尺度

発声音が登録語中の当該語とどれくらい似ているか (スコア値)、またそのときに次に似ている語との差はどれくらいか (類似度) などである。

* 短く鋭い雑音。

(2) 通信系パラメータ

通信系のパラメータは、マイク入力されたところから認識装置へ入力するまでの径路に関するものである。

(a) マイク種別および位置

単一指向性マイクと狭指向性接話型マイクでは、狭指向性接話型マイク利用の方が認識率がよく、特に高騒音下 (80~90 ホーン) では、単一指向性マイクは極端に認識率が低下する。

マイク位置は口もと 1~3 cm 位に置く必要がある。それ以上離すと認識率の劣化がみられる。

(b) 無線関連

現用の構内無線の規格内でよいかということである。主な項目として帯域幅、歪率、SN比などである。現行の帯域幅は 300~3 kHz であるが認識率はあまり変化なかった。現行の歪率は 10% まで許容されている。歪みの質にも依存すると思われるが 5% を越えるとなりに認識率に影響があるとみられるが、はっきりしたことは判明していない。現行の SN 比は 30 db 以上であるが、オーダからして無視できると考えた。

(3) 話者系パラメータ

話者の発声のばらつきは音声そのものの本質からむもので、難しい面を持っている。発声音の長短、強弱、高低、抑揚、訛り、風邪をひいている等々色々なものがからんでくる。発声音の強弱は、 ± 5 db の変動なら認識率にあまり影響を与えない。また操作語に対する発声として、できるだけ互いに区別され易い発声音を選べば認識率は向上するはずである。いずれにしても、装置そのものの性能に反映しているものと考えられる面が多い。

(4) 環境系パラメータ

話者近辺での環境については、騒音、雨、風、温度などが問題となろう。

(a) 騒音

ヤード内のベース騒音はかなり高く、特に入換機関車近辺では高い。背景騒音のレベルは、登録時と認識時に合っていないと認識率は急に悪くなる。

環境系に関しては、今後行われる現地試験に待つところが多い。

(5) システムに依存するパラメータ

音声入力適用されるシステム特性によるものである。操作語の数、そのシンタックス、登録語のセット数などである。これらは音声入力操作を含んだシステムとして、バランスのとれたシステムになるように

選ぶことになる。

4.3 システム目標の達成に対する評価

前節での実験の結果、今回のシステム環境では、入力音声に対する認識率は95%位であり、誤認率は1%位であることが推察された。

以上の結果を踏まえて考察すると、4.1の第1項の目標はそのまま実現可能とみられる。第2項に関しては確認応答が採用された。すなわち現場機器の動作を伴う操作については、入力後正当なものかどうかの音声応答を行う。それに対する確認のOKかNGで返答する。このOKかNGに対応する操作語として、お互いに判別のよいものを選ぶことにより（今回はオーライと訂正）、誤認識をほとんど零にすることが可能と考える。

本章で述べた実験は室内実験であり、特に音声認識というばらつきが多い要素を含んだものであるから、今後現場での実験の回数を重ねることにより、実験結果の確度を高める必要がある。

5. 今後の試験計画と将来展望

既述したように、本システムは55年度末より実用化試験を行っている。実際に車両の入換を行うことにより各種試験を行い、それまでの室内試験をふまえて、主に以下の事柄を明確化する予定である。

(1) 実際の構内無線を利用しての音声入力の認識率が95%以上で、誤認率を0.3%以内にすることができるか？ この場合雨、風、ヤードの高騒音なども普通の頻度で起きるとする。

(2) 標準パターン登録時の状態はどうあるべきか？

(3) 作業中の再登録は無くてもだいじょうぶか？

(4) 作業能率は本システムを使用しない場合より、よくなるか？
などである。

本システムは入換作業の近代化に関するものである。無線を利用することにより、入換動車（機関車）

を操車担当自らが操縦すれば、さらに安全で簡易なシステムとなると考える。

また、本システムでの進路の設定は、入換作業のみを対象としたものであるが、今後あらゆる進路設定の操作に適用可能と考える。その線にそっての研究も進められている。

6. むすび

国鉄の各システムにおける音声入力方式の利用は、まだその緒についたばかりである。音声は人間にとって最も自然な入力方式であるので、音声認識装置の技術革新がさらに進むことにより、今後色々な分野で利用されることになると思われる。そこで、音声入力を利用する側から以下の整理が必要と考える。

(1) 適用に対応する性能を評価できる技術

一言でいえば、使用環境条件が与えられたら、それに対応する装置と性能を示すことができる。

(2) システム的な対応の取り方の技術

適用システムでの最も適切な音声入力の範囲、方式を示すことができる。

(3) 障害物（騒音等）の除去の技術

該当システムで予想される、音声入力方式にとって望ましくない環境、方式を別個のものに置換えたりする。

参 考 文 献

- 1) 「入換作業システム改善の研究」報告書、日本鉄道運転協会 (1973)。
- 2) 大石他：無線を利用した進路設定システム (RF-RC)、第17回鉄道サイバネティクス論文集、pp. 119-123 (1980)。
- 3) 千葉他：音声認識システム、「大型プロジェクト・パターン情報処理システム」論文集、pp. 157-165 (1980)。
- 4) 鈴木：ユーザーからみた音声認識装置の評価、電子技術 21 巻 12 号、pp. 12-15 (1979)。
(昭和 55 年 12 月 3 日受付)