

音声翻訳

情報環境論(6)

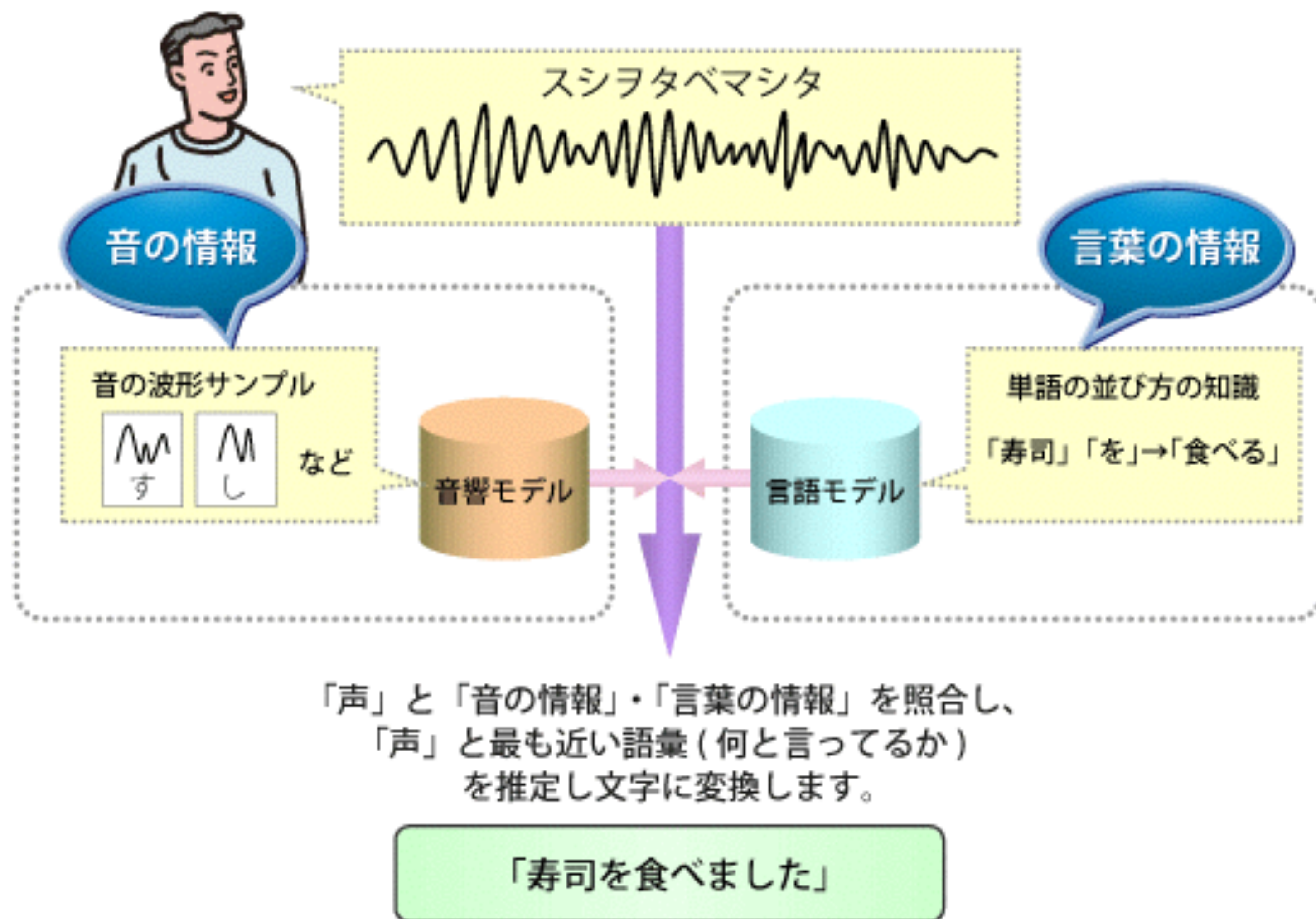
音声翻訳技術の近年における進化

- 実用化（もしくは実用的なサービス）の増加
 - ✓ はなして翻訳（NTTドコモ）
 - ✓ Skype Translator
 - ✓ VoiceTra+（情報通信総合研究機構：NICT）
 - ✓ Google翻訳（スマホ版）

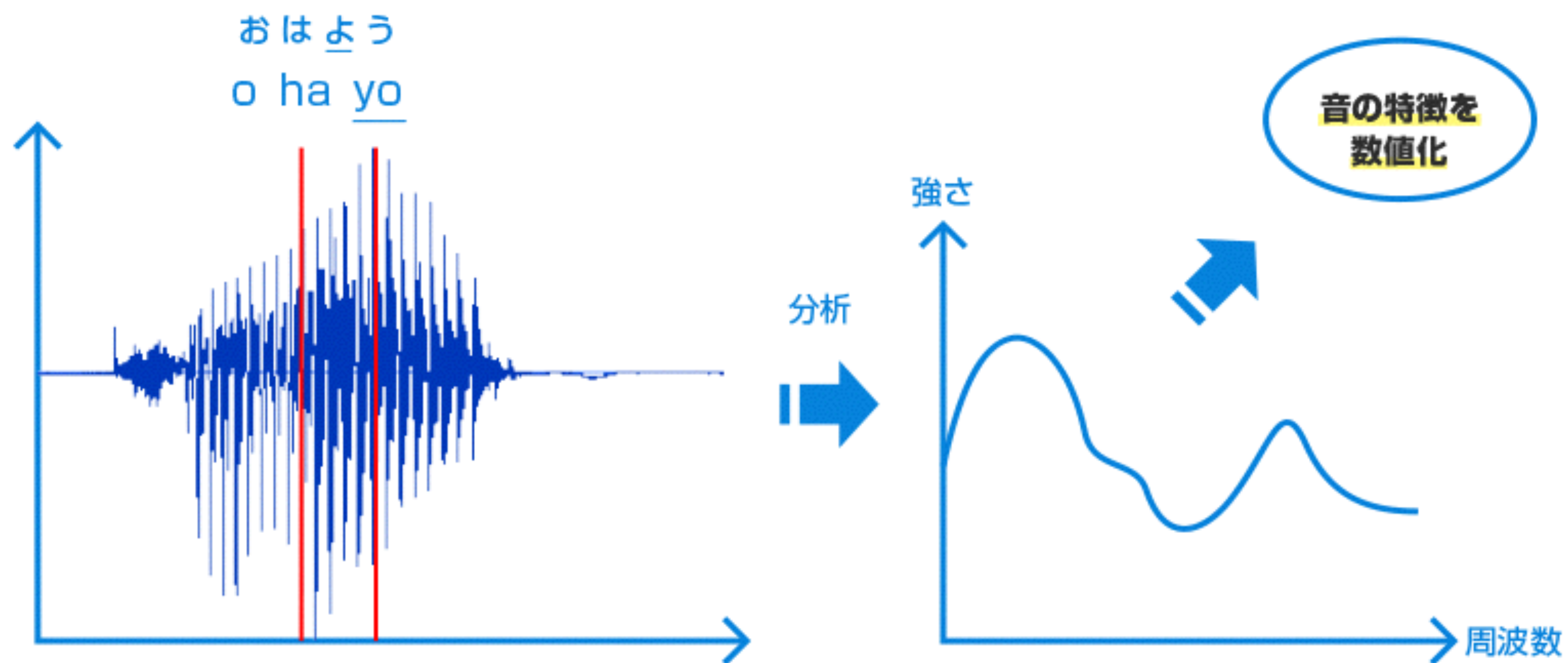
音声翻訳の技術

- 音声翻訳 = 音声認識 + 機械翻訳 + 音声合成
 - ✓ 音声認識：音声を変換
 - ✓ 機械翻訳：テキストを他の言語のテキストに変換
 - ✓ 音声合成：テキストを音声に変換

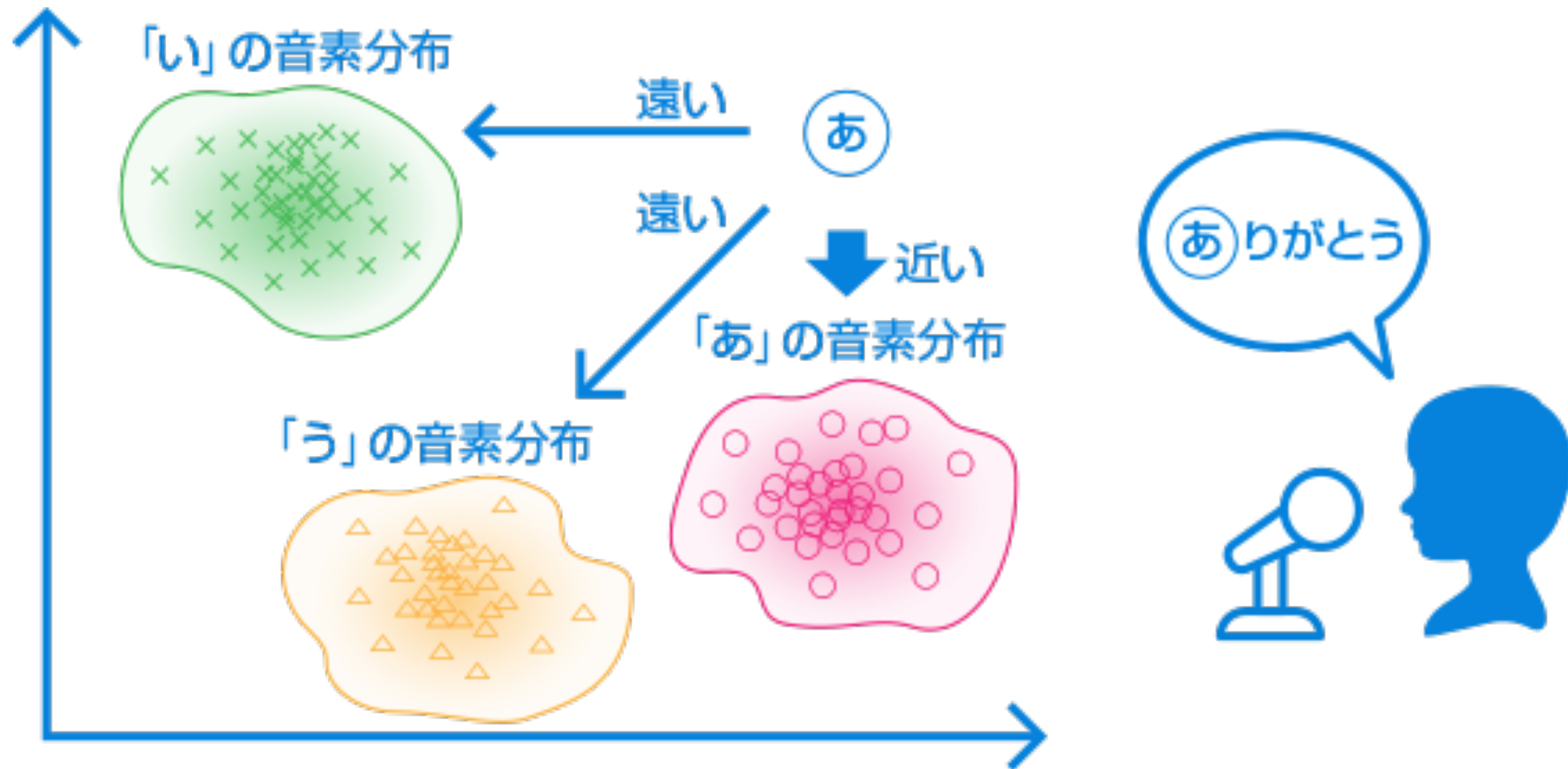
音声認識



音響モデルの構築



音響モデルの利用



単語音響モデル

- 音素列 → 単語にするための辞書

単語	読み	音素列
哀れ	あわれ	a w a r e
哀願	あいがん	a i g a N
愛	あい	a i
遭っ	あっ	a t t a

言語モデル

- 単語間の出現確率を記録 → 統計モデル

第1語	第2語	第3語	
メロス	— は	激怒 (4.2553%)	← 「走れメロス」の中でこの組み合わせで出現する確率
		走った (2.1277%)	
		来ます (2.1277%)	
		村 (4.2553%)	
		単純 (2.1277%)	

機械翻訳

- 機械翻訳 = 自動翻訳 or 翻訳支援
→ 以下では自動翻訳に限定

翻訳の基本的なプロセス

1. 原文書の意図を解釈

2. 意図を目的言語で表現

非常に多くの知識を必要とする

✓ 文法 (grammar)

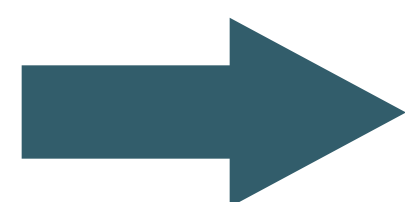
✓ 構文 (syntax)

✓ 意味 (semantics)

✓ 成句 (idioms)

✓ 背景となる文化

などなど...



コンピュータには非常に困難

機械翻訳の手法

- ルールベース
- 統計ベース

ルールベース翻訳

- 言語規則に則って言語を変換する
 1. 原言語文の構文解析
 2. 構文木の変換
 3. 目的言語文の生成

形態素解析 (Morphological Analysis)

- テキストを形態素に分割 + 品詞と原型を付与
- 形態素 = 意味を持つ最小単位
 - ✓ 内容形態素 : 単独で利用され意味を持つ形態素
 - ✓ 拘束形態素 : 内容形態素に付随して現れる形態素
 - ✓ 機能形態素 : 文法的な機能を持つが意味を持たない形態素

形態素解析の手法

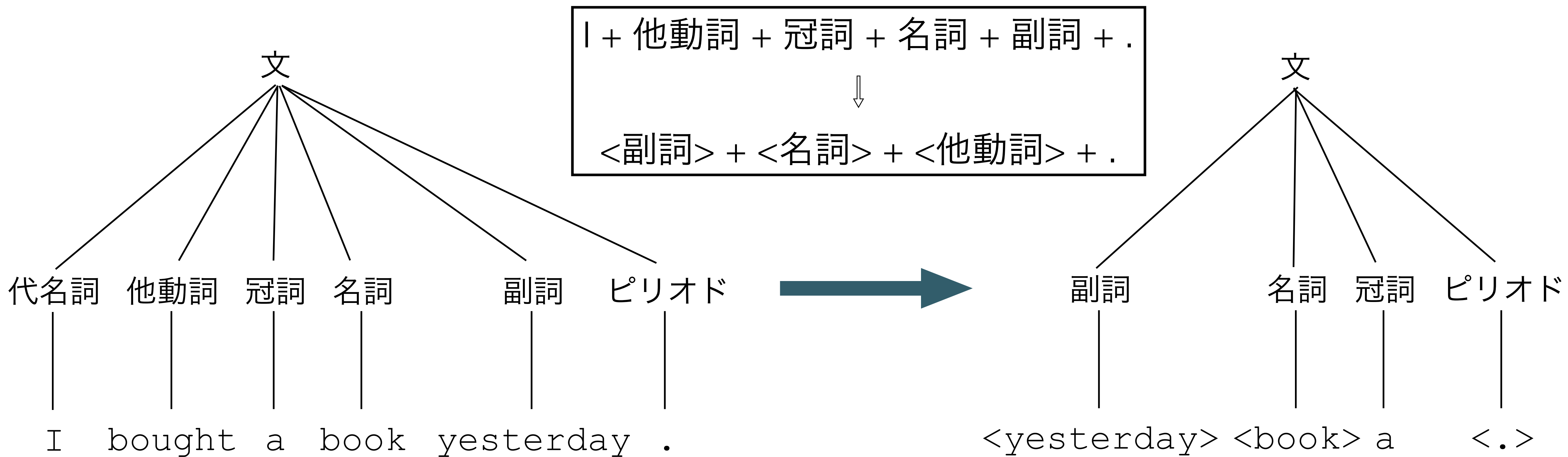
- 規則に基づく手法
 - ✓ 単語の辞書 + 文法テーブル
- 統計的手法
 - ✓ n-gram + 確率モデル (HMM / CRF / ベイジアンネット)

構文解析

- 係り受けの解析
- 構文木の生成

構文木の変換

変換規則



目的言語文の生成

- ・ 構文木から目的言語のテキストを生成する（生成規則）

I sold the computer last week.

→ <last week> <sold> <computer>。

→ 先週本を売った。

I wrote a book last year.

→ <last year> <wrote> <book>。

→ 去年本を書いた。

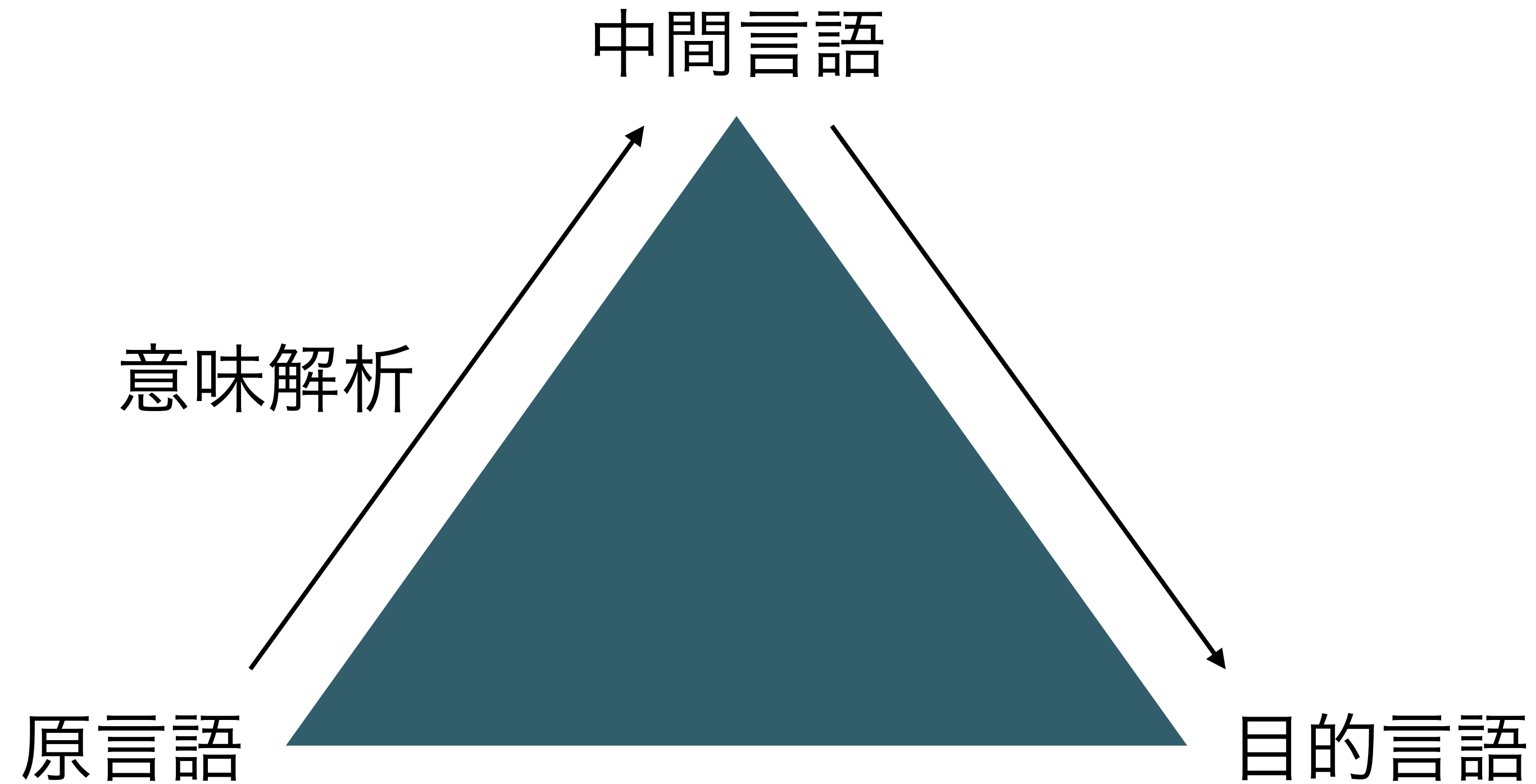
I sent a letter the day before yesterday.

→ <the day before yesterday> <sent> <letter>。

→ 一昨日手紙を送った。

中間言語法（ルールベース翻訳の一種）

- いったん言語に非依存な中間言語表現に変換



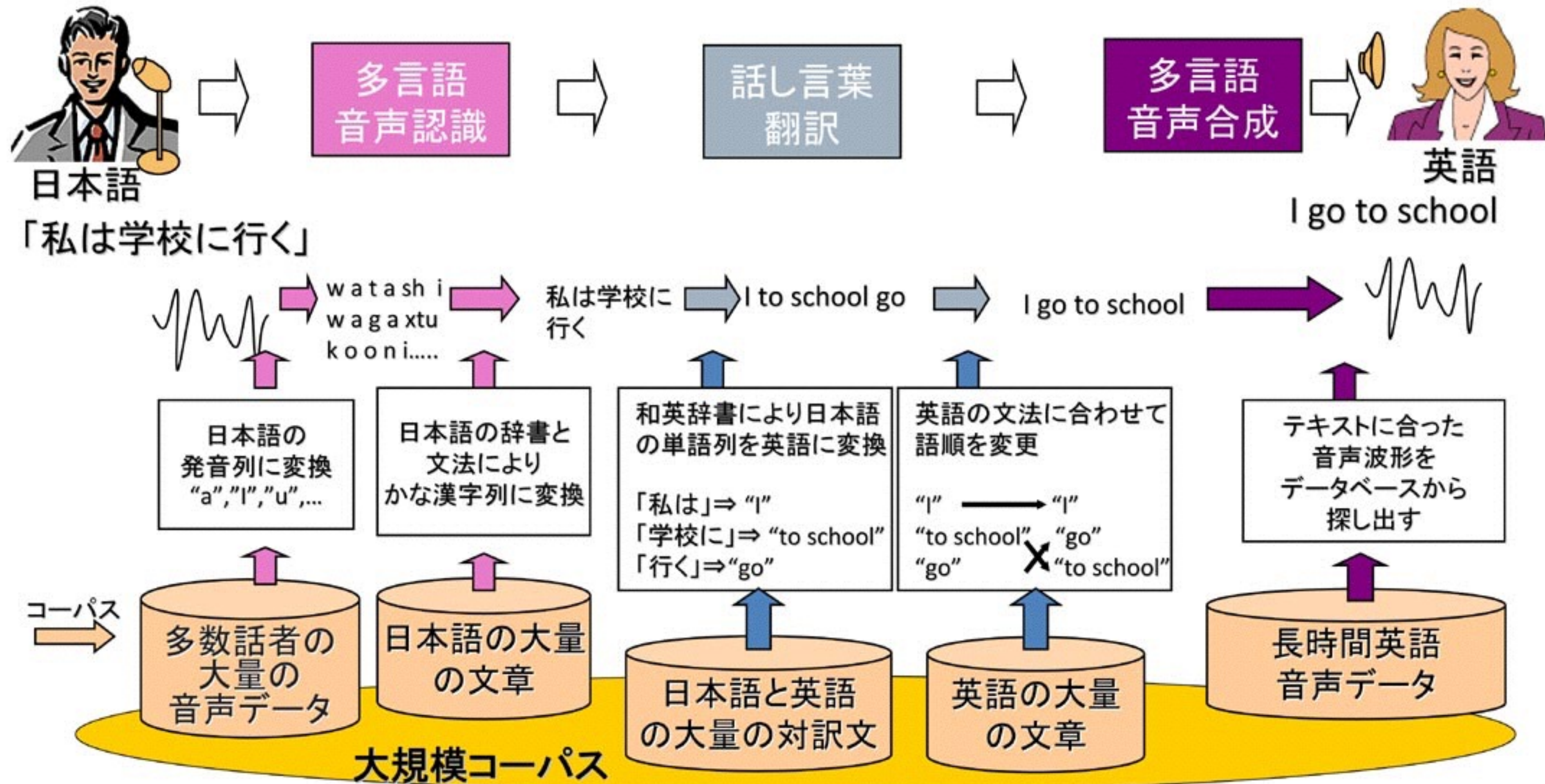
統計ベース翻訳

- 単語対のデータベース＋統計処理 → 意味処理は行わない

音声合成

- 波形接続型音声合成
 - ✓ 録音された人間の音声を利用
 - ✓ 音素のデータベースを利用
 - ✓ 音の欠損（接続部分など）が出る場合がある
- フォルマント合成
 - ✓ 人工的に音声波形を作る

音声翻訳の流れ（まとめ）



音声翻訳の現状

- ロマンズ語系諸言語 ⇔ 英語 ⇔ ゲルマン系諸言語
→ 実用レベル
- 日本語 ⇔ 韓国語
→ 80～90%
- 英語 ⇒ 日本語 > 日本語 ⇒ 英語

スマートフォンによる音声翻訳

大量のコーパスデータ
+
膨大な計算量

→ 大型のコンピュータの必要性



スマホ ←————→ クラウド
高速なネットワーク