

# 日本語音声の持つ非日本語話者への 情報伝達性にかかわる要因の実験的検討

## Experimental Study on a Potential Factor Related to Communicability of Japanese Speech Sounds to non-Japanese Speakers

宮本 真希<sup>†</sup>, 日高 昇平<sup>†</sup>  
Maki Miyamoto, Shohei Hidaka

<sup>†</sup> 北陸先端科学技術大学院大学  
Japan Advanced Institute of Science and Technology  
mymt.makiron@jaist.ac.jp

### 概要

同じ言語を共有していない話者間での音声による情報伝達にはふつう困難が伴う。これは両者の間で音声とその指示内容を結び付ける規則を共有していないためと考えられる。そこで、本研究では音声とその指示内容の結びつきが完全には恣意的でないといわれるオノマトペであれば、同じ言語を共有していない話者間でも情報伝達が可能ではないかと考え、話し手と聞き手が音声の指す内容をそれぞれ選択肢から選ぶという限定的な状況でのオノマトペの情報伝達性について検証した。

キーワード：言語の恣意性、プーバ・キキ効果、オノマトペ、情報伝達

い場合の情報伝達では、音声と内容を結び付ける規則が話し手の使用する言語と聞き手の使用する言語では異なっている状態となっていることから、話し手が音声化した内容と、聞き手がその音声から想起する内容が一致することはなく、情報伝達は失敗に終わるといえる。そこで、本研究では、話し手が音声化した内容と、聞き手が話し手の音声から推測した内容が一致した場合を伝達の成功と定義し、話し手の言語における音声と内容の結びつきの規則を知らない聞き手が、話し手の音声から結びつく内容を推測できるかどうかに関心を当てる。

### 1. 背景と目的

#### 1.1 音声による情報伝達の問題

日本人と外国人の会話に見られるように、話し手と聞き手が同じ言語を共有していない場合の情報伝達は、「言葉の壁」により一般的には難しいとされる。その「言葉の壁」の要因のひとつとして、言語の恣意性 [1] が考えられる。言語の恣意性とは、ある言語音声に対して特定の内容が結びつかなければならないという必然性がない性質のことをいう。すなわち、ある音声とその音声の指す内容との結びつきは、ひとつの言語体系内では、おおむね一意に定まっているが、別の言語体系ではその音声に対して異なる内容が結びついている、もしくは結びつく内容が存在しない可能性があり、これが互いに異なる言語を使用する話者間での、音声による情報伝達を困難にしているといえるだろう。

話し手と聞き手が同じ言語を共有している場合の理想的な情報伝達では、音声と内容を結び付ける規則を話し手と聞き手で共有している状態となっており、話し手が音声化した内容と、聞き手がその音声から想起した内容が伝達の前後で一致していた場合、話し手が音声で伝えたかった内容は聞き手に伝わったといえる [2]。一方、話し手と聞き手が同じ言語を共有していな

#### 1.2 本研究の目的と設定

本研究では、聞き手が話し手の言語における音声と指示内容の関係についての特定言語に固有の事前知識を持たないときに、話し手の音声のみからその指示内容を推測できるか（情報伝達性）を調べ、その要因を明らかにすることを目的とする。本研究においては、話し手を日本語話者、聞き手を日本語の知識を有さない非日本語話者、話し手が発した音声は日本語話者が理解できる音声（日本語音声）とした。また、話し手が音声で伝えたい内容と聞き手が音声から想起する内容を可視化するために、話し手があらかじめ決められた内容を音声化し、聞き手がその音声から推測して既定の選択肢から選ぶという、限定的な状況下における情報伝達を考えることにした。

#### 1.3 アプローチ

音声による伝達が成功しない要因は、音声とその指示内容の結びつきが恣意的であるためであるとしたが、これに反する知見としてプーバ・キキ効果 [3] がある。プーバ・キキ効果は、呈示された音声に対する事前知識がないにもかかわらず、その音声の指す指示対象を複数の選択肢から選択させると、チャンスレベルを超

えて選択が偏るという現象で、音声とその指示内容の関係は完全に恣意的とはいえず、何らかの関係性があることが示唆されている。そこで本研究では、このブーバ・キキ効果から着想を得て、音声とその指示内容の結びつきが強いとされる語があるとする、その語に対する事前知識がない状態で音声に対して指示内容を選択させると、多くの人がその音声と結びついている指示内容を選択する（選択に偏りが生じる）のではないかと考えた。

音声とその指示内容の結びつきが強いとされる語にオノマトペ（擬音語・擬態語）がある [4]。このことから、日本語オノマトペであれば、チャンスレベルを超えてその音声の指す内容が、日本語の知識を有さない非日本語話者にでも推測でき、内容の伝達が可能であると予想した。

#### 1.4 実験の流れ

本研究では、音声の指示内容として平面図形を使用し、日本語オノマトペ音声を聞いた日本語話者と非日本語話者が、音声の指示内容として同じ図形を選択できるかどうかを検証する。実験1では、日本語話者に日本語オノマトペ音声を聞かせ、その音声から日本語話者が連想した図形をその音声のさす正解の図形として設定した。次に、実験2として非日本語話者に日本語話者が聞いたものと同一の日本語オノマトペ音声を聞かせ、その音声の指す図形を推測させた。

## 2. 実験用オノマトペの選定

本研究では、清水ら [5] の方法を参考に、オノマトペ様の語を機械的に作成し、日本語母語話者3名に日本語としての容認度判断をさせ、3名中2名以上が選択した語を、実験で使用する実験用オノマトペを選定した。選定のための語群は、143個の日本語音韻（アからワ、濁音半濁音、拗音付き「ャ、ユ、ヨ」、撥音「ン」、促音「ッ」、長音「ー」、外来語表記に用いる仮名 [6]）をABAB型（e.g. ドキドキ、キュンキュン）になるよう機械的に組み合わせで作成し、合計20449語とした。ABAB型としたのは、日本語オノマトペでは最も一般的な形態とされるためである [4]。

作成した語群の選定作業は、Excelシート上で行った。Excelシートは、語の入ったセルをダブルクリックすることで着色されるよう設定し、選択・解除が容易にで

きるよう配慮した。視認性の観点から20449語は10分割してExcelシート上に表示させ、1シート当たり約2000語とし、合計10シートとした。

選定に参加したのは日本人大学院生3名（男性、平均年齢22.67歳、SD=0.47）であった。選定のための教示は3種類用意し、教示1「日本語としてありそうだと思う語を選んでください」、教示2「まんがや小説に出てきそうだと思う語を選んでください」、教示3「自分がその語から思い浮かべるイメージは、その語を用いれば他の日本語話者にも伝えられそうだと思う語を選んでください」とした。各参加者には、前の指示によって選択された語の中から、次に指示された語を選択させ、参加者ごとに2段階で語の数を減らす方法で行った。特に教示1による選定では語数が多いことから、参加者には無理のないペースで作業するよう説明し、選定作業期間は最大一週間とした。

選定の結果、各教示段階において、3名中2名以上が選択した語は、教示1では249語、教示2では122語、教示3では33語となった。この結果から、教示1で選定された249語ですでに実験可能な語数であると判断し、この249語を実験用オノマトペとして採用した。これらの語の中には一般的にオノマトペとは思われない語も含まれたが、日本語オノマトペの典型的な形態をもつことから、これらの語もオノマトペとなりうる可能性も考え、実験に使用することにした。

## 3. 実験1: 日本語話者による正解の設定

実験1では、選定した実験用日本語オノマトペの音声とその音声の指していると推測される内容を選択させ、その対応付けを日本語話者による正解の対応付けとすることを目的とした。

**実験刺激** 実験に使用する音声は、2で選定した語群を読み上げることで作成した。読み上げにおいては、プロソディによる影響を考慮し、人の声による読み上げと、合成音声による読み上げの2種類の方法で行った。合成音声による読み上げは、Google翻訳 [7] の読み上げ機能を使用し、一語ずつ読み上げさせたものを録音した。Google翻訳による読み上げ音声は女性の声質に近かったことから、人の声による読み上げは女性の日本語母語話者とし、一語ずつ読み上げさせたものを録音した。実験では249語のうち、ランダムに選択した半数の語を人の声による読み上げ、残りの半数を合成音声による読み上げとした。また、音声セットはA、B

の2種類を用意し、A セットで人の声による読み上げ音声とした語は、B セットでは合成音声による読み上げ音声とした。どちらのセットも、呈示順は順序効果を考慮し、それぞれランダム化した。

また、本研究では音声の指す内容を平面図形としていことから、実験参加者に選択肢として呈示する平面図形のセットを作成した。図形セットを構成する図形には、山口 & 椎名 [8] が作成・選定した 36 個の図形を使用する。視認性を高めるために、これらの図形の輪郭線が元の図形よりも太くなるよう筆者が加工を施したものを本研究では使用することにした。これらの図形に付属するラベルは除去し、形状が類似すると筆者が判断した図形同士を隣接させて、3 行 12 列となるよう配置しなおしたものを実験 1 の図形セットとした (図 1)。この図形セットを横方向の A4 用紙に 5 行 2 列となるよう 10 セット配置し、印刷したものを実験 1 の回答用紙とした。

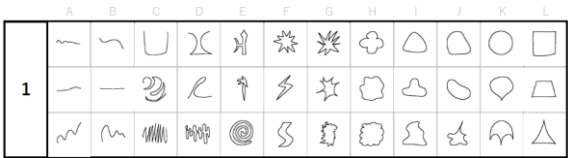


図 1 実験 1 で使用した図形セット

**実験参加者** 実験 1 の実験参加者は、日本人大学院生 12 名(男性 10 名, 女性 2 名, 平均年齢 24.17 歳, SD=1.46)であった。

**実験手続き** 実験参加者 12 名には、実験用オノマトペの音声を聞き、その音声から推測される図形を回答用紙の図形セットの中から 1 つだけ選択して、丸で囲むよう指示した。音声はコンピュータ画面上に表示された音声プレイヤーを参加者自身が再生し、コンピュータに接続したヘッドホンを通して呈示された。参加者への教示は、「まず音声を聞き、あなたがその音声から連想するイメージに最も近い図形を、印刷された図形リストの中から一つ選んで、丸で囲んでいってください。あなたの直感に基づき、一つだけ図形を選んでください。」とした。実験終了後、手作業による集計を行った。

**実験結果** 本研究では、各語における図形の選択確率からエントロピーを算出し、実験参加者間において、一つの語に対してどの程度共通して図形が選択されているか表す指標として使用する。本研究の実験においては、エントロピーの値が小さいほど参加者間でより共通した図形が選択されていると考えられる。エントロピーは非負値をとり、その絶対値に本質的な意味は

ないため、以下の結果は最大のエントロピーを 1、最小のエントロピーを 0 と正規化したエントロピーを用いた(以降「エントロピー」は正規化エントロピーを指す)。表 1 に実験 1 の結果から算出したエントロピーの値とその語に対応する図形および参加者による図形の選択割合、図 2 にエントロピーのグラフを示す。表 1 では、語をエントロピーの小さい順(形状選択の一致率の高い順)に配置し、その語に対して複数人が選択した図形の選択割合の高い上位 2 位までを示した。参加者間での図形の選択の一致率が高い語は、日本語話者が日常的に使用すると考えられる日本語オノマトペ(e.g., グルグル, クネクネ)である傾向がみられた。

| Rank | Word   | Normalized Entropy | Shape (Rate)                                                                        |                                                                                     |
|------|--------|--------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| 1    | グルグル   | 0                  |    | -                                                                                   |
| 2    | モアモア   | 0.256152145        |    |    |
| 3    | クネクネ   | 0.290338346        |    |    |
| 4    | モフモフ   | 0.290338346        |    |    |
| 5    | カチカチ   | 0.331585582        |    |    |
| 6    | ザアザア   | 0.336828837        |    | *                                                                                   |
| 7    | コロコロ   | 0.349133127        |    |    |
| 8    | トゲトゲ   | 0.386177402        |    |    |
| 9    | ブスブス   | 0.395623618        |    |    |
| 10   | カクカク   | 0.453509509        |    |    |
| 11   | ネバネバ   | 0.460262115        |    | *                                                                                   |
| 12   | ゴロゴロ   | 0.5                |    |    |
| 13   | ゴアゴア   | 0.528942946        |  |  |
| 14   | ガオガオ   | 0.54319642         |  | *                                                                                   |
| 15   | クシャクシャ | 0.54319642         |  | *                                                                                   |
| 16   | ナゾナゾ   | 0.54319642         |  | *                                                                                   |
| 17   | ジトジト   | 0.54319642         |  | *                                                                                   |
| 18   | シマシマ   | 0.54319642         |  | *                                                                                   |
| 19   | ズキズキ   | 0.544169094        |  |  |
| 20   | ゴツゴツ   | 0.546490491        |  |  |
| 21   | ピリピリ   | 0.546490491        |  |  |

表1 実験1でのエントロピーの値と選択された図形  
エントロピーの小さい順に21位までを掲載  
\*は2人以上の選択がなかったことを示す

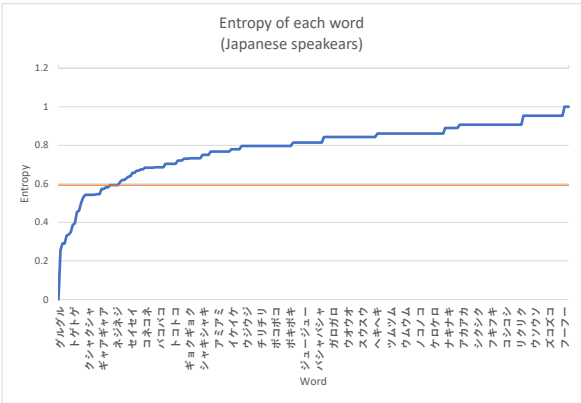


図2 実験1での各語のエントロピー

また、人の音声と合成音声による選択の違いについて、エントロピーの相関係数を求めたところ、中程度の正の相関関係( $r = 0.422084$ )が認められたことから、人の音声による図形を選択確率が高くなるにしたがって、合成音声による選択確率も高くなる傾向があるといえる。このことから、人の音声と合成音声の違いによる図形選択への影響は大きくないものと考えられる。

4. 実験2: 非日本語話者による内容推測

実験3では、実験2と同様、非日本語話者に実験用日本語オノマトペの音声を聞かせ、その音声が指していると推測される図形を図形セットから選択させることで、音声からどの図形が連想されているかを検証した。

実験刺激 実験に使用する音声刺激は実験1で使したものと同じものを使用した。

実験2では、実験参加者が耳慣れない日本語音声を聞き続けることになることから、参加者の負担を考慮し、図形セットの図形数を削減することにした。実験1の結果をもとに、エントロピーの小さい方から30位まで(正規化エントロピーの値では0.59まで)の語について、各図形を選択回数を算出し、選択回数が多かった図形の上位21位までで、すべての回答の92.22%をカバーしていることから、これらの図形を選択肢として採用することにした。選定した21個の図形には、極めて形状の近い図形が2個含まれていたことから、1つを除いて、図形は全20個とした。これら20個の図形は、実験1と同様に類似図形が隣接させて、4行5列となるよう配置し、実験2の図形セットとした(図3)。この図形セットを横方向のB4用紙に5行5列となるよ

う 25 セット配置し、印刷したものを実験 2 の回答用紙とした。

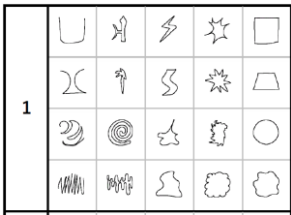


図 3 実験 2 で使用した図形セット

**実験参加者** 実験 2 の実験参加者は、大学院生 27 名（男性 14 名，女性 13 名，平均年齢 27.70 歳，SD=4.64）であり，すべて日本国外から来日している外国人留学生であった。参加者の出身国の内訳は，タイ：2 名，ミャンマー：2 名，中国：6 名，インドネシア：1 名，バングラデシュ：3 名，ケニア：1 名，ベトナム：5 名，インド：4 名，ラオス：1 名，マレーシア：1 名であった。参加者のうち半数となる 13 名は，日本語学習歴 1 年未満であった。

**実験手続き** 実験参加者 27 名には，実験 1 と同様に実験用オノマトペの音声聞き、その音声から推測される図形を回答用紙の図形セットの中から 1 つだけ選択して，丸で囲むよう指示した。音声はコンピュータ画面上に表示された音声プレイヤーを参加者自身が再生し，コンピュータに接続したヘッドホンを通して呈示された。参加者への教示は，「まず音声を聞き，あなたがその音声から連想するイメージに最も近い図形を，印刷された図形リストの中から一つ選んで，丸で囲んでください。あなたの直感に基づき，一つだけ図形を選んでください。」とした。インストラクションはすべて英語で行った。実験終了後，手作業による集計を行った。

**実験結果** 実験 1 と同様に，各語についてのエントロピーを算出した。表 2 に実験 1 の結果から算出したエントロピーの値とその語に対応する図形および参加者による図形の選択割合，図 4 にエントロピーのグラフを示す。表 2 では，語をエントロピーの小さい順に配置し，その語に対して選択された図形の選択割合の高い上位 2 位までを示した。実験 2 においては，参加者間での図形の選択割合は，最大で 48%程度にとどまった。

また，人の音声と合成音声による選択の違いについて，エントロピーの相関係数を求めたところ，相関係数は認められなかった( $r = 0.046051$ )ことから，人の音声と合成音声の違いが図形選択に影響した可能性が

示唆された。

| Rank | Word   | Normalized Entropy | Shape (Rate)   |                |
|------|--------|--------------------|----------------|----------------|
| 1    | オウオウ   | 0.618230706        | ○<br>(0.481)   | ⬮<br>(0.111)   |
| 2    | ミニミニ   | 0.671817962        | ↑<br>(0.296)   | ⬮ ⬮<br>(0.148) |
| 3    | ジロジロ   | 0.672076658        | ○<br>(0.370)   | ⬮ ⬮<br>(0.111) |
| 4    | イシイシ   | 0.680217745        | ↑<br>(0.259)   | ⬮ ⬮<br>(0.222) |
| 5    | スエスエ   | 0.73190651         | ⬮<br>(0.296)   | ⬮<br>(0.148)   |
| 6    | ライライ   | 0.745634835        | ○<br>(0.333)   | ⬮<br>(0.185)   |
| 7    | チリチリ   | 0.748285366        | ↑<br>(0.222)   | ⬮<br>(0.185)   |
| 8    | オラオラ   | 0.751936025        | ○<br>(0.296)   | ⬮<br>(0.148)   |
| 9    | ネコネコ   | 0.751936025        | ⬮<br>(0.296)   | ⬮ ⬮<br>(0.111) |
| 10   | オレオレ   | 0.754194979        | ○<br>(0.333)   | ⬮<br>(0.111)   |
| 11   | ケチケチ   | 0.761754995        | ⬮ ⬮<br>(0.222) | ↑ ⬮<br>(0.111) |
| 12   | コロコロ   | 0.761754995        | ○<br>(0.333)   | ⬮<br>(0.148)   |
| 13   | クセクセ   | 0.767314753        | ⬮ ⬮<br>(0.185) | ⬮<br>(0.148)   |
| 14   | ゴトゴト   | 0.767895825        | ⬮<br>(0.259)   | ⬮<br>(0.111)   |
| 15   | アキアキ   | 0.769315011        | ⬮<br>(0.222)   | ⬮<br>(0.148)   |
| 16   | モフモフ   | 0.781784511        | ⬮<br>(0.333)   | *              |
| 17   | パコパコ   | 0.786925212        | ⬮<br>(0.259)   | ⬮<br>(0.185)   |
| 18   | カツカツ   | 0.788344398        | ⬮<br>(0.185)   | ⬮<br>(0.148)   |
| 19   | キャーキャー | 0.795904414        | ↑<br>(0.185)   | ⬮ ⬮<br>(0.148) |
| 20   | イキイキ   | 0.796904542        | ↑<br>(0.222)   | ⬮<br>(0.148)   |

表 2 実験 2 でのエントロピーの値と選択された図形  
小さい順に 20 位までを掲載  
\*は 3 人以上の選択がなかったことを示す

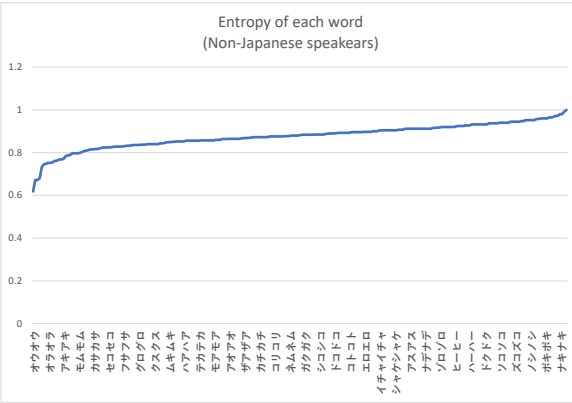


図 4 実験 2 での各語のエントロピー

5. 日本語話者・非日本語話者の選択一致度の関係

実験 1 および実験 2 の結果から、日本語話者が高い割合で対応付けた語と図形の組の中で、非日本語話者が比較的高い割合で同じ対応付けとしたのは、「コロコロ」と「モフモフ」のみであった。

また、全 249 語について、日本語話者と非日本語話者の図形選択における正規化エントロピーの相関係数を求めたところ、相関関係は認められなかった( $r = -0.022337$ ) (図 5)。この結果は、必ずしも日本語話者で一致度の高いオノマトペが、非日本語話者での一致度が高いわけではないことを示唆する。

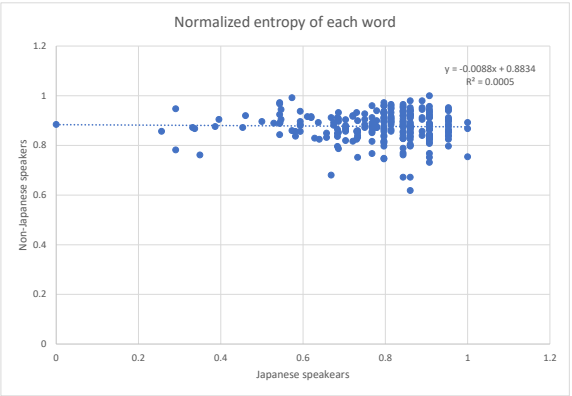


図 5 日本語話者と非日本語話者の正規化エントロピーの相関

6. 結論

現時点でのデータ分析結果では、日本語話者間の一致度から非日本語話者間の一致度を予測するような関係性は見られなかった。大会での発表では、主に実験 1

および実験 2 の結果を分析することで推定される情報伝達性に関連する音声・音韻的な特徴を特定し、報告する。

文献

[1] Saussure, F. De. (1916) Course in General Linguistics.  
[2] 池上嘉彦 (1984) 記号論への招待, 岩波書店  
[3] Ramachandran, V. S., & Hubbard, E. M. (2001) Synaesthesia--a window into perception, thought and language., Journal of consciousness studies 8(12), 3-34.  
[4] 田守育啓 & ローレンス・スコウラップ (1999) オノマトペ: 形態と意味 (Vol. 6), くろしお出版  
[5] 清水祐一郎, 土斐崎龍一, & 坂本真樹 (2014) オノマトペごとの微細な印象を推定するシステム, 人工知能学会論文誌 29(1), 41-52.  
[6] 文化庁 外来語の表記「外来語の表記」に用いる仮名と符号の表 (2021 年 4 月 20 日時点で閲覧可) [https://www.bunka.go.jp/kokugo\\_nihongo/sisaku/joho/joho/kijun/naikaku/gairai/honbun01.html](https://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kijun/naikaku/gairai/honbun01.html)  
[7] Google 翻訳 <https://translate.google.co.jp/> (2021 年 4 月 20 日時点で閲覧可)  
[8] 山口由衣 & 椎名健 (2005) 手描き図形を媒体とした感情伝達, 図書館情報メディア研究 3(2), 15-23