

発話支援のための予測変換を用いた文字入力方法の検討

Study of Character Input Using Word Prediction for Utterance Support

濱井裕也† 水町光徳†
Yuya Hamai Mitsunori Mizumachi

†九州工業大学
Kyushu Institute of Technology

中藤良久† 松井謙二‡
Yoshihisa Nakatoh Kenji Matsui

‡大阪工業大学
Osaka Institute of Technology

1. はじめに

我々が生活する上で、音声を用いたコミュニケーションは大変重要なものである。しかし、先天性や後天性の障害や事故等により、音声を用いたコミュニケーションが困難になる人は多く存在している。音声を用いたコミュニケーションが困難になる理由としては、聴覚の障害が原因の場合と、構音機能や、発声機能の障害が原因となり発話ができない場合（発話障害）がある[1]。このような発話障害者の主なコミュニケーション方法としては、筆談、手話等が挙げられる。しかし、筆談や手話は会得する為の訓練や筆記用具を持ち歩く必要がある等のデメリットがある。現在では、新しい発話支援システムとして、発話障害者の唇の動きから発話内容を認識するシステムや[2]、情報端末で文字入力を行い、入力された内容を音声合成にかけ発話させる、日本語音声支援による意思伝達アプリ「Globalvoice 音声支援」[3]等がある。

本稿では、情報端末を用いた発話支援を目的とし、文字を入力して音声合成により発話を行う方法をいかに早く、効率的に行うかの検討を行う。特にユーザが入力した文字を元に文字列を予測する変換方法である予測変換[4][5]に着目し、効率的な予測変換を行うための検討を行う。本稿では、予測変換を効率的に行うため、1)日常会話でよく用いられる挨拶文等で構成された定型文辞書と、2)単語における品詞情報を用いたクラス 2-gram を用いる。

評価では、発話を想定した文字入力方法と、クラス 2-gram を用いた場合に予測変換の精度や速度はどのように変化するかについて検討を行ったので報告する。

2. 想定する発話支援システムの全体構成

本システムは、入力を行う入力部、入力部からの入力を元に予測変換を行う予測変換部、予測変換部でユーザが選択した結果を音声合成して出力する音声合成部から構成されている。

入力部は予測変換部にローマ字の情報を送る部分であり、より素早い会話実現に重要な要素である。スマートフォン等の情報端末では、備え付けてある GUI を用いて文字を入力する方法が一般的であるが、本システムでは、Linux 上のライン入力による入力方法を用いたため、iPhone のフリック入力などと比較して、使い勝手がやや劣ると考えられる。予測変換部は、単文節予測変換と定型文予測変換を行う部分である。予測変換は、現在スマートフォン等の情報端末の変換方法として多く用いられている。予測変換部の詳細については3章で述べる。音声合成部は、予測変換後の結果を音声として出力する部分である。音声合成語の音声はユーザの好みにあわせて、ある程度選べるよういくつかのパターンを用意する。ただし、本稿では音声合成部分の評価は行わない。

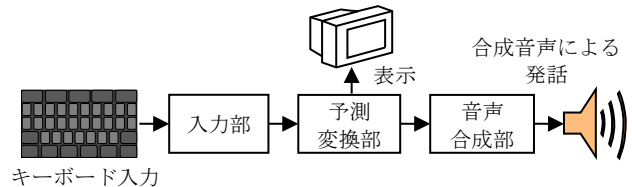


図1:全体構成

3. 提案する予測変換処理の詳細

ここでは、提案する予測変換部の詳細を述べていく。予測変換とはユーザの文字入力から、ユーザが希望している文字列を予測して提示する変換方法である。予測変換部の詳細図を図2に示す。

ユーザが入力した文字は、辞書の検索に用いられる。また、辞書には文字列だけではなく、品詞情報と頻度情報も含まれており、これらは並び替えや次期候補表示に使われる。並び替えの部分には、辞書から検索にマッチした結果の文字列と品詞情報、頻度情報が送られ、頻度の高い順に並び替え、表示、選択部にてユーザに提示される。クラス 2-gram には前回の出力結果から、品詞情報のみが抜き取られて送られ、この情報をもとに、頻度情報の変更や、次期候補の表示を行う。次項より、1)辞書、2)頻度情報、3)クラス 2-gram の詳細を述べていく。

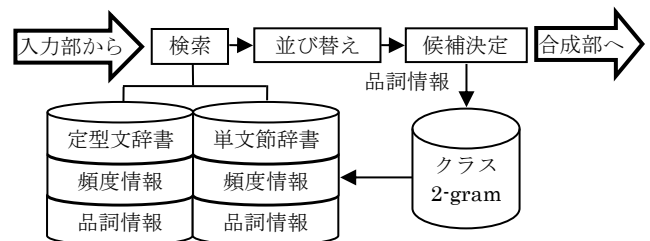


図2:予測変換部の詳細図

3.1 予測変換辞書構成

辞書は単文節辞書と定型文辞書があり、入力文字を各々の辞書で検索にかけることにより、文字列の候補を得ている。それぞれの辞書について以下に述べる。

3.1.1 単文節辞書

単文節辞書の作成には主に SKK 辞書 (M サイズ) [6] を用いた。しかし、SKK 辞書単体では名詞の候補を多くすることはできるが、その他の品詞が少ない。そこで、英会話学習サイトの例文の日本語訳を「茶筌」[7]を用いて形態素解析し、解析後の文字列とその品詞情報を辞書に追加することで内容を充実させた。形態素解析とは、文章を言語で意味を持つ最小単位に分解し、それぞれの品詞について解析を行うことである。これにより単文節辞書の総エントリ数は 13,798 となった。以下、単文節辞書の構成の詳細を述べていく。

単文節辞書にはそれぞれの単語について 1)番号、2)検索用の読み、3)結果の文字列、4)品詞情報が登録されている。表1に単文節辞書の具体例を示す。1)は頻度情報の変更のために用いられる数値である。2)はすべてローマ字で表記されており、ユーザの文字入力もすべてローマ字に限定している。ユーザの文字入力をローマ字に限定した理由は、少ない文字入力数で候補を表示することができ、更に入力誤りの訂正も行い易いからである。2)の作成には「kakasi」[8]を用い、必要入力数を抑える為にすべて訓合式に統一した。最終的に検索結果としてユーザに提示されるのは3)の文字列であり、検索結果にマッチした文字列は候補として最大5つ表示される。4)の品詞情報は、頻度情報の変更と、次期候補表示に利用される。

表1:単文節辞書

番号	検索用の読み	結果	品詞情報
5992	katamuki	傾き	名詞
5993	katamusubi	片結び	名詞
5994	katame	片目	名詞
5995	katamen	片面	名詞
...

3.1.2 定型文辞書

定型文辞書とは、挨拶文などの長い文字列が含まれた辞書である。表2に定型文辞書の具体例を示す。定型文辞書を単文節辞書とは別に作成した理由としては、決まった挨拶などの文を入力する際に、単文節辞書単体では入力に手間がかかってしまうと考えたからである。

定型文辞書には、日常会話でよく使うと予測できる文を中心に、日本語学習者向けサイトの例文を用いて作成した。例文数は約550文である。辞書自体の構成は単文節辞書と変わらないが、品詞情報の部分が単文節辞書のそれと異なっている。定型文辞書の品詞情報は、語頭に用いられる文は「その文の語尾」、語尾に用いられる文の場合は「その文の語頭」の品詞情報となっている。例として、「であると考える」という文は語尾に接続する文であるので、品詞情報は語頭の「で」の品詞である助詞となる。ただし、語尾、語頭等関係がない言い切りの形の文の場合は、品詞情報は感動詞として登録した。

表2:定型文辞書

番号	検索用の読み	結果	品詞情報
0	gomennasai	ごめんなさい	感動詞
1	suimasan	すいません	感動詞
2	mousiwakenaidesu	申し訳ないです	名詞
3	owabisimasu	お詫びします	名詞
...

3.2 頻度情報

頻度情報は単文節辞書、定型文辞書それぞれにテキストデータとして存在している。表3のように、頻度情報には1)番号、2)頻度が含まれており、2)の数値の大小を用いて、ユーザに表示する予測変換結果の並び替えを行う。

具体的には、ユーザによって発話内容が決定されると、決定された文字列の「番号」と同じ番号の頻度が1増加する。これにより、選択されることが多い文字列の頻度は多

くなり、上位に出現しやすくなる。しかし、頻度情報に基づき並び替えられ、表示された候補の中にユーザの意図している文字列が含まれていない場合も考えられる。その場合は、候補を選択せずに決定キーを入力することで、現在表示されている候補の頻度をすべて半減させることができるように工夫している。これにより、もう一度同じ入力を行った場合に表示される候補が変動し、自分の意図する結果が表示される確率が高くなると考えた。このように頻度は、ユーザの癖を学習するために変動するようになっている。

表3:頻度情報

番号	頻度
0	1
1	5
2	8
...	...

3.3 クラス2-gramを用いた変換方法

ここでは、ユーザの意図する結果が候補に出る確率を上げるために使用したクラス2-gramについて説明する。

3.3.1 2-gram

2-gramは、2つの文字列に対して確率を定義するものである[9]。2-gramの具体例を「私は海が好きだ」という文章について示す。この文章を形態素に分解すると、「私/は/海/が/好き/だ」となる。さらにこの結果を2-gramについて分解すると、「私/は」「は/海」「海/が」...と表4のようになる。

2-gramでは、「私/は」のような2個を単位とする文字列の組み合わせを「共起」関係にあるとする。このとき、「は/海」のような意味を持たない可能性が高い共起を「ノイズ(ノイズデータ)」と呼ぶ。また、テキスト全体での任意の「共起」が現れる回数を「共起頻度」と呼ぶ。

この共起頻度を数値化してその結果を集計することで、「よく使われる表現から推測されるテキストの傾向」のような、テキストの持つある側面を把握することができる。

表4:2-gramの例

私	は	海	が	好き	だ
私/は					
	は/海				
		海/が			
			が/好き		
				好き/だ	

3.3.2 クラス2-gram

通常の2-gramは単語の共起関係を用いるのに対し、クラス2-gramは品詞の共起関係を用いる。そのため、通常の2-gramに対してデータ総数が圧倒的に少なく、利用する際の処理の負担も少ない。本稿のシステムは情報端末上で動作することを目的としているためクラス2-gramを用いた。以下、クラス2-gramの詳細について述べていく。

まず、日本語の品詞について表5にまとめる。ここで、自立語とは単独で文節を構成できる品詞のことであり、付属語とは単独で文節を構成できない品詞のことである。本来、日本語の品詞は表5の10種類であるが、形容動詞については品詞として認めない場合もあるため、本稿では形容動詞を抜いた9種類とした。

表 5:日本語の品詞

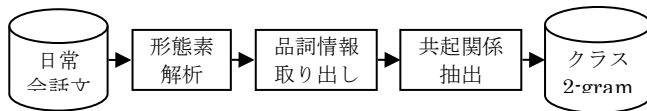
活用の種類	自立語	付属語
活用あり	動詞	助動詞
	形容詞	
	形容動詞	
活用なし	名詞	助詞
	連体詞	
	副詞	
	接続詞	
	感動詞	

表 6:実際のクラス 2-gram

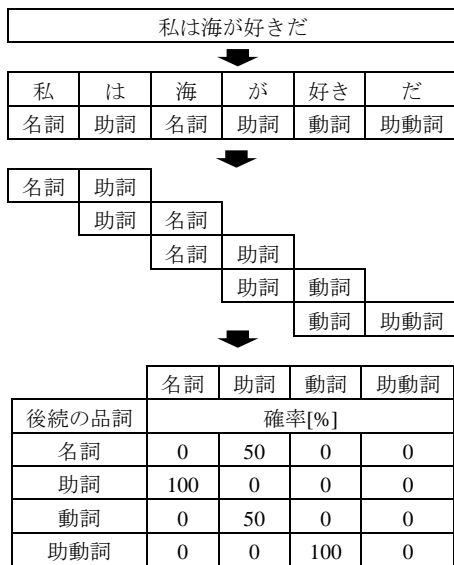
後続の品詞	名詞	動詞	形容詞
	確率[%]		
接続詞	0.061318	0.117256	0
名詞	19.56055	13.30858	32.83396
動詞	7.235565	7.777995	16.22971
助動詞	11.17016	46.82431	27.8402
感動詞	0.132856	0.175884	0
助詞	60.98109	30.03713	20.09987
副詞	0.224834	0.390854	0.249688
接頭詞	0.194175	0.234512	0.249688
形容詞	0.255493	0.605824	1.498127
連体詞	0.183955	0.527653	0.998752

クラス 2-gram の実際の作成方法と具体例を図 3 に示す。通常の 2-gram が形態素解析後の文字列の共起関係を調査するのにに対し、クラス 2-gram は形態素解析後の結果から品詞情報のみを取り出し、品詞の共起関係を調査し作成する。これにより、品詞の並びの規則性を得ることができる。

表 6 は図 3 で示した処理を実際に日常会話例文 4000 文に対して行い作成したクラス 2-gram の一部である。このクラス 2-gram を参照することで、それぞれの品詞の後にどの品詞がくる確率が高くなるかがわかる。この品詞間の確率を用いた機能については以降の項で述べる。



(a) 作成方法



(b) 具体例

図 3:クラス 2-gram の作成方法と具体例

3.3.3 次期候補表示

先ほど作成したクラス 2-gram を用いた機能の一つ目として、次期候補の表示がある。次期候補表示は、ユーザが発話内容を決定後、システムが自動的に次の文字列を予測し、ユーザに表示するものである。

次期候補として表示させるのは、出現頻度が高い品詞二種類のみである。例えば、前回の入力「名詞」であった場合は表 6 より、次期候補として「助詞」「名詞」が表示される。特に助詞は確率が 60%と高いため、助詞を優先して表示させる。

3.3.4 頻度情報変更

クラス 2-gram を利用した機能の二つ目として、頻度情報の変更がある。頻度情報の変更には、前回の出力結果の品詞情報を利用する。例として、前回の結果が「私」であった場合について述べる。まず、「私」は「名詞」なので、表 7 の名詞の列からそれぞれの品詞が後続する確率を読み取る。そして、その確率を直接それぞれの品詞を持つ文字列の頻度に乗算する。これにより、今回の場合は必然的に「助詞」が他の品詞よりも候補として表示されやすくなり、確率の低い「接続詞」は候補としてあらわれにくくなる。

4. 評価実験

ここでは、今回提案した予測変換について行った評価実験の手法と、結果を示す。

4.1 評価実験方法

本稿では、評価実験として「入力に必要な入力の数 (100 文字)」と「入力時の速度 (1min)」の 2 つについて行い、それぞれの結果を「入力必要数の比較」と「速度の比較」に示した。また、特に今回提案した「クラス 2-gram を利用した追加機能」を予測変換に追加した場合どのような効果があるかについて注目して評価を行った。

実験に用いた文章は「第 86 回日本語ワープロ検定試験準 2 級」の検定問題を使用し、比較対象は、連文節変換方法として主に使用されている「Microsoft Office IME 2010」と、現在広く普及しているスマートフォン「iPhone」の予測変換を選んだ。また、本予測変換と Microsoft Office IME の入力には Microsoft 社の「Wired Keyboard 600」を使用し、iPhone の入力は本体の GUI をそのまま用いた。

それぞれの評価項目ごとの被験者数を表 7 に示す。

表7:評価項目ごとの人数

評価項目	被験者数
入力必要数の比較	1名 (筆者本人)
速度の比較	5名 (筆者含む)

4.2 実験結果

4.2.1 入力必要数の比較

入力必要数比較結果を表8に示す。本予測変換は、単体で評価した場合とクラス 2-gram を追加して評価した場合とで別の結果として示した。表8から、本予測変換は他の入力方法に比べて入力必要数の大幅な減少がみられる。特にクラス 2-gram を追加した場合は、従来のキーボード入力の半分の入力数で入力が可能となり、iPhone に比べても4割の減少がみられた。

表8:入力必要数の比較結果

入力手法	入力数 (/100文字)
本予測変換	119
本予測変換+クラス 2-gram	102
iPhone	171
Microsoft Office IME	221

4.2.2 速度の比較

文字入力の速度を比較した実験結果を表9に示す。入力速度の数値は、大きいほど速く文字入力できることを表しており、入力必要数の結果と同様に、本予測変換は単体の場合とクラス 2-gram を追加した場合とで別の結果として示した。表9の結果から、筆者以外の被験者については、本予測変換を用いても、他の入力手段の入力速度に勝つことはなかった。また、クラス 2-gram を追加した場合に、追加しない場合よりも速度が遅くなってしまった被験者もいた。

表9:文字入力速度の比較実験結果

入力手法	入力速度 (/min)				
	被験者				
	1	2	3	4	筆者
本予測変換	41	29	42	41	81
本予測変換+クラス 2-gram	35	33	34	57	85
iPhone	84	65	79	67	77
Microsoft Office IME	101	57	93	76	98

5.考察

本予測変換を用いることで従来の入力方法に比べて入力必要数の減少がみられた。これは、定型文と単文節の辞書を混在させたことや、クラス 2-gram による頻度情報の変化、次期候補予測の実装による効果であると考えた。特に、クラス 2-gram による頻度情報の変化、次期候補予測を用いた場合と用いない場合では必要入力数に1割程度の差がみられ、このことからクラス 2-gram を予測変換に追加することは変換効率を上げるのに有効であると考えられる。

しかし、文字入力速度では従来の入力方法よりも劣っていた。これは、実験終了後のアンケートに「予測変換後の結果が見えにくい」と被験者全員が答えたことから、Linux 上のライン入力が原因と考えられた。事実、Linux 上での

入力に慣れた筆者のみ、入力速度が iPhone の結果よりも速くなった(筆者は iPhone ユーザではないため入力に慣れていない)。

また、本予測変換の特徴の一つである「訓令式ローマ字入力」も原因であると考えられる。訓令式ローマ字入力は変換誤りや、その修正時間、入力時間を減少するために実装されたものであるが、「自分の入力がローマ字で表示されると確認しにくい」や、「訓令式での入力は慣れていない」等の意見もアンケートに見られた。

さらに、単文節辞書の利用が被験者に「入力が難しい」と感じさせた可能性がある。これは、普段我々は会話する際に文節を意識することはほぼ無く、与えられた文章を文節に区切ることも意識しない限りは難しい。

6. 結論

我々は、先天性や後天性の構音障害により発話が困難な人のための文字入力を用いた発話支援システムの検討を行った。具体的には、普段持ち歩くことが多いスマートフォン等の情報端末で文字入力を行い、入力された文字列を予測変換にかけ、音声合成を用いて変換結果を発話させるというシステムである。本稿では、予測変換効率の向上のため、日常会話文を解析して作成したクラス 2-gram を使い、次期候補表示機能と頻度情報変更機能を作成し、予測変換に搭載した。

評価実験の結果、入力必要数が従来の入力方法より4割減少し、変換効率の向上が見られた。しかし、入力速度は従来法には及ばなかったため、今後はクラス 2-gram の改良や、より会話に近い辞書の作成、GUI の整備等を行っていく予定である。

謝辞

本研究の一部は、文部科学省学術研究助成基金助成金 (NO.24500664) の援助を受けて行われた。

<参考文献>

- [1] 厚生労働省社会・援護局障害保健福祉部企画課、平成18年身体障害児・者実態調査結果、pp.4、pp.25
- [2] 若手 ICT 研究者等育成型研究開発、発話障害者のコミュニケーション支援のための携帯電話用読唇システムの開発、九州工業大学大学院 齊藤剛史
- [3] Globalvoice 音声支援 HOYA サービス株式会社
- [4] 天野 真家, 森 健一、漢字・日本語処理技術の発展: 日本語ワードプロセッサの誕生とその歴史(<連載>日本の情報処理技術の足跡)、pp.3-5 情報処理 43(11)、1217-1225、2002-11-15
- [5] Toshiyuki Masui. POBox: An Efficient Text Input Method for Handheld and Ubiquitous Computers. In Proceedings of the International Symposium on Handheld and Ubiquitous Computing (HUC'99), pp. 2-3, September 1999.
- [6] SKK Openlab <http://openlab.ring.gr.jp/skk/dic-ja.html>
- [7] 形態素解析ソフト <http://chasen-legacy.sourceforge.jp/>
- [8] 漢字かなまじり文をひらがな文やローマ字文に変換することを目的として作成したプログラムと辞書の総称 <http://kakasi.namazu.org/>
- [9] 山田崇仁、N-gram 方式を利用した漢字文献の分析、立命館白川静記念東洋文字文化研究所紀要 (1)、1-23、2007-03、立命館大学白川静記念東洋文字文化研究所、pp1-3