

# 「対話のこぼ」の実時間性を可能としている構造 —話すと同時に消えるのに精神的負担を感じない構造を探る—

市川 薫 (千葉大学 大学院融合科学研究科, ichikawa@1964.jukuin.keio.ac.jp)

## Spoken language structure that enables real-time dialogue: In search of light mental burden

Akira Ichikawa (Graduate School of Advanced Integration Science, Chiba University, Japan)

### 要約

ヒトであることの最大の特徴の一つは「言語」を持つこととされているが、その言語の原点は、話すと同時に（実時間性）消えてゆく（揮発性）にもかかわらず楽に対話ができる（軽精神的負担性）「対話のこぼ」にある。聴覚言語の音声、視覚言語の手話、触覚言語の指点字などの共通性や非共通性、それらを支えている諸現象など様々な角度から分析を行ない、日本語音声を例に実時間対話を可能にしている構造を明らかにする。脳の省エネの必要性が軽精神的負担性で実時間性を可能にしているという仮説も併せ、これらの分析結果を俯瞰的に見ることにより、実に巧妙なシステムが浮かび上がる。この構造は、ヒトに優しい実時間対話型ヒューマンインタフェース実現などの工学応用が期待される。

### Abstract

One of the greatest characteristics of human beings is to have a “language”. The origin of language lies in a face-to-face real-time dialogue that can be lightly performed mentally, even though it disappears as soon as it is spoken. We analyzed from various view-points such as commonality and or non-commonality of auditory language (voice), visual language (sign language), and tactile language (finger braille), and also various phenomena that support them. As a result, we clarified a structure that enables real-time dialogue by taking Japanese spoken dialogue as an example. We also examined the hypothesis that the need for energy saving in the brain enables a light mental burden and real-time nature of language. By looking at the analysis results from a bird’s-eye view, a truly sophisticated system emerges. From this structure, engineering applications such as the realization of a human-friendly real-time interactive human interface are expected.

### キーワード

対話のこぼ, 実時間性, 揮発性, 軽精神的負担性, 多角的アプローチ

### 1. はじめに

AIの進歩による音声認識の性能向上を利用した対話システムが注目されているが、実時間に近い応答特性は、コンピュータの高速化や大容量化、通信技術の発達を利用し、疑似的に実現しているのが実態であろう。ヒトの対実時間話を可能にしている仕掛けは、従来の多くの研究が取り上げている確率的視点や言語理論からのアプローチでは必ずしも成功していない。本稿では対話における発話を実体化している物理信号の中から手掛かりを探り、一つのモデルを示す。

我々の周りには文字情報が溢れており、言語を考える場合、どうしても「書記言語」のイメージとなることが多い。しかし日常対話では「話しこぼ」を使っており、話す内容に思いを巡らすことはあっても、話すこと自体は無意識と言って良いほど気楽に使っている。おそらく地上生活を始めた人類の祖先は、捕食動物からの危険を避けるために集団で活動し、相互に一瞬に意思を伝えること（実時間性）が必須であることから生まれた特性であろう。従って言語は元々対話に用いられる「話しこぼ」が言語の原点であり、その中に言語の本質的特性が多く存在していると考えられる。この点に注目し、本稿では、言語一般と区別するために「話しこぼ」を「対話の

こぼ」と表記し（以下さらに「こぼ」と略す場合がある）、負担を感じさせない対話を可能としている特性を多角的に検討する。なお「こぼ」というと「音声」のみを連想するが、ここで「多角的」というのは、「手話」なども含むという意味である。

対話は交換される情報（話す内容など）と交換手段からなっている。前者は主に言語記号レベルの情報であり、後者は主に音声波形などの物理信号レベルの情報からなっている。言語記号情報にも、例えば挨拶や質問と返事の関係など情報交換のための仕掛け（隣接ペアなど）が存在するが、実時間性の機能は小さい。本稿では実時間性<sup>(1)</sup>を実現する機能を主に担う物理信号レベルに注目する。極めて多様な領域と関連する課題であるが、優れた数多くの先行研究の知見と合わせて俯瞰的に見ることにより、実に巧妙な有機的組織に構成されていることが浮かび上がってくる。

その巧妙なプロセスをどの様に実現しているのか、そのプロセスをヒトはどのように獲得したのか（自然言語の獲得）、乳幼児は短期間でどのように言葉を学習するのか、など今後の研究に期待するところも大きい。

なお話す内容や特に日本語でのスタイルに関しては極めて多様な要素が関係し、主に語用論などの分野で扱われており、本稿の対象からは外すことにする。

## 2. ことば

### 2.1 言語としての条件

「ことば」は、意図を伝える道具として、意味表現機能を持つ単位と、その組み合わせで文意(命題)を表現する階層構造を持っていることが基本的に重要である。本稿では音声日本語を対象に検討を進めるが、英語や手話などの対話のことばも同様と考える。

#### 2.1.1 2重分節性

言語は意味表現機能を持つ最小単位である「形態素」<sup>(2)</sup>(第1次分節)と、それを構成している意味を持たない有限種の最小単位である「音素」<sup>(3)</sup>(第2次分節)からなっている。この構造により無限の表現力を持つことが、ヒトの言語の特徴である(Martinet, 2008)。以下本稿ではこの視点からの文法を理論的文法<sup>(4)</sup>と仮称する。

なお形態素の種類や、その組み合わせ方は言語により異なる。コミュニティを囲む環境の違いなどにより生活に必要な概念が異なる為である。

#### 2.1.2 階層構造

言語は、言語学(理論的文法)の視点からは「音素→形態素→文→談話→対話」のように階層構造を持ち、より複雑な情報を表現することが出来ると考えられている。その具体的な組み合わせ方は言語により異なる。

一方、物理レベルからの視点(話しことば)では理論的文法とは異なり、実体としての階層構造に「文節」<sup>(5)</sup>(アクセント節とも)や音節(単独で発声可能な最小単位)という単位を考え、「音→音節→文節→文→談話→対話」という見方がある(日本語の学校文法など)。

本稿の大きな特徴は、「対話のことば」の本質を解明するためには後者の視点が本質的であると考えている点にある。なお手話などには文法構造に未解明な点も多いため、文節相当の単位として「文節レベル単位」と仮称することとする。

#### 2.1.3 恣意性

スイスの言語学者フェルディナン・ド・ソシュール(F. de Saussure)が導入した概念であり、言語の表現形式(signifiant)と意味内容(signifié)の間には必然的な結びつきがないという特徴を指す。

#### 2.1.4 転位性、創造性、生産性

言語は、思考を豊かにする機能を備えている。そのためには転移性・創造性・生産性が重要である(石黒他, 1996)。転位性とは、時間的・空間的に離れた場所で起こる出来事について叙述することのできる特徴である。また、創造性・生産性は、今までに一度も話題にしたことのない内容や、事実と異なる内容(夢、嘘、おとぎ話、など)を述べることのできる特徴である。

#### 2.1.5 母語

言語獲得臨界期までに獲得した言語という視点と、アイデンティティとしての自己の言葉という社会的視点の異なる2種類がある。本稿では主に前者を対象とする。なお臨界期をどの様に定義するかについては様々な視点が考えられ、各臨界期前後や非母語話者との比較からも有益な情報がもたらされる。

## 2.2 ことば(メディア)と知覚(モダリティ)

対話のことば(メディア)として知覚(モダリティ)できる感覚には、聴覚、視覚、触覚が存在する。触覚は多様な感覚(圧力、力覚、痛み、温度、など)が組み合わせられている。「味覚」や「嗅覚」は化学的感覚で応答時間特性などが実時間コミュニケーションには向かないため利用されていない。<sup>(6)</sup>したがって、具体的にモダリティとメディアの組合せとして成立する「ことば」としては、聴覚言語(音声)、視覚言語(手話)、触覚言語(指点字<sup>(7)</sup>、触手話、触指文字、握り点字など)が主に用いられている(表1)(市川・長嶋, 2021)。

表1: 主な「ことば」の構造

モダリティ	聴覚言語	視覚言語	触覚言語			
メディア	音声	手話	指点字	触手話	触指文字	握り点字
言語特性	自然言語の体系					
	実時間対話言語					
	揮発性言語					
発信器官	口・声道	手指・身体	指	手指	指	指
受信器官	耳	目	指	手指	指	指
音韻情報	スペクトル形状	手型・位置・動き	指の組合せ	手型・位置・動き	指形状	指の組合せ
プロソディ	時間構造(リズム)					
	パワー	動作の大きさ	打点の強度	動作の大きさ		握り強さ
	表情	非手指動作*	-	身体動作	-	-
	基本周波数	-	-	-	-	-
わたり	調音結合	調音結合	-	調音結合		-

出典: 市川・長嶋(2021)。

注: \* 非手指動作: 顔の表情や身体動作など。手指による手話表現と同時並行的に表出が可能である。

### 3. 対話のことばに見られる特徴

#### 3.1 実時間性、揮発性、軽精神的負担性

先に述べたように、対話では話す瞬間に(実時間性)その言葉は消えてゆく(揮発性)にも拘らず、話す動作自体には殆ど負担を感じない(軽精神的負担性)という大きな特徴を持っている。コミュニケーション上の様々な課題に対処するためには、この特性が何に由来しているかを解明することが、その出発点となる。

#### 3.2 記号レベルと物理信号レベル

「ことば」は時間軸に沿って展開する情報で、表現対象と表現手段からなり、後者はさらに記号レベル情報(言語情報)と物理信号レベル情報(音声波形など)からなる。言語情報は、物理レベルの信号の存在を前提とした記号レベルの情報により表現され機能する。物理信号レベルの情報は主に音韻情報(音声では主に周波数スペクトル構造)とプロソディ<sup>(8)</sup>からなる。後者はさらに言語的プロソディと情動的プロソディからなり、両者の処理プロセスは異なると考えられている(皆川他, 2011)。

#### 3.3 伝達内容情報と伝達支援情報

「対話のことば」の持つ情報には、伝える情報内容(伝達内容情報)を表現する機能と伝達を支援する機能(伝達支援情報)からなる。伝達内容情報は、言語情報では主に言語記号レベルの、発話者情報などは主に物理信号レベルの情報で表現されているが、ことばにより差異が存在する。個々の伝達内容情報は記号レベルか物理レベルで行われるかは、メディアやモダリティの夫々の特徴を活かしており、様々である。

#### 3.4 プロソディと文理解

文を読んでいくと、通常なら文が終わると思われる先に未だ文が続いており、文頭に戻って解釈をし直して読まない文末にたどり着けないような文は、ガーデンパス文と呼ばれ

る。しかしこのような文章でも音声で聞くと、ほとんどの場合一度で内容を理解することができる(市川, 2011)。文字提示では、単語区間の判断や単語の意味把握に有利な漢字仮名混じり文でも、13.9%の理解率に過ぎないのに対し、音声聴取では66.7%であった。この違いは、プロソディ情報が音声には存在することによる。

またプロソディ情報が欠けた音声を聞いて、どの程度内容を聞き取ることができるかを調べた。ほとんどの人が読んだことのない新刊のエッセー集から、音声で約2分半程度となるものを取り上げ、プロソディ情報の組合せがさまざまに欠落した合成音声を聞かせ、後から内容について20問の質問を行い、正解率を調べた。数字を転記しながら聞く条件(二重課題法)に設定し、実験を行った(図1)(北原他, 1977)。プロソディ情報が欠けるにつれて理解度が低下することが判る。手話や指点字でも、定量的実験結果はないが、経験的にほぼ同程度の低下(80%→40%)であった。

### 4. リズム

表1に示したように、モダリティの違いがメディアの特徴に表れているが、プロソディの中ではリズムが共通な情報構造であることが判る。「対話のことば」が時間軸を中心に展開することを考えると、リズム構造が共通の基盤として存在していることを示唆している。

定型発達児の音声喃語と先天性聴覚障がい児の手話喃語の出現時期にあまり差がないことや(鳥越, 2006)、胎内への音響信号は400 Hz程度以下までということから(志村他, 1989)、胎児はプロソディとしての言語のリズムをおもに獲得している可能性が高い。

#### 4.1 一体的リズムと分析的リズム

言語を含むヒトのすべての運動のリズムには、真似には練習を要する約420 ms以上のもの(分析的リズム: analytic、以下図中などではK1と略記)と、練習無しに真似のできる約330 ms以下のもの(一体的リズム<sup>(9)</sup>: holistic、以下K2と略記)

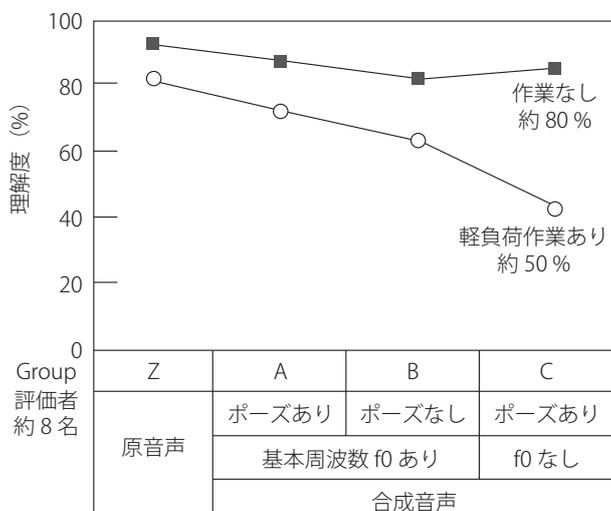


図1：プロソディの効果  
出典：市川(2011)。

の2種類がある(河野, 1998; 2007)。両者は神経心理学的に異質な組織で、分析的処理は一体的処理を前提として同時並行的に処理できるが、分析的処理のために一時的に一体的処理を抑制し、抑制が緩むと一体的処理が復活する生得的リズム処理機構であると指摘されている。

2.1の冒頭で、意図を伝える道具としての「ことば」には、意味表現機能を持つ単位と、それを組み合わせで文意を表現する階層構造の存在が重要であることを指摘したが、後に明らかになるように、一体的リズムが前者を、分析的リズムが後者を実現する基本構造となっている。

#### 4.2 時間基準点

時間的情報である「対話のことば」では、時間構造の基準点が重要になる。

##### 4.2.1 いりわり

「わり」<sup>(10)</sup>は音韻や単語間の移行区間を指す用語である。例えば手話を例に挙げると、後続の単語が決まることにより移行が可能となり、後続単語が開始されたことになる。この事実から「わり」の開始時刻が重要であることに注目し、特

に「いりわり」と呼ぶ。半母音など単独で音韻を形成している区間とは異なる概念である。

##### 4.2.2 事象関連電位 N400

音声では文脈的に逸脱したアクセスの際のミスマッチ反応として概ね400 ms後に脳波の事象関連電位(ERP) N400が観察され(Kutas and Hillyard, 1980)、手話でも観察されている(田中他, 2008)。N400計測の始点は文節レベル区間の始点(いりわりの頭部)で(図2)(大森他, 2009)、心的辞書へのアクセスとミスマッチの判断処理、一体的リズム、分析的リズムの処理が同時並行的に始まることを意味している。

#### 4.3 一体的リズム

##### 4.3.1 一体的リズムの機能

一体的リズムが継続している間は、構成単位(音節、拍など)は形態素レベルを超えて一つにまとめられ、文節レベル単位としての処理が可能となる。

##### 4.3.2 様々な言葉の一体的リズム

一体的リズムに関するさまざまな報告をまとめて示す(表

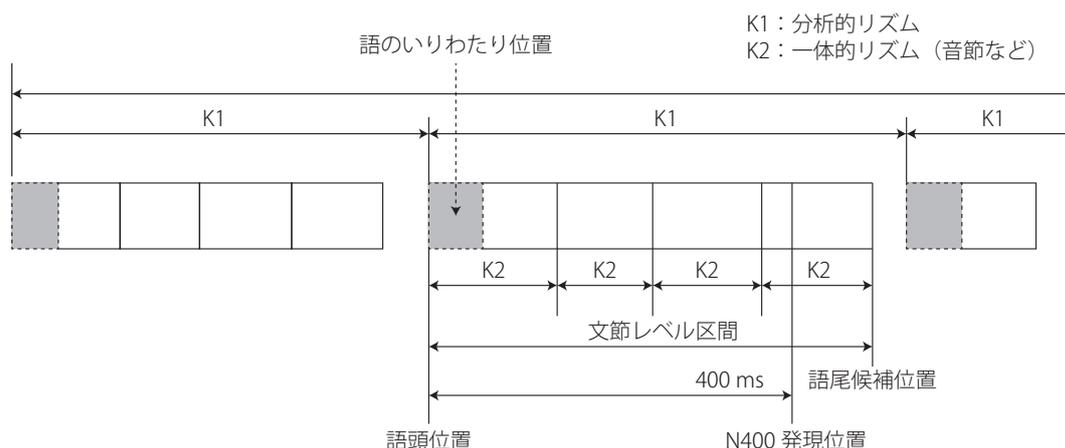


図2：リズムと事象関連電位N400の関係

表2：一体的リズムの例

メディア・モダリティ		一体化リズム(msec) (330以下)	報告者	
言語	音声(聴覚)	日本語	平均145	河野
			141 ~ 169	藤崎・杉藤
		140 ~ 186	Kinoshita et al.	
		200以下	中島	
	英語	214 ~ 269	Tauroza et al.	
		スペイン語 中国語 フィンランド語 インドネシア語	330以下	河野
手話(視覚)	日本手話	33 ~ 283 (平均175)	長嶋・市川	
		約180 (約5.5 Hz)	市川	
指字(触覚)	日本語	330以下	市川・堂坂	
音楽	ピアノ(聴覚)	日本人ピアニスト	330以下(180 BPM以上)	市川・堀内

出典：市川・長嶋(2021)。

2) (市川・長嶋, 2021)。

日本語では2拍続いてもおおむね一体的リズムの約330 msに収まることから、日本語のリズムとして知られている2拍の単位フット (ビート) も条件を満たす (河野, 1998; 2007)。音節や拍の内部の時間構造は言語 (Hiroya, 2013) や方言 (窪 蘭, 2017) で異なる。なお脳磁図 (MEG) を用いた脳機能計測により200 ms程度の $\theta$ 帯域 (一体的リズムの範囲になる) の活動の観測から、音声知覚の単位が音韻ではなく音節や拍のレベルであることが示唆されている (市川・長嶋, 2021; Low and Poeppel, 2007)。

手話では高精細・高精度の手話言語データベース (長嶋 他, 2018) の手話の分析結果は平均175 msであった (市川他, 2019b)。また文字を含まない絵画を見せ、ろう者3名が手話で自由に説明した手の動きを周波数分析したところ、5.5Hz付近 (約180 ms) にピークが現れたが、非母語話者である健聴の手話暦1.5年程度2名の手話文にはピークが見られなかった (市川他, 1998)。なお手の動作は遅いが、調音器官の数が時並列的に複数存在 (両手や顔の表情など) する一方で、組合せの複雑さに対する制約もあり、手話も音声も1分間の表出語彙数はほぼ同じになっていると考えられる。

点字書籍や点字ディスプレイの点字は基本的に文節の分かち書き形式で表示されており、分かち書きの間 (ポーズ) が認知可能なように構成されている。点字に熟達した視覚障がい者が読む状況を観察すると、1つずつデジタル的に読むのではなく、一方の手の指先を滑らして点の凹凸形状を物理信号の波としてアナログ的に読んでいる。他方の手の指先で先読みをしており、予測をしていると思われる。この読み方は音声の聞き手の認知処理手順と類似した構造となっている。書記形式であるが、点字自体の読み取りリズムも熟達者では1分間に約250文字 (240 ms)、最高速度は600～700文字である (牟田口, 2012)。いずれも一体的リズム内に位置している。

#### 4.3.3 音楽とこぼ

語学教育を目的に、代表的ないくつかの「こぼ」のリズムに関する研究として相対的変動率 (nPVI: normalized Pairwise Variability Index) を用いて分析された例がある (Grabe and Low, 2002)。一方こぼと音楽のリズム構造には深い関係が見られている (大串, 2010)。日本語やラテン系など音節間隔の等時性の高い平板的リズムが母語である演奏家の演奏リズムに比較し、変化の大きいストレスリズムの英語などゲルマン系母語の演奏家のリズム変動幅は大きい。そこで日本語母語話者のみではあるが、3名のピアノ演奏者の分析を行った (市川他, 2019a)。結果は等時的で、平均で約330 ms以下 (180 BPM以上) で、「こぼ」の「一体的リズム」の領域に対応していた。ヒトの「こぼ」と「音楽」のリズムの処理系は共通であり、すべての運動のリズムは共通である (河野, 1998) という見解に一致していることを示唆している。

#### 4.4 分析的リズムと文の係り受け構造 (階層構造)

発話は揮発し、発話後に修正することが出来ないにもかかわらず、録音した発話文のプロソディ (分析的リズムを反

映) を分析すると、発話文の構造が現れる。音声では、基本周波数 $f_0$ パターンの変化位置は、藤崎モデルにおける声帯に対する指令パルスの位置 (リズム) に対応するため、 $f_0$ パターンから文構造が得られる (小松・市川, 1992)。手話では、単語への「いりわり」 (大森他, 2009) 開始時点の間隔 (分析的リズム) を短いものから順次結合すると文の構造が現れる (平山他, 2001)。指点字では打点の間隔を短い順に結合すると分析的リズムの時間領域 (420 ms以上) に階層構造が現れる (Miyagi et al., 2005)。 $f_0$ 経路で得られた4か国の音声言語の文構造の例を図3に示す (小松・市川, 1992)。

### 5. 軽精神的負担性と実時間構造

揮発性メディアで実時間対話が楽に実現できている (軽精神的負担性) のは、対話処理の負担が意識に登らないような巧妙な仕掛けが存在しているためと考えられる。そこには、それを可能にしている感覚器や発声器官、脳などの処理構造、「こぼ」の構造などの存在がある。

#### 5.1 ウェーバー・フェヒナーの法則

例えば男女や子供の声の特性は、発声器官のサイズの違いによる共振周波数 (フォルマント周波数 $F$ など) や声の高さ (基本周波数 $f_0$ ) が大きく異なるにもかかわらず円滑な対話が可能である。感覚は対数を近似した特性となっている (ウェーバー・フェヒナーの法則)。典型的な正規化処理としては除算を用いるが、入力信号が感覚器で対数に近似変換されていれば、減算処理で近似可能となり、神経生理レベルでの実時間処理が容易になる。上記の例では、音声の男女の各種特性は絶対値では大きな差があるが、時間特性に注目し、時間方向の差分をとると、特性の違いはほぼ正規化される。

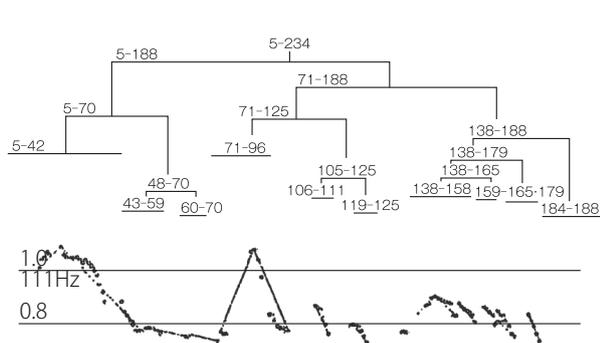
#### 5.2 一体的リズムによる実時間セグメンテーション

連続発声された音声から文節レベル区間が切り出される (セグメンテーション) のかが問題であった。音素連続の確率的データを用いているという従来の見方は、物理レベルの音を記号の音素に変換する処理や、確率的情報を乳幼児が獲得する学習量の問題、発話未まで処理が出来ない、などの課題が存在していた。しかし、本稿で提案するように、一体的リズムに注目することによりこの問題は解決される。一体的リズムが継続している間は一つの文節レベルの単位と見做し、一体的リズムが途切れた時点でセグメンテーションされると考えることが出来る。

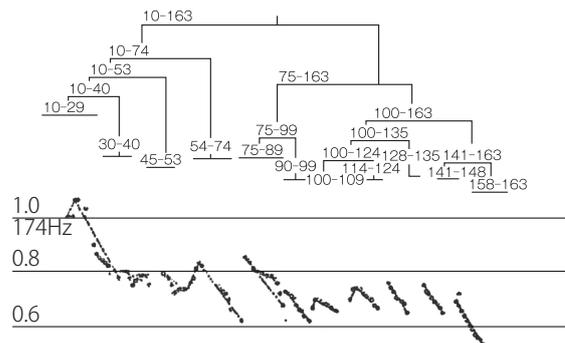
#### 5.3 予定・予告・予期・予測

##### 5.3.1 発話計画

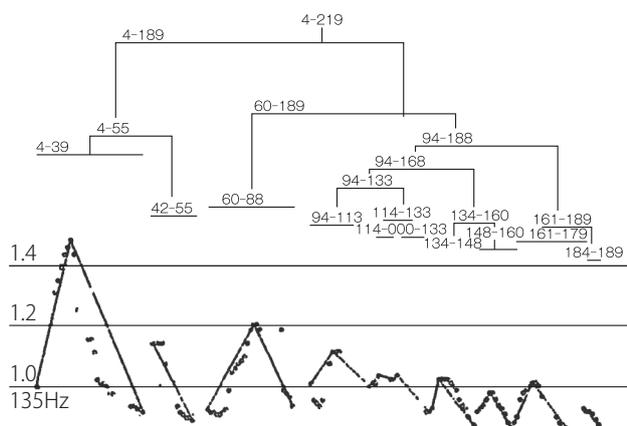
音声、手話、指点字のプロソディから構文木が得られたことは、対話のこぼでは発話の開始時点で発話計画 (Willem, 1993) が立案され、その発話計画による「予定」に従って発話が行われたことを示唆している。「思考は自己との対話である」 (内言) <sup>(11)</sup> という見方に従えば (Vygotsky, 1962)、何を発話しようかという思考プロセスに伴い生成されていると考えられる。予定に従って発話が行われているということは、聞き手の立場からは、予定に基づいた「予告」情報が存在すると



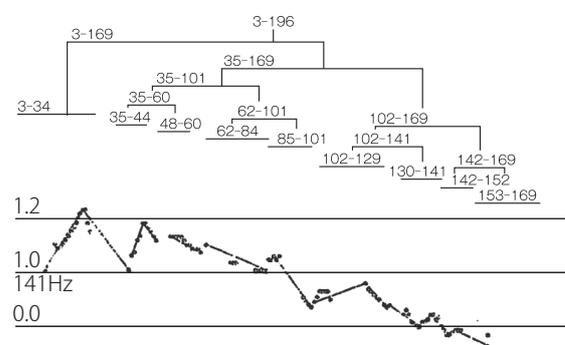
(a) 日本語：音声認識は難しく21世紀でも実現するかどうか分かりません



(b) 米語：Speech recognition won't be solved until the twenty first century.



(c) 仏語：Je pense que la reconnaissance de la parole continue sera difficilement résolue en moins de 25 ans.



(d) 独語：Automatische Spracherkennung wird wohl nicht in den nächsten 20 Jahren gelöst werden.

図3：音声のプロソディから得られた文構造の例  
出典：小松・市川(1992)。

いうことになる。ただし予告情報がどのように存在するかは、モダリティとメディアの特性や、言語構造の違い、コミュニティの習慣などにより異なる。立案とともに予定されると仮定すると、音声では発話開始とともに基本周波数のパターン形状は決まり予告されていることがわかる (Oohashi et al., 2010)。

また実際に大量の実音声を集めた「日本語話し言葉コーパス」を分析した報告では、文節レベルに対応するアクセント成分APを5から30まで含むような長文の基本周波数 $f_0$ パターンを分析し話者毎に正規化すると、平叙文の発話文末ではほぼ同じ値に収束する傾向が示されている (前川・西川, 2019)。この事実も発話開始の時点で、終了時点が予定され、基本周波数が制御されていることを示唆している。実際の音声では基本周波数以外にパワーや時間構造 (分析的リズム等) などのプロソディ情報が同時に存在しているため、予告能力はさらに高くなると考えられる。

### 5.3.2 理解計画

N400は文脈的に逸脱したアクセスの際のミスマッチに反応して生じる。この事実は、聞き手は構文木 (階層構造) を生成し予測している可能性を示唆している。言い換えれば聞き

手は入力から「理解計画」(仮称) のような階層構造を実時間で予測生成し、入力との整合性をチェックしていることになる。

### 5.4 話者交替

母語話者による円滑な対話を見てみると、日本語でも英語でも話者交替では高頻度で発話に重複現象が見られる。重複発話や共話現象<sup>(12)</sup>が生じるということは、聞き手は発話を最後まで聞くことなく予測、理解し (Oohashi et al., 2010)、応答の発話計画を立案する余裕が得られていることを示している。話し手は聞き手が重複して応答を始めたことにより、聞き手が発話権を引き取ったことと了解し継続発話の発話計画を立てる必要がなくなったと判断することが可能になる。重複発話が円滑な対話の実現に本質的に重要であることを示している。

母語対話では、重複発話の生じる先行発話の区間は話者移行的確場 (TRP: Transition-Relevance Place) と呼ばれ、多数の言語の音声対話で重複発話の存在が確認されている (Stivers et al., 2009)。音声では相互の発話が衝突し対話の妨害になるため話者交替の予告を利用し、短いTRPを実現しているが、手話では相互に妨害しないためTRPとしての重複区間が長く

なる傾向にある。日本手話では、話者交替時に文末に旧情報（同一文内で表出済みの情報）を再表出することが多く、この区間で重複が生じる。旧情報表出によりTRPを用意していると考えられる（斎藤他，2009）。

日本語母語音声話者間での通常対話ではTRPより早く重複が生じると母語話者は不自然に感じ、避けることになる（榎本，2009）。このような現象をTRP制約と呼ぶことにする。まったく文法構造の異なる英語でもTRP制約が機能することが観察されている（滝沢他，2013）。

一方、日本語非母語話者の中国語話者や英語非母語話者の日本人話者は、それぞれの非母語の言語レベルの情報に関して十分に学習しているにもかかわらず、TRP制約が機能しないことも実験的に報告されている（市川他，2016）。これらの例から、言語情報自体には予告情報は存在せず、母語話者では言語情報を実体化するために不可分に存在する物理情報のプロソディにあることを示唆している（市川他，2016）。母語話者のTRP制約の獲得時期は、日本語音声の場合、プロソディの予告情報の獲得は3歳程度で見られるが（Kajikawa et al., 2004）、TRPとして利用可能な語彙が獲得されている4～5歳と推定される。TRP制約は非母語話者では獲得が難しいことも示している。なおTRPを文法的に一般的に定義することは困難で、それぞれの言語コミュニティ毎に認知実験で確認することが求められる。

### 5.5 発話実体と理論的文法の階層構造の不一致

「ことば」は音波（音声）や身体動作（手話、指点字など）の物理信号により実体化されるが、理論的文法と発話実体の階層構造は対応しない。「意味」を持たない単位である「音素」に対応する物理レベルの信号は「音」であるが、単独では発音できない「音」（子音など）もあり、発音可能な最小単位として複数の「音」からなる「音節」（例えば子音+母音）という単位が物理レベルでは存在することになる。

そのため、一つの音素のみからなる形態素も存在し、複数の「形態素」を含む「音節」も存在可能となる。顕著な例として

動詞の語幹と活用語尾の関係をみてみると、例えば「書く」が「kak・u」のように音節「く」の中の子音と母音で別の形態素に分かれて対応していることが判る。

さらに日本語では、語と語の関係を示す格助詞などの形態素は、語の本体との形態素列からなる「文節レベル単位」構造を形成する。同様に英語のような冠詞や前置詞などのある言語でも一体的に纏められ、語ではなく「文節レベル」で纏まった発話がなされる。このため理論的文法レベルの階層構造と、実対話を反映している物理的実体レベルの階層構造の間にはズレが生じることになる（図4）。

なお「文節レベル単位」には係り先の具体的情報が不足しているが、「分析的リズム」が実時間で指定している。したがって理論的文法情報（文節レベル単位に内蔵）と階層構造（分析的リズム情報として）は既に反映されていることになり、言語処理は不要となる（図5）（例：「昨日買った本をなくした」の「昨日」は「買った」と「なくした」のどちらにかかるのか？）。

階層構造の各層間の不一致はさらに多くの利点を持つ。一体的リズムによる実時間処理を可能にするだけではなく、機能語を含む文節レベルの構造は語単独よりも長い音素列となるため、組み合わせの確率は低くなり、心的辞書の探索範囲も小さくてすむことになる。この点でも実時間性に有利になっている。さらには、このことは、発音の自由度や認知の揺らぎへの耐性の点でも有利となる。また各層間の交差は、言語情報を反映した調音結合による発話文の一体感を生じ（5.7参照）、自然性向上による精神的負担が軽減されるなどの積極的意義が存在していると見るべきである。

### 5.6 説明モデルとしての文法理論

見方を換えれば、理論レベルの文法は言葉の構造を解釈するための説明モデルということになる。いわば、電子が電磁方程式を解くことなく電子機器の中で機能し、同方程式は電気領域の技術者が機器を設計する上での説明モデルのような位置付けと言えよう。

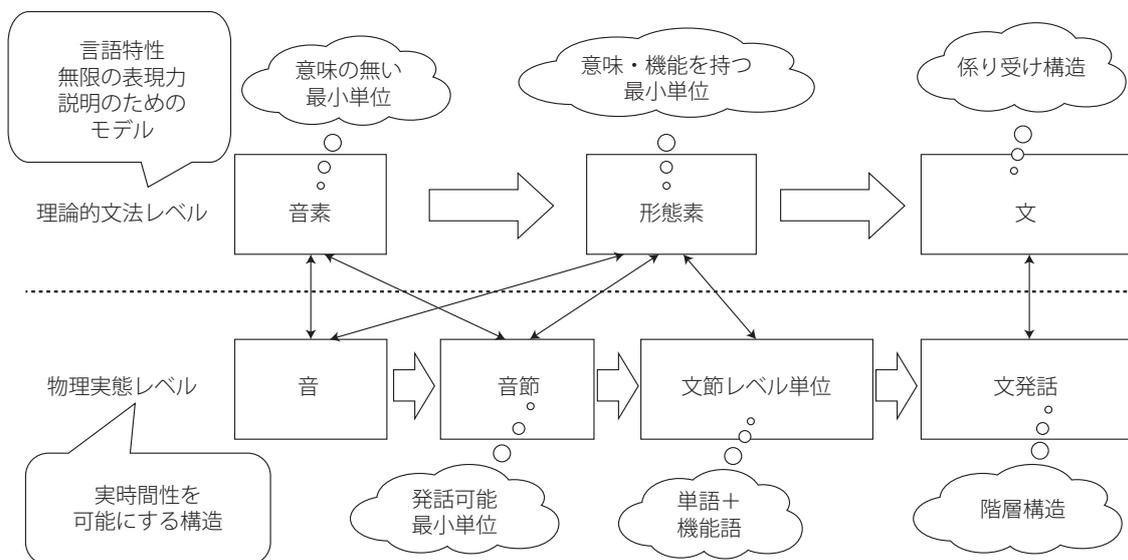


図4：言語理論と物理的実体の階層のズレ

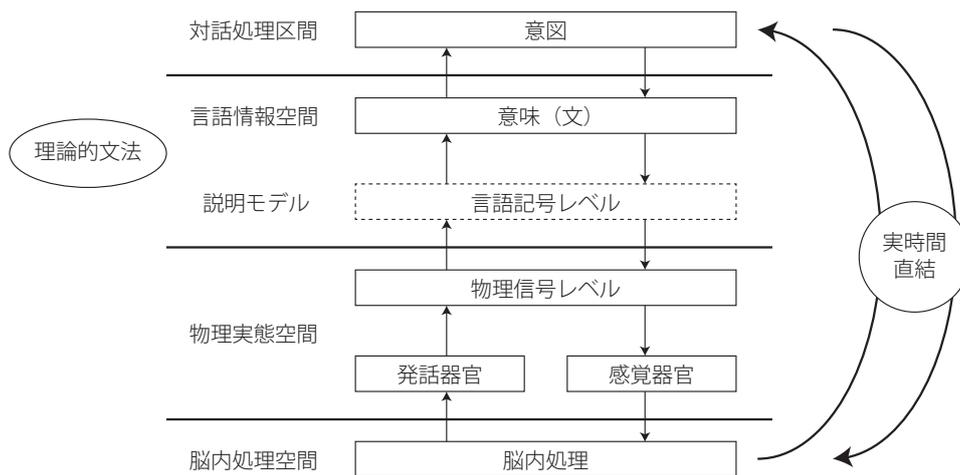


図5：対話のことばの実時間構造

### 5.7 発話運動の自然な制約の重要性

身体的・生理的制約による発話情報の変形はむしろ精神的負担を軽減させる効果が認められる。音声ではスペクトルに、手話は表出位置に怠け現象が現れるが、音声合成や手話で正確に表現すると、受け手はむしろ疲労感を感じる。このことは6.2で触れる運動理論の妥当性を示唆している。

また時間方向では自然で安定な条件である躍度<sup>(13)</sup>最小条件も重要である。例えば、手話CGにおいて同じ手話単語間隔で単純に単語間を滑らかに接続するよりも躍度最小条件を考慮した接続の方がろう者(手話母語話者)にはゆっくりと感じられ、精神的負担が軽くなる(山崎他, 2008)。

### 5.8 発話者情報

語彙の使用頻度の偏りなどの個人差は発話者が既知であるかどうかの頻度情報として心的辞書に反映され、アクセス優先度が上がり、精神的負担軽減と安心感を与える効果も存在することが示唆される。

#### 5.8.1 方言

発話者の大きな情報源は方言である。語彙、概念、プロソディなど、様々な側面に重要な情報が比較的安定に存在している。プロソディには、方言によるアクセントやイントネーションの違いが現れ、また声の高さや大きさ、リズムなどにも個性が現れる(森他, 2014)。

#### 5.8.2 発話器官

個人の発話器官のサイズは固定しており、方言などによっても安定で、個性が保たれる(森他, 2014)。音声では特に鼻腔は個人差が大きく、鼻音は個性が現れる。手話では、方言の影響が大きいこと他に身体的あるいは運動特性の特徴などに、また指文字や触手話などの触覚言語では身体的あるいは運動特性の特徴や手触りなどに個人差が現れる。

#### 5.8.3 意図表現

発話者の強調等の意図や感情は、語彙の選択やプロソディによる表現(時間構造や表現の大きさ・強さなど)の両者の組

合せで表現される(森他, 2014; 高松, 2018)。手話のコミュニケーションでは、表情は非手指動作として大きな意図表現情報を持っている。音声でも基本周波数やその大きさ・強さの他に表情にも表れる。

### 6. 軽精神的負担性と脳の構造

以上に述べてきた事項以外に、精神的負担を軽減する構造として脳の省エネルギーが重要と考えられる。

#### 6.1 省エネルギー構造

ヒトの脳は身体全体の体積の約2%に過ぎないが、消費エネルギーは約20%に達しており、生命の維持活動のためには脳の省エネルギーが極めて重要になる。

- 身長が急速に伸びる時期：  
ヒトの身長は12歳頃から17歳頃に急速に伸びる。これは受胎から10歳頃までの間は脳の発育にエネルギーを優先的に配分しているためと考えられている。
- 母語基盤の獲得時期：  
音声言語では、受胎後約24～30週で聴覚は機能を始め、おもに母親の声を聞いて母語のプロソディの獲得が始まり、誕生後18ヶ月ころまでに母語獲得のテンプレートとして構成されていく(Uchida-Ota et al., 2019)。ほぼ同時期に発話器官である口(声道)のサイズなども急速に発育するが、その後発育速度は半減する(Vorperian et al., 2005)。

これらの現象は、この期間に発話のハード(声道など)やソフト(音韻やプロソディなど)も母語に適応し、脳の負担を軽減させるために、使用の可能性の低い非母語への対応機能は落とされることを示唆している(仮称：母語獲得基盤の臨界期)。

#### 6.2 運動理論

言語知覚には生成系の関与が必要であるという仮説(Liberman, 1967)における主要な理論の一つで、発話器官を制御する脳からの指令が参照されているという考え方であ

る。様々な議論や支持する実験、修正理論などの報告があり、生成系が関与するという基本構造は妥当と考える。

例えば、自己の発話においては発話運動と聴覚からの音声との間の相関は極めて高いと考えられ、省エネの視点からもこの情報を活用していないとは考えられない。また小さな現象であるが、破裂子音では声帯発振現象の口の狭めによる流体力学的負荷変動により基本周波数が変化する（マイクロプロソディ）。合成音声にこの現象を付与すると自然な音質になるが明瞭性は改善されない。この変化は、発話の運動指令は関与していない音響信号レベルでのローカルな現象のため明瞭度が改善されないと解釈される。なお運動理論に関する詳しい議論は関連文献に譲る（廣谷，2017）。

### 6.3 脳の運動制御と言語処理への拡張・転用の可能性

#### 6.3.1 言語処理とブローカ野

ブローカ野が損傷すると発話の障害が出るため、発話の中核と考えられてきたが、さらに様々なことばの処理に関係し、その基盤と考えられてきている。また身体運動の階層制御が、言語のリズムや階層構造処理の前駆体となっている可能性が注目される（Zaccarella et al., 2021）。

#### 6.3.2 運動の階層制御と言葉の階層構造

脳の省エネの視点からも身体部位の運動機能の階層構造処理と言語の階層処理の関連の解明は興味深い課題である。階層処理の共用へと進むと考えるのも自然であろう。音声为非線形な写像の音響空間を経由するのに対し、手話言語は身体運動と直結しており、研究例はほとんどないが、有意義な情報をもたらしてくれることが期待される。

## 7 「対話のことば」の存在領域と共通基盤構造

### 7.1 各メディア特有の現象

夫々の言葉は利点を活かし欠点をカバーするように実現している。例えば、5.4 で述べたように、音声と手話では衝突回

避の戦略が異なっている。

### 7.2 ことばの共通基盤構造

発話意図に基づく発話計画では、一体的リズムと分析的リズムの2種類のリズム構造の組み合わせが共通基盤となり、さらに夫々のメディアの特性に合わせて変換し、表出していることになる（図6）。一方、発話中に意図が変わったり、環境や相手からの応答に影響されたりすると分析的リズムは変動し、構文構造に影響する。

### 7.3 誤解の効用

電話の開発過程で明らかになっているように、実時間性や揮発性により日常対話では了解率は8割程度である。100%でないことが誤解の生まれる原因でもあるが、その誤解が新しいアイデアを生むことが多い。会議を開き、議論することのプラス面を支える現象と考えれば、「対話のことば」の持つ「誤解の効用」と考えるべき特性と言えよう。

## 8 おわりに

### 8.1 まとめ—対話の実時間性を可能にしている構造

言語の原点が「対話のことば」にあることに留意して、検討した結果を以下に纏める。

- 円滑な対話には、語レベルの持つ意味情報と文構造情報の実時間による認知が不可欠である。
- 対話の実時間機能の検討には、言語の特徴を理論的に説明する（理論的文法）2重分節性などの階層構造の視点よりも、発話の実体を表現している音節や文節レベルの情報（自立語と機能語など複数の形態素の組み合わせからなる単位、以下「文節」と記す）などによる階層構造の視点（例えば日本語音声では「学校文法（橋本文法）」）が有効である。
- プロソディ情報、特にリズム情報が不可欠である。
- プロソディのリズム（一体的リズム・分析的リズム）が、話

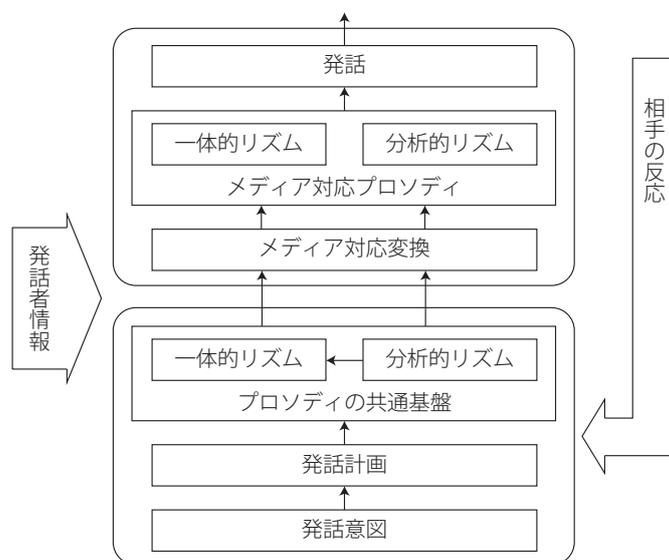


図6：発話の共通基盤と構造

し手と聞き手双方における「対話のことば」の共通基盤となっている。

- 発話文の文頭で発話計画が立案され、発話のプロソディに反映される。
- 発話計画は話し手の予定情報であり、聞き手にとっての予告情報となる。
- リズムは、対話に用いられている夫々のメディアやモダリティの特性を反映した情報に変換され、実体化されている。
- 一体的リズムにより、発話文から文節が実時間でセグメンテーションされる。
- 心的辞書の構造としては、文節を基本とするPDPモデル等が想定される。
- 文節に内蔵している他の文節との関係に関する形式情報（例えば格助詞など）だけでは階層構造情報としては不十分である。
- 一方、分析的リズムは、文構造（階層構造）情報である文節の間の関係の深さ情報を持つ。
- 文節に不足している関係先の具体的位置情報を「分析的リズム」が指定し、補完、実時間で階層構造が得られる。
- 発話文からの文節のセグメンテーションと心的辞書へのアクセス処理に拠る語レベルの意味の認知、分析的リズムにより補完される文構造処理、ミスマッチ判定の聞き手における処理などは同時に開始され、並行して進行する。
- 聞き手は改めての言語の記号処理を要しない。
- 「理論的文法」は言語情報の説明理論である。
- その他に、以下のような実時間性や軽精神的負担性をサポートする効果が存在する。
- 言語の理論的文法の階層構造と物理レベルの階層構造との複雑な対応関係は、発話を言語情報の側面から一体化した自然な発話として纏める効果と、多様なレベル間の組み合わせによる制約からの心的辞書の探索範囲限定効果が存在する。
- 話者情報も心的辞書探索の範囲を限定する効果を持つ。
- 身体的・生理的制約による発話情報の変形（調音結合など）も、精神的負担を軽減させる効果を持つ。
- 母語では、予告情報やTRPが話者交替における重複発話を可能とし、対話者双方の精神的負担を軽減している。

## 8.2 残された課題の例

本稿のモデルの検証には多くの課題が残されている。例えば、促音は無音状態の情報で音声波形の入力は無く、その長さによっては一体的リズムが途絶える可能性が存在する。これに対しては、入力に対して発話器官の制御情報が心的辞書から誘発（運動理論）され、処理されているという仮説が考えられる。逆に言えば、この確認は運動理論の検証にもつながるだろう。

手話に関しては学校文法レベルの構造が未解明で、音声の文節の持つ機能がどのような構造で実体化しているのか、などの課題が残されている。例えば、音声ではメディアの特徴である0に変換されて表出されたように、視覚言語の特徴を活かして動作の大きさ情報を補助的に活用している可能性の検討も必要であろう。文節レベルの単位や、発音可能な最小

単位である「音節」よりさらに小さく、単独では発音が難しい「音素」（第2次分節）などの形成プロセスの検討も必要である。

乳幼児の言語獲得・学習のプロセスも不明な点が多い。先天性全盲ろう児の言語教育の試みでは、言語情報が無くても物理的情報の世界でカテゴリー化（概念化）が可能なが見出されている（梅津他, 2000）。更にモノの概念化からモノとコトをセットとした単命題単位での概念化に進んだと考えられよう。しかし日常生活環境下で表れる様々な現象として、単命題では表現が出来ない高次の命題に対する概念化（階層化）が必要となる。言語や数式に対する階層処理は運動制御の階層機能を拡張し、獲得したと考えられるが、その確認も今後の課題である。

## 謝辞

障がい者や高齢者のコミュニケーション支援（市川・長嶋, 2021）に関する検討を通して、特に工学院大学の長嶋祐二教授には対話の実時間性や手話との関係、ミスマッチ反応のN400等数多くの情報・見解や本稿への貴重なコメントを頂いた。豊田工業大学の原大介教授には言語理論に関して、国立リハビリテーションセンター研究所幕内充室長には脳のブローカ野における言語や運動の階層構造処理に関して情報を頂き、また、ことばのリズムとの類似性に関する音楽の分析には千葉大学堀内靖雄准教授のご協力を頂いた。これらの議論が本稿発想の基盤となった。

## 注

- (1) 実時間性：対話における明確な定義はない。多くのデータは精神物理学の立場からの心理実験データである。心理的には対話に遅れを感じない程度の遅延の範囲であるが、感覚器で受信してから内容を理解するまでの間には意識に登らない膨大なステップの処理が行われている。定型発達の成人が、雑音の無い条件で日常会話のようなタスク下で、了解度85%以上を達成できる遅延条件（電話システムなどの目安）なども参考になろう。
- (2) 形態素：モノやコト、語の関係を示す助詞や前置詞、活用語尾など、言語情報としての機能を持つものなど。
- (3) 音素／音韻：言語レベルの同一の単位に対する名称であるが、「音素」は「意味を区別する働きを持った最も小さな音の単位」として、「音韻」はその言語の「体系」の中での機能に注目した単位として音声機器などのシステムを対象とする工学系などで使われる傾向がある。
- (4) 文法：言語の持つべき機能などを出きるだけ幅広くかつ簡潔に理論的に説明することを目的とするもの（理論的文法と仮称）と、実体の構造を示すもの（例えば学校文法など）がある。
- (5) 文節：不自然にならないように文を区切った音声上の単位で、橋本新吉の用語。主に一つの自立語と機能語（助詞など）からなり、アクセント句とも呼ばれる。言語学の視点からは批判がある。
- (6) 感覚の情報量：一般に人が感覚器官を通じて外界から受ける情報量は、視覚83%、聴覚11%、触覚1.5%、臭覚3.5%、味覚1%と言われている（教育機器編集委員会, 1972）。

- (7) 指点字：点字を構成している6点を左右の手の3本の指に対応付けて鍵盤に見立て、打点して伝える方式。
- (8) プロソディ：単語あるいはそれ以上の時間的に広い範囲に関係し、時間的変化や標準的値からの偏差が持つ物理的情報。
- (9) 引用文献(河野, 1998)では「holistic」に「全体的」という用語を用いているが、第二次分節を「一体に纏め」て第一次分節として知覚を可能にしているリズムなので、本稿では以下「一体的リズム」という用語を用いることとする。
- (10) わたり：元々音声領域の用語で、手話でも流用。本稿では音声と機能が同じ手話の現象を英語にならい「調音」「調音器官」「調音結合」「わたり」などと音声と共通の用語を使う。
- (11) 内言(inner speech/covert speech)と外言(overt speech)：内言とは発話を伴わない思考の道具としての言語、外言とは発話を伴う伝達の道具としての社会的言語、というヴィゴツキー(Vygotsky, L. S.)の学説(皆川, 2015)。
- (12) 共話現象：聞き手が話し手の発話を予測してほぼ同時に同じ発話を行う現象。
- (13) 躍度：加速度の微分値。

#### 引用文献

- 榎本美香(2009). 日本語における聞き手の話者移行適格場の認知メカニズム. ひつじ研究叢書(言語編) 69.
- Grabe, E. and Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Laboratory Phonology*, Vol. 7, 515-546.
- 平山望武・堀内靖雄・市川薫(2001). 日本手話における時間構造の分析. *ヒューマンインタフェース学会論文誌*, No. 3, No. 3, 9-14.
- 廣谷定男(編著), 日本音響学会編(2017). 間くと話すの脳科学. 音響サイエンスシリーズ17. コロナ社.
- Hiroya, S. (2013). Non-negative temporal decomposition of speech parameters by multiplicative update rules. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 10, 2108-2117.
- 市川薫(2011). 対話のことばの科学—プロソディが支えるコミュニケーション—. 早稲田大学出版部.
- 市川薫・長嶋祐二(編著), 日本音響学会編(2021). 音声コミュニケーションと障がい者. 音響サイエンスシリーズ22. コロナ社.
- 市川薫・長嶋祐二・堀内靖雄・原大介(2019a). 「一体的リズム」と「分析的リズム」. *電子情報通信学会ヒューマンコミュニケーション基礎研究会資料*, HCS2019-30, 7-12.
- 市川薫・長嶋祐二・堀内靖雄・原大介・酒向慎司(2019b). 超高齢化時代が対話システムに求める物理層の基盤的特性. *人工知能学会 言語・音声理解と対話処理研究会(第87回)*, SIG-SLUD-B902-16, 80-85.
- 市川薫・田中裕史・堀内靖雄(1998). 手話のリズム. *日本手話学会代24回大会論文集*, 6-9.
- 市川薫・大橋浩輝・菊池英明・仲真紀子・堀内靖雄・黒岩眞吾(2016). 母語話者の話者交替タイミング. *科学技術研究*,

- Vol. 5, No. 1, 113-122.
- 石黒昭博・山内信幸・赤楚治之・北林利治・菊田千春・伊藤徳文・須川精緻・川本裕未(1996). *現代の言語学*. 金星堂.
- Kajikawa, S., Amano, S., Kondo, T. (2004). Speech overlap in Japanese mother-child conversations. *Journal of Child Language*, Vol. 31, No. 1, 215-230.
- 北原義典・武田昌一・市川薫・東倉洋一(1977). 音声言語認知における韻律の役割. *電子情報通信学会論文誌D*, Vol. J70-D, No. 11, 2095-2101.
- 小松昭男・市川薫(1992). 音声による会話のための自然言語処理. *システム・制御・情報*, Vol. 36, No. 7, 447-452.
- 河野守夫(1998). モーラ、音節、リズムの心理言語学的考察. *音声研究*, Vol. 2, No. 1, 16-24.
- 河野守夫(編著)(2007). ことばの認知としくみ, 4.4リズムとイントネーションの認知と生成. 三省堂書店.
- 窪園晴夫(2017). 通じない日本語—世代差・地域差からみることばの不思議—. 平凡社新書.
- Kutas, M. and Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, Vol. 207, No. 4427, 203-205.
- 教育機器編集委員会(編)(1972). *産業教育機器システム便覧*. 日科技連出版社.
- Lieberman, A. M. (1967). Perception of the speech code. *Psychological Review*, Vol. 74, No. 6, 431-461.
- Low, H. and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, Vol. 54, No. 6, 1001-1010.
- 前川喜久雄・西川賢哉(2019). 「日本語話し言葉コーパス」への音質情報付与と予備的分析. *言語資源活用ワークショップ2019発表論文集*, 205-221.
- Martinet, A. (2008). *Éléments de linguistique générale, 5th ed.* Armand Colin.
- 皆川泰代(2015). 内言語機能. *脳科学事典*.
- 皆川泰代・直井望・小嶋祥三・山本淳一(2011). プロソディ処理の聴覚野機能—発達とその障がい—. *認知神経科学*, Vol. 13, No. 1, 56-63.
- Miyagi, M., Nishida, M., Horiuchi, Y., and Ichikawa, A. (2005). Prosody rule for time structure of finger Braille., *Guidelines on Tactile and Haptic Interactions*, Saskatoon, Canada. *Proceedings of GOTH1 705*, 32-36.
- 森大毅・前川喜久雄・粕谷英樹(2014). 音声は何を伝えているか. *音響サイエンスシリーズ12*, コロナ社.
- 牟田口辰己(2012). 点字読み熟達者の読速度に関する研究—読速度の左右差に焦点を当てて—. *特殊教育学研究*, Vol. 50, No. 4, 343-352.
- 長嶋祐二・原大介・堀内靖雄・酒向慎司・渡辺桂子・菊澤洋子・加藤直人・市川薫(2018). 多様な研究分野に利用可能な超高精細・高精度手話言語データベースの開発. *言語資源活用ワークショップ2018発表論文集*, 148-155.
- Oohashi, H., Ohsuga, T., Horiuchi, Y., Kikuchi, H., and Ichikawa, A. (2010). Prosody, supporting real-time conversation. P2b-07, *Speech Prosody 2010*.

- 
- 大串健吾 (2010). 音楽のリズムと言語のリズム. 電子情報通信学会 音声研究会資料, SP2009-149, 7-12.
- 大森健志・寺内美奈・長嶋祐二 (2009). 手話母語話者による手話の語彙認知構造の実験的考察. 電子情報通信学会技報, Vol. 108, No. 489, 1-6.
- 斎藤涼子・堀内靖雄・黒岩慎吾 (2009). 話者交替規則に基づく日本手話のオーバーラップ現象の分析. 第35回日本手話学会大会, 25-28.
- 志村洋子・山内逸郎・福原博篤 (1989). 胎児をとりまく音環境. 騒音制御, Vol. 13, No. 4, 15-19.
- 鳥越隆士 (2006). 手話言語の発達と言語環境. 心理学評論, Vol. 49, No. 1, 180-191.
- Stivers, T., Enfield, B., Englert, N. J., Hayashi, P., C. M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., and Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of National Academy of Sciences of the United States of America*, Vol. 106, No. 26, 10587-10592.
- 高松美也子 (2018). 音訳の読語過程における間(ま)による読語リズムの形成過程の分析. 早稲田大学学位申請論文.
- 滝沢恵子・菊池英明・大橋浩輝・市川熹 (2013). 日本人学習者の英語対話における流暢性に関する研究. 日本音響学会 2013 春季研究発表会予稿集, 1-Q-49c.
- 田中久弥・宮本一郎・長嶋祐二 (2008). 事象関連電位 N400 計測に基づく日本手話理解における意味処理分析. 計測自動制御学会論文集, Vol. 44, 768-775.
- Uchida-Ota, M., Arimitsu, T., Dan, I., Ikeda, K., Takahashi, T., and Minagawa, Y. (2019). Material speech shapes the cerebral frontotemporal network in neonates: A hemodynamic functional connectivity study. *Developmental Cognitive Neuroscience*, Vol. 39, 100701.
- 梅津八三, 梅津八三の仕事刊行会 (編) (2000). 盲ろう二重障害者の言語形成についての心理学的研究, 心理学, 梅津八三の仕事, 第2巻, 春風社.
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., and Yandell, B. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, Vol. 117, No. 1, 338-350.
- Vygotsky, L. S., 柴田義松(訳) (1962). 思考と言語(上). 明治図書出版.
- Willem, J. (1993). *Speaking: From intention to articulation*. MIT Press.
- 山崎志織・堀内靖雄・西田昌史・黒岩眞吾・市川熹 (2008). 3DCGによる手話文アニメーション合成システムの構築と評価. 電子情報通信学会技術研究報告, WIT2007-77, 31-36.
- Zaccarella, E., Papitto, G., and Friederica, A. D. (2021). Language and action in Broca's area: Computational differentiation and cortical segregation. *Brain and Cognition*, Vol. 147, 105651.

(受稿：2021年4月1日 受理：2021年4月16日)