

認知モデルを利用した個人の音韻意識の推定：音声フィルタを用いた評価手法の検討

Estimating individual phonological awareness using cognitive models: an evaluation method using audio filters

西川 純平[†], 森田 純哉[†]

Jumpei Nishikawa, Junya Morita

[†] 静岡大学

Shizuoka University

nishikawa.jumpei.16@shizuoka.ac.jp

概要

Language development is supported by phonological awareness, which is related to attention to phonological aspects of spoken language. We aim to develop a system that supports phonological awareness formation using cognitive models. The user's phonological awareness is estimated as a preference for a model through interaction with a system equipped with multiple cognitive models. This paper examines an experimental design with adult subjects to evaluate this estimation method. Audio filters are applied to the system output (word utterances) to simulate an immature language learner. Using the system with audio filters, we test the hypothesis that audio filters affect model preferences. Preliminary results with two participants indicate the influence of the filters on the user's model choice. Further experiments with more participants are needed to test the hypothesis.

キーワード：認知モデリング, 音韻意識, ACT-R

1. はじめに

言語学習者（子どもや第二言語学習者）は、その言語学習の過程においてさまざまな困難に直面する。言語が発達する過程の一部は、音声言語における音素やリズムなど音韻的側面への注意に関係する音韻意識という能力に支えられるとされる。音韻意識のような直接観察できない人の内部プロセスを理解・予測するためには計算機モデルの構築が有効である。このような考えから、著者らは、計算機モデルを活用した音韻意識の形成を支援するシステムの開発に取り組んできた。本研究では、この長期的目標の実現に向け、著者らが提案してきた計算機モデルを利用した個人の音韻意識を推定する手法を評価するための実験デザイン

を検討する。

2. システム

先行研究において、認知アーキテクチャACT-R[1]を用いた音韻意識のモデルがすでに実装されている[2]。このモデルは、ACT-Rが持つ一般的な記憶検索のメカニズムを音韻意識と対応づけることで、未熟な音韻意識に対応づけられるしりとり中の誤りを表現した。具体的には、モデルが持つモーラ¹の記憶には類似度が設定される。しり通りのように、モーラをキューとして単語を検索する際には、モーラ間で設定された類似度により、「似た音を取り違える」誤りが発生する。さらに、このモデルはACT-Rのパラメータを調整することでバリエーションを持つ。たとえば、モーラ間の類似度計算方法や、類似度の影響の大きさに対応する係数 P の値が操作されている。このパラメータ調整により、一部のモデルは、子音脱落のような子どもに見られる特定の誤りに対応づけられている。

図1は音韻意識形成支援システムの概観である。このシステムには、先行研究に基づいた複数の音韻意識モデルのバリエーションが搭載される。システムは、単語選択制のしりとり（モデルの提示する選択肢から適切な単語を選び取ることを提供する。ユーザの音韻意識が、システムとの単語選択制しりとりを通して推定されると想定する。たとえば、子音脱落のような音韻処理の困難を抱える学習者は、自身と同様の傾向を有するモデルの誤りに気づかず、そのモデルへの選好が生じることが考えられる。

¹音の単位のひとつ。継続時間によって定義される[3]。日本語はこれを単位として処理される。

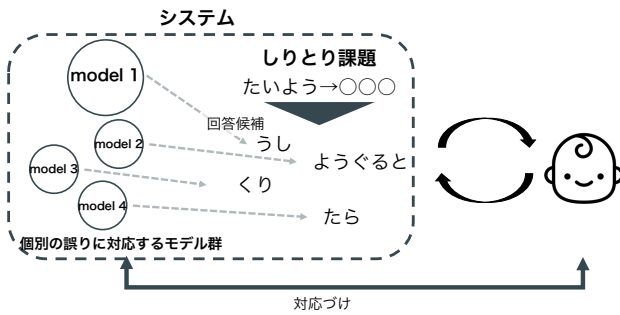


図1 システムのイメージ。ユーザが単語選択制しりとりをプレイするなかで音韻意識が推定される。単語は例。

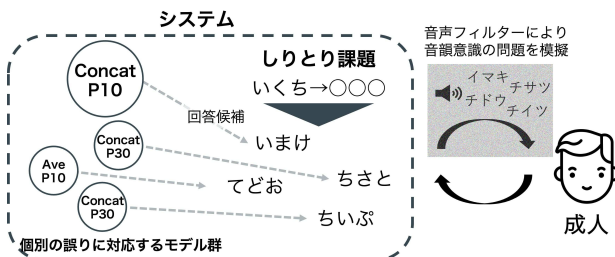


図2 実験コンセプト。無意味単語による単語選択制しりとり。システムの出力にフィルタをかけることで、個人の音韻意識の特性を模擬する。音声フィルタの例はConcatベースフィルタを適用したもの。

3. 実験デザイン

図2に実験のコンセプトを示す。この実験において、モデルのパラメータはモーラ類似度の計算方法で2水準(Concat vs. Ave)²と類似度の影響度合 P で2水準(10 vs. 30)で変更され、計4体のモデルがシステムに搭載された。さらに、この実験では、リクルートコストの観点から、言語学習者(子どもや第二言語学習者)ではなく、日本語を母語とする成人を対象とする。参加者がシステムと行う単語選択制しり通りの最中において、システムの出力(モデルが提案する単語)に音声フィルタを適用することで、音韻処理(とくに聞き取り)が未成熟な言語学習者の状態を模擬する。この設定に基づいて、音声フィルタ(つまり学習者個人の特性)がモデルの選好に影響するという仮説を検証する。

モデルは、100種類のモーラを3つ結合することで生成された無意味単語2,000個を語彙として有すると仮定する³。単語選択制しりとりにおいてモデルが提

²詳細は先行研究[2]を参照。類似度計算のためモーラをベクトル化する際、モーラに含まれる音素を結合(Concat)するか、平均(Ave)するかが異なる。

³濁音・半濁音・拗音を含む。語頭に存在し得ない「ん」は、生成に利用したモーラから除外した。モーラの出現頻度は、日本語の

案する回答候補単語は、音声ファイルを再生することでシステムから出力される。モデルの語彙2,000単語についてテキスト読み上げサービスAmazon Polly⁴を利用して音声ファイル(mp3)を用意した。

システムの出力(モデルの回答候補単語の読み上げ)に音声フィルタが適用されることは、誤った発声の音声ファイルを用いることで実現できる。音声ファイルを作成する際、正しく語彙を発声する音声ファイルに調整を加え、語彙を誤って発声する音声ファイルを用意する。今回の実験では、複数の学習者個人の特性におけるモデルの選好を調査するため、2種類の発声の誤りパターンを用意した。誤った発声の音声ファイルは、モデルの語彙2,000単語について、語頭1モーラおよび語尾1モーラを置換することで得られる。モーラ置換規則は以下の手順で作成される。

1. 2つのモーラの全組み合わせについて類似度を計算し類似度ランキングを作成する。この類似度計算はモデルが持つ類似度テーブルに等しく2種類作成される(Concat/Ave)。
2. 日本語辞書[4]に収録された全単語から音素の出現頻度を数え、モーラの頻度ランキングを作成する。
3. モーラ類似度ランキングの上位から順にモーラのペアを取り出し、モーラ頻度ランキング低順位のモーラから高順位のモーラへ置換する規則を追加する。ただし、連鎖する変換を防ぐため、現在の高順位モーラから変換する規則がすでに存在する場合は、現在のモーラペアを無視する。

モーラ置換規則の例を表1に示す。表2は、この規則によって作成された誤発声の例である。正しく発声される音声ファイルとふたつの誤って発声される音声ファイルについて、ひらがな表記と国際音声字母表記を示した。なお、日本語のモーラ(ひらがな)に対する国際音声字母は、日本語辞典に記載の対応表[5]を参考にした。

4. 予備実験

これまでに述べた実験設定に基づいた予備的な調査について報告する。静岡大学情報学部男子学生2名が参加した。音声フィルタ(に模擬される学習者個人の特性)がモデルの選好に影響するという仮説を検証するために、2種類の音声フィルタについてそれぞれ30分間の単語選択制しりとりを実施し、システム中の各モデルが選択された回数を調査した。

辞書[4]に登場する頻度に基づく。

⁴<https://aws.amazon.com/de/polly/>

表1 誤発声ファイルを作成するためのモーラ置換規則の例. 2種類のモーラ類似度に基づくふたつの置換規則について, それぞれ上位5つを例として示す.

順位	Concat	Ave
1	じゅ→しゅ	ひゃ→は
2	によ→にゅ	ぎゃ→ぎゃ
3	みゅ→みゅ	ぎゅ→ぎゅ
4	りよ→りゅ	ぎょ→ぎょ
5	ず→ず	ひよ→ほ
...

表2 音声ファイルの例

フィルタタイプ	ひらがな	IPA
オリジナル	てどお	te do o
Concat ベースフィルタ	ちどう	tʃi do u
Ave ベースフィルタ	てどう	te do u

結果を図3に示す. 各系列はモデルに対応する. モデルが持つ類似度テーブルごとに値を集約(類似度の影響度合 P について合算)した結果を示している. 横軸に参加者および音声フィルタの条件が並ぶ. 縦軸は参加者の誤答選択数がモデルの誤答数に占める割合を示す. つまり図の最右にある青バーは, Concat ベースフィルタ条件の30分間のしりとりにおいて, 参加者AがConcatテーブルのモデルを選択しつつ誤答した回数を分子と, 同条件でConcatテーブルのモデルが誤答した回数を分母とした割合の値である.

この図から, フィルタ間でモデルの選択されやすさが異なることが読み取れる. つまり, Concat条件におけるモデル間の差より, Ave条件におけるモデル間の差が大きい. このことは音声フィルタ(に模擬される学習者個人の特徴)がモデルの選好に影響するという仮説と整合的である.

5. まとめ

本稿では, 個別化認知モデルを用いた音韻意識形成支援システムの実現に向け, 音声フィルタの適用によって成人を参加者としてシステムの実現性を評価する実験デザインを検討した. 本稿の結果は予備的なものであり, 参加者数を増やした実験による検証が必要である.

今後は, 音声フィルタの設定を吟味することも重要になる. 困難を抱える子どもの特性(自閉スペクトラム症の特異な知覚)を再現する研究[6]を参考に, 知見に基づくフィルタを作成することで, より現実の困難に対応づいた検証が可能になる. これらの実験の結

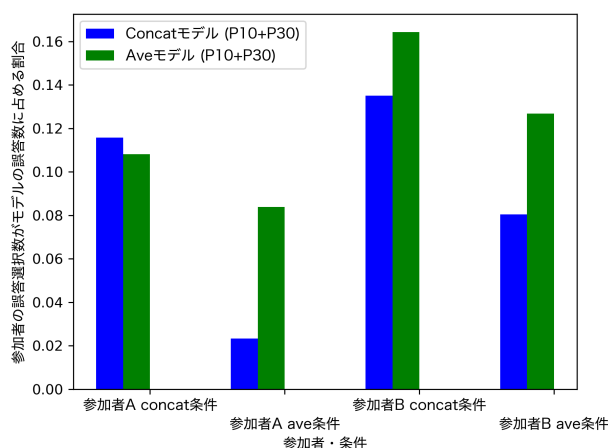


図3 予備実験の結果. 縦軸は参加者の誤答選択数がモデルの誤答数に占める割合を示す.

果を踏まえてシステムのブラッシュアップをした後, システムが本来対象とする言語学習者を対象とするシステム評価実験を実施する.

文献

- [1] Anderson, John R, (2007) "How can the human mind occur in the physical universe?": Oxford University Press.
- [2] Nishikawa, Jumpei and Morita, Junya, (2022) "Cognitive model of phonological awareness focusing on errors and formation process through Shiritori," *Advanced Robotics*, Vol. 36, No. 5-6, pp. 318-331.
- [3] Port, Robert F., Dalby, Jonathan, and O'Dell, Michael, (1987) "Evidence for mora timing in Japanese," *The Journal of the Acoustical Society of America*, Vol. 81, No. 5, pp. 1574-1585.
- [4] 天野成昭・小林哲生, (2008) "基本語データベース: 語義別単語親密度", 学習研究社.
- [5] 松村明, (2006) "大辞林第三版", 三省堂.
- [6] 長井志江, (2016) "自閉スペクトラム症の特異な視覚世界を再現する知覚体験シミュレータ", *精神看護*, Vol. 19, No. 1, pp.59-63.