

# 音声認識読影レポートシステム 「Natural-Report™」の開発

A development of the Voice Input Reporting System "Natural-Report™"

井桁 嘉一<sup>1)</sup> Yoshikazu Igeta  
北村 稔<sup>1)</sup> Minoru Kitamura

友寄 克亮<sup>2)</sup> Katsuaki Tomoyori  
伊藤 芳三<sup>2)</sup> Yoshizou Itou

<sup>1)</sup>株式会社日立メディコ 技術研究所

<sup>2)</sup>株式会社日立メディコ 医療情報システム本部

近年のパソコンの普及、インターネット技術の進展に伴い医療分野においても各種情報の電子化が加速している。また、画像診断装置から出力する情報も画像情報をはじめ患者情報や検査結果情報と幅が広がっており、現場の医師の意見にも読影環境に関する要望が多くなってきている。その理由として最近の診断装置は処理スピードが速いことから一日の撮影枚数も多く、その日に要求される読影件数が増えており、読影医の仕事量が非常に増えていることがある。そのため、放射線科読影医の読影レポート作成は多くの時間を割いて行われており、作成効率向上は大きな課題となっている。そこで読影レポート作成時間の短縮を目的に新しい技術「音声認識技術」を読影ワークフローへ適用することで、ディクテーションによる所見入力とレポート作成のためのパソコン制御を音声入力で可能とした読影レポートシステム「Natural-Report™」を開発した。また、認識率を実用化まで高めるために、「放射線医向け音声認識辞書」を開発し認識率95%を実現した。

Recent popularization of personal computers and development of Internet technologies have been accelerating electronization of various information even in the medical care field. In addition, information to be output from image diagnostic systems has been expanded to wider fields such as image data as well as patient data and study resultant data, and more demands for the image reading environment from clinical physicians are increasing. The reasons are that high-speed data processing capability of the recent diagnostic systems allows acquisition of a great number of images per day, and great increase in number of images to be diagnosed in the same day gives heavy workload on the image diagnosing doctors. Consequently, image-diagnosing doctors in the radiology department must spend a lot of time to diagnose images and describe diagnostic reports, therefore improvement in reporting efficiency has now become a serious problem. For the purpose of reduction in time to describe diagnostic reports, a new technology, We have applied "Voice Recognition Technology" to the image diagnosing workflow and have developed the Voice Input Reporting System "Natural-Report™" that is able to allow dictatorial input of clinical finding and PC control for description of reports by voice inputting system.

To improve recognition accuracy up to the practical level, we have developed "Voice Input Dictionary for Radiology" and have achieved its recognition rate of 95%.

**Key Words:** Voice Recognition, Voice Input Dictionary for Radiology, DICOM SR

## 1. はじめに

読影レポートは放射線読影医の最終的な依頼医への出力であり、撮影画像とともに重要な役割を果たしている。従来、読影レポートは手書きの書類としての保存が多かったが、最近では電子化した読影レポートが多く使用されている。しかし、その作成にあたっては情報の入力に時間を要し、特に所見入力では専門用語の入力や日本語の変換に代表されるキーボード入力に多大の時間がかかるため、必然的に読影医の一日の仕事においてレポート作成時間が大半を占めることになり、作成時間を短縮しレポート作成効率を向上させたいという要望を数多く受けた。一方、読影レポートの作成について米国の状況では、読影医の所見の口述記録をレポート化する

トランスクリイバーによる作成が一般的に広まっているため、読影医の入力作業とその時間の負担は日本に比べて少ない。

そこで、今回日本語によるレポート作成の時間短縮を実現するため、音声入力が可能な読影レポートシステムを開発した。また、IHE(Integrating the Healthcare Enterprise)で行われている病院内の部門間情報交換連携を実現するため、レポート作成時、必要な患者情報や画像情報を医療情報システムやモダリティーから取得し情報の共有化を図れるよう、読影レポートの新しいDICOM規格である構造化レポート(Structured Reporting)を採用し、このレポート規格に準拠したレポート入力装置を開発した。

## 2. 現状分析

読影医にとって使い易い読影レポートシステムを開発するためには、実際にレポートを作成している現場を取材し、読影レポート作成のワークフローを調査し、読影医分析(ユーザ分析)に基づいた仕様を検討する必要がある<sup>1)</sup>。そこで実際に読影をしている読影室においてユーザ分析を行った。

### (1) 実験概要

東京慈恵会医科大学柏病院放射線部では読影レポートの電子化を以前より推進しており、既に数万件単位で読影レポートが保存されている。このレポート作成には読影レポートシステムが導入されており、サーバ1台、入力端末7台が稼働し、読影医は全員レポート作成システムで読影レポートを作成している。そこで慈恵医大放射線科の協力を得て、読影ワークフローの分析を行った。実験は「ビデオ分析」と「文脈による質問法」を用いて行った。

「ビデオ観察による分析」は実際の業務をビデオで撮影し、繰り返し見直ししながら問題点を抽出する方法である。今回は読影中の医師の視線移動、キーボード操作回数、操作時間などを観察のパラメータとして評価を行った。また、同時に読影医がアプリケーションを使って仕事をしている様子を観察

しながらリアルタイムにその場で「なぜ今そのキーを使ったか」と言うような質問をし、リアルタイムでなければ得られない情報を集める「文脈による質問法」を併用しユーザ分析を行った。実験の様子を図1に、レポート作成状況を図2に示す。読影医は読影レポート入力装置のキーボード、ディスプレイ、シャーカステンの画像の3点を繰り返し、視点移動しながらレポートを作成する。この様子をビデオ撮影し分析、また質問を行いリアルタイムで行動理由などを記録分析した。図3にレポート作成画面の例を示す。

### (2) ビデオ分析結果

操作履歴を調べることで、使い易さや使わない機能などがよくわかる。本実験によりわかったことを次に記す。

- (a) 所見入力手順：所見入力手順は各人により異なる。所見の入力を、画像をじっくり見た後一気に書く。また、画像を見ながら同時平行で所見を書く、の2種類の記載方法に大きく分けられるが、全体的な時間の差異はあまりない。
- (b) 入力時間：作成に必要な操作時間を測定した結果の例を図4に示す。レポート作成時間は多くの場合1件あたり3分から5分程度であり、正常所見と異常所見では異常所見に要する時間が長い。また1件のレポート作成に



図1：実験の様子

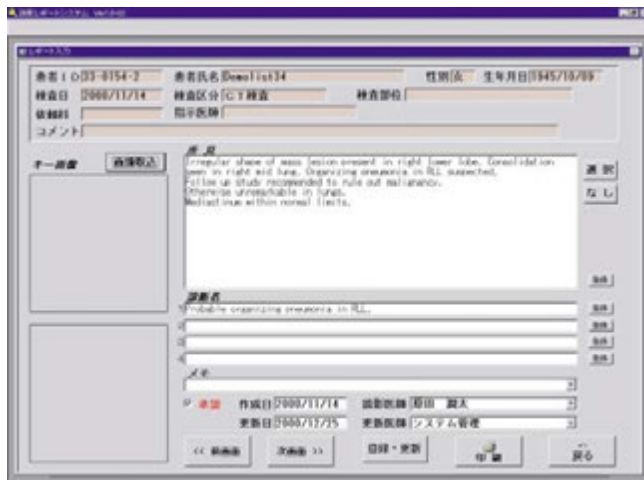


図3：読影レポートの入力画面



図2：レポート作成状況

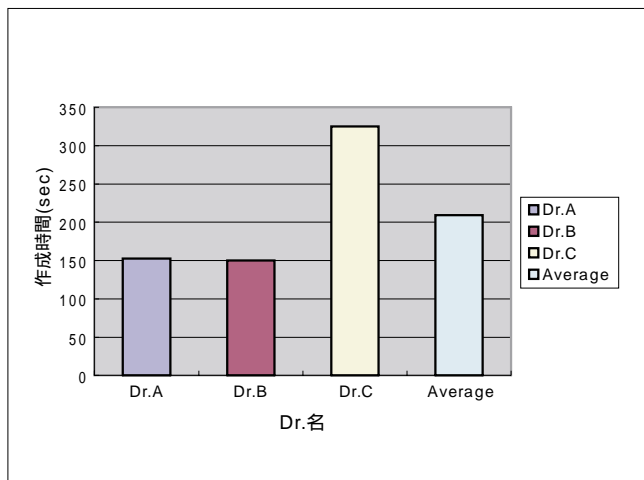


図4：レポート作成作業時間例

おける操作時間割合は所見入力操作が長く、約3分の1を占めていることがわかる。正常所見はテンプレートをを用いる場合が多い。

(c) 操作時間分析

表1および図5に1件あたりのレポート作成における操作時間の分析結果例を示す。表1は操作時間の測定結果、図5は読影医の視線の移動回数を示している。この結果から次のことがわかる。

レポートを書くにあたりキーボードとディスプレイを見る回数が多く、画像を見る時間が比較的短い。これは一日に処理する読影レポートの量より1件あたりに使う時間は限られていることから、1件の作成時間の中でワープロ操作である「所見作成に要するキーボード操作」と「ディスプレイを見る時間」が多くなっていることを意味する。この傾向は他の施設でも同様であった。

(3) 文脈質問法による結果

操作途中および操作後に読影医から聴取した意見をまとめ、日頃の仕事のワークフローにおける問題点を分析した結果を表2に示す。また所見入力についての内容では、キーボードとマウスの移動による文書入力には抵抗があり、もっと簡単に短時間で所見を作成したいこと、日本語変換に時間がかかり専門用語の入力に辞書機能が弱いので時間を要することなどを改めて確認することができた。

3. 開発システム概要

音声認識を用いた読影レポートを実現するにあたり、次のことを考慮し開発を行った。

- i) 音声認識率は効率向上を目標とすることから95%を確保し、キーボード入力の訂正にかかる手間の同等以下でなければ実用的ではない。
- ii) 読影のワークフローに沿った自然な操作を実現すること。
- iii) 病院内の情報システムやモダリティーと連携が取れること。

本システムの動作環境を表3に示す。システム構成は1台のパソコン上でのスタンドアロンの運用や1台のレポートサーバに複数台のレポートクライアント端末の組み合わせが可能である。また、表4に音声入力読影レポートシステムの主な機能を示す。本システムの特徴の第一は、不特定な話者に対しても事前の音声登録や訓練が必要でなく、使用を始めたい時から直に音声入力ができることにある。これにより使い始めのやり易さが向上した。第二は放射線科向けの音声認識辞書を既存の10万件以上の所見データからデータマイニング技術により抽出したので、読影所見に特化した高い認識率を確保することができた。第三は、本システムはRISやPACSがなくてもモダリティーから直接撮影画像をDICOM受信することにより、患者のID番号を初めとする個人情報をそのままレポートへ反映することができる。また、DICOM受信した画像データを自動的にJPEGデータに変換することができるため、レポートへの画像張り付けやその画像へのアノテーションをシェーマとして書き込みが行えるので質の高いレポートを作成することができる。

表1：レポート作成時間内訳

操作内容	操作時間 (s)			
	Dr.A	Dr.B	Dr.C	平均
患者番号検査区分入力他	16.9	20.2	14.5	17.2
フィルム観察	23.8	40.5	27.0	30.4
テンプレート入力	7.5	7.8	5.5	6.9
所見入力、診断名入力	44.3	64	154	87
総合時間	92.5	132	201	142

表2：ユーザの意見例

項目	内容
番号の手入力	各種番号の手入力は間違いのもと
画像連携	レポートへのキー画像張り付け
GUIについて	操作しやすい画面配置と表現がほしい
所見入力	日本語変換が大変なので英語で書いてしまう

表3：動作環境

ハードウェア	CPU	Pentium <sup>®</sup> III 850MHz以上
	必要メモリ	196MB以上
	必要HD	空き20MB以上
OS環境	Windows <sup>®</sup> 98/2000/NT4.0/XP	

表4：主な機能

機能	内容
音声入力機能	所見入力、患者番号、生年月日、登録、印刷、など
各種検索機能	未読影検査の検索、患者検索、ID番号検索など
KEY画像添付機能	DICOM受信した画像を添付可能
シェーマ機能	キー画像にアノテーション付加
DICOM SR	DICOM SR形式の保存が可能

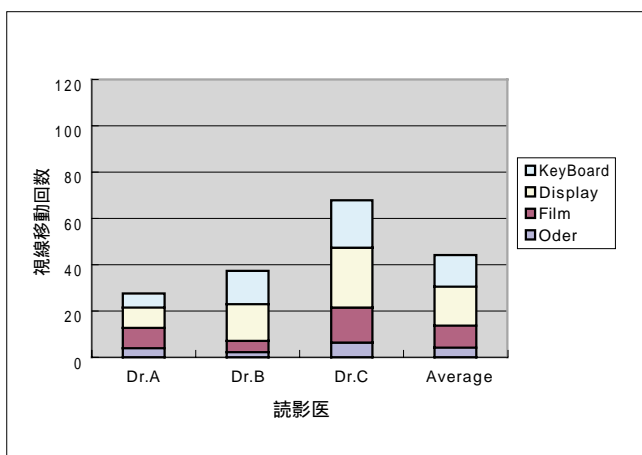


図5：レポート1件あたりの操作内訳

### (1) 放射線医向け音声認識辞書

音声入力をレポート作成で実用するには、高い認識率を確保する必要がある。そのため放射線読影レポートでは今回新たに放射線医向けの音声認識辞書を開発した。開発にあたり、複数の病院のご協力により10万件余の所見データを収集した。収集した複数の病院の膨大なデータを分析し、各病院における特殊な言葉や言いまわしを排除、あるいは標準的な医学用語への置き換えなどを行ってデータベースを完成させた<sup>2)</sup>。

辞書開発では、収集した所見データをテキストマイニングにより、単語に分解した後、所見の中に出現する単語の相関関係や出現確率を求め、これを考慮した音声認識辞書を作成した。図6にテキストマイニングによる辞書作成の概略を示す。

### (2) ワークフローに沿った音声入力

読影レポートのワークフローに沿って、従来からのキーボード入力やマウス操作に加えて音声による操作を可能とし、キータッチレスでの入力も可能とした。音声入力を使用すると、たとえばメニュー選択を行う時もメニューボックスにある多くのサイドメニューへマウスを移動しクリックする動作がなくなり、「とろろく」と言う直接の司令で操作することができるため、キーボードやマウス操作のわずらわしさから開放される。音声入力が可能なのは患者番号入力、生年月日入力、性別入力、診断名、所見、コメント、登録、印刷などの制御コマンド入力である。これら一連の操作が音声で可能になり、操作がキーボードレスで実現できるシステムである。

### (3) DICOM画像受信とDICOM SR

本読影レポートシステムはDICOM受信が可能である。従来は、患者情報を自動的にレポートへ反映させるためにはRISやPACSとの連動が必要であり、それらのシステムをもつ施設での運用しか可能ではなかったが、本システムは独自にDICOM画像の受信を行えることから、PACSのない施設においても、モダリティーから画像を送信することで患者情報をレポートに反映でき、患者情報入力の手間が省略でき、かつ正確な情報入力が可能となる。また作成したレポートはDICOM SR(supplement23)の構造化レポートとしてデータベースに登録することができ、DICOMデータとしての扱いが

可能であり、今後病院内外の情報システムとのデータ連携を行うこともDICOM準拠の仕様のもと可能となり、国際的な部門間連携を目指して行われているIHEの動きにも対応していくことが可能である。

### (4) 画像張り付け・アノテーション記入

DICOM受信機能によりモダリティーから送られて来たキー画像は自動的にJPEG変換されJPEG保存されます。この画像をキー画像としてレポートの中に張り込むことがカットアンドペーストで可能であり、かつ選択した画像には文字や線画を貼ることが可能で、より詳しいレポートが必要な場合にはこれらの機能により詳細なレポートを書くことができるため、レポートの質の向上が図れる。

図7に開発した音声入力読影レポートの所見入力画面を示す。画面の上側は主に患者情報および検査情報記入欄となっており、その下に所見入力および診断結果、コメントなどのディクテーションエリアとなっている。また、画面の右側はキー画像貼付やシェーマ機能などの入力エリアとなっている。

## 4. 操作性評価

開発した音声入力読影レポートの操作性評価を行った。評価は次の項目とした。

- ・ 所見入力の認識率
- ・ ワークフローの操作改善

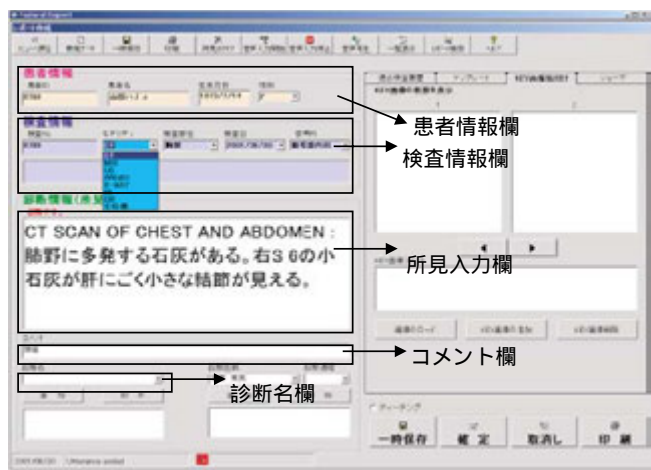


図7：所見入力画面

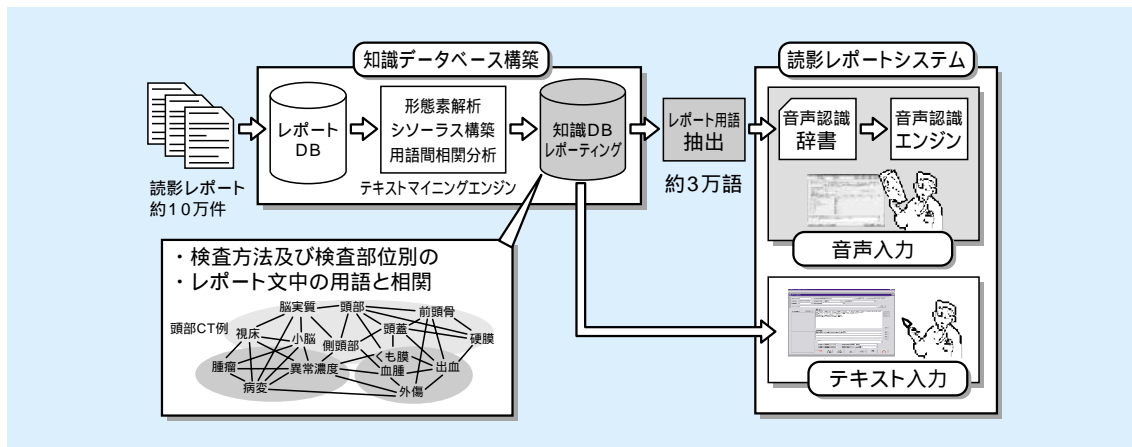


図6：テキストマイニング概略

### (1) 所見入力認識率

読影レポートの開発目的の第1である所見入力の操作効率の向上に関し、特に所見入力の認識率を評価した。評価は慈恵医大柏病院で運用中の読影レポートシステムへ音声入力システムを組み込み、実際の運用においてワークフローに沿った実運用を行い、その中での認識率を調査した。

読影室で複数の読影医が通常の業務の中で所見入力を行い、この様子および入力画面をビデオで記録し、ディクテーションの内容と文字認識の内容とを比較し認識率として調査した。図8に読影室でのレポート作成の様子を示す。今までの読影と同様に読影画像を正面のシャカステンへかけ、これを見ながら読影を行う。音声入力はヘッドセットマイクを用いて実施する。また読影は同時に複数の読影医が行う。なお周囲の雑音は通常どおりにして、特に静けさを管理しなかった。認識率の結果を図9に示す。読影所見は主にMRI、X線CT、RI、超音波診断装置の画像診断である。読影医により多少違いがあるが95%の認識率を実現し、実用に際してはほぼ問題がないことを確認することができた。認識率の違いの原因は導入まもないことから、音声入力、ディクテーション入力に対する慣れ不慣れがあり、それがそのまま認識率へ反映されていることによる。音声入力は導入当時慣れがある程度必要であるが日常使用する上で徐々に解消されていくと考えられる。



図8：音声入力の様子

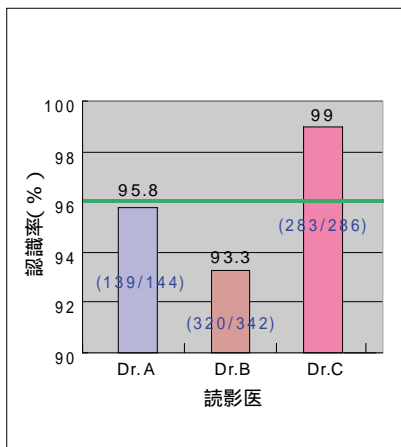


図9：読影医別認識率

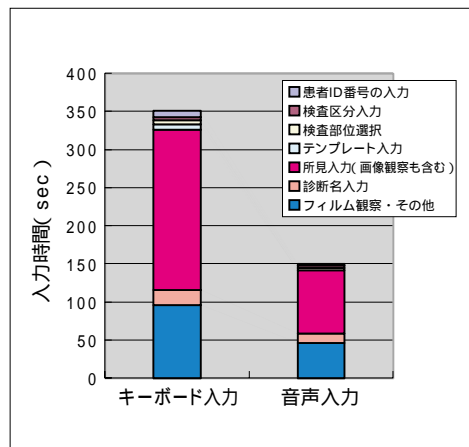


図10：所見入力時間比較

### (2) ワークフローの操作性改善

音声入力を所見入力以外の読影レポート作成ワークフローへ適用した。表5に従来システムと本システムのワークフローにおける違いを示す。本システムでは、キーボードを使用せず声だけで入力および操作を行うことができる。そのため声による操作を行いながら別の作業を行うことができ、作業の効率も上がっている。図10に所見を作成するに当たっての従来と本システムにおける作成時間を比較したグラフを示す。本レポートでは作成時間が半減していることがわかる。

また、実際に運用している医師からは手や腕の疲れ、肩こりなど肉体的な負担が軽減したという効果も聞いている。

## 5. まとめ

音声認識読影レポートシステム「Natural-Report™」を開発し、実際の読影現場へ持ち込み、臨床での有用性を評価した。

今回開発した音声入力は話者毎のトレーニングが不要であり、また放射線医向け音声認識辞書を開発搭載したことにより、所見入力認識率95%を確保し実用的な利用が可能となった。またDICOM受信機能を持ちPACSのない環境でも直接画像診断装置から画像や患者情報を取り込めるので、作成効率の向上と入力間違いなどのリスクをほとんどなくすることができるので、施設での幅広い運用を提供することができる。

## 謝辞

本読影レポートの開発にあたり、東京慈恵会医科大学柏病院 原田助教授、東北大学大学院医学研究科 石橋助教授、湘南鎌倉総合病院 安田先生、世田谷中央病院 入江先生ほかのご協力を頂きましたことを感謝いたします。

- 1 Pentiumは米国Intel Corp.の登録商標です。
- 2 Windowsは米国Microsoft Corp.の登録商標です。

## 参考文献

- 1) 黒須正明, ほか: ユーザ工学入門, 共立出版, 1999.
- 2) 平松慶博 編: 実用画像レポートの書き方. 123-166, メディカルレビュー社, 1998.

表5：読影操作フロー比較

	ワークフロー	従来システム	本システム
1	患者番号入力	キーボード入力	音声入力
2	患者名入力	キーボード入力	音声入力
3	生年月日入力	キーボード入力	音声入力
4	性別入力	キーボード入力	音声入力
5	検索	キーボード入力	音声操作
6	検査番号入力	キーボード入力	音声入力
7	モダリティ入力	キーボード入力	音声入力
8	検査部位入力	キーボード入力	音声入力
9	所見入力	キーボード入力	音声入力
10	診断名入力	キーボード入力	音声入力
11	保存	マウス操作	音声操作
12	印刷	マウス操作	音声操作