

音韻論的記述への計算モデルのアプローチ —音便変化のモデリングによる検証—

A computational approach to phonological changes

岸山 健^{1*}

Takeshi Kishiyama¹

¹ 東京大学大学院総合文化研究科言語情報科学専攻

¹ Language and Information Sciences, The University of Tokyo

要旨: 本研究は計算モデルが自身で表現できない音韻論的な記述を再現できるかを検証した。人が音声から音素を知覚する過程の計算モデルとして、近年の研究では hidden Markov model (HMM) が利用されている。HMM は音素と音素の関係を音素間の遷移確率として表現できるが、遷移確率のみでは歴史的な音便変化を表現できない。こうした表現できない記述を計算モデルが再現できるかは議論の余地があった。そこで本研究では「表現できない音韻論的な記述」をモデル同士のデータの生成と識別の過程で再現できるかを検証した。実験 1 では /kawuta/ → [kawta] を生成時の操作で実現させ、対応するデータをモデルに識別させた。実験の結果、[kawta] に対し /kaQta/ と促音が挿入されたため、計算モデルで音便変化が再現可能であることを示した。実験 2 では同様の実験をウ音便に対しても行った結果、同様に音便変化を再現できることを示した。

1 はじめに

本研究が検証する問は「計算モデルは自身で表現できない音韻論的な記述を再現可能か¹」である。心理言語学で人の音声知覚として提案・検証されてきたモデル [1] は近年、計算機上で動作する計算モデルとして検証されている [2–5]。しかしながら、主流の計算モデルは音韻論的な記述の一部を表現できないため、計算モデルが音韻論に与える貢献は限定的に見える。そこで本研究は計算モデルに与える入力を制御し、表現できない記述である歴史的な音便変化が再現可能か検証した。

以下に本章の構成を示す。はじめに音声知覚の計算モデルの概要を示した後、計算モデルが内部で表現できる記述と表現すべき記述を例示する。次に歴史的な音便変化の概要を示し、計算モデルが内部で表現できない、すべきでない理由を挙げる。最後に計算モデル外の要素で音便変化を再現する際の課題、検証方法をまとめる。

1.1 音声知覚の計算モデル

心理言語学では音声知覚モデル、つまり音声から音素への推論過程が提案・検証されてきた。中には計算モデルに知覚現象を再現させ、仮定している推論過程の心理的実在性、つまり認知的な過程としての妥当性を検証した研究がある。検証された知覚現象の例として、音素の典型例周辺で音声の識別が困難になる現象 (perceptual magnet effects) が挙げられる [6]。この現象を再現して心理的実在性の支持を得た計算モデル [7] は次の単純ベイズ分類器 (naive Bayes classifier) と等しい [8]。

単純ベイズ分類器は音素と音声をパラメータとして、音声から音素への推論をアルゴリズムとして定義する。例として、音素カテゴリー c は確率 $P(c)$ 、音声 S は音素 c ごとの確率 $P(S|c)$ としてそれぞれパラメータで定義される²。そして新規の音声 S_{new} から音素 \hat{c} への推論は

$$\hat{c} = \arg \max_c P(c)P(S_{\text{new}}|c) \quad (1)$$

となる。式 (1) では音声 S_{new} に対して $P(c)$ と $P(S_{\text{new}}|c)$ の積を最大化する音素 c を最尤の候補 \hat{c}

*連絡先: 東京大学総合文化研究科言語情報科学専攻
〒153-8902 東京都目黒区駒場 3-8-1
E-mail: kishiyama.t@gmail.com

¹本研究において表現という用語は記述 $A \rightarrow B$ をパラメータやアルゴリズムとしてモデルに定義することを示す。対して、再現という用語はモデルに A を与えて B を出力させ、記述 $A \rightarrow B$ を実現させることを示す。

² $P(S)$ は音響情報の分布を示す確率分布となるため、例えば二つのフォルマントで分布を表現する場合は二次元のガウス分布となる。その確率分布を音素 c ごとに定義したものが条件付き確率 $P(S|c)$ となる。

として求めている．しかしながら単純ベイズ分類器は音素を $P(c)$ としか定義しないため，音素間の関係を含む音韻論的な記述は表現できない．

対照的に，音素間の関係を遷移確率 [9] として表現できるモデルには隠れマルコフモデル (hidden Markov model: 以下 HMM と略称する) がある．HMM は一つ前の音素 c_{t-1} から現在の音素 c への遷移確率を $P(c|c_{t-1})$ としてパラメータで定義する．そして新規の音声から音素への推論は

$$\hat{c} = \arg \max_c P(c|c_{t-1})P(S_{\text{new}}|c) \quad (2)$$

となる．単純ベイズ分類器の式 (1) で $P(c)$ だった部分が $P(c|c_{t-1})$ となっており，音素配列のパターンが表現可能となっている．これを実現する過程は並列的なアルゴリズムで定義され [10]，知覚レベルでの母音の錯覚 [1, 11] を再現して心理的実在性の支持を得ている [5]．

HMM が再現した母音の錯覚とは，日本語母語話者が母語にない音素列 /ebzo/ を聞いたとき，/b-/z/間に母音を挿入して知覚する現象である．HMM は再現実験で妥当性が検証されているものの [2-4]，音素間の関係は $P(c|c_{t-1})$ としか表現できない．知覚レベルの母音の錯覚の再現に必要な記述は子音連続を認めない $P(/z/||b/) = 0$ 程度であり [5]，HMM が表現できる記述に属する．対して，2 音素間の関係だけでは 3 音素間の関係や音節構造を表現できず，これが HMM が表現できない記述に相当する．

記述が表現可能に加え，表現すべきか否かという問題もある．例として，母音の錯覚はその定義から知覚レベルで発生するため [1]，音声知覚の計算モデルで表現すべき現象である．対照的に，日本語の歴史的な音便変化は HMM で原因を表現できないだけでなく，表現すべきでない記述である．そこで次に音便変化の概要と，計算モデルで表現すべきでない理由を述べる．

1.2 音便変化と計算モデル

知覚の計算モデルが内部で表現すべきでない記述として，本研究では歴史的な音便変化を挙げる．音便変化にはもともと音形が [kawu] であったとされる動詞「買う」が「買った」となる促音便や「買った」となるウ音便などがある [12]．以下では促音を /Q/ として [13]，促音便の変化を /kawuta/ → /kaQta/ と表現する．なお促音化には諸説あり，仮に促音化を /kawta/ → /kaQta/ と表現しても本研究は /kawuta/ → /kawta/ → /kaQta/ の変化を再現するため両方の議論を含む．次に先行研

究で検証された母音の錯覚 /ebzo/ → /ebuzo/ と上に挙げた促音化の違いを 2 つ示す．

第一に，母音の錯覚と異なり歴史的な音便変化は知覚のみが原因ではない．母音の錯覚 (/ebzo/ → /ebuzo/) は変化の前後が知覚レベルで弁別できなくなるため [1]，計算モデルで原因を表現すべき記述である．対して /kawuta/ と /kaQta/ のそれぞれから生成された音声を日本語母語話者が弁別できないという報告はなく，仮に音便変化の前後で弁別できるとすれば知覚のみが原因とはならない．したがって歴史的な音便変化は音声知覚の計算モデルのみで表現すべきでなく，この点で促音便の変化は母音の錯覚と異なる．

加えて，記述の原因が HMM で定義できない点も母音の錯覚のケースと異なっている．母音の錯覚では c_t が /z/ の時点で $P(c_{t-1} = /u/)P(/z/||u/)$ > $P(c_{t-1} = /b/)P(/z/||b/)$ ならば c_{t-1} に /u/ が選択されるため (図 1 参照)，表現は $P(/z/||b/) = 0$ かつ $P(/z/||u/) \neq 0$ で十分である [5]．対して促音化 (/kawuta/ → /kaQta/) は $P(/u/||w/) = 0$ か $P(/t/||u/) = 0$ でなければ /wut/ を抑制できず，どちらも非現実的な表現である³．しかしながら，モデルの外部である入力で [kawta] と母音の削除が起きた場合は促音化を次のように説明できる．

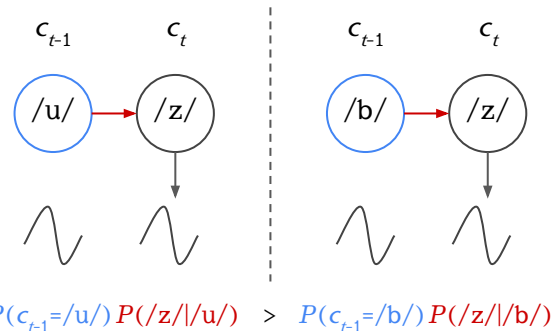


図 1: 左のグラフは $P(c_{t-1} = /u/)P(/z/||u/)$ ，右のグラフは $P(c_{t-1} = /b/)P(/z/||b/)$ を示す． $P(c_{t-1} = /u/)$ と $P(c_{t-1} = /b/)$ が 0 でない条件では $P(/z/||b/) = 0$ かつ $P(/z/||u/) \neq 0$ で $c_{t-1} = /u/$ が最尤となり，母音の挿入が再現される．

促音化の背景には内在時間長の短さによる母音の削除が仮定されている [12]．話者が意識できるほど大きくないものの，高母音の /i/ や /u/ は低母音と比較して発音に至るまでの時間 (内在時間長) が短い．仮に調音時に母音が削除される場合はモデルへ

³HMM は 3 音素間の遷移確率も定義できるが，2 音素間で得られている経験的な支持 [9] がないため認知モデルとしては不適切となる．

の入力の時点で /kawuta/ → [kawta] となる．そこで /t/ を知覚した時点で $P(c_{t-1} = /Q/)P(/t/|/Q/)$ > $P(c_{t-1} = /w/)P(/t/|/w/)$ ならば知覚の過程で /kawuta/ → [kawta] → /kaQta/ と記述を再現できる．

以上をまとめると，促音便の変化は母音の挿入と二つの点で異なる．一つは知覚のみで促音便を説明できない点であり，もう一つは HMM の定義で原因を表現できない点である．さらに，母音を削除した入力を与えれば促音化が再現可能なものの，母音 /i/ が削除されて「持ちて (motite)」が「持って (motte)」となる [12] ほど自明な促音化ではない．また，「買う」の変化は促音便だけでなくウ音便もあり，この点の再現も音韻論との関係を考察する材料となる．そこで本研究では母音 /u/ や子音 /w/ の持続時間を操作して音便変化を再現可能か検証した．最後に検証する問と実験の大枠を記述した後，各実験の詳細に移る．

1.3 本研究

母音の削除により促音化は /kawuta/ → [kawta] → /kaQta/ と再現される可能性を述べた．しかしながら，促音化 ([kawta] → /kaQta/) だけでなく母音の挿入 ([kawta] → /kawuta/) も選択されうするため，促音化が起きるとは限らない．そこで削除に対して促音化が発生するという音韻論的な記述を計算モデルが再現可能か本研究では検証する．実験 1 では促音便の再現を検証し，実験 2 ではウ音便の再現を検証する．

加えて，内在時間長の差が完全な削除を引き起こさない条件でも音便変化が起きるのが不明である．仮に母音が削除されない条件で再現できる場合，母音の削除なしで促音化をモデルが再現可能となる．内在時間長の値を検証する目的には音の持続時間の表現に弱い HMM では不適切であるため，本研究では明示的に持続時間をモデリングできる隠れセミマルコフモデル (hidden semi-Markov model: 以下 HSMM と略称する) を用いる．

補足的に，内在時間長の違いをモデルが学習する場合の挙動も考察する．産出時の削除や短縮といったノイズにより高母音の持続時間は短くなるため，それを聞く話者の母音長の知識も更新される可能性がある．そのような学習によるモデルの更新は産出時だけでなく知覚時にも利用する知識として反映されうる．本研究では (i) 内在時間長の短さをノイズとする場合と (ii) HSMM のモデルのパラメータを更新する場合を「学習の有無」という要因として検証する．モデルの知識として高母音を短く更新させたときでも促音化が再現可能

か実験し，内在時間長の差が産出時のノイズかモデルの知識かの検証材料にする．

2 実験 1: 促音便の検証

実験 1 では母音の持続時間を操作し，計算モデルによる促音便の再現を検証した．以下に実験の手順を述べ，結果の報告と議論へ移る．

2.1 モデル

実験ではモデルにデータを与える過程があるため，本研究ではモデルを以下の手順で作成した．まずモデルである HSMM は条件付き確率分布 3 つ (ガウス分布，ポアソン分布，そして遷移確率) から構成される．以下にそれぞれの概要と作成方法を述べる．

ガウス分布は各音における音の尤度 $P(S|c)$ で表現される．本研究では $P(S|c)$ を二次元のガウス分布として作成した．音素ごとの分布を可視化する目的で，各音素のガウス分布から 100 個ずつサンプリングした値を図 2 に示した．例として /a/ の分布を見ると，x が 1, y が -1.5 周辺で最も尤度が高くなっている．次にこの分布を作成した手順を述べる．

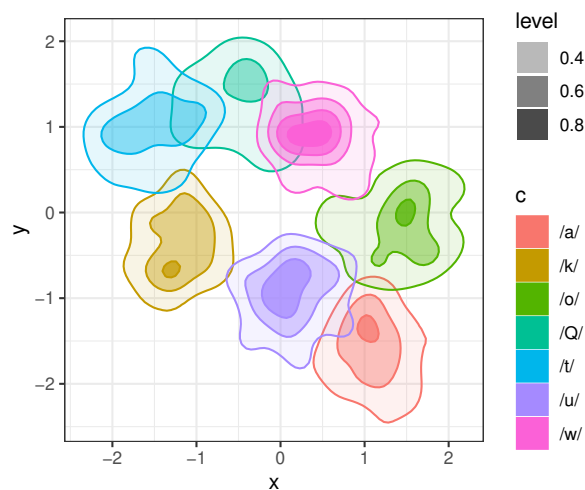


図 2: 各音素で 100 件ずつサンプリングした二次元の値を入力とし，カーネル密度推定した結果で分布の確率を濃度・等高線で表現した．

図 2 のガウス分布を作成するために，まずは各音素を表 1 のように 6 次元で表現した [14]．それぞれの値を標準化した後，UMAP (Uniform Manifold Approximation and Projection) で 2 次元に圧縮した [15]．この過程により，もともと 6 次元だった

情報は2次元となり、それぞれが各音素のガウス分布の平均となる。なお、UMAPと二次元ガウス分布に必要なパラメータは(1) /kawta/に対応するデータに対し /kaQta/が認識される度合いと(2)分散の小ささを最大化するよう最適化した [16].

表 1: 各音素 c を表現する 6次元の数値

c	F1	F2	閉鎖	歯茎	軟口蓋	S.
/a/	800	1600	0	0	0	3
/u/	300	1200	0	0	0	3
/o/	550	800	0	0	0	3
/w/	350	1000	0	0	0	2
/k/	0	0	1	0	1	1
/t/	0	0	1	1	0	1
/Q/	0	0	0	0	0	0

(注: 先行研究 [14] で記述されていた母音・半母音でフォルマントの値を本研究は利用した。なお/w/は各フォルマントで遷移があるものの、実験のため高母音に先行する場合のF2を平均した。最適とされたF1と実験を選択し、子音は弁別可能な調音点の情報を二値で持たせた。聞こえ度 (Sonority) は順序とした。)

HSMMLに必要な確率分布の2つ目である「各音素の持続時間を表現する分布」としてポアソン分布がある。音素ごとのポアソン分布はパラメータ λ をそれぞれ持ち、本研究では母音と促音に9、半母音と子音に3の持続時間を仮定した。本研究で促音/Q/と母音の持続時間を等しく仮定している理由は過度な促音挿入を抑制するためである。仮に子音と同様に持続時間を短く設定した場合、母音を削除した結果の区間で/Q/の尤度を上げる結果となり、促音化に対し過度に有利な設定となる。半母音の持続時間は反対の理由から子音と同じ持続時間に設定した。

HSMMLに必要な最後の確率分布は音素間の遷移確率 $P(c|c_{t-1})$ である。母音を c_{t-1} とした場合、自身へ遷移する確率を0.1、他の音素へ遷移する確率を0.15とした。また子音の/w/, /k/と/t/を c_{t-1} とした場合、自身への遷移確率を0、母音/a/, /u/, /o/への遷移を等しい確率で表現した。最後に促音/Q/を c_{t-1} とした場合、/k/と/t/のみへの遷移を等しい確率で表現した。なお、オンセットに促音/Q/が来ないことは該当する初期確率を0と設定して表現した。

以上のガウス分布、ポアソン分布、そして遷移確率でHSMMLを構成した。次にモデルを利用したデータの生成・識別を含めた実験の手順を述べる。

2.2 手順

実験1では上述したモデルにデータを生成させ (/kawuta/ → [kawta]), そのデータをモデルに識別させて促音便の再現を検証した。生成の過程が依存するのは (i) 生成させる音素列と (ii) 母音長, (iii) 識別側のモデルのパラメータ更新の3点である。以下に例として生成させる音素列を /kawuta/, 実験設定として母音削除の条件の手順を挙げる。

まず、最初の音素/k/に対応するデータとして、/k/のガウス分布 (図2) に基づきモデルは2次元のデータを生成する。ただし、生成するデータ数は/k/のポアソン分布のパラメータである3となる。この過程を各音素で繰り返してデータを生成させる。ただし母音削除の条件である場合、/u/に対応するデータ数は0となる。以上の過程がデータの生成であり、モデルへの入力の子音が3、母音が2あるため27行2列のデータとなる。このデータを識別させると27の音素列となるため、/kaQta/のように重複を削った。

実験1では母音の長さは10パターンあり、母音の長さが0のときに母音削除の水準となる。また、モデル更新の要因では (i) 母音の削除が産出時のノイズとして表現される場合 (production) と、(ii) その削除をモデルが学習する場合 (update) がある。1回の試行では上記の全組み合わせを実施し、全20回の試行を回した。

最後に組み合わせ内での手順の例として、母音削除条件で母音削除を産出時のノイズとして表現する場合の過程を記述する。まず生成側のモデルでは音素列と実験設定にあったデータを生成し、識別側のモデルはデータを音素列へ変換する。この過程を一つの組み合わせで20回繰り返し、音素ごとに識別された回数をカウントする。すると一つの条件では20を上限とした音素列の識別頻度が得られる。

以上の実験の結果、母音の持続時間が0か1かで顕著な差があったため、結果では該当した条件のみを抽出した分析を報告する。

2.3 結果

促音化が起きる場合は識別結果が /kaQta/ となるのに対し、起きない場合は /kawuta/ となる。実験1の分析では母音の削除に対する促音化の有無をもとに計算モデルで促音化が再現可能か検証する。そのため、促音化された回数から促音化されていない回数を引いた値を促音化の指標 y として分析する。分析にあたっては、 y に対する時間長の学習と母音の削除の有無、そして2要因の交互作用を検証した。

図3に条件ごとの促音化の指標を示す。左枠は母音の持続時間長が生成のときのみ反映された条件で、右枠には知覚側のモデルで時間長が学習されたケースを示す。枠内の横軸は母音の有無を二値で示し、縦軸には促音化の指標を示している。

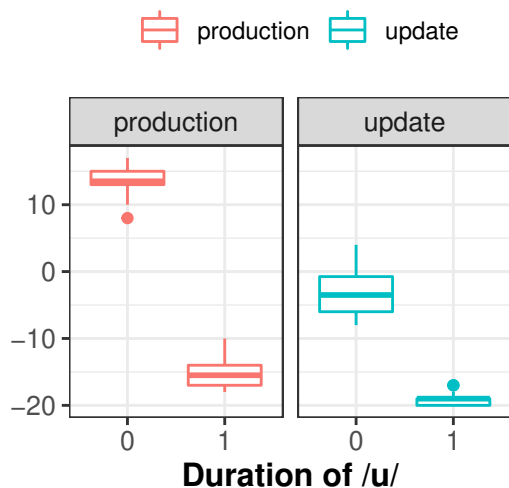


図3: 縦軸の促音化の指標が0のとき、促音化の発生が起きる割合と起きない割合が等しくなる。なお、-20のとき、20回全てで/kawuta/と推論されたことになる。

線形モデルによる分析の結果、母音削除の主効果が見られた ($p < .001$)。したがって、持続時間の学習の有無にかかわらず、削除されている条件で/kawuta/の結果が多かった。また、モデルによる学習 (update 条件) には負の効果があり ($p < .001$)、母音削除とモデル更新の交互作用もあった ($p < .001$)。更新のあるケースでは母音削除条件でも指標が0周辺にあるため、/kawuta/は優先されなくなる。結果的に、母音を削除しモデルを更新しない条件でもっとも促音化が再現された。

2.4 議論

実験1で検証した課題の1つは母音の削除により [kawuta] となった場合、促音化の発生が自明でない点であった。この課題に対し計算モデルは促音化を再現する結果となった。促音化の記述 /kawuta/ → /kaQta/ は3音素間の関係の定義が必要だが、本研究のモデルは2音素間しか表現できない。この表現できない記述を内在時間長による母音の削除という入力の操作で再現した。したがって、計算モデル自体が表現できない音韻論的な記述もモデル間の生成と識別の過程で再現可能となる。次に再現が発生した過程を考察する。

まず母音の削除は [kawta] を生成させた。そこで c_t に /t/ を知覚した時点で c_{t-1} には /Q/ と /w/ が候補として残っている。ここでモデルは母音の錯覚の過程と同様、モデルは $P(c_{t-1} = /Q/)P(/t//Q/)$ と $P(c_{t-1} = /w/)P(/t//w/)$ を比較する。本研究では $P(c_{t-1} = /Q/)$ と $P(c_{t-1} = /w/)$ 、そして $P(/t//Q/)$ は0にならないのに対し、 $P(/t//w/)$ は0となる。したがってより尤もらしい /kaQta/ が再現された。

本研究の課題の2つ目は内在時間長の差が削除を引き起こさない条件でも音便変化が起きるかであった。この点に関し、削除が起きない場合でも促音化が発生する試行は見られた。しかしながら、促音化の指標が0を超えておらず、削除した条件のように多数とはならなかった。

補足的に、内在時間長の違いをモデルが学習する場合の挙動も考察した。内在時間長の表現を産出時 (production 条件) のみでなく知覚のモデルでも更新する場合 (update 条件) は促音化が発生するものの、産出時のノイズとした条件までは発生しなかった。しかしながらパラメータを調整した結果 (付録A参照)、どちらの条件でも母音を削除すれば促音化が発生し、交互作用は見られなかった。したがって、内在時間長の更新は有効な検証とはなっていない。

以上をまとめると、実験1では計算モデルによる促音化の再現を支持した。次に同じモデルとパラメータを利用し、異なる現象であるウ音便が再現可能か検証した。

3 実験2: ウ音便の検証

実験2では子音の持続時間を操作し、計算モデルによるウ音便の再現を検証した。以下に実験の手順を述べ、結果の報告と議論へ移る。

3.1 モデル、手順

実験2では実験1と同じくガウス分布、ポアソン分布、そして遷移確率で HSMM を構成した。また実験1と同様、モデルにデータを生成させ、そのデータを識別させてウ音便の再現を検証した。なお実験1と異なり、持続時間を操作する対象は母音/u/ではなく子音/w/である。データの取得方法も実験1と同じであり、子音の持続時間が0か1かで顕著な差があったため、結果では該当した条件のみを抽出した分析を報告する。

3.2 結果, 議論

ウ音便が発生する場合は識別結果が /kauta/ となるのに対し, 起きない場合は /kawuta/ となる. 実験 2 の分析では子音の削除に対するウ音便の有無にもとづき, 計算モデルでウ音便が再現可能か検証する. そのため, ウ音便となった回数からウ音便とならなかった回数を引いた値をウ音便の指標 y として分析する. 分析にあたっては, y に対する時間長の学習と子音の削除の有無, そして 2 要因の交互作用を検証した.

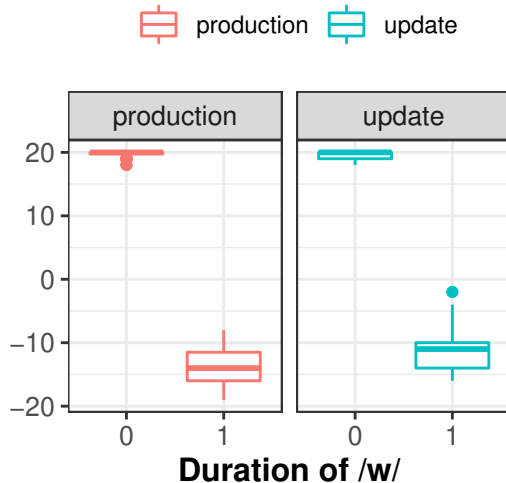


図 4: 実験 1 と同様, 音便変化は再現されている. ただし実験 1 と異なり, モデル更新に関わらずウ音便は顕著に再現されている.

線形モデルによる分析の結果, 母音削除の主効果が見られた ($p < .001$). したがって, 持続時間の学習の有無にかかわらず, 削除されている条件で /kauta/ の結果が多かった. また, モデルによる学習 (update 条件) には効果がなく, 母音の削除とモデルの更新の有無の交互作用があった ($p < .01$). 結果的に実験 1 と同様, 削除の条件で音便変化が再現された.

実験 1 と異なり, 子音の削除により [kauta] となった場合において /kauta/ の再現は明らかである. 次に本研究の実験結果を考慮して計算モデルと音韻論的な記述の関係を議論する.

4 総合議論

歴史的な促音化 (/kawuta/ → /kaQta/) は計算モデルで表現できない記述であった. 音素 /wut/ を抑制して /Qt/ を生成するためには $P(/u//w/)$ か $P(/t//u/)$ が 0 である必要があるものの, どちら

も非現実的である. たしかに 3 音素間の遷移確率による /wut/ の抑制は技術的に可能だが, 2 音素間で得られている経験的な支持は得られておらず, 認知モデルとしては妥当ではない. そこで本研究ではモデルの外部である入力で /kawuta/ → [kawta] となる母音の削除という過程 [12] を要因として操作した. 実験 1 では促音便が再現され, 子音を削除した実験 2 ではウ音便も再現可能であった.

上記の結果から, 計算モデルは自身が表現できない音韻論的な記述を再現可能と結論づける. また, 従来の計算モデルは人の認知機能の解明を目的として利用されてきたが, 通時的な変化を再現する目的でも計利用可能であることを示唆する. その際, 計算モデルで再現可能かは定義に依存しないものの, 今回の計算モデルには少なくとも 3 つの課題が残っている.

第一に, 本研究で扱った現象は限定的である. 動詞の種類は一つだけであり, 同じパラメータのセットを利用して他の動詞の変化も説明可能か検証が必要である. また, 本研究で再現された /kauta/ は再び変化が起き /koota/ となり, 削除という外部要因で /koota/ を再現するのは容易である. しかしながら, 現象ごとに音形データを操作するのは現実的ではないため, 音素列から音形を生成する抽象的な操作が必要となる.

第二に, 本研究のモデルはパラメータの課題が多い. 今回は先行研究に基づきガウス分布やポアソン分布のパラメータを決定したが, その方法では記述が限定的な言語に対して効果が期待できない. また, パラメータの設定で実験結果が変わる点も課題である. 補足 A で示すように他の設定でも再現できる手順を踏むか, 音声コーパスから直接パラメータを取得して解決できる. しかしながら, データが不足している言語に対しては有効な手段がない. したがってパラメータの設定には記述かデータのどちらかが必要になるが, 常に得られるとは限らない点が今後の課題となる.

最後の課題は仮定した持続時間のモデルが本当に必要だったか検証されていない点である. HSMM は HMM と異なり, 音素の持続時間のためだけにパラメータを持つ. 他方, HMM は持続時間専用のパラメータを持たないため, 音素数だけ HSMM はパラメータが多い. この差が音便変化の再現に必要なだったのかを検証すれば過度に複雑なモデルの提案を避けられる. この点の検証も今後の課題としたい.

A パラメータ調整の影響

本研究のパラメータ調整は (1) /kawta/ に対応するデータに対し /kaQta/ が認識される度合いと (2) 分散の小ささを最大化するよう最適化した。この際、母音削除を知覚モデルが学習しない条件 (update) で調整したが、更新する条件で最適化した場合も追加で検証した。仮に実験結果が調整時の設定に依存する場合、母音削除を学習する条件で最適化した場合は実験 1 の結果とは逆の効果が期待される。追実験の結果、逆の効果は見られず、パラメータ調整の結果に依存した結果でないと示唆される。

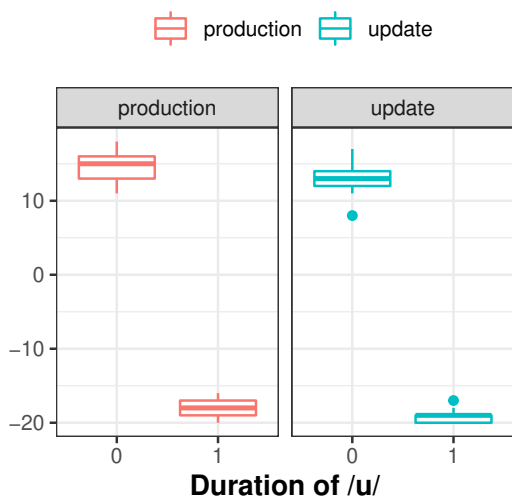


図 5:

参考文献

- [1] Emmanuel Dupoux, Erika Parlato, Sonia Frota, Yuki Hirose, and Sharon Peperkamp. Where do illusory vowels come from? *Journal of Memory and Language*, Vol. 64, No. 3, pp. 199–210, 2011.
- [2] Adriana Guevara-Rukoz. *Decoding perceptual vowel epenthesis: Experiments & Modelling*. PhD thesis, Ecole Normale Supérieure (ENS), 2018.
- [3] Thomas Schatz and Naomi H Feldman. Neural network vs. HMM speech recognition systems as models of human cross-linguistic phonetic perception. In *Proceedings of the Conference on Cognitive Computational Neuroscience*, 2018.
- [4] Jian Gong, Martin Cooke, and Maria Luisa García Lecumberri. A quantitative model of first language influence in second language consonant learning. *Speech Communication*, Vol. 69, pp. 17–30, 2015.
- [5] 岸山健. 母音の錯覚から見る並列処理の心理的実在性. 言語・音声理解と対話処理研究会, Vol. B5, No. 01, pp. 6–11, 2021.
- [6] Patricia K Kuhl. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & psychophysics*, Vol. 50, No. 2, pp. 93–107, 1991.
- [7] Francisco Lacerda. The perceptual-magnet effect: An emergent consequence of exemplar-based phonetic memory. In *Proceedings of the XIIIth international congress of phonetic sciences*, Vol. 2, pp. 140–147. Stockholm University Stockholm, 1995.
- [8] Naomi H Feldman, Thomas L Griffiths, and James L Morgan. The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological review*, Vol. 116, No. 4, pp. 752–782, 2009.
- [9] Alexander Kilpatrick, Shigeto Kawahara, Rikke Bundgaard-Nielsen, Brett Baker, and Janet Fletcher. Japanese perceptual epenthesis is modulated by transitional probability. *Language and Speech*, pp. 1–21, 2020.
- [10] G David Forney. The Viterbi Algorithm. *Proceedings of the IEEE*, Vol. 61, No. 3, pp. 268–278, 1973.
- [11] Emmanuel Dupoux, K. Kakehi, Y. Hirose, C. Pallier, and J. Mehler. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 25, pp. 1568–1578, 1999.
- [12] 窪蘭晴夫. 日本語の音声. 現代言語学入門, 1999.
- [13] 前川喜久雄, 籠宮隆之, 小磯花絵, 小椋秀樹, 菊池英明. 日本語話し言葉コーパスの設計 (<特集> 音声研究関連データベースの動向). *音声研究*, Vol. 4, No. 2, pp. 51–61, 2000.
- [14] 粕谷英樹, 鈴木久喜, 城戸健一. 年齢, 性別による日本語 5 母音のピッチ周波数とホルマント周波数の変化. *日本音響学会誌*, Vol. 24, No. 6, pp. 355–364, 1968.
- [15] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- [16] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 2623–2631, 2019.