

平成24年10月31日判決言渡 同日原本領収 裁判所書記官

平成23年(行ケ)第10427号 審決取消請求事件

口頭弁論終結日 平成24年10月10日

判 決

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 原 | 告 | ア | バ | イ | ア | イ | ン | コ | ー | ポ | レ | ー | テ | ッ | ド |
| 同 | 訴 | 訟 | 代 | 理 | 人 | 弁 | 理 | 士 | 岡 | 部 | 讓 |   |   |   |   |
|   |   |   |   |   |   | 吉 | 澤 | 弘 | 司 |   |   |   |   |   |   |
|   |   |   |   |   |   | 三 | 山 | 勝 | 巳 |   |   |   |   |   |   |
|   |   |   |   |   |   | 濱 | 口 | 岳 | 久 |   |   |   |   |   |   |
| 被 | 告 | 特 | 許 | 庁 | 長 | 官 |   |   |   |   |   |   |   |   |   |
| 同 | 指 | 定 | 代 | 理 | 人 | 新 | 川 | 圭 | 二 |   |   |   |   |   |   |
|   |   |   |   |   |   | 竹 | 井 | 文 | 雄 |   |   |   |   |   |   |
|   |   |   |   |   |   | 田 | 部 | 元 | 史 |   |   |   |   |   |   |
|   |   |   |   |   |   | 守 | 屋 | 友 | 宏 |   |   |   |   |   |   |

主 文

- 1 原告の請求を棄却する。
- 2 訴訟費用は原告の負担とする。
- 3 この判決に対する上告及び上告受理の申立てのための付加期間を30日と定める。

事実及び理由

第1 請求

特許庁が不服2009-23761号事件について平成23年8月9日にした審決を取り消す。

第2 事案の概要

本件は、原告が、後記1のとおりの手続において、特許請求の範囲の記載を後記2とする本件出願に対する拒絶査定不服審判の請求について、特許庁が同請求は成

り立たないとした別紙審決書（写し）の本件審決（その理由の要旨は後記3のとおり）には、後記4の取消事由があると主張して、その取消しを求める事案である。

### 1 特許庁における手続の経緯

(1) 原告は、平成17年6月23日、発明の名称を「視覚的な監視チャンネルを有する対話式音声処理のための方法および装置」とする特許を出願した（特願2005-183062。パリ条約による優先権主張日：平成16年（2004年）6月23日（アメリカ合衆国）。請求項の数10）。

特許庁は、平成21年7月28日付けで拒絶査定をしたため、原告は、同年12月3日、これに対する不服の審判を請求した。

(2) 特許庁は、これを不服2009-23761号事件として審理し、平成23年8月9日、「本件審判の請求は、成り立たない。」との本件審決をし、その審決謄本は、同月22日、原告に送達された。

### 2 特許請求の範囲の記載

本件審決が判断の対象とした特許請求の範囲の請求項1の記載（平成21年7月6日付け手続補正書（甲6）による補正後のもの）は、以下のとおりである。なお、「／」は、原文における改行箇所を示す（以下、特許請求の範囲の請求項1に記載された発明を「本願発明」といい、本願発明に係る明細書（甲5, 6）を、図面を含めて「本件明細書」という。）。

ユーザとIVR（対話式音声応答）システムの間での対話を監視する方法であって、  
／複数のコマンドを有するIVRスクリプトに従って前記ユーザからの音声通信を  
処理するステップと、  
／前記IVRスクリプトに基づいてエージェントに前記音声通信の視覚表示を提示するステップと、  
を含み、  
／前記視覚表示が、前記音声通信と実質上同時に提示されかつ前記ユーザから取得された情報を取り込むために少なくとも1つのフィールドを有し、  
／前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される方法

### 3 本件審決の理由の要旨

(1) 本件審決の理由は、要するに、本願発明は、後記アの引用例に記載された発明及び後記イないしエの周知例 1 ないし 3 に記載された周知技術に基づいて、当業者が容易に発明をすることができたものであるから、特許法 29 条 2 項の規定により、特許を受けることができない、というものである。

ア 引用例：特開平 10-322450 号公報（甲 1）

イ 周知例 1：特開 2002-215670 号公報（甲 2）

ウ 周知例 2：特開 2003-5778 号公報（甲 3）

エ 周知例 3：特開 2004-32742 号公報（甲 4）

(2) 本件審決が認定した引用例に記載された発明（以下「引用発明」という。）並びに本願発明と引用発明との一致点及び相違点は、次のとおりである。

ア 引用発明：利用者と対話型音声認識システムの間での対話における音声認識方法であって、前記利用者からの音声の音声認識ステップと、オペレータに前記音声認識結果の表示を提示するステップと、を含み、前記表示が、前記音声と実質上同時に提示されかつ前記利用者の発声した音声の認識結果を表示するために少なくとも 1 つの入力欄を有し、前記入力欄内に表示された前記認識結果が前記オペレータによって訂正される方法

イ 一致点：ユーザと I V R（対話式音声応答）システムの間での方法であって、前記ユーザからの音声通信を処理するステップと、エージェントに前記音声通信の視覚表示を提示するステップと、を含み、前記視覚表示が、前記音声通信と実質上同時に提示されかつ前記ユーザから取得された情報を取り込むために少なくとも 1 つのフィールドを有し、前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される方法

ウ 相違点 1：「方法」に関し、本願発明が「対話を監視する方法」であるのに対し、引用発明は「対話における音声認識方法」である点

エ 相違点 2：「音声通信を処理するステップ」に関し、本願発明では「複数のコマンドを有する I V R スクリプトに従って」と限定しているのに対し、引用発明

ではそのような限定がない点

オ 相違点3：「音声通信の視覚表示を提示するステップ」に関し、本願発明では「前記IVRスクリプトに基づいて」と限定しているのに対し、引用発明ではそのような限定がない点

#### 4 取消事由

本願発明の容易想到性に係る判断の誤り

- (1) 引用発明の認定の誤り
- (2) 一致点及び相違点の認定の誤り
- (3) 相違点1に係る判断の誤り

### 第3 当事者の主張

〔原告の主張〕

- (1) 引用発明の認定の誤りについて

ア 引用例の【0018】【図3】によると、引用例において、「音声認識システム」及び「オペレータ」は、別個の構成要素として開示されているものである。

また、引用例の【0022】【0023】によると、引用例に記載された発明において、音声認識システムが利用者から「発呼」を受け、音声ガイダンスを利用者に提供した後、利用者はオペレータに向かって発声を行い、オペレータは利用者に向かって復唱するものである。その際、利用者とオペレータとの間の音声の一連のやりとりにおいて、音声認識システムは、利用者にガイダンスを提供し、利用者又はオペレータからの発声をオペレータ又は利用者に受け渡す処理を行うにすぎない。

「対話」（向かい合って話し合うこと。また、その話）及び「ガイダンス」（不慣れで事情の分からない者に対して、初歩的な説明をすること。案内。手引き）の一般的な意味からすると、「ガイダンス」は「対話」を構成するものではないから、音声認識システムは、ガイダンスによってその対話の導入案内をするにすぎない。

また、引用例に記載された発明の解決課題や効果からしても、同発明は、オペレータが発話し、利用者と対話することを前提とするものである。

したがって、引用例に記載された発明において、「利用者」が「対話」する相手は「オペレータ」であり、利用者は音声認識システム自体と「対話」するものではなく、同発明の音声認識システムは、「対話型音声認識システム」ではないから、本件審決が、引用例に記載された発明について、「利用者と対話型音声認識システムの間での対話における音声認識方法であって」と認定した点は、誤りである。

イ 本件審決は、引用例の【0027】ないし【0030】によると、引用例に記載された発明において、「オペレータが発声した音声の認識結果により上記入力欄B1の認識結果は訂正されるものとするが、引用例【0027】【0030】は、「主制御部」により「入力欄B1」の認識結果が訂正されるものと開示しているのであるから、正しくは、「オペレータが発声した音声の認識結果に基づいて主制御部により上記入力欄B1の認識結果は訂正されるもの」というべきである。

ウ 以上によれば、本件審決の引用発明の認定は、誤りである。

(2) 一致点及び相違点の認定の誤りについて

ア 一致点について

(ア) 本件審決は、引用例に記載された発明の「利用者」「オペレータ」が、それぞれ本願発明の「ユーザ」「エージェント」に相当することを前提として、本願発明と引用例に記載された発明とが、「エージェントに前記音声通信の視覚表示を提示するステップ」を有する点で一致すると認定した。

しかしながら、「オペレータ」(機械を)操作・運転する人。計算機類の操作者、無線通信士・電話交換手など)及び「エージェント」(ユーザの連続した操作を必要とせず、自律的に情報収集や状況判断を行い、適切な処理動作を実行できる機能)の一般的な意味からすると、「エージェント」は、自立的に状況判断を行う機能を有するものであり、単に機械を操作する「オペレータ」よりも狭いか、あるいはそれとは異なる概念の構成要素を示す用語である。また、引用例に記載された発明の「オペレータ」は、全ての対話について、その当事者として「常に」関与する構成要素であるのに対し、本願明細書【0008】において、「このようにして、

エージェントは、発呼者とのIVRスクリプトの対話を監視し、必要な場合、介入することができる。」と記載されているように、「エージェント」は、「対話」における第三者として適時に関与する構成要素である。

したがって、引用例に記載された発明の「オペレータ」と、本願発明の「エージェント」とは、文言上の意味及び処理の役割が異なり、相違する構成要素であって、「オペレータ」が「エージェント」に相当するとの本件審決の認定は誤りである。

そうすると、「エージェントに前記音声通信の視覚表示を提示するステップ」を有する点についても、本願発明と引用例に記載された発明の一致点であるとした本件審決の認定は誤りである。

(イ) 引用例に記載された発明の「オペレータ」は、本願発明の「エージェント」に相当しないことからすると、引用例に記載された発明において、「フィールド内に取り込まれた情報」は、「音声処理システムの主制御部」によって自動的に更新されるものであって、「オペレータ」によって更新されるものではない。

また、引用例に記載された発明は、オペレータの作業の負荷を軽減することができる音声認識システム、コールセンタシステム、音声認識方法及び記録媒体を提供することを課題とするものであるから、オペレータが情報を逐一訂正するような構成は、オペレータの負担をむしろ増加させるものである。

したがって、引用例に記載された発明において、フィールド内に取り込まれた情報は「オペレータ」によっても更新されるものではないから、「前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される」点についても、本願発明と引用例に記載された発明の一致点であるとした本件審決の認定は誤りである。

(ウ) 以上によれば、本件審決の一致点の認定は、誤りである。

イ 相違点の認定の誤りについて

前記のとおり、本件審決の一致点の認定は誤りであって、本願発明と引用例に記載された発明との相違点としては、相違点1ないし3のほか、以下の相違点4及び

5についても認定されるべきである。

(ア) 相違点4：本願発明は「エージェントに前記音声通信の視覚表示を提示するステップ」を有するのに対し、引用例に記載された発明は同事項を有さない点

(イ) 相違点5：本願発明は「前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される」構成を有するのに対し、引用例に記載された発明は同事項を有さない点

(3) 相違点1に係る判断の誤りについて

ア 本件審決は、引用例に記載された発明の音声認識方法では、音声認識システムとの対話における利用者の音声の認識結果の精度は常に比較判断されているから、対話が「監視されている」と表現することができるとして、同発明の「対話における音声認識方法」と本願発明の「対話を監視する方法」との間に実質的な差異があるとは認められないとした。

しかしながら、引用例に記載された発明において、「対話」は利用者と音声認識システム自体との間ではなく、利用者とオペレータとの間に存在するから、「利用者とオペレータとの間の対話における音声認識方法」と「ユーザとIVR（対話式音声応答）システムとの間の対話を監視する方法」とでは、対話を実行する構成要素が異なるものである。

また、本願発明は、対話の当事者であるユーザとIVRシステムではなく、対話における第三者であるエージェントが対話を監視するのに対し、引用例に記載された発明では、対話の当事者は利用者とオペレータであり、対話における第三者である音声認識システムが音声認識を行うものであるから、本願発明の「対話を監視する方法」と引用例に記載された発明の「対話における音声認識方法」とでは、対話に対して、監視のための処理又は音声認識の処理を施す主体も異なるものである。

仮に、引用例に記載された発明における対話が利用者と音声認識システムとの間に成立していたとしても、本願発明は対話における第三者であるエージェントが対話を監視するのに対して、引用例に記載された発明は対話の当事者である音声認識



システムが対話を音声認識するものであるから、やはり、本願発明の「対話を監視する方法」と引用例に記載された発明の「対話における音声認識方法」とでは、対話に対して監視のための処理又は音声認識の処理を施す主体が異なるものといえることができる。

さらに、「監視（する）」（警戒して見張ること、不都合な事の起こらぬように見張ること）及び「音声認識」（人間の声などを、コンピューターに認識させること。音声の特徴から発声者を識別したり、話し言葉を文字列に変換したりする機能。コンピューターを用いて音声による指示を識別、判断すること。また、それにより電子機器を操作・制御すること）の一般的な意味からすると、「監視する方法」と「音声認識方法」とは全く異なる概念の方法といえることができるから、これらを同一視することはできない。

したがって、本願発明と引用発明とでは、「対話」の当事者となる構成要素、「対話」に処理を施す主体及び「対話」に対して施される処理内容のいずれも異なるものであるから、本件審決が、引用例に記載された発明の「対話における音声認識方法」と本願発明の「対話を監視する方法」との間に実質的な差異はないとした判断は誤りである。

イ 以上によれば、本件審決の相違点 1 に係る判断は誤りである。

#### (4) 小括

本件審決は、以上のとおり、引用例に記載された発明の認定並びに本願発明と引用例に記載された発明との一致点及び相違点の認定を誤り、相違点 1 に係る判断を誤ったものといわざるを得ず、したがって、本願発明は、引用発明及び周知技術に基づいて、当業者が容易に発明をすることができたものといえることはできない。

〔被告の主張〕

#### (1) 引用発明の認定の誤りについて

ア 引用例の【0022】【0034】【0036】【0039】によると、利用者は、引用発明の音声認識システムが提供するガイダンスによる質問若しくは要



求に応答する形で、返事を発声しており、音声認識システムと利用者との間で実質的に意味のある対話が行われていることは明らかである。

また、引用例の【0023】【0035】【0037】【0041】は、上記対話における利用者の発声に対するオペレータの発声の具体例について開示するが、これらは、いずれも「オペレータは」「復唱する」ものとされており、オペレータは利用者の返事をほぼオウム返しに繰り返しているにすぎない。このような、相手の返事を繰り返すだけの音声のやり取りが実質的に意味のある対話を構成するものではない。

イ したがって、引用発明が、利用者と対話型音声認識システムとの間の対話における音声認識方法であるとする本件審決の認定に誤りはない。

(2) 一致点及び相違点の認定の誤りについて

ア 一致点について

(ア) コールセンタシステムの技術分野において、「エージェント」という語句が、引用発明の「オペレータ」と同様、顧客サービス要員、業務代行者等のように、ユーザからの呼に対応する「人」の意味で用いられることは周知である（乙1～4）。

本願発明における「エージェント」も、「聞取りのみ」の音声回線を任意に選択することが可能であること、ドロップダウン・リストからターゲットを選択する必要があるだけであることなどからすると、引用発明の「オペレータ」と同様、ユーザからの呼に対応する「人」であることは明らかである。

したがって、引用発明の「オペレータ」が本願発明の「エージェント」に相当するとした本件審決の認定に誤りはないから、「エージェントに前記音声通信の視覚表示を提示するステップ」を有する点についても一致点であるとした本件審決の認定も、同様に、誤りはないというべきである。

(イ) 引用発明の「オペレータ」が本願発明の「エージェント」に相当し、引用発明において、入力欄内に表示された認識結果がオペレータによって訂正される以上、「前記フィールド内に取り込まれた前記情報が前記エージェントによって更新

される」点についても一致点とした本件審決の認定に誤りはない。

イ 相違点について

前記のとおり、本件審決の一致点の認定に誤りがない以上、相違点の認定についても、同様に、誤りはない。原告が主張する相違点4及び5は、存在しない。

(3) 相違点1に係る判断の誤りについて

ア 引用発明において、「対話」は利用者と音声認識システムとの間で行われているから、ユーザとIVRシステムとの間で「対話」を行う本願発明との間で、「対話」を行う構成要素については相違しない。

ところで、本願発明の「対話を監視する方法」において、エージェントが「更新」を行うためには、その前に、更新されるべき情報がIVRシステムによってエージェントに提示されなければならないことは明らかであるから、対話に対する処理は、「フィールド内に取り込まれた前記情報が前記エージェントによって更新される」ことのみならず、IVRシステムによって「エージェントに前記音声通信の視覚表示を提示する」ことをも含むというべきである。

したがって、本願発明において、「対話」に対して処理を施す主体は、「IVRシステム」及び「エージェント」であるというべきである。

一方、引用発明の「対話における音声認識方法」においては、表示された音声認識結果がオペレータによって「訂正」されることにより、より精度の高い音声認識を実現することができるものであるから、オペレータによる「訂正」も、音声認識の1つのステップであるというべきである。そして、対話に対する処理は、対話型音声認識システムが「オペレータに前記音声認識結果の表示を提示する」ことや「入力欄内に表示された前記認識結果が前記オペレータによって訂正される」ことをも含むものというべきであるから、このような対話に対する処理は、対話が監視されていると表現できるものである。

したがって、引用発明において、「対話」に対して処理を施す主体は、「対話型音声認識システム」及び「オペレータ」であるということができ、本願発明及び引

用発明において、「対話」に処理を施す主体は相違するものではない。

仮に、引用発明における「処理」が、「対話」を行う構成要素以外、すなわち、「対話」における第三者からの処理のみを意味するものであるとしても、当該「処理」は、本願発明では、「前記フィールド内に取り込まれた前記情報」の「更新」に相当するものであって、これを施す主体は「エージェント」であるのに対し、引用発明においては、「前記入力欄内に表示された前記認識結果」の「訂正」であって、これを施す主体は「オペレータ」であるから、本願発明と引用発明とにおいて、「対話」に処理を施す主体に何ら相違はない。また、本願発明における「監視」「更新」と引用発明における「音声認識」「訂正」との文言上の相違が存在するとしても、両発明における「対話」に対して施す処理内容の間に何ら相違はない。

イ 以上によれば、本件審決の相違点1に係る判断に誤りはない。

#### (4) 小括

本件審決の引用発明の認定、本願発明と引用発明との一致点及び相違点の認定並びに相違点1に係る判断には、以上のとおり、誤りはなく、したがって、本願発明は、引用発明及び周知技術に基づいて、当業者が容易に発明をすることができたものというべきである。

### 第4 当裁判所の判断

#### 1 本願発明について

本願発明の特許請求の範囲は、前記第2の2に記載のとおりであるところ、本件明細書（甲5，6）には、おおむね次の記載がある。

##### (1) 技術分野

本願発明は、コールセンタ又は他のコール処理システムに関し、発呼者とコールセンタにおける対話式音声応答システムとの間の対話を監視するための方法及び装置に関する発明である（【0001】）。

##### (2) 背景技術

顧客と情報を交換するためのコールセンタでは、IVR（interactive voice

response：対話式音声応答）システムを使用することが多く，記録されたメッセージの形で発呼者に情報を提供し，記録された照会に対する発呼者からのキーパッド又は音声応答により情報を取得する（【0002】）。

I V Rシステムには，エラーを訂正するための機構であるデバッグ・ツールを有するものもあるが，デバック・ツールは，通常通話の進行に応じて実時間で変更を加えることはできない。また，現在利用可能なデバッグ・ツールは，I V Rシステムと発呼者との間の対話を中断し，変更し，訂正するトランザクションをエージェントが監視することはできない（【0004】）。

### (3) 発明が解決しようとする課題

I V Rシステムにおいて，システムを監視又はデバッグするための改良技術，実行中のI V Rアプリケーションの視覚的な解釈を提供するI V Rシステムへの視覚的なインターフェースを提供することが求められている（【0005】）。

### (4) 課題を解決するための手段

本願発明は，I V Rシステムに視覚的なインターフェースを提供するものである。これにより，動作中のI V Rアプリケーションの視覚的な解釈を提供し，発呼者などのユーザとI V Rシステムとの間の対話が監視可能となる。発呼者からの電話通話などの音声通信は，複数のコマンドを有するI V Rスクリプトに従って従来方法で処理される。本願発明は，I V Rスクリプトに基づいて，その音声通信の視覚的な表示をエージェントに提示する（【0006】）。

発呼者がI V Rシステムで会話すると，視覚表示中の1つ又は複数のフィールドが発呼者の発話を取り込むことができる。エージェントは，発話を取り込まれた視覚表示中のフィールドを任意選択で再検討又は更新することができる。この監視機能により，エージェントはI V Rスクリプトのフローを変更し，又は音声通信に介入することが任意選択で可能になる（【0007】）。

### (5) 発明を実施するための最良の形態

本願発明のエージェント端末は，ワークステーション及び電話を備えており，I

VRユニット，発呼者装置及びエージェント端末との間で対話が行われる（【0012】【図1】【図2】）。

発呼者は，VXMLアプリケーションを実行しているIVRユニットと対話する。エージェントは，音声チャンネルによって発呼者とVXMLアプリケーションとの間の対話を聞くことができる。さらに，エージェントは，VXMLアプリケーションの状態を示す視覚チャンネルによって発呼者とVXMLアプリケーションとの間の対話を見ることができる。視覚チャンネルは，例えば，ウェブ・ページとして表示することができる（【0014】【図7】）。

エージェント・ワークステーションからウェブ・サーバに，現在のIVRアプリケーションの視覚的なバージョンを求める要求が到着すると，IVRアプリケーション・エンジンは，適切なページを生成する。そのページは，標準のウェブ接続を介してエージェント・ワークステーションのウェブ・ブラウザに送られる。エージェント・ワークステーションにより応答が入力されると，その結果がウェブ・サーバにポストされ，IVRアプリケーション・エンジンに通知され，適切な措置を講ずる。エージェントはまた，エージェントのヘッドホンと，発呼者及びIVRの音声チャンネルとの間に「聞取りのみ」の音声回線を任意選択で有する。このようにすると，エージェントは，IVRダイアログの進捗と方向について決定を行い，必要な場合，訂正を行うことができる（【0015】【図2】）。

エージェント・ワークステーションのエージェントに，本願発明の視覚的なチャンネルを提示するためのインターフェースの例としては，発呼者から得られる情報が取り込まれる1つ又は複数のフィールドを含むインターフェースを挙げるることができる（【0020】【図4】）。

本願発明の実行中，エージェントは，発呼者とIVRの間の音声対話をスイッチ観察機能を介して監視することができる。あるフィールドで発呼者からの入力が終わった場合，VXMLインタプリタは，その入力データをエージェントのディスプレイに送り，エージェントがその入力を指定変更するのを少しの間待つ。エージェ

ントは、例えば、代替のテキストを適切な視覚的なダイアログ入力項目に入力することによって指定変更することができる。所与のフィールドに対応する承認ボタンをクリックすることにより、フィールドのコンテンツをVXMLインタプリタに対して確定し、次いで、通常のVXMLフォーム解釈に従って入力処理に進む。エージェントが入力フィールドを変更せず、承認ボタンを押す場合、元のコンテンツがVXMLインタプリタで使用される。フォームのサブミットボタンにより、VXMLインタプリタに対してフォーム全体のコンテンツがサブミットされ、次いで、通常のVXMLフォーム解釈に従って入力処理に進む（【0021】）。

本願発明は、通話フロー問題の理解を助けるために、通話中に何がダイアログで行われているかをエージェントにわかるようにするVXMLアプリケーション・デバッガを提供する。アプリケーションの配置中に、視覚チャンネルは、ASR文法が不完全又は指示メッセージが間違っていて指示しているという問題点を突き止めやすくすることができる。発呼者との実時間の対話において、自動ダイアログの実行時にエラーを繰り返し受けた発呼者に対してエージェントが聴取できるようにすることにより、視覚チャンネルは、継続が危うくなった自動セッションを支援することができる。エージェントは、ASRの誤りを訂正し、又はその通話を取り上げることができる。さらに、本願発明の視覚チャンネルは、エージェントのシャドウにより音声認識を正確にすることができる（【0031】）。

## 2 引用発明の認定の誤りについて

### (1) 引用例の記載

引用例（甲1）には、おおむね次の記載がある。

#### ア 発明の属する技術分野

引用例に記載された発明は、利用者の音声を認識する対話型の音声認識システムにおいて、オペレータの負担を軽減することができるコールセンタシステムに関する発明である（【0001】）。

#### イ 発明が解決しようとする課題



コールセンタシステムにおいて、オペレータが復唱した音声を認識するシステムの場合、利用者が正しく発声してもオペレータが間違って復唱した場合、発声を繰り返す必要が生じる（【0003】）。

引用例に記載された発明は、オペレータの作業の負荷を軽減することができる音声認識システム、コールセンタシステム、音声認識方法及び記録媒体を提供することを目的とする（【0004】）。

#### ウ 課題を解決するための手段

引用例に記載された発明は、利用者からの電話をオペレータが受け付けるコールセンタシステムにおいて、利用者の音声を受信する受信手段と、受信された利用者の音声を分析し、第1の認識結果と当該認識結果に対する第1の認識確率とを生成する第1の認識手段と、第1の認識手段により生成された第1の認識結果をオペレータに提示する手段と、利用者の音声をオペレータに供給する手段と、オペレータの音声を入力する入力手段と、入力されたオペレータの音声を分析し、第2の認識結果と当該認識結果に対する第2の認識確率とを生成する第2の認識手段と、第1と第2の認識手段により生成された第1と第2の認識結果が一致するか否かを判別する手段と、第1と第2の認識結果が一致すると判別された場合、第1と第2の認識結果を最終的な認識結果として決定し、一致しないと判別された場合、第1と第2の認識確率を比較して最終的な認識結果を選択する手段とを備える（【0011】）。

引用例に記載された発明は、利用者の音声とオペレータの音声との双方の認識処理を行い、それらの認識結果が異なる場合、より認識確率の高い方を最終的な認識結果として採用するため、より認識精度の高いコールセンタシステムを実現することができる。また、オペレータは、表示された認識結果を参照しながら利用者の用件を復唱することができるため、オペレータの作業の負荷を削減することができる（【0012】）。

#### エ 発明の実施の形態



(ア) 引用例に記載された発明は、商品に関する問い合わせ、修理依頼等を電話で受け付けるサポートセンタに設置された音声認識システム等において実施することができる（【0017】）。

引用例に記載された発明は、回線制御部、オペレータ操作部、音声認識部、主制御部及びデータベースを有する。回線制御部は、利用者の電話からの音声を公衆回線を介して受信し、音声認識部に転送する。また、回線制御部は、回線の接続・切断等も行う。オペレータ操作部は、マイク、キーボード、スピーカ、表示装置等を有する（【0018】【図1】）。

音声認識部は、音声データをA/D変換部でデジタルデータに変換し、認識処理を行う（【0019】）。

主制御部は、音声認識部からの利用者の音声とオペレータの音声の認識結果の比較処理を行い、また、この音声認識システム全体を制御する（【0020】）。

引用例に記載された発明は、利用者の音声の認識結果を暫定的な認識結果としてオペレータに提示し、オペレータによる人為的ミスを削減するとともに、利用者の音声の認識結果とオペレータの音声の認識結果とを比較することにより、その認識精度を高めることができる音声認識システムである（【0021】）。

(イ) 引用例に記載された発明では、まず、音声認識システムの回線制御部が、公衆回線を介して利用者からの電話（発呼）を受信する。これに応答して、主制御部は、「どのような御用件ですか？」等の用件の入力（発声）を促す旨のガイダンスを送信する。利用者は、これに応答して、「故障修理をお願いします。」等の用件を発声する。主制御部は、この音声の受信に応答して、オペレータ操作部のスピーカを介してオペレータに利用者の音声を供給するとともに音声認識部に受信した音声の認識を指示する（【0022】）。

音声認識部は、この指示に従い、利用者の音声を認識し、認識結果とその尤度を主制御部に送信する。主制御部は、認識結果に対応する画面（例えば、認識結果が「故障修理」の場合、故障修理受付画面）をオペレータ操作部の表示装置に表示す

る。オペレータは、「故障修理ですね。」のように、利用者からの音声を示す事項を復唱する。オペレータは、表示画面より認識結果が正しいと判断した場合、その表示画面（例えば、故障修理受付画面）を参照しながら復唱することができる。オペレータの音声は、オペレータ操作部のマイクを介して主制御部に送られる（【0023】）。

主制御部は、回線制御部と公衆回線を介してオペレータの音声を利用者に供給するとともに音声認識部に音声の認識を指示する。音声認識部は、この指示に従い、オペレータの音声を認識し、認識結果とその尤度を主制御部に送信する（【0024】）。

主制御部は、利用者の音声の認識結果及び尤度と、オペレータの音声の認識結果及び尤度とを比較し、最終的な認識結果を生成する（【0025】）。

(ウ) 主制御部は、利用者の音声の認識結果とオペレータの音声の認識結果とが一致するか否かを判別する。一致すると判別した場合、最終的な認識結果として、初めに認識した利用者の音声の認識結果を選択し、処理を終了する。一致しないと判別した場合、利用者の音声の認識確度の方が高いと、最終的な認識結果として、利用者の音声の認識結果を選択し、処理を終了する。また、オペレータの音声の認識確度の方が高いと、最終的な認識結果として、オペレータの音声の認識結果を選択する（【0026】）。

上記比較処理において、最終的な認識結果として、オペレータの認識結果を選択した場合、主制御部は、オペレータ操作部の表示装置に表示されている画面を、オペレータの音声の認識結果に対応する画面に変更し、初めに認識された利用者の音声の認識結果を訂正する（【0027】）。

このようにして、利用者とオペレータの双方の音声を認識し、それらの認識結果を比較し、認識精度の高い方を選択することにより、より正確な認識結果を取得することができる。また、初めに認識した利用者の音声の認識結果を暫定的な認識結果としてオペレータに提示することにより、オペレータがその認識結果を参照して

利用者の音声の内容を復唱することができる。これにより、聞き間違い等の人為的ミスを防ぎ、オペレータの作業の負荷を削減することができる（【0028】）。

(エ) 1回の用件において、利用者からの入力（発声）が複数回必要な場合、前記の処理が繰り返される。例えば、利用者からの用件が「故障修理」の要求であった場合、主制御部は、オペレータ操作部の表示装置に複数の入力項目を備える画面を表示し、製品名を尋ねる旨のガイダンスを利用者に送信する。これに応答して、利用者は製品名を発声する。主制御部は、この音声の受信に応答して、オペレータに利用者の音声を供給するとともに音声認識部に受信した音声の認識を指示する。音声認識部は、主制御部からの指示を受けて、この製品名の認識処理を実行し、認識結果と尤度を主制御部に渡す（【0029】【図5】）。

主制御部は、画面の製品名の入力欄に認識結果を表示する。オペレータは、スピーカからの利用者の音声（製品名）を復唱する。この際、オペレータは、表示された認識結果を正しいと判断した場合、画面の製品名の入力欄に表示された製品名を参照しながら復唱することができる。主制御部は、復唱された製品名を利用者に供給するとともに音声認識部に音声の認識を指示する。主制御部は、2つの認識結果に対して比較処理を行い、最終的な認識結果を決定する。ここで、2つの認識結果が異なり、最終的な認識結果をオペレータが発声した製品名の認識結果とした場合、表示画面の製品名の入力欄の認識結果を訂正する（【0030】【図5】）。

## (2) 本件審決の引用発明の認定の当否

本件審決における引用発明の認定のうち、引用例に記載された発明が「利用者と対話型音声認識システムの間での対話における音声認識方法」として認定した点及び「前記入力欄内に表示された前記認識結果が前記オペレータによって訂正される」として認定した点を除くその余の認定については、当事者間に争いが無い。

ア 「利用者と対話型音声認識システムの間での対話における音声認識方法」として認定した点について

(ア) 前記(1)によると、引用例に記載された発明の音声認識システムは、コール

センタシステムで音声認識システムから入力（発声）を促す旨のガイダンスを利用者に送信し、このガイダンスに応答して利用者が発声するものである。

また、前記(1)によると、引用例に記載された発明の音声認識システムは、オペレータ操作部、音声認識部及び主制御部等を有し、オペレータ操作部は、マイク、スピーカ、表示装置等を備え、オペレータは、オペレータ操作部の表示装置に表示された利用者の音声認識結果を参照するとともに、オペレータ操作部のスピーカからの利用者の音声を復唱するものである。

そして、引用例に記載された発明の音声認識システムにおいて、オペレータが発声した音声は音声認識部により認識され、主制御部により、利用者の音声認識結果とオペレータの音声認識結果が比較されて、認識確度の高い方が選択され、オペレータの音声認識結果が最終的な認識結果とされた場合、表示画面の入力欄に表示された初めに認識した利用者の音声認識結果は、オペレータの音声認識結果により訂正されるものである。このように、初めに認識した利用者の音声の認識結果を暫定的な認識結果としてオペレータに提示することにより、オペレータがその認識結果を参照して利用者の音声の内容を復唱することができ、聞き間違い等の人為的ミスを防ぎ、オペレータの作業の負荷を削減するという課題を解決するものである。

(イ) 前記(ア)のとおり、引用例に記載された発明は、コールセンタシステムにおいて、音声認識システムから入力（発声）を促す旨のガイダンスを利用者に送信し、このガイダンスに応答して利用者が発声するものである。

そうすると、引用例に記載された発明の音声認識システムは、本願発明における「IVR（対話式音声応答）システム」と同様、「利用者」と「対話」するものであるというべきであるから、引用例（【0001】）に記載されているとおり、「対話型の音声認識システム」であると認められる。

(ウ) この点について、原告は、引用例に記載された発明において、利用者が対話する相手はオペレータであり、利用者は音声認識システム自体と対話するものではなく、対話型音声認識システムであるとはいえないなどと主張する。

しかしながら、前記1(2)によると、本願発明が前提とするIVRシステムとは、「記録されたメッセージの形で発呼者に情報を提供し、記録された照会に対する発呼者からのキーボードまたは音声応答により情報を取得する」システム、すなわち、利用者との対話するシステムを意味するものであるところ、引用例に記載された発明も、コールセンタシステムにおいて、音声認識システムから入力（発声）を促す旨のガイダンスを利用者に送信し、このガイダンスに応答して利用者が発声するものであるから、利用者との対話するものということができる。

したがって、原告の前記主張は採用できない。

(エ) よって、本件審決が、引用例に記載された発明が「利用者との対話型音声認識システムの間での対話における音声認識方法」とであると認定した点に誤りはない。

イ 「前記入力欄内に表示された前記認識結果が前記オペレータによって訂正される」と認定した点について

(ア) 前記ア(ア)のとおり、引用例に記載された発明の音声認識システムにおいて、主制御部により、利用者の音声認識結果とオペレータの音声認識結果とが比較され、オペレータの音声認識結果が最終的な認識結果とされた場合、表示画面の入力欄に表示された初めに認識した利用者の音声認識結果は、オペレータの音声認識結果により訂正されるものである。そして、オペレータの発声によってオペレータが発声した音声の認識結果が示され、場合によっては利用者の音声認識結果がオペレータの音声認識結果に訂正されることになるのであるから、引用例に記載された発明は、「前記入力欄内に表示された前記認識結果が前記オペレータによって訂正される」構成を有するものというべきである。

(イ) この点について、原告は、引用例に記載された発明では、「オペレータが発声した音声の認識結果に基づいて主制御部により上記入力欄B1の認識結果は訂正され」ているにすぎず、引用例が「前記入力欄内に表示された前記認識結果が前記オペレータによって訂正される」ことを開示しているとした本件審決の認定は誤りであるなどと主張する。

確かに、引用例に記載された発明では、主制御部により、利用者の音声認識結果とオペレータの音声認識結果とが比較され、オペレータの音声認識結果が最終的な認識結果とされた場合、オペレータの音声認識結果に訂正されるものである。

しかしながら、引用例に記載された発明では、前記のとおり、オペレータは利用者の音声認識結果を確認し、当該認識結果が誤りであると判断した場合、利用者の音声認識結果とは異なる内容を発声することにより、訂正を行うものであって、オペレータによる訂正を前提として、オペレータの発声を認識した主制御部が訂正を行うにすぎないものである。

したがって、原告の前記主張は採用できない。

(ウ) よって、本件審決が、引用例に記載された発明が、「前記入力欄内に表示された前記認識結果が前記オペレータによって訂正される」と認定した点に誤りはない。

### 3 一致点及び相違点の認定の誤りについて

#### (1) 一致点について

本件審決における一致点の認定のうち、引用発明の「オペレータ」が本願発明の「エージェント」に相当する点並びに本願発明と引用発明とが「エージェントに前記音声通信の視覚表示を提示するステップ」を有する点及び「前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される」とした点を除くその余の認定については、当事者間に争いが無い。

ア 引用発明の「オペレータ」が本願発明の「エージェント」に相当する点について

(ア) 前記1(5)によると、本願発明の「エージェント」とは、ワークステーションに表示されるウェブ・ページによって、発呼者とVXMLアプリケーションの間の対話を見ることができ、ワークステーションに表示された発呼者からの入力データを指定変更することができるものである。

(イ) 前記のとおり、引用発明において、利用者と対話するのはオペレータでは



なく対話型音声認識システムであるところ、引用発明の「オペレータ」は、利用者の音声認識結果の表示を見ることができ、訂正することができるものであるから、本願発明におけるエージェントと同様の機能を果たしているものといえることができる。

(ウ) この点について、原告は、一般的に、エージェントは、自立的に状況判断を行う機能を有するものであり、単に機械を操作するオペレータよりも狭いか、あるいはそれとは異なる概念の構成要素を示す用語であるところ、引用発明のオペレータは全ての対話についてその当事者として常に関与するのに対し、エージェントは対話における第三者として適時に関与するものであるから、引用発明のオペレータは、本願発明のエージェントとは、その文言上の意味及び処理の役割が異なり、全く相違する構成要素といえるなどと主張する。

しかしながら、引用発明におけるオペレータが、本願発明のエージェントと同様に、利用者の音声認識結果について自立的に判断し、訂正を行うものであることは、前記のとおりである。

また、引用発明において、利用者と対話するのは、オペレータではなく対話型音声認識システムであるから、引用発明のオペレータは、全ての対話についてその当事者として常に関与するものではない。のみならず、本願発明において、エージェントが対話における第三者として適時に関与する構成要素であることを特定する記載はないから、原告の主張は、本願発明の特許請求の範囲の記載に基づくものではない。

したがって、原告の前記主張は採用できない。

(エ) よって、引用発明の「オペレータ」は、本願発明における「エージェント」に相当するものといえることができ、この点に関する本件審決の認定に誤りはない。

イ 「エージェントに前記音声通信の視覚表示を提示するステップ」を有するとした点及び「前記フィールド内に取り込まれた前記情報が前記エージェントによっ



て更新される」とした点について

(ア) 原告は、「エージェントに前記音声通信の視覚表示を提示するステップ」を有するとした点及び「前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される」とした点についても一致点とした本件審決の認定は誤りであると主張する。

しかしながら、原告の主張は、引用発明のオペレータが本願発明のエージェントに相当しないことを前提とするものであり、その前提自体が誤りであることは前記アのとおりであるから、失当である。

また、原告は、引用発明は、オペレータの作業の負荷を軽減することが解決課題であるところ、オペレータが情報を逐一訂正するような構成は、オペレータの負担をむしろ増加させるものであって、課題を解決することは不可能であると主張する。

しかしながら、引用発明は、初めに認識した利用者の音声の認識結果を暫定的な認識結果としてオペレータに提示することにより、オペレータがその認識結果を参照して利用者の音声の内容を復唱することができ、聞き間違い等的人為的ミスを防ぎ、オペレータの作業の負荷を削減するという課題を解決するものであるから、引用発明において、オペレータが情報を逐一訂正することが、上記課題の解決の妨げとなるとはいえない。

したがって、原告の前記主張はいずれも採用できない。

(イ) よって、本願発明と引用発明とが「エージェントに前記音声通信の視覚表示を提示するステップ」を有するとした点及び「前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される」とした点に係る本件審決の認定に誤りはない。

(2) 相違点について

原告は、本件審決の一致点の認定が誤りであることを前提として、相違点の認定もまた、誤りであると主張するが、その前提自体に理由がないことは、前記(1)のとおりである。

したがって、本件審決の相違点の認定に誤りはない。

#### 4 相違点1に係る判断の誤りについて

##### (1) 本願発明における対話の監視について

本願発明における「ユーザとIVR（対話式音声応答）システムの間での対話を監視する方法」とは、「複数のコマンドを有するIVRスクリプトに従って前記ユーザからの音声通信を処理」し、「前記IVRスクリプトに基づいてエージェントに前記音声通信の視覚表示を提示」し、「前記視覚表示が、前記音声通信と実質上同時に提示されかつ前記ユーザから取得された情報を取り込むために少なくとも1つのフィールドを有し、前記フィールド内に取り込まれた前記情報が前記エージェントによって更新される」ものである。

また、本願明細書の【0006】の記載を併せ考えると、「ユーザとIVR（対話式音声応答）システムの間での対話を監視する」とは、エージェントが、視覚表示（視覚的なインターフェース）によって、ユーザとIVRシステムとの間の対話を監視することを意味するものと解される。

##### (2) 引用発明の音声認識方法について

ア 引用発明では、利用者と対話型音声認識システムとの間で対話が行われるところ、オペレータには利用者の音声認識結果の表示が提示されるものである。

また、引用発明において、オペレータは、オペレータ操作部の表示装置に表示された利用者の音声認識結果を参照するとともに、オペレータ操作部のスピーカから利用者の音声を復唱するものである。

イ 引用発明において、オペレータに提示される前記アの音声認識結果の表示は、利用者と対話型音声認識システムとの間の対話の内容であり、オペレータは、音声認識結果の表示を監視し、復唱のために発声する際に音声認識結果を確認し、認識結果に誤りがあると判断した場合には、これを訂正するものであるから、引用発明において、オペレータは、提示される利用者からの音声の音声認識結果の表示によって、利用者と対話型音声認識システムとの間の対話を監視しているものと認めら

れる。

(3) 相違点1に係る判断の是非について

ア 原告は、本願発明では、対話における第三者であるエージェントが対話を監視するのに対して、引用発明は対話の当事者である音声認識システムが対話を音声認識するものであるから、本願発明と引用発明とでは、対話に対して監視のための処理又は音声認識の処理を施す主体が異なるものであるなどと主張する。

この点について、本件審決は、引用例には、利用者とオペレータの双方の音声を認識し、それらの認識結果を比較し、認識精度の高い方を選択することにより、より正確な認識結果を取得することができると記載されていることなどから、引用発明の音声認識方法では、音声認識システムとの対話における利用者の音声の認識結果の精度は常に比較判断されており、このような状況は、対話が「監視されている」と表現することができるとして、引用発明の「対話における音声認識方法」と本願発明の「対話を監視する方法」との間に実質的な差異があるとは認められないと判断しているところ、引用発明において、「利用者とオペレータの双方の音声を認識し、それらの認識結果を比較し、認識精度の高い方を選択すること」、すなわち「音声認識システムとの対話における利用者の音声の認識結果の精度」を常に比較判断することは、「オペレータ」（本願発明の「エージェント」に相当する。）ではなく、「音声認識システム」により行われるから、「音声認識システムとの対話における利用者の音声の認識結果の精度」を常に比較判断することをもって、本願発明の「対話を監視する方法」との間に実質的な差異があるとはいえないとした本件審決の説示は、措辞不適切であるといわざるを得ない。

しかしながら、引用発明において、「利用者とオペレータの双方の音声を認識し、それらの認識結果を比較し、認識精度の高い方を選択すること」を行う前段階において、前記(2)のとおり、オペレータは、利用者と対話型音声認識システムとの間の対話の内容である音声認識結果の表示を監視し、復唱のために発声する際に音声認識結果を確認し、認識結果に誤りがあると判断した場合には、これを訂正するもの

であるから、引用発明において、オペレータは、提示される利用者からの音声の音声認識結果の表示によって、利用者と対話型音声認識システムとの間の対話を監視しているものと認められる。

したがって、相違点1の構成、すなわち、引用発明の「対話における音声認識方法」と本願発明の「対話を監視する方法」との間には、実質的な差異があると認めすることはできない。

イ 原告は、引用発明において、対話は利用者とオペレータとの間に存在するから、本願発明と引用発明とでは、対話を実行する構成要素が異なるものであると主張するが、引用発明においても、対話は利用者と対話型音声認識システムとの間で行われるから、原告の主張はその前提自体が誤りである。

ウ 以上によると、本願発明は、引用発明及び周知技術に基づいて、当業者が容易に発明をすることができたものであるとした本件審決の判断は、その結論において正当であるものというべきである。

## 5 結論

以上の次第であるから、原告の請求は棄却されるべきものである。

### 知的財産高等裁判所第4部

裁判長裁判官 土 肥 章 大

裁判官 井 上 泰 人

裁判官 荒 井 章 光