

TR-IT-0350

ATR-MATRIX 制御プログラム仕様書

松田 猛† 西野 敦士 横尾 昭男
Takeshi MATSUDA† Atsushi NISHINO Akio YOKOO

2000.2.2

内容概要

音声翻訳システム (ATR-MATRIX) を制御するプログラム、「ATR-MATRIX 制御プログラム」の基本設計と仕様概要について説明する。

エイ・ティ・アール音声翻訳通信研究所

ATR Interpreting Telecommunications Research Laboratories

©エイ・ティ・アール音声翻訳通信研究所 2000

©2000 by ATR Interpreting Telecommunications Research Laboratories

目次

本ドキュメントについて.....	3
このバージョンの特徴.....	3
1 <i>ATR-MATRIX</i> 概略.....	4
<i>ATR-MATRIX</i> の機能.....	4
機能とコントローラーについて.....	4
<i>CSTAR-II</i> 国際実験.....	5
2 システム構成概略.....	6
<i>MATRIX</i> 制御プログラムシステム構成図.....	6
<i>CSTAR-II</i> 国際実験システム構成図.....	7
3 メインコントローラー概略.....	8
<i>SipMainController</i> 構成概略図.....	9
4 サテライトコントローラー1： <i>SPREC</i> コントローラー概略.....	10
<i>SipSprecController</i> 構成概略図.....	11
5 サテライトコントローラー2： <i>TDMT</i> コントローラー概略.....	12
<i>SipTdmController</i> 構成概略図.....	13
6 サテライトコントローラー3： <i>CHATR</i> コントローラー概略.....	14
<i>SipChatrController</i> 構成概略図.....	16
7 サテライトコントローラー4： <i>GUI</i> コントローラー概略.....	17
<i>SipGuiController</i> 構成概略図.....	18
8 サテライトコントローラー5： <i>CONNECT</i> コントローラー概略.....	19
<i>SipConnectController</i> 構成概略図.....	20
9 サテライトコントローラー6： <i>SOUND</i> コントローラー概略.....	21
<i>SipSoundController</i> 構成概略図.....	22
索引.....	23

本ドキュメントについて

ATR-MATRIX は日本語から英語、ドイツ語、韓国語、中国語へ、また、英語から日本語と言った異なる言語間での人間同士のコミュニケーションを可能にする音声翻訳統合システムです。ATR 音声翻訳通信研究所で開発された基礎技術:「音声認識」,「言語翻訳」,「音声合成」の機能を統合し、デモンストレーション、研究開発支援などを行うことを目的として開発されました。

本ドキュメントは「ATR-MATRIX version 2.5.0」のソフトウェアの中の「制御プログラム」に関するソフトウェア仕様書です。このドキュメントは ATR-MATRIX の保守、管理、修正、機能追加などで、直接 ATR-MATRIX ソフトウェアを更新するに当たり有効な資料となるものとして作成されました。

このドキュメントでは ATR-MATRIX の中でもシステム全体の流れを制御管理するプログラム「ATR-MATRIX 制御プログラム」について解説しております。

このバージョンの特徴

1999 年 7 月 22 日に行われた、CSTAR-II 国際実験の為に数多くの追加機能、詳細仕様の変更等を行いました。無論その他の用途のデモンストレーションにも対応が可能です。

1 ATR-MATRIX 概略

ATR-MATRIX の機能

ATR-MATRIX には以下の機能があります。

- 音声認識機能
- 言語翻訳機能
- 音声合成機能
- 通信機能
- 音声認識結果、翻訳結果表示機能
- ユーザーインターフェースによる処理制御機能
- 処理タイミング発信音発声機能

ATR-MATRIX では、これらの機能を統合的、同期的に処理する必要があります。そのため処理の混乱、制御の不具合を処理するため各機能の総合管理する機能として、

- 統合処理管理機能
を設けました。

機能とコントローラーについて

ATR-MATRIX では各機能を実現するために、それぞれの機能ごとに「コントローラー」と呼ばれるプログラムがあります。そのプログラムは単体で、動作するプログラムであり、それぞれの機能を実現する為の情報を管理しています。また、独立性が高く他のコントローラーの動作に依存しません。統合処理管理機能以外のこれらのコントローラーを「サテライトコントローラー」と呼びます。これに対し、すべてのサテライトコントローラーの中央に位置し、すべてのサテライトコントローラーの統合処理管理機能を実現するコントローラーを「メインコントローラー」と呼びます。

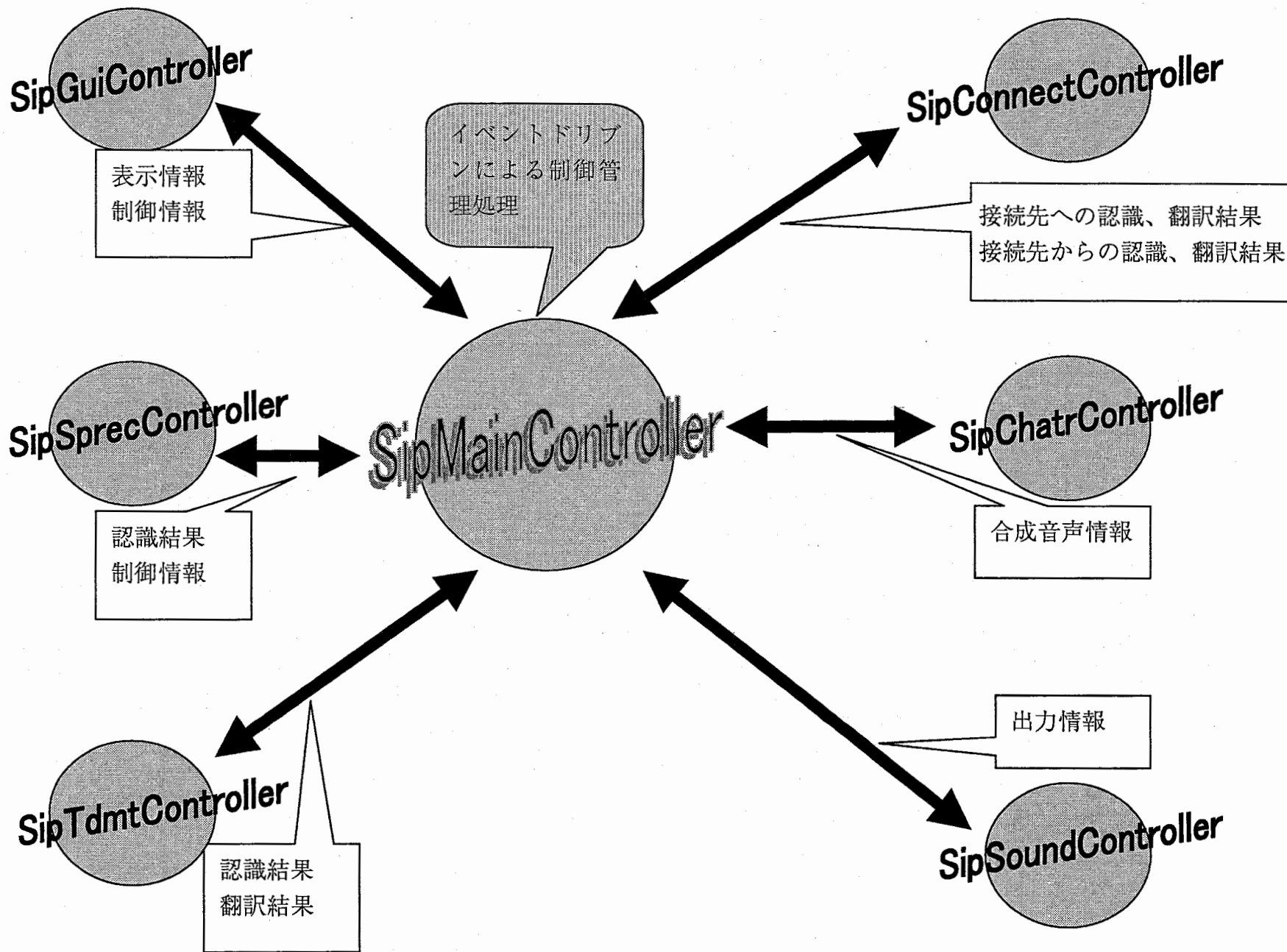
すべてのサテライトコントローラーは、メインコントローラー以外との通信を基本的には行いません。(通信機能は除く) また、それぞれの機能にはそれぞれの管理するモジュールの名前がついています。

■ 音声認識機能	SipSprecController
■ 言語翻訳機能	SipTdmController
■ 音声合成機能	SipChatrController
■ 通信機能	SipConnectController
■ 音声認識結果、翻訳結果表示機能	SipGuiController
■ ユーザーインターフェースによる処理制御機能	同上
■ 処理タイミング発信音発声機能	SipSoundController
■ 統合処理管理機能	SipMainController

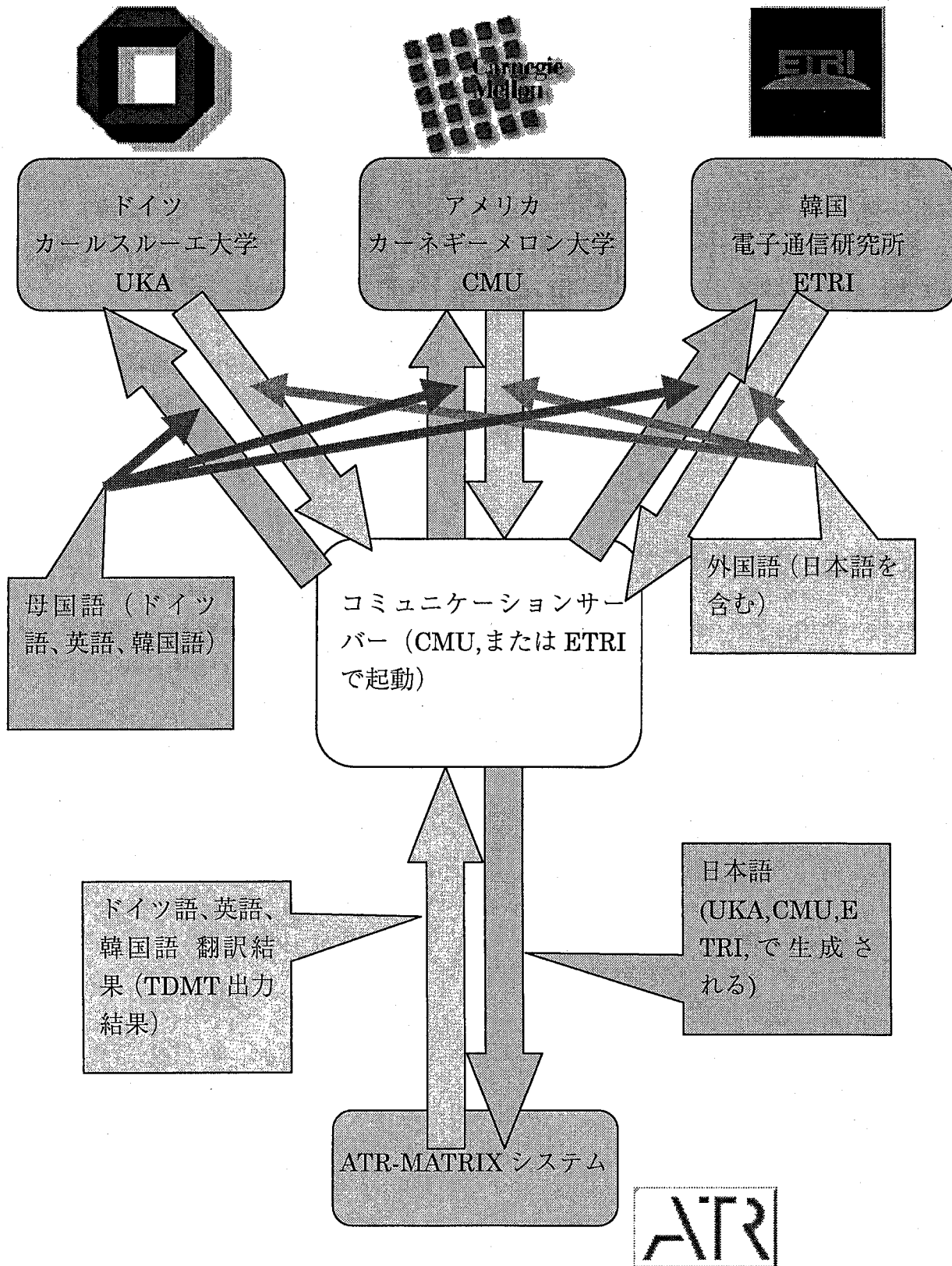
各機能についての詳細は後述いたします。上述のコントローラーの関係については、MATRIX 制御プログラムシステム構成図 (→P6) を参照して下さい。

CSTAR-II 国際実験

1999年7月22日、ATR 音声翻訳通信研究所、CMU (カーネギーメロン大学：米国)、ETRI (韓国電子通信研究所)、UKA (カールスルーエ大学：ドイツ) 等の世界の各研究機関の間で音声翻訳国際実験「CSTAR-II 国際実験」が行われました。各国研究機関の音声翻訳システムを繋いだこの国際実験において、ATRの音声翻訳システムとしてATR-MATRIX が使用されました。「CSTAR-II 国際実験」のシステム構成はCSTAR-II 国際実験システム構成図 (→P7) を参照して下さい。



CSTAR-II 国際実験システム構成図



3 メインコントローラー概略

メインコントローラーは、ATR-MATRIXの中心的役割を果たすコントローラーです。すべてのサテライトコントローラーはメインコントローラーから起動され、メインコントローラーの子プロセスとして動作します。

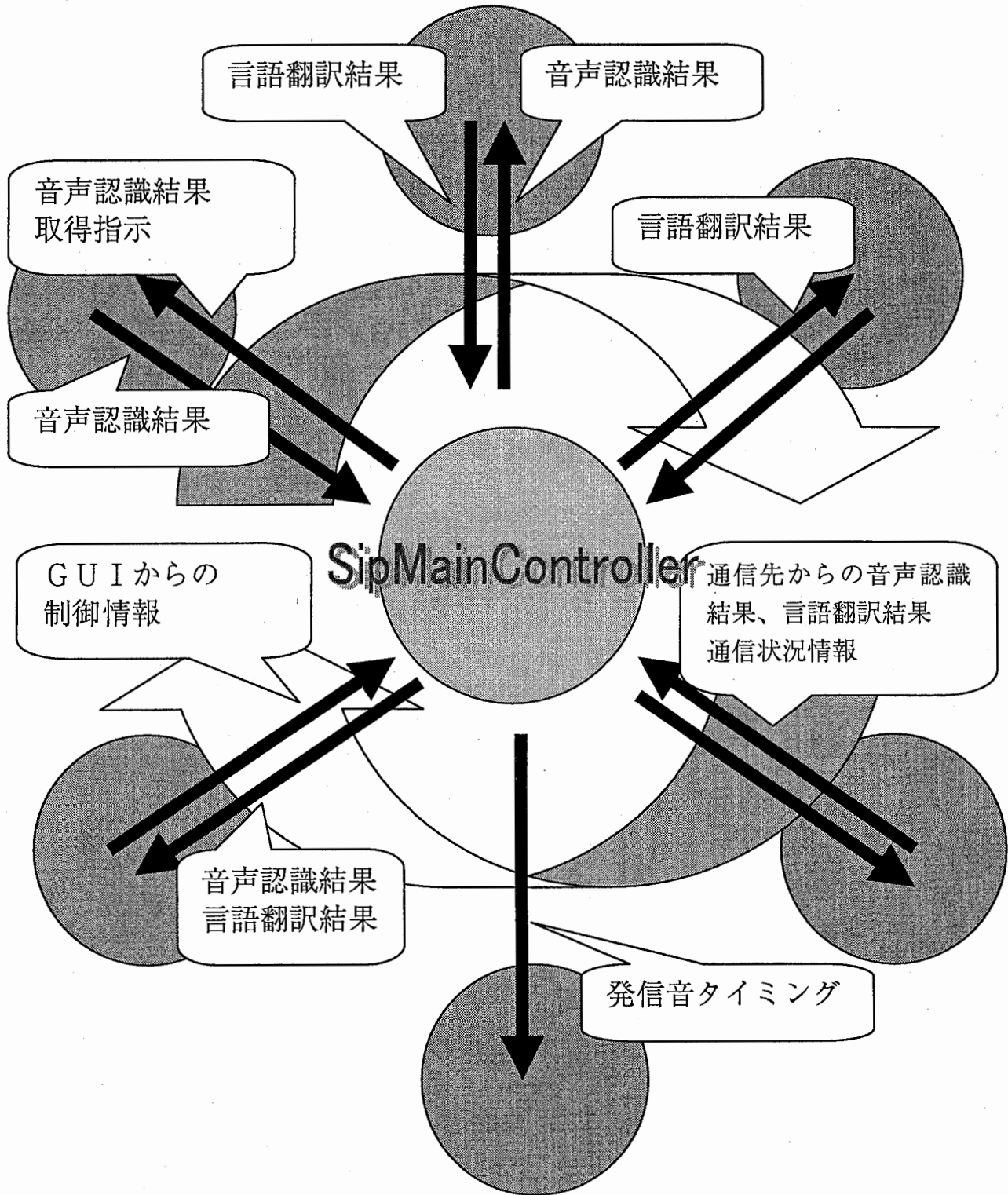
サテライトコントローラーからの全ての出力データは一部を除いてすべてメインコントローラーに渡されます。メインコントローラーは「子プロセスからのデータ受信」をイベントとするイベントドリブン方式で動作しており、何のデータも受信されないときは、待機状態になっています。

データを送信するすべてのサテライトコントローラーのデータ受信をイベントとしているため、イベントフラグとして、すべてのサテライトコントローラーのファイルディスクリプターの値を配列に持ち(子プロセス起動時に取得)、Poll(C/C++言語) ファンクションで、監視をしています。

メインコントローラーは、コンフィグレーションファイル(松田 猛他著「ATR-MATRIX Configuration File仕様書」 TR-IT-0351)の設定により子プロセスに必要な情報を与え起動、もしくは起動させないといった判断処理を行っております。ATR-MATRIX全体の処理の流れを把握し、必要なデータの流れを作るのがメインコントローラーの主な役割です。

メインコントローラー、サテライトコントローラー間で通信されるデータはすべて、「パケット」と呼ばれる定型可変長のデータセットです。このパケットには、その情報がどこからきたのか、どこへいくのか、また、どういう種類のデータなのかを保存しています。このパケットデータはメインコントローラーを介するときにログファイルに保存され、全体としてどのような処理の流れがあったのかを知ることができます。

SipMainController 構成概略図



4 サテライトコントローラー1: SPRECコントローラー概略

SPRECコントローラーは音声入力モジュール「ATRSPREC」の制御、管理を行うコントローラーです。

このコントローラーの以下の主な機能があります。

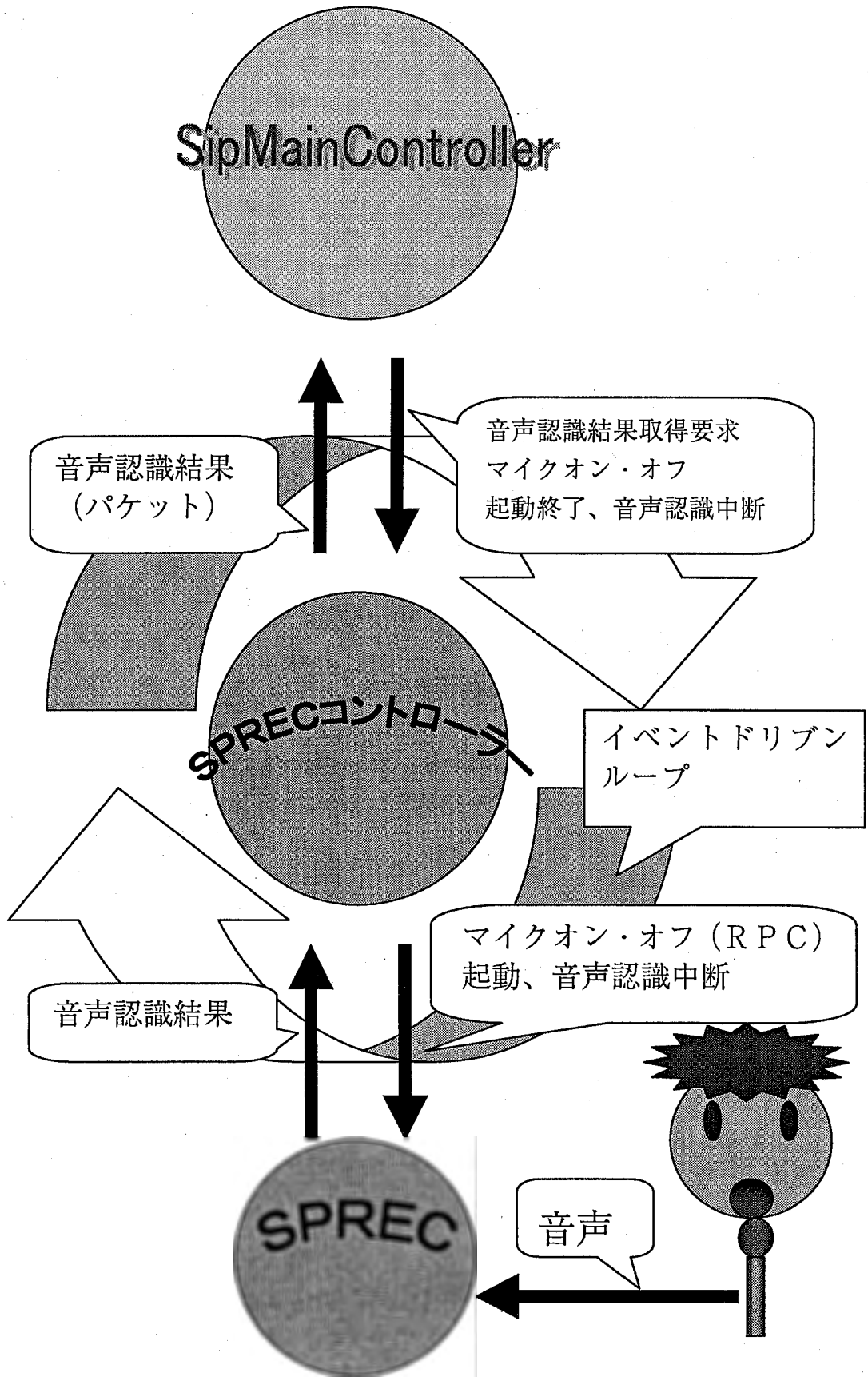
- ATRSPRECの起動・終了
- ATRSPRECからの認識結果の取得
- 取得した複数認識結果の保存
- 認識結果のメインコントローラーへの送信
- ATRSPRECのマイクオン・オフ

SPRECコントローラーはメインコントローラーから子プロセスとして起動されます。起動されたSPRECコントローラーはコンフィグレーションファイルの値にしたがって、ATRSPRECを子プロセスとして起動し、ATRSPRECに対して入出力を行うファイルディスクリプタを取得します。

SPRECコントローラーは、メインコントローラーからの指示に関わらず、SPRECから出力された結果を、自らのバッファ内に保存します。(最大 30 文の保存が可能です。それ以上の認識結果は破棄されます。)

SPRECコントローラーがメインコントローラーからの「認識結果受信要求」の packets を受けた場合、バッファリングしている 1 「認識結果」を packet データに編集しメインコントローラーへ、他の付加情報とともに送信します。

SPRECコントローラーはメインコントローラーからのマイクオン・オフの指示に従い、ATRSPRECの処理状況に関わりなくマイクオン・オフ信号をATRSPRECに送信します。この制御処理にはRPC (リモート・プロシジャーコール) を使用します。



5 サテライトコントローラー2：TDMTコントローラー概略

TDMTコントローラーは言語翻訳モジュール「TDMT-MULTI」の制御、管理を行うコントローラーです。

このコントローラーの以下の主な機能があります。

- TDMT-MULTIの起動・終了
- TDMT-MULTIへの認識結果の入力
- TDMT-MULTIからの翻訳結果の取得
- 翻訳結果のメインコントローラーへの送信
- TDMT-MULTIへの「翻訳方向」の指示

TDMTコントローラーはメインコントローラーから子プロセスとして起動されます。起動されたTDMTコントローラーはコンフィグレーションファイルの値にしたがって、TDMT-MULTIを子プロセスとして起動し、TDMT-MULTIに対して入出力を行うファイルディスクリプタを取得します。

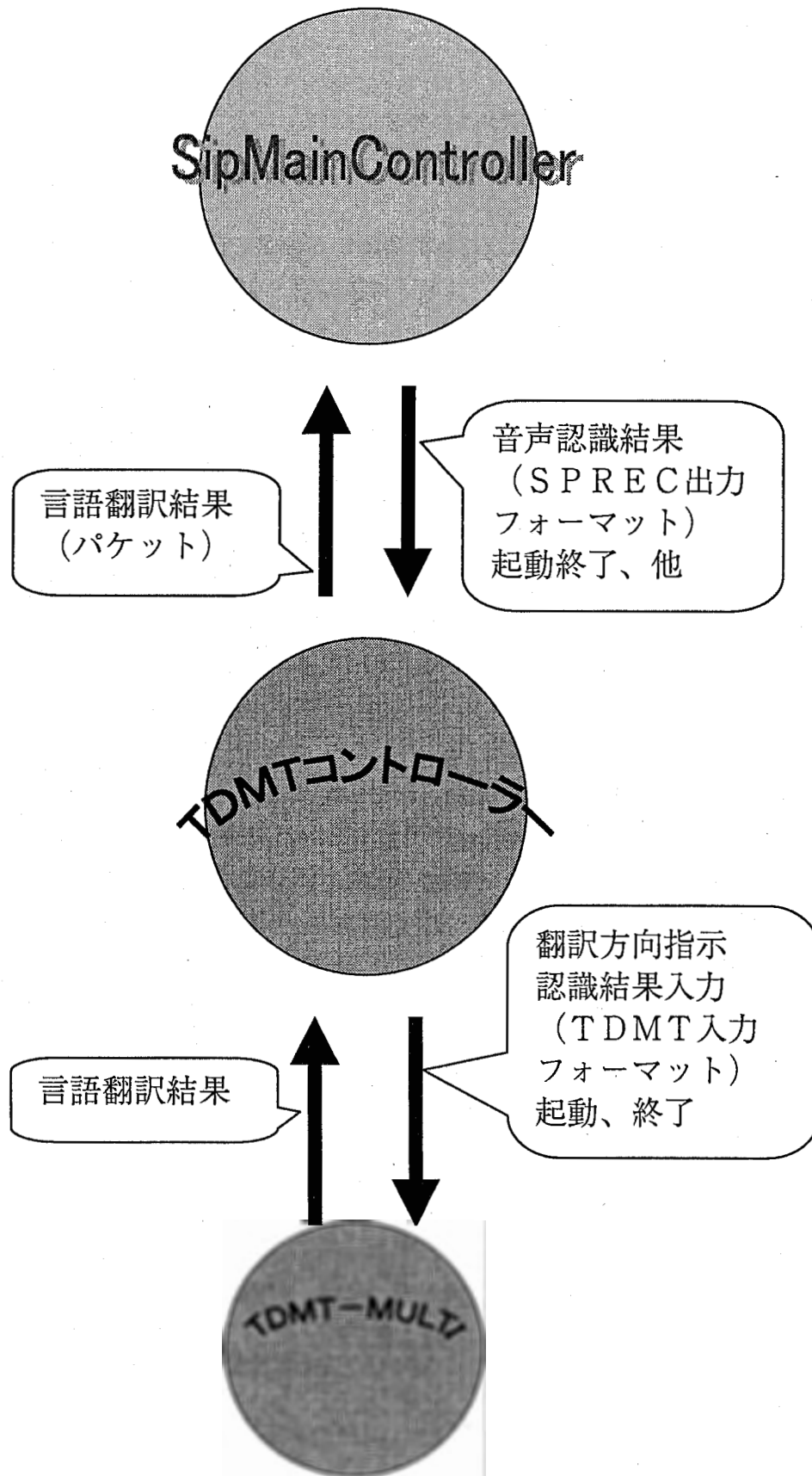
TDMTコントローラーは、メインコントローラーからの指示に関わらず、TDMTから出力された結果を、自らのバッファ内に保存します。

(1言語入力→多言語翻訳の場合、多言語分保存します。)

TDMTコントローラーはメインコントローラーから認識結果を受け取ると、コンフィグレーションファイルに記述されている翻訳言語(複数)に翻訳します。TDMTは同時に多言語への翻訳ができないためTDMTコントローラーはTDMT-MULTIに対して「翻訳方向指示」「認識結果「入力」「翻訳結果取得」の処理をくり返し行います。

TDMTコントローラーはすべての翻訳処理が終わったときにメインコントローラーに対して翻訳結果をパケットデータにして送信します。

SipTdmController 構成概略図



6 サテライトコントローラー3：CHATRコントローラー概略

CHATRコントローラーは音声合成モジュール「CHATR」の制御、管理を行うコントローラーです。

このコントローラーの以下の主な機能があります。

- CHATRの起動・終了
- CHATRへの翻訳結果の入力
- CHATRへの話者決定支持（日本語男性・女性／ 英語男性・女性／ 韓国語男性・女性／ドイツ語男性・女性／中国語男性・女性）

CHATRコントローラーはメインコントローラーから子プロセスとして起動されます。起動されたCHATRコントローラーはコンフィグレーションファイルの値にしたがって、CHATRを子プロセスとして起動し、CHATRに対して入出力を行うファイルディスクリプタを取得します。

CHATRコントローラーは、メインコントローラーから翻訳結果を受け取ると、翻訳結果を音声合成用データに変換します（松田 猛他著「ATR-MATRIX データフォーマット仕様書」TR-IT-0349 参照）。現在日本語以外の音声合成の場合、言語フィルター（G FILTER）を使用しています。この言語フィルターの入力データは翻訳結果であるTDMTの出力のなかにあり、その部分を抽出して入力データを作成します。

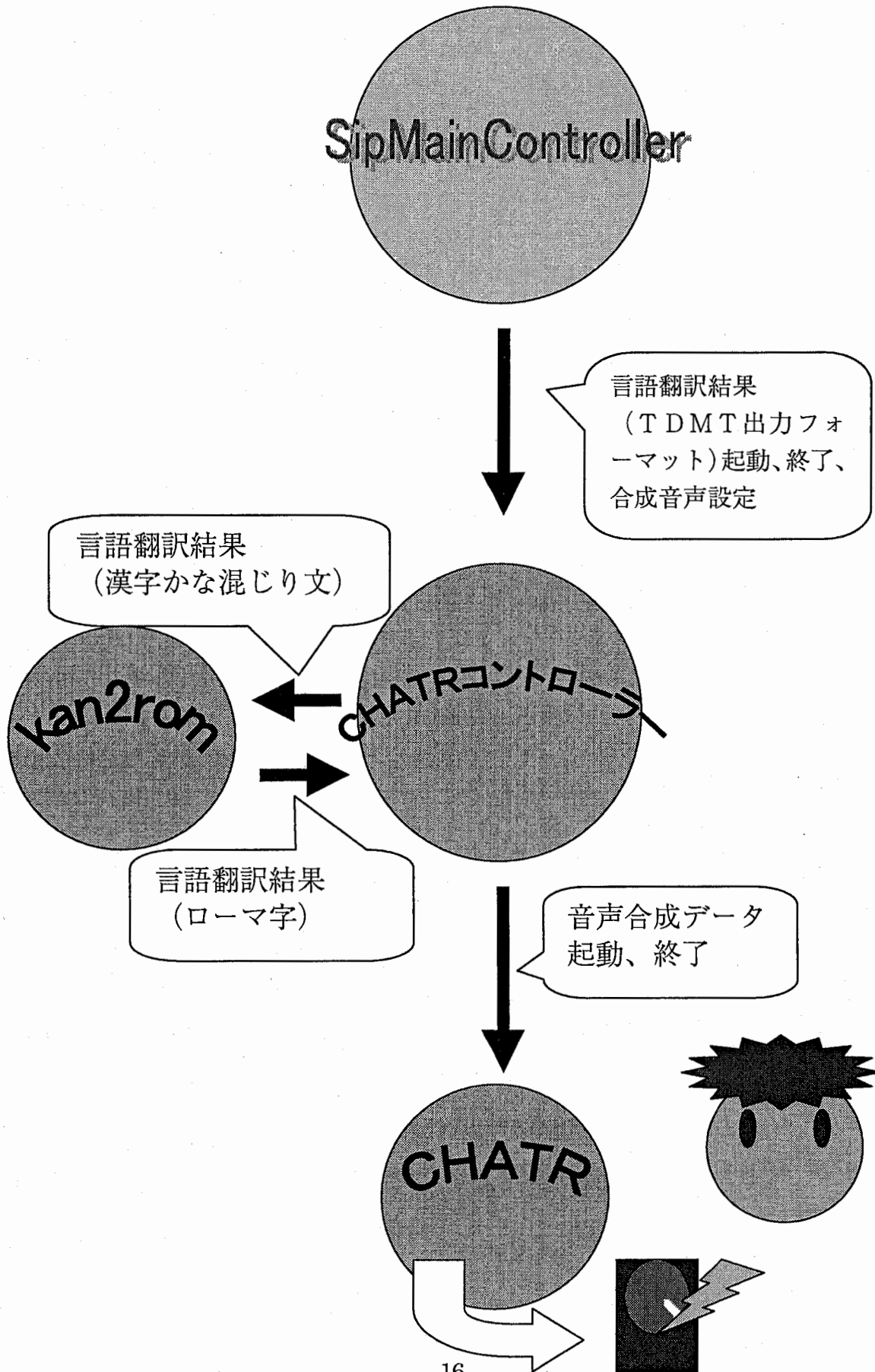
ただし、日本語の音声合成においては使用される条件により、2つのCHATR用入力データ作成方法があります。CHATRでは「漢字かな混じり文」をそのまま入力データとして使用できないため、「ローマ字」に変換する必要があります。1つの方法は、上述の他言語の場合と同じく言語フィルター（G FILTER）を使用する方法。もう一つは、「kan2rom」モジュールを使用する方法です。「kan2rom」は「かな漢字混じり文」をローマ字に変換するソフトウェアです。

現在この使い分けは、

- ATR-MATRIX がスタンドアロンの状態で起動している場合：言語フィルター（G FILTER）
- ATR-MATRIX が他のATR-MATRIX とソケットを介してサーバークライアント形式で接続している場合（→P19）：言語フィルター（G FILTER）
- ATR-MATRIX がコミュニケーションサーバー（→P19）を介してATR-MATRIX 以外のシステムと接続している場合：kan2rom
- ATR-MATRIX がコミュニケーションサーバーを介してATR-MATRIX と接続している場合：両方使用可能。ユーザーの用途により切り替える。

となっています。

version 2.5.0 から、GUI からのメインコントローラーを介しての合成音声出力調整が可能となりました。GUI のメニューからインタラクティブに音声合成する言語を指定することが可能です。指定外の言語のデータがメインコントローラーから送られた場合、そのデータをCHATRコントローラーは処理しません（音声合成を行いません）。



7 サテライトコントローラー4：GUIコントローラー概略

GUIコントローラーはGUIモジュール「MATRIXGUI」の制御、ATR-MATRIXそのものの管理を行うコントローラーです。

このコントローラーの以下の主な機能があります。

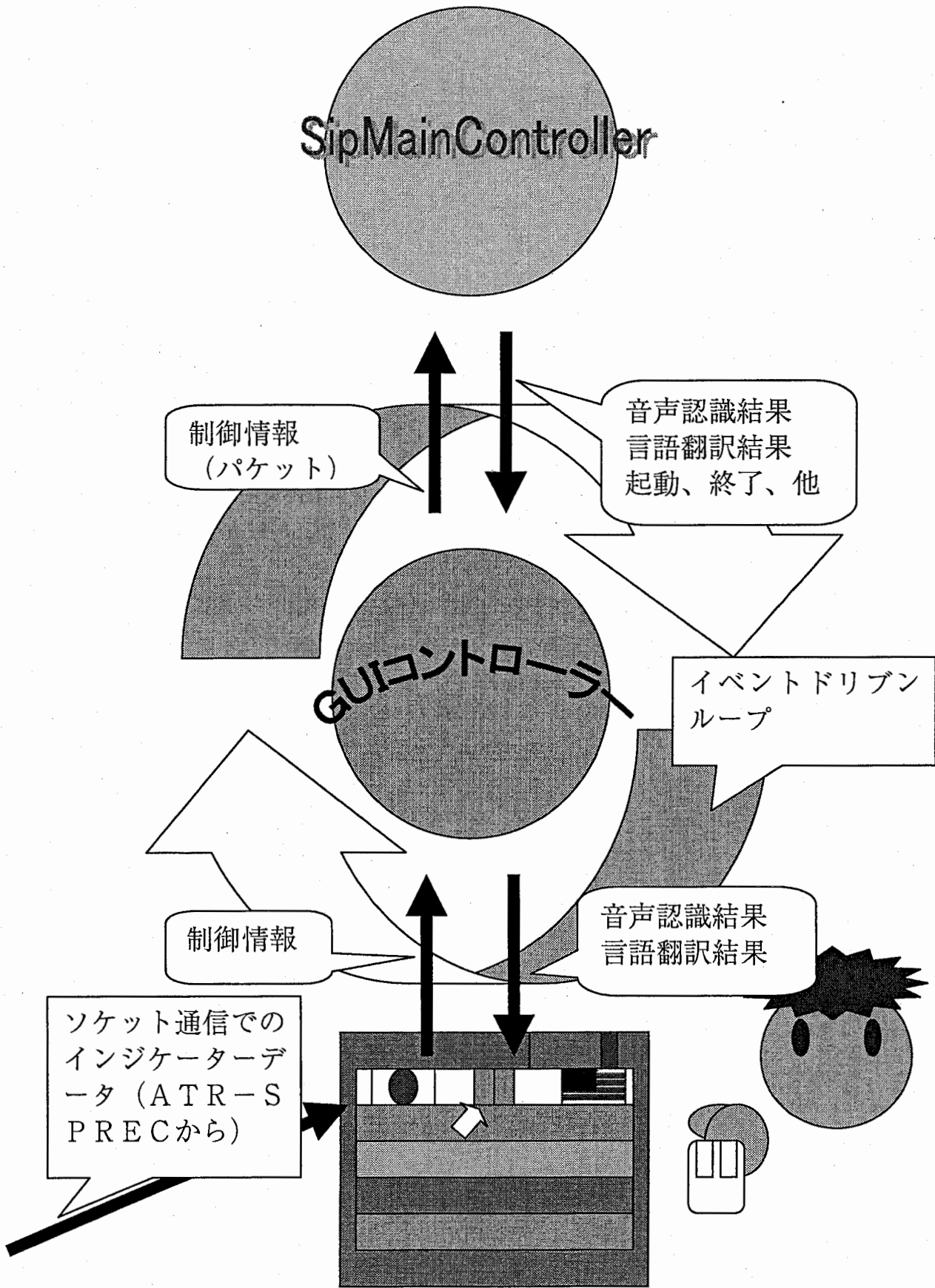
- GUIの起動・終了
- GUIへの認識結果の入力
- GUIへの翻訳結果の入力
- GUIからの終了指示の取得
- GUIからのマイクオン・オフ指示の取得
- メインコントローラーからの認識結果の取得
- メインコントローラーからの翻訳結果の取得
- ●GUIからの合成音声出力言語設定 (追加)
- ●GUIからのデータ無視指示 (追加)

GUIコントローラーはメインコントローラーから子プロセスとして起動されます。起動されたGUIコントローラーはコンフィグレーションファイルの値にしたがって、GUIを子プロセスとして起動し、GUIに対して入出力を行うファイルディスクリプタを取得します。GUIコントローラーは、メインコントローラーから認識結果、翻訳結果を受け取ると、認識結果、翻訳結果をGUIに送信します。

GUIから、マイクオン・オフ情報をGUIコントローラーが取得したときは、メインコントローラーに対して、制御パケットを送信します。

GUIから、MATRIXシステム終了要求をGUIコントローラーが受けた場合メインコントローラーに対して、制御パケットを送ります。

GUIから、合成させる音声（言語）を任意に選ぶ事ができます。但し、接続状況に応じて日本語、英語しか選択できない場合があります。（ATR-MATRIX以外のシステムと接続している場合など）また、単独でシステムが動いている場合でも、合成音声話者データが無い場合はもちろん合成音声は出ません。（但し、指定をする事はできます）



8 サテライトコントローラー 5 : CONNECT コントローラー概略

CONNECTコントローラーは他のシステムとの接続の制御、管理を行うコントローラーです。

CONNECTコントローラーは以下のシステムに接続通信制御を行うことができます。

- ATR-MATRIX
- コミュニケーションサーバー (CMU)

ATR-MATRIX同士で接続する場合、CONNECTコントローラーは相手側ATR-MATRIXのCONNECTコントローラーと通信制御を行い、データの送受信を行います。

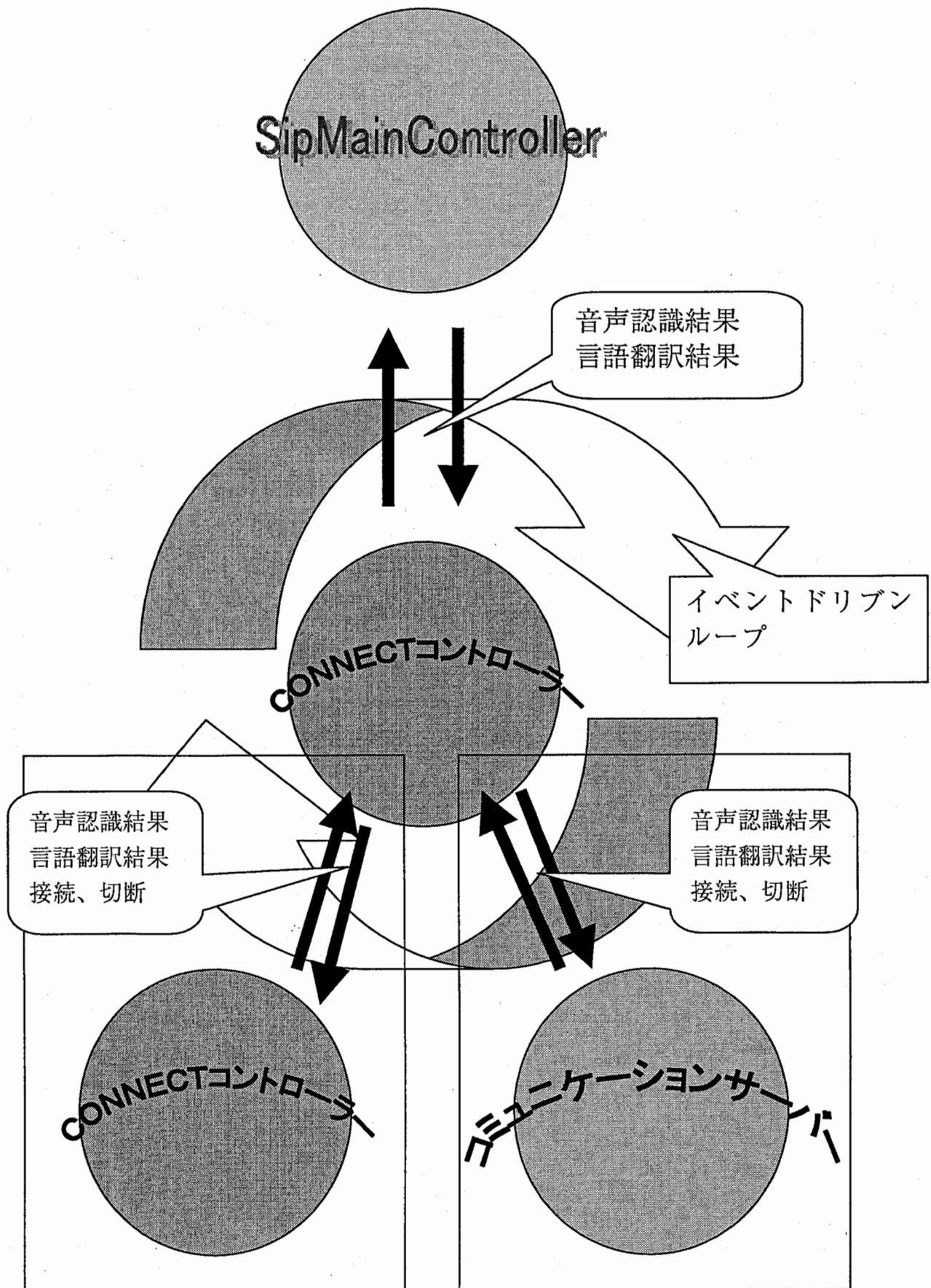
コミュニケーションサーバー (以下CS) は「音声翻訳国際研究コンソーシアム・C-STAR II」の共同実験のためにCMU (米国) で開発された通信サーバーシステムです。ATR-MATRIXはこの場合CSと通信制御を行います。CSの仕様については別途資料を参照ください。

このコントローラーの以下の主な機能があります。

- コネクション確立
- コネクション切断
- 接続先システムへのデータ送信
- 接続先システムからのデータ送信

CONNECTコントローラーはメインコントローラーから子プロセスとして起動されます。起動されたCONNECTコントローラーはコンフィグレーションファイルの値にしたがって、ATR-MATRIXまたは、CSとコネクションの確立を行います。コネクションの確立が成立したときメインコントローラーに対して、その旨の制御パケットを送信します。

CONNECTコントローラーは、メインコントローラーから認識結果、翻訳結果を受け取ると、認識結果、翻訳結果を ATR-MATRIX またはCSに送信します。また、CONNECTコントローラーは、接続先のシステムから認識結果、翻訳結果を受け取るとそれをメインコントローラーに送信します。



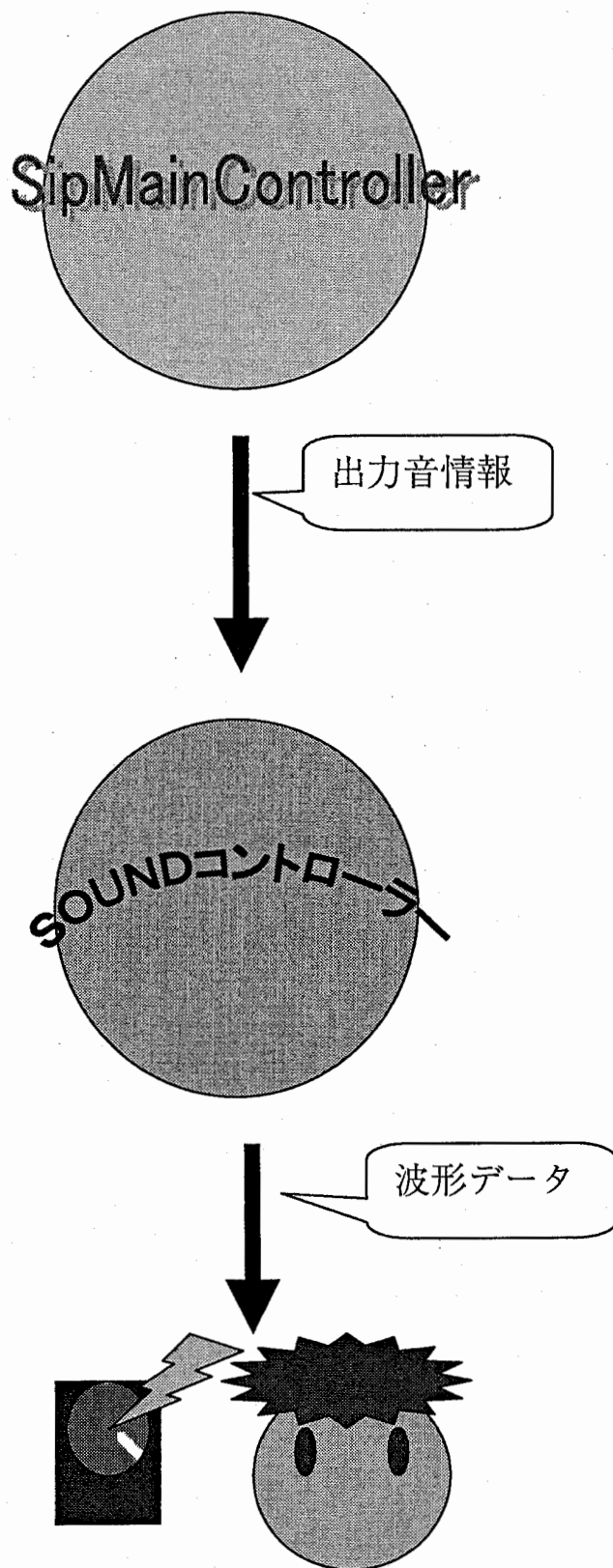
9 サテライトコントローラー6：SOUND コントローラー概略ローラー

SOUNDコントローラーはATR-MATRIXシステム独自で音や声を鳴らす処理を行うコントローラーです。

SOUNDコントローラーは現在、以下のタイミングのときに音を発声しています。

- メインコントローラーが音声認識結果を取得した。
- メインコントローラーが言語翻訳結果を取得した。
- メインコントローラーがCHATRコントローラーに合成音声データを送信した。

SOUNDコントローラーはメインコントローラーから子プロセスとして起動されます。起動されたSOUNDコントローラーはコンフィグレーションファイルの値にしたがって、メインコントローラーからの指示で音を鳴らします。



索引

A

ATRSPREC 10

C

CHATR 14, 21

CHATRコントローラー 15

Configuration File 8

CONNECTコントローラー 19

CS 19

G

GUI 15

GUI 17

GUIコントローラー 17

K

kan2rom 14

R

RPC 10

S

SipChatrController 5, 16

SipConnectController 5, 20

SipGuiController 5, 18

SipMainController 5, 9

SipSoundController 5, 22

SipSprecController 5, 11

SipTdmController 5, 13

SOUNDコントローラー 21

SPRECコントローラー 10

T

TDMT-MULTI 12

TDMTコントローラー 12

い

イベントドリブン方式 8

え

英語 3, 7, 14

お

音声合成 3, 14

音声合成機能 4, 5

音声認識 3, 21

音声認識機能 4, 5

音声認識結果 4

音声認識結果、翻訳結果表示機能 4, 5

か

カーネギーメロン大学 5

かな漢字混じり文 14

韓国語 3, 7, 14

き

起動 10, 14, 17

け

言語フィルター 14

言語翻訳 3, 12, 21

こ	
子プロセス	8, 10, 12, 14, 17, 19, 21
コミュニケーションサーバー	19
コントローラー	4, 8, 10, 12, 14, 17, 19, 21
コンフィグレーションファイル	8, 10, 12, 14, 17, 19, 21

さ	
サテライトコントローラー	4, 8, 10, 12, 17, 19, 21

し	
自然言語翻訳機能	4, 5
処理タイミング発信音発声機能	4, 5

ち	
中国語	3, 14

つ	
通信機能	4, 5

と	
ドイツ語	3, 7, 14
統合処理管理機能	4, 5

に	
日本語	3, 7, 14

認識結果 4, 5, 10, 12, 17, 19

は	
パケット	8, 10, 12, 17, 19

ふ	
ファイルディスクリプタ	8, 14, 17

ほ	
翻訳結果	4, 5, 12, 14, 17, 19
翻訳方向	12

め	
メインコントローラー	4, 8, 10, 12, 14, 17, 19, 21

ゆ	
ユーザーインターフェースによる処理制御機能	4, 5

り	
リモート・プロシジャーコール	10

ろ	
ローマ字	14
ログファイル	8
