

1st Electronic edition

Sophia University
Working Papers
in
Phonetics

Vol.26 2020

Phonetics Laboratory

Sophia University

序

松井さなえ

今年も例年のように Sophia University Working Papers in Phonetics 2020 が出版され、大変嬉しく思います。この論文集は上智大学の音声学研究室に関わる方々の研究成果をまとめたもので、その内容は音声学、音韻論、言語障害学、応用言語学などの分野にまたがります。

本年の論文集は、ダジャレに関するもの、音声単語認識に関するもの、室長雑記から成っています。

音声研では勉強会を開催しており、2019年度は心理言語学、統計学をテーマとして学生や教員が毎週発表を行っていました。2020年度は、4月の時点で上智大学が全面的にオンラインで授業を実施することを決定したため、音声研でもオンラインで統計学の勉強会を行っております。

刊行に際し、藤本雅子先生、田嶋圭一先生、吉田夏也先生、事務の蓑島さんに心より感謝申し上げます。

2020年4月

目次

序

松井 さなえ

1. ダジャレの面白さの機械学習

渋谷 美佳 1

2. バイリンガル話者の音声単語認識における音韻的近傍密度効果

松井 さなえ 22

室長雑記

COVID-19 と教育・研究

北原 真冬 31

ダジャレの面白さの機械学習

渋谷 美佳

上智大学外国語学部 英語学科

E-mail: chakamin0000@gmail.com

1. 序論

1.1. 契機

人間は何を「面白い」とみなし、何を「面白くない」とみなすのか。このことに興味を持ち始めたのはダジャレが好きになったからである。私が「ダジャレ、おやじギャグ」などと呼ばれる韻を踏む言葉遊びが含まれる文章に興味を持ち始めたのは大学1年生の頃だった。幼少期から落語やお笑い番組、CMを通して当たり前のようにダジャレをインプットしていたからか、人との会話の中でダジャレを発見してはそれを指摘し、時には面白い視点だと評価され、時にはつまらないと批判されていた。

しかし、自分が面白いと思って言った渾身のダジャレも周りからの受けは良くない時も多く、面白さの基準は人によって異なることを感じていた。その中でどのような指標が人間の言葉遊びの評価に影響を与えているのかに疑問を持つようになった。それを明らかにすることによってより多くの人から面白いと思われるようなダジャレを作りたいと思った。

大学で言語学や音韻論を履修し、韻を踏んでいるダジャレは音韻論との関係があると考えようになった。そしてダジャレを音韻論の視点から研究をしたいと思い、人間の面白さの指標を数値化することでダジャレを体系的に分析したいと考えた。

1.2. 先行研究

言葉遊びの中でも音を重視しているダジャレの「面白さ」を数値化したいと考えていたところで見つけた先行研究が「子音の音韻類似性及びSVMを用いた駄洒落検出方法」(谷津・荒木, 2016)である。谷津・荒木(2016, p.875)はシステムがユーモアを理解するためには話者によるユーモアを検出すること、ユーモアにスコア付けを行うこと、そしてスコアに基づいた適切な応答または対応動作が必要であるとした。そこで、サポートベクトルマシン(以後SVM)を用いて子音の類似度によって一般文の中からダジャレ文を検出する方法を明らかにしていた。

彼らのSVMにおいては、4つの規則に基づいた指標を用いて分析を行っていたが、その中で効果が表れたのは2つのみであった。そのため、子音間の類似度によってダジャレ文を検出することは困難であり、母音を用いた指標が必要だと述べられていた。その結果、正答率は90.9%となっていた。また、ダジャレを複数のブログサイトから抽出していたため、データの重複が起り、重複データの削除に苦戦していた。

これに加え、「複数の言語資源を用いたユーモアを含む対話システム」(吉田・萩原, 2014)ではユーモアの一つとしてダジャレを挙げ、ダジャレを発話するシステムを開発しているが、音韻的に類似するものを発話するだけで、必ずしも面白いダジャレが生成できているシステムとは言えないも

のだった。また、吉田・萩原(2014)の研究はきわめて工学的で、「面白さ」という概念に対する真摯な追求が欠如していると感じた。

1.3. ダジャレに含まれる成分

1.3.1. Shadow の存在の発見

先行研究ではダジャレにおいて基盤となる一般的な語または節を「種表現」、種表現に近付けてダジャレとなるために音韻的・語彙的に変化が起こっている語または節を「変形表現」とし、ダジャレの成分を2つに分類していた。しかし、多数のダジャレの分析を進めているうちに、ダジャレには種表現と変形表現の間に、表層に現れない裏の語または節があるということを発見した。この「変形表現の裏」はダジャレを成立させるために特別に変形された表現であり、独立した状態では意味を持たないこともある。本研究では種表現を「Target」、変形表現を「Paronomastic」、変形表現の裏を「Shadow」と呼ぶ。

1.3.2. TPS 分類法

具体例を示すことでこの分類方法の有意性を示す。

(1)「あの指揮者の服、しきしやい(色彩)豊かだね」の場合

あの 指揮者 の 服 、 しきしやい(色彩) 豊かだ ね

T P S

→Targetの「指揮者」に合わせるために本来はShadowの「色彩」であるが形を変えて

Paronomasticは「しきしやい」となっている。

このように、ダジャレによってはTarget、Pronomastic、Shadowの形が全て異なっている場合があることから、TargetとParonomasticの他にShadowについて考慮することは有意義であると言える。Shadowはダジャレにのみ現れる特別な表現方法で、これを含む分類法は先行研究のダジャレ分類法では区別できていなかった部分を分類することのできる新しいメソッドである。これを「TPS分類法」と名付け、本研究で利用していくことにした。

1.4. ダジャレの分類方法

ダジャレは構造的に複数の種類があり、その分類方法は様々だが、本研究ではまず先行研究の谷津・荒木(2016)に倣い「併置型駄洒落」と「重畳型駄洒落」に分類することにした。併置型駄洒落とは、「文内に2つの音韻的に類似した区間(種表現及び変形表現)を持つ」ダジャレのことである。中でも変形表現(=本研究で言うParonomastic)の違いによってダジャレを「Perfect puns」と「Imperfect puns」に分類している。この概念は、本研究で提案する「Shadow」の形式を確認することでより明確にそれぞれを定義することができる。TPS分類法を用いるとPerfect punsはParonomasticとShadowが完全一致しているダジャレ、Imperfect punsはParonomasticとShadowが一致していないダジャレであるということが出来る。それに対し、重畳型駄洒落とは文内に音韻的に類似して対となる音素列が存在しないダジャレのことである。

1.5. 本研究の目的

先行研究ではダジャレの子音の類似度からダジャレの面白さを判別するシステムやダジャレを発話するシステムを開発していた。しかし、子音の類似度だけでなく、母音の類似度も判断材料として必要であるという課題があった。また、ダジャレの分類方法が曖昧で、併置型/重畳型では分類しきれないダジャレが多数存在していた。そしてその原因はダジャレの表層には表れていない変形表現の裏である「Shadow」について考慮されていない分類法だったからだと分かった。また、ダジャレの「面白さ」に影響を与えるファクターは音韻的情報だけではないと考えられ、音韻的情報のみからダジャレの面白さを判断することは難しい。したがって、ダジャレの面白さに関して、現在までのダジャレの機械学習の研究では未解明の部分が多く残る。

本研究では人間の思うダジャレの面白さを機械に教え込むに当たり、日本語話者に対してアンケートを行った。アンケート結果と共に音韻的・形態的なさまざまな情報を機械に学習させるために、まず本研究の問いとして以下の3つを設定した。

- i. ダジャレの表層に現れている Target-Paronomastic 間の距離がダジャレの面白さに一番大きく影響を与えており、表層に現れていない Shadow に関連している Paronomastic-Shadow 間、Target-Shadow 間の距離は Target-Paronomastic 間の距離程には面白さに影響を与えない。
- ii. T-P の語尾が類似している方が語頭や語中が一致している時よりも面白く感じる。
- iii. 距離を測る語同士の子音の類似度と母音の類似度がダジャレに与える影響は異なっている。
- iv. ダジャレの面白さは音韻的類似度以外にも文章としての意味の有無や事前に必要な知識の有無などからの影響も受けている。

また、本研究でダジャレを分類する際には谷津・荒木(2016)の使用していた併置型/重畳型の分類法と本研究の 1.3.2 で提案する新たな分類方法である「TPS 分類法」を併用して分類する。

2. 調査方法

2.1. ダジャレに含まれる成分の解明

2.1.1. 併置/重畳分類法と TPS 分類法を併用したダジャレの分類

1.3.1.でも述べたとおり、今回の研究では Target、Paronomastic、Shadow の3項目に分けて分解を行うことにしている。谷津・荒木(2016)の提案するダジャレの分類法と本研究で提案する TPS 分類法を併用し、いくつかのダジャレを分類していく。

(2) 「ふとんがふっとんだ」の場合

ふとん が ふっとん だ
T P/S

→文内に2つの音韻的に類似した区間を持ち、その対が完全一致していないため、併置型駄洒落(Imperfect pun)と判断される。Paronomastic と Shadow が一致しているが、Target と Paronomastic/Shadow を構成する音素は異なっている。

(3) 「潤いを売る甥」の場合

潤い を 売る 甥
T P/S

→文内に2つの音韻的に類似した区間を持ち、その対が完全一致しているため、併置型駄洒落(Perfect pun)と判断される。Paronomastic と Shadow が一致しており、Target と Paronomastic/Shadow を構成する音素も一致している。

(4)「床の間を汚したのは、とこの間抜けだ！」の場合

床の間 を 汚し た の は とこの間抜け(どこの間抜け) だ !
T P S

→文内に2つの音韻的に類似した区間を持ち、その対が完全一致していないため、併置型駄洒落(Imperfect pun)と判断される。表層に現れる Paronomastic は本来 Shadow の形をしているが、この文章をダジャレにするために変形している。

(5)「クローズアップ玄米」の場合

クローズアップ 玄米(現代)
T/P S

→文内に2つの音韻的に類似した区間は持たず、「玄米」と「現代」を重ねることでダジャレ文になっている。そのため、これは重畳型駄洒落と判断される。Target と Paronomastic が一致している。しかし、Paronomastic の本来の形は Shadow の表す「現代」であるので Paronomastic と Shadow は異なっている。

例(2)と例(3)、例(4)と(5)は併置/重畳の分類で方法では一致しているが、TPS の側面から分析すると同じ種類には分類されないことが分かる。先行研究ではダジャレを3種類に分類していたが、併置型/重畳型の分類方法では分類しきれないダジャレが存在していることから、本研究では TPS 分類法を兼用することでその側面を捉えている。

2.1.2. T-P-S の距離

仮説 iii にもある「ダジャレの面白さに与える影響は子音と母音で異なっている」ことを明らかにするためには、比較対象を用意する必要がある。距離を量るファクターとして妥当なものを検討した結果、子音と母音を同時に使用して類似度を測っているレーベンシュタイン距離を使用することにした。これはある文字列と別の文字列の最小の編集距離を求めて数値化したものであり、字列間の距離を量るにあたって代表的な手法である。文字の挿入・削除・置換のコストをそれぞれ1として、2つの単語がどのように変化しているかを計算する。

2.2. アンケートの作成

2.2.1. 対象者

より普遍的なダジャレの面白さを明らかにするため、本アンケートは対象者を絞らず、老若男女から被験者を募集することにした。しかし、小学生以下は中学生以上よりも集中を続けることが難しい可能性があることと、ネット環境下で正しくアンケート結果を送信できない可能性があることから、アンケートを取る量を中学生以上の被験者に対して行う質問の量に比べて少なくし、紙媒体でのアンケートを作成することにした。それに対し中学生以上からは不特定多数のデータが必要であることから、ネット上でのアンケートを用いることで広くデータを集められるようにした。な

お、被験者には謝礼無しで協力していただいた。アンケートの詳しい作成方法に関しては 2.2.3 で述べる。また、本研究において小学生の被験者を「子供」とし、中学生以上の被験者を「大人」と呼ぶことにする。

被験者は大人 162 名、子供 12 名を合わせた 174 名である。1 人の子供の結果からほぼすべてのダジャレに対して「どこがダジャレになっているのかが分からない」と回答している被験者が子供の被験者の中にいた。その被験者は日本語話者だが、両親の母国語が韓国語で、小学校はインターナショナルスクールに通っているため、家庭では韓国語と英語、学校でも英語を使用している。そのため、今回の実験のデータとしては採用しないことにした。

出身地/一番長く住んでいる地域の分け方に関して地方は八地方区分等に従った。なぜなら、方言が使用されているダジャレがアンケートにて使用されていたからである。被験者の詳しい内訳は以下のとおりである。

表 1. 被験者の属性の分布

		被験者(N=174)
性別	男性	59 (うち8名は子供)
	女性	112 (うち3名は子供)
	その他	2
年齢	幅	6 - 60
	中央値	21
一番長く住んでいた地域	北海道・東北地方	4
	関東地方	120
	中部地方	12
	近畿・中国・四国地方	9
	九州・沖縄地方	18
	その他	1
	海外	9
英語レベル	C2	0
	C1	23
	B2	81
	B1	32
	A2	11
	A1	9
	分からない	17

2.2.2. 調査するダジャレの情報

アンケートに使用するダジャレはまず『これをダジャレで言えますか?—ダジャレ練習帳 (2の巻)』(多治家, 2001)より無作為に 700 個、自作のダジャレから無作為に 50 個選出し、計 750 個選出した。本研究ではダジャレを音韻的な視点から研究するため、それぞれ以下の音韻的な要素を抽出した。

- ① Target/Paronomastic/Shadow の読み方をカタカナに変換
- ② ①で入力したカタカナをヘボン式ローマ字に変換(ただし、長音や促音などは子音と母音の類似度を測るために手動で変換)
- ③ Target/Paronomastic/Shadow のモーラ数
- ④ Target-Paronomastic(以降 T-P)間の相違点、Paronomastic-Shadow(以降 P-S)間の相違点、Shadow-Target(以降 S-T)間の相違点の数値化
- ⑤ まったく新しく追加されているモーラ数
- ⑥ 語尾一致度(T-Pのみ計算)
- ⑦ アクセントずれの有無(有→0、無→1)
- ⑧ バウンダリーずれの有無(有→1、無→0)
- ⑨ Shadowの参照先が文章内にあるか、文章外にあるか(内→0、外→1)
- ⑩ 文章として意味が成立しているか(成立している→0、成立していない→1)

- ⑪ ダジャレの中に固有名詞や歌など、事前知っておく必要のある情報があるか
(有→0、無→1)
- ⑫ 感嘆詞追加の有無(有→1、無→0)
- ⑬ 方言の有無(有→0、無→1)
- ⑭ 関連語変換の有無(有→1、無→0)

指標④は、T-P、P-S、S-Tの相違点を子音と母音に分け、その Distance Matrix を数値化したものである。母音の Distance Matrix は「ダジャレから見る母音の近似性」(川原・篠原, 2009) から、子音の Distance Matrix は「The role of psychoacoustic similarity in Japanese puns: A corpus study」(Kawahara & Shinohara, 2009)を参照した。それぞれの数値は以下のとおりである。

左から表 2. 母音の類似度(川原・篠原(2009) より引用)

表 3. 子音の類似度(Kawahara & Shinohara (2009)より引用)

	a	e	o	i	u
a	0	0.63	0.47	1.38	1.29
e		0	1.35	0.53	1.82
o			0	2.18	0.65
i				0	0.49
u					0

	θ	p	b	φ	m	w	t	tʃ	s	j	z	d	j	n	r	k	g	h
θ	0	.7	.24	.77	2.5	6.3	.64	.45	.36	.37	0	0	0	1.5	4.6	1.4	.46	3.7
p		0	8.5	5.6	0	0	1.1	0	1.1	0	0	0	0	0	.37	1.1	.24	.77
b			0	1.6	4.7	5.2	.30	0	0	0	.42	1.1	0	.35	0	.25	.65	2.9
φ				0	2.3	0	0	0	.81	.84	0	0	0	0	0	2.7	0	4.2
m					0	0	.64	0	.53	0	0	0	0	8.9	2.1	0	.34	1.7
w						0	0	0	0	0	0	1.2	2.1	1.5	0	.53	0	4.5
t							0	.57	.90	0	0	7.6	0	.62	.44	.87	.14	.94
tʃ								0	.95	7.9	0	0	13	0	.47	.23	0	.49
s									0	4.7	11	.2	0	1.6	.37	.73	0	.78
j										0	0	0	6.8	.54	0	.38	0	.41
z											0	1.7	1.2	0	.62	0	0	0
d												0	.39	1.1	4.0	0	.39	.42
j													0	2.0	.72	0	.93	0
n														0	4.1	.25	1.7	1.1
r															0	.36	1.2	.38
k																0	8.0	1.7
g																	0	0
h																		0

この時、ひとつ前の音素が複製された場合、他の音素に変更される時よりも影響が小さいと考えた。そこで、表の中での Distance Matrix 最小値に近い数値をどの音素が複製された場合でも 0.3 追加するものとした。

促音の追加も一定して 0.3 と定めた。これは「促音/tt/の知覚: アクセント型と促音・非促音語の音響的特長による違い」(大深, 2003)において促音の有無によって知覚に違いがあり、その差は有意であったとあるという結果をもとにしている。なお、このとき子音と母音の Distance Matrix はそれぞれ独立したものであり、相関関係がないということを前提としている。

また、指標⑥「語尾一致度」は T-P 間において音素が変化している地点を語頭からの距離で計算して仮説 ii を検証する。P-S 間、S-T 間での計算を行わなかったのは、相違点の距離が重要視されているのは表層に現れている Target と Paronomastic の距離であって、表層に現れていない Shadow との距離はダジャレの面白さに影響を与えにくいと考えたためである。これは仮説 i の考えを前提としている。

計算をするにあたり、まず語尾のほうが面白いという考えを数値化するために語尾が一致しているほうがスコアが高くなるような計算方法にすべく、以下のようなモデルを考えた。

P 形全体の長さを n、P 形における変化点の語頭からの位置を p とすると、p より後ろの部分の長さは(n-p)となる。なお、どちらもモーラ数で測り、このとき全体の長さ n というのは計算の関係上

た。n=T または P のモーラ数で大きい方:である 5 (P)、p=相違が起こっている場所(一番語尾に近い方): 語頭から 3 モーラ目である。式[2]に代入すると、

$$y = \text{LOG}[5 * (5 - 3) / (3 + 1)]$$

$$= 0.397940009 \quad \text{となる。}$$

2.2.3. 指標④⑤⑥をもとにしたダジャレの選出

本アンケートではダジャレの音韻的情報である指標④⑤⑥に重きをおいて調査する。そのため、④⑤⑥のデータを抽出し、T-P、P-S、S-T 間での数値を見て選出することにした。750 個のダジャレを分析して集計したところ、以下のような結果となった。

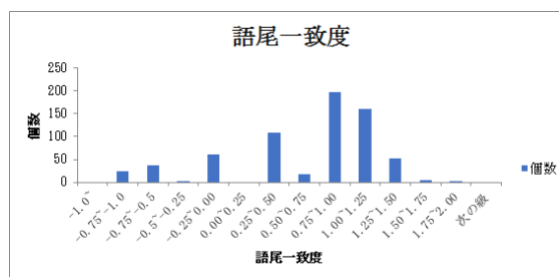


図 2. 指標⑥語尾一致度のヒストグラム(左から T-P, P-S, S-T)

子音と母音の類似度、そして追加モーラの数は T-P、P-S、S-T いずれにおいても 0 となるダジャレがほとんどだった。そのため、アンケート用にダジャレを選出する際には指標④、指標⑤が 0 の時と 0 より大きい時に分けることにした。また、指標⑥である語尾一致度は正の数のほうが多いものの、負の数も一定数存在していた。そのため、語尾一致度に関しては 0 以下の時と 0 より大きい時で分けることにした。

選出する際に学習データとして学習させにくいと考えられる、Paronomastic が関連語に変換されているダジャレ(指標④において 1=有になっているダジャレ)は学習データから削除した。例えば「アメリカに行って USA(憂さ)晴らし」や「レッドカードはあかん(「レッド」とあかんの「あか」)」などがあり、計 13 個だった。

これを除いたデータから、TP に注目したとき、PS に注目したとき、ST に注目した時の指標④～⑥の値をそれぞれ取り出した。条件を満たすものが全てあった場合、 $(2 * 2 * 2) * 3 * 3 = 144$ 個のデータが選出される予定だった。しかし、条件を満たすダジャレが 1 つもないことがあった。例えば、TP に注目した時に、母音は完全一致、子音は不一致、追加モーラあり、語尾一致度は 0 以上という条件を満たすダジャレは一つもなかった。このとき、「追加モーラがある」という条件を満たすものがなかったために選出できなかった。これはこの条件を満たそうとするとダジャレ文からかけ離れてしまうから無かったということが考えられた。以上のことが原因で最終的に選出されたのは 123 個となった。

なお、選出する際に筆者が 1 つのダジャレを誤って入れてしまったために、TP に注目して選出されたダジャレが 1 つ多くなってしまっている。

また、2.2.1 で述べたとおり、子供は大人よりも長時間集中力を維持することが難しいと考えられることから 30 個のダジャレに絞った。選出の際に、指標⑩の「ダジャレの中に固有名詞や歌など、事前に知っておく必要のある情報があるか」は子供の場合事前知識が少ない可能性があることから、できるだけ事前知識がなくても分かるようなダジャレを選出するようにした。

2.2.4. 特別に選出したダジャレ

選出したダジャレの中には、ある点を考察するために意図的に選出されたダジャレがあり、それは重畳型ダジャレ「い列車ーい！（いらっしやーい）」である。今まで小学生と会話をするときにダジャレを言うと歓迎されることが多く、ダジャレを頻繁に使用してきた。しかし、重畳型ダジャレを発した際に子供からそれをダジャレと認めてもらえないことがほとんどであった。そのため、重畳型ダジャレは本当にダジャレとして認めてもらえないのかを本アンケートで検証すべく、意図的に重畳型ダジャレ「い列車ーい！（いらっしやーい）」を採用した。

2.2.5. アンケートの作成方法

2.2.5.1. 大人を対象にしたアンケート

大人を対象にしたアンケートは「Google Form」を利用してアンケートを作成した。そのため、スマートフォンやパソコンなど、ネット環境が整っているデバイスならばどこからでもアクセスできる状態であった。被験者が回答している際に質問に答えられるわけではないため、回答方法を細かく書いた。

アンケートではまず基本情報として以下の項目を聞く設問を用意した。

- ① 性別(男/女/回答しない/その他)
- ② 年齢
- ③ 出身地(国内の場合都道府県名まで、国外の場合は国名を記入)
- ④ 一番長く住んでいる地域(ただし、日本国内の場合都道府県名まで、国外の場合は国名を記入)
- ⑤ 英語レベル(CEFR)

これらを必須で回答してもらった。

①②は性別/年齢によって面白いと思われやすいダジャレがあるのかを明らかにするために聞いた。設問の中に方言を含むダジャレがあったため、その方言に親しみがあることがダジャレを面白いと思うことに影響しているのかを明らかにするために③④を聞いた。同様に、英語を含むダジャレが含まれていることから⑤において英語レベルを聞いた。⑤は英語の資格試験結果と CEFR 対応表 (文部科学省, 2019)を基に解答してもらった。

ダジャレの面白さは深く考えるのではなく、直感で答えてもらうように指示を出した。会話の中でダジャレが発話された際に深く考えることは少なく、そのときの直感で面白さを判断することの方が多く考えたため、できるだけ普段の会話に近づけるためである。選択肢は 4 つにし、選択肢①を「面白い」、選択肢②を「面白くない」、選択肢③を「聞いたことがあるので面白くない」、選択肢④を「どこがダジャレになっているのか分からない」とした。選択肢②～④は、最終的なデータとしては同じ「面白くない」に分別される。これは教師データにする際に 0/1 の 2 項に分別し

た方がより「面白い」「面白くない」の区別がはっきりしたデータになると考えたからである。

「面白い」は選択肢を1つのみにしたのは理由が2つある。1つ目は、「面白い」には英語で言う「interesting」や「funny」など、人やダジャレによって様々な解釈があるが、敢えて線引きをせず、結果からどのような面白さを聴衆が求めているのかを考察しようと思ったからである。2つ目は、「やや面白い」などの項目を準備してしまうと0/1で区別できないからである。

2.2.5.2. 子供を対象にしたアンケート

子供を対象にしたアンケートでは紙媒体で筆者が普段小学生の面倒を見るボランティアに行っている団体の活動後に行った。被験者を教室に集めて解答方法の説明をし、解答方法や質問事項があればその都度答えた。

2.3. SVM(サポートベクトルマシン)

本研究はダジャレの面白さを教え込むシステムとして、川原・荒木(2016)が使用していたSVMを使用する。これは教師あり機械学習手法で、多数あるデータを種類分けするための「マージン」を様々な方法で引くためのシステムである。ソフトウェアR[3.5.2]を用いて外部パッケージ「kernlab」を組み込んで計算を行う。これは「サポートベクトルマシンの考え方」(馬場, 2018)を参考にして行っている。なお、事前に必要な情報としてTPSそれぞれの2.1.2の方法で計算したレーベンシュタイン距離(T-P, P-S, S-T)、2.2.2で述べた14種の指標のデータをcsvデータで読み込んでいる状態であることを前提としている。

```
[3] svmM_dajareXYZ <- ksvm(fun_no_fun ~ X+Y+Z, data=dajare,  
                           type="C-svc",kernel="rbfdot", kpar=list(sigma=1))
```

式[3]はSVMの学習モデルの式である。「kernlab」の中にカーネル係数は9種類組み込まれており、自由に選択をすることができたが、SVMを試用した際に一番正答率に変化が現れやすかった「rbfdot」を使用することにした。rbfdotはRBF(Radical Basis Function)カーネルのことで、学習データの与える範囲を明らかにするときに利用される関数のことである。また、sigmaの値は大きくすればするほど面白い/面白くないの境目が明確になる。sigmaは0~100に設定することができるが、25以降は正答率に変化は現れなかった。sigmaを増加させることによって結果が変わる様子は以下のとおりである。

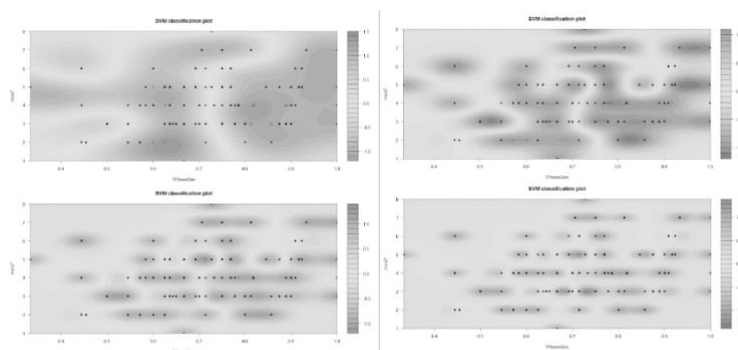


図3. X. sigmaを増加させることによって変わるSVMの結果
(左上: sigma=1, 右上: sigma=5, 左下: sigma=10, 右下: sigma=20)

グラフの濃い灰色の部分が「面白い」で、薄い灰色の部分が「面白くない」であるが、 σ の値が増加するごとにダジャレ 1 つ 1 つ固有の面白さが影響しているようで、より汎用性の高いダジャレの面白さの定義付けが難しいといえる。そのため、本研究では $\sigma=1$ に設定した。

1 行目にある「svmM_dajareXYZ」は SVM の結果に付けられる名前である。このときは「M のモーラ数を基盤にして指標 X、Y、Z を組み込んだ SVM」であろうということを意味している。また、2 行目にある XYZ は SVM に教え込む際に使用する指標で、教え込む指標によって数も変わる。指標は何個でも追加することができるが次元が増えてしまうので 3 次元以上はグラフにすることができなくなる。

この SVM の結果が正しいかを検証するためには以下の計算を行う。なお、この計算は「R 言語で SVM(Support Vector Machine)による分類学習」(菊池, 2012)を参考にして行った。

```
[4] rowdata <- nrow(dajare)
     random_dajare <- sample(rowdata, rowdata*0.5)
     dajare_training <- dajare[random_dajare, ]
     dajare_predicting <-dajare[-random_dajare, ]
[5] resultM_predictXYZ <- predict(svmM_dajareXYZ, dajare_predicting)
[6] table(resultM_predictXYZ, dajare_predicting$fun_no_fun)
```

式 [4]ではまずランダムにダジャレを検出するように行数を検出し、その中からランダムに行番号を検出している。ランダムに検出するデータ数は全体の半分である $123 \text{ 個} \times 0.5 = 61.5 \div 62$ 個にした。それらをもとに学習データと予測データを作成している。学習データと予測データの割合を考えたときに、学習と予測結果の精度をどちらもできるだけ上げるために比率は 1:1 にしている。

式[5]では SVM の結果と予測データをもとに学習させている。「resultM_predictXYZ」の命名方法も「svmM_dajareXYZ」と同様である。

式[6]では予測結果と正解を比較した表を作成している。正解数をもとに正答率を計算することができる。以上の計算で SVM の計算とその正答率を明らかにすることができる。

3. 調査結果

3.1. アンケート

3.1.1. 読み方の変換

アンケートの集計結果からレーベンシュタイン距離を測るためにモーラ数を数えやすい特殊ローマ字を使用して読み方をすべて変換し、それにバウンダリー位置を「#」、アクセント位置をアクセントの付く音素の直後に「^」で書き加えた。

Paronomastic はそれ自体では意味をなしていない語も多く、その背景には Shadow があることから、Paronomastic のバウンダリー位置は音が異なっても Shadow に統一した。また、固有名詞は一つの単語として捉えるとした。具体例を示すと以下のとおりである。

(7) 「始皇帝死んで思考停止」の場合

始皇帝 死ん で 思考 停止
T P/S

3.1.2. 面白さ評価

大人を対象にしたアンケートの面白さ評価の集計結果は以下のとおりとなった。

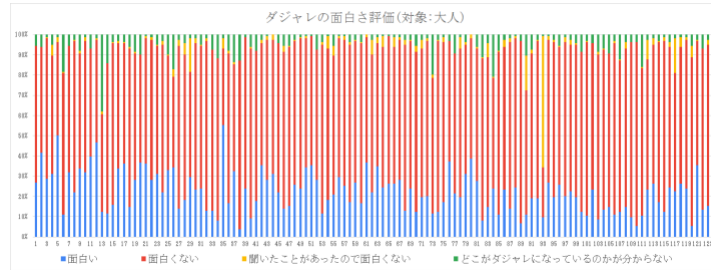


図 4. アンケート結果「ダジャレの面白さ評価」(対象: 大人)

SVM の教師データとするにあたって、「面白い」と「面白くない」の境界を決めるべく、「面白い」の合計値から中央値と平均値を測ったところ、中央値は 36、平均値 36.4796748 となった。よって、「面白い」の合計値が 36 を超えたものを「面白い」ダジャレとし、36 以下のものを「面白くない」ダジャレにすることにした。

子供を対象にしたアンケートの面白さ評価の集計結果は以下のとおりである。

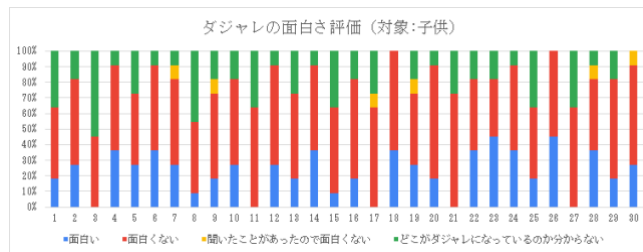


図 5. ダジャレの面白さ評価(対象: 子供)

3.1.3. 男女別・年代別・地域別評価傾向

図 6 は男女別(大人の被験者のみ)に面白さの評価傾向を示したグラフである。「面白い」の評価も「面白くない」の評価も限りなく等しいことが分かる。

図 7 は年代別(大人の被験者のみ)に面白さの評価傾向を示したグラフである。被験者のほとんどは 20 代だったため、数を分散させるために 23 歳以下と、24 歳以上に分類した。よって 10 代から 23 歳、24 歳から 30 代、40 代、50~60 代に分類された。10 代の最年少者が 16 歳、30 代の最年長者が 31 歳だったためにこのような分類方法になっている。

図 8 は地域別(大人の被験者のみ)に面白さの評価傾向を示したグラフである。この分析は一番長く住んでいる地域のデータをもとに行った。2.2.1 で述べたとおり、出身地/一番長く住んでいる地域の分け方に関して地方は八地方区分等に従った。また、北海道、中国地方、四国地方は被験者の数が極めて少なかったため、北海道は東北地方と、中国・四国地方は近畿地方と合併した。

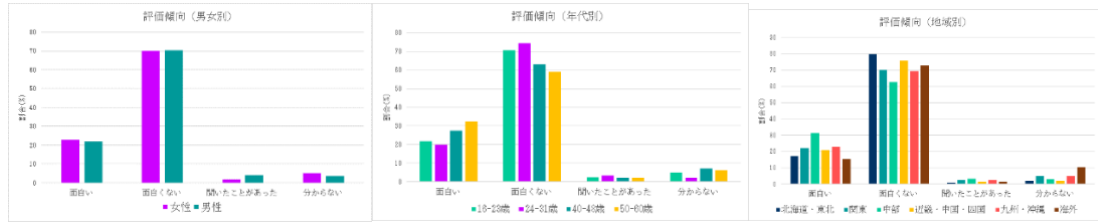


図 6. 評価傾向(男女別) 図 7. 評価傾向(年代別) 図 8. 評価傾向(地域別)

3.1.4. 特別に選出したダジャレ

2.2.4 で述べた、特別に選出した 重畳型ダジャレ「い列車ーい！（いらっしゅーい）」(対象：子供)のこれらの面白さ評価の結果は以下のとおりである。

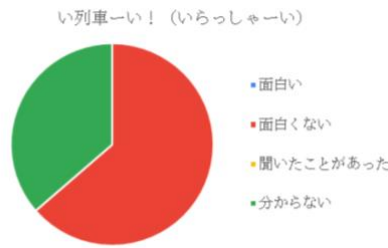


図 9. ③の面白さ評価

3.1.5. 評価の高かったダジャレ、低かったダジャレ

本アンケートで使用した 123 個のダジャレの中でひとときわ「面白い」と評価されたものと評価されなかったダジャレは以下のとおりである。

表 4. ダジャレランキング

トップ3	面白さ合計値	ワースト3	面白さ合計値
野口英世の愚痴ひで一よ	89	ズボンらなひとだなあ	6
始皇帝死んで、思考停止	81	CDを貸しーでー	9
井伊直弼と言ひ直すけ？	76	グリーン列車にビッググリーンした	11

3.1.6. 音韻的でないダジャレの指標に注目した面白さの差

指標⑦～⑭の音韻的でないダジャレの指標に注目して面白さの差を検出した。指標ごとに値が 0 のダジャレと 1 のダジャレ(指標⑦の場合アクセントずれがあるダジャレとアクセントずれがないダジャレ)に分け、面白さの差を全体の中央値を取ることによって調べた。なお、2.2.3 で述べた通り、Paronomastic が関連語に変換されているダジャレはアンケート用に選出したダジャレに含まれていないため、指標⑭「関連語変換の有無」に関しても検出していない。指標⑦～⑬の 0/1 に分けた面白さの中央値は以下のとおりである。

表 5. 指標⑦～⑬の面白さの差(中央値)

中央値	0	1
指標⑦: アクセントずれの有無(有→0、無→1)	32	41.5
指標⑧: バウンダリーずれの有無(有→1、無→0)	36	36
指標⑨: Shadowの参照先が(文章内→0、文章外→1)	36	38
指標⑩: 文章として意味が成立している(→0、いない→1)	37	24
指標⑪: ダジャレ内に事前知識が必要なものがあるか(有→0、無→1)	32	36
指標⑫: 感嘆詞追加の有無(有→1、無→0)	36	31.5
指標⑬: 方言の有無(有→0、無→1)	35	36

3.2. SVM

3.2.1. 学習データに適した指標

SVMの結果を掲載する際に指標の名前が冗長になることを防ぐために、これ以降の図では省略語を使用している。その対応表は次のとおりである。

表 6. 省略語対応表

	省略名	対応する意味
モーラ数	moraT	Targetのモーラ数
	moraP	Paronomasticのモーラ数
	moraS	Shadowのモーラ数
語の類似度	TPlevenSim	T-P間のレーベンシュタイン距離
	PSlevenSim	P-S間のレーベンシュタイン距離
	STlevenSim	S-T間のレーベンシュタイン距離
	TPCons	T-P間の子音の類似度
	PSCons	P-S間の子音の類似度
	STCons	S-T間の子音の類似度
	TPVowel	T-P間の母音の類似度
	PSVowel	P-S間の母音の類似度
	STVowel	S-T間の母音の類似度
	Distance	T-P間の語尾一致度
0/1で測る指標	Accent=Acc	アクセント変更の有無
	Boundary=Bnd	バウンダリーずれの有無
	Dialect=Dial	方言の有無
	Exclamation=Ex	感嘆詞追加の有無
	Meaningfulness=Mean	文章として意味が成立しているか
	Motoneta=Moto	ダジャレ内に事前に必要な情報があるか
	Reference=Ref	Shadowの参照先が文章内/外か

2.3 で述べた SVM に学習させ、その正答率を検証するに当たって使用した式は[3][5][6]である。これらの結果をもとに明らかになった正答率は以下を図 10 に表す。ここではモーラ数と語の類似度を学習させた。結果から分かるように、学習させるモーラ数を増やせば増やすほど正答率は上昇し、どの類似度においても TPS 全てのモーラ数を学習させたときが一番高い正答率となっている。

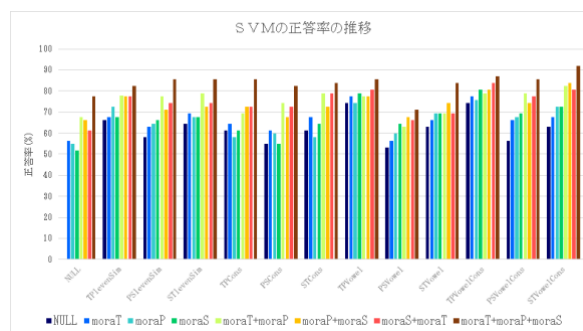


図 10. SVM 正答率の推移(モーラ数と語の類似度を学習)

モーラ数のみを学習させたところ、表層に現れる moraT と moraP が一番正答率を上げており、中でも moraT と moraP では moraP の方が正答率が高いことが分かる。また、全てを組み込んだ moraT+moraP+moraS が一番高いスコアとなっているため、学習に適しているといえる。

また、語の類似度を学習させたものである。Distance は元から T-P 間でのみ計算しているため、この結果からも全体的に TP が関わる数値が含まれているときにスコアが高いことが分かる。ST に関するスコアが TP に関するスコアをわずかに上回っているときもあるが、TP のスコアのほうがより差を大きくして上回っていることから、TP に関するスコアは安定的に高い正答率を保つことができると考えられる。よって、学習させるのに適しているのは TP に関するスコアであると言える。また、子音の類似度 TPCons、母音の類似度 TPVowel は両方足した TPVowelCons の方が正答率は上がっている。よって、学習に適しているのは、TPlevenSim、TPVowelCons、Distance であるといえる。

次に、0/1 で測っている 7 つの指標を学習させた際の正答率を調べたところ、正答率向上により影響を与えているの Meaningfulness と Motoneta であることが分かった。

以上の結果から、正答率に一番大きく影響を与えているのは、mora 数においては TPS 全てのモーラ数、語の類似度においては TPllevenSim、TPVowelCons、Distance、そして 0/1 で測る指標の中では Meaningfulness と Motoneta で、これらが学習に適していることが分かった。これら 3 つのファクターを掛け合わせて一番正答率の高くなる学習データを明らかにした結果が以下の表である。マス内の数値は正答率で、小数点第 3 位以下は切り捨てている。なお、この SVM はどれも必ず moraT+moraP+moraS を学習させている。つまり、1 マス目の NULL と NULL の場合でもモーラ数の学習データのみ組み込んだ場合の正答率が記載されているということになる。

表 7. SVM に学習させた指標とその正答率の対応表(単位=%)

	NULL	TPllevenSim	TPVowelCons	Distance
NULL	69.354	80.645	80.645	75.806
Mean	77.419	87.096	83.87	80.645
Moto	80.645	88.709	83.87	85.483
Mean+Moto	83.87	91.935	85.483	88.709

この結果から、一番正答率を上げることのできる学習データは

- ①TPS 全てのモーラ数
- ②T-P 間のレーベンシュタイン距離
- ③文章として意味が成立しているか
- ④ダジャレ内に事前に必要な情報があるか

の 4 つを学習させたデータであり、その正答率は 91.935%であることが明らかになった。参考値として、方法論が非常に似ている川原・荒木(2016)の SVM の正答率は 90.9%であり、その結果を 1%以上超える結果となった。

4. 考察

ここで、第 1 章で述べた本研究の仮説を再掲する。

- i. ダジャレの表層に現れている Target-Paronomastic 間の距離がダジャレの面白さに一番大きく影響を与えており、表層に現れていない Shadow に関連している Paronomastic-Shadow 間、Target-Shadow 間の距離は Target-Paronomastic 間の距離程には面白さに影響を与えない。
- ii. T-P の語尾が類似している方が語頭や語中が一致している時よりも面白く感じる。
- iii. 距離を測る語同士の子音の類似度と母音の類似度がダジャレに与える影響は異なっている。
- iv. ダジャレの面白さは音韻的類似度以外にも文章としての意味の有無や事前に必要な知識の有無などからの影響も受けている。

4.1. アンケート

4.1.1. 大人の結果から

全体的にどのダジャレも「面白い」と評価された割合はあまり高くなかった。この結果は「ダジャレは寒い」というダジャレに対する印象を示す典型的な表現を後押しするものとなった。また、「面白い」の割合が低かった原因として考えられるのは、文脈が存在しない状態でダジャレを掲示したから、ということも考えられる。普段の会話の中でダジャレが発話される際は、会話の中の単語を利用してダジャレが生成されることが多い。しかし、本研究でのアンケートではダジャレのみを羅列したため、文脈が存在しない状態であった。ダジャレの面白さに影響を与える指標の中で「文章として意味が成立しているか」が選ばれていたように、聞き手側は会話の流れの中で生じるダジャレにより面白いと感じるのだと考察した。

3.1.3 の結果からも分かるように、男女の違いという視点から考察すると、面白さの評価の傾向に男女差は無かった。よって、ダジャレの面白さに性差は影響を与えていないということが分かった。

年代別に評価傾向を見た場合、年代が上昇するにつれて「面白い」と評価しやすい、ということが分かった。ダジャレが別名「おやじギャグ」といわれるのも、年代が上がるにつれてダジャレをより面白いと思うようになり、使用頻度が上がるからだと考えた。また、16-23 歳が 24-31 歳よりも多く「面白い」と評価しているのは、近年 1.2.6 で述べたような昔の「洒落」に近いような言語遊戯が若者言葉として頻繁に使用されているからだと考えられる。

地域による傾向の違いに大きな差は見られなかったが、一番長く住んでいる地域が日本国外の人は「どこがダジャレになっているのかが分からない」としている人がほかの地域に比べて非常に多かった。ダジャレは日本国内に長く居住し、日本語に触れる機会が多い人により面白くと評価されやすいことが分かった。

また、2.2.5.1 において『「面白い」には英語で言う「interesting」や「funny」など、人やダジャレによって様々な解釈があるが、敢えて線引きをせず、結果からどのような面白さを聴衆が求めているのかを考察したい』と述べた。3.1.5 の評価の高かったダジャレ、低かったダジャレから考察できたことは文脈が存在しない状態の場合においては「funny」といえるようなクスッと笑えるくだらないダジャレよりも「interesting」といえるような秀逸で巧みなダジャレの方が面白いと思われている傾向にあると考察した。しかし、文脈が存在する場合においてはこの結果は変化すると考えられる。

4.1.2. 子供の結果から

子供の被験者も大人と同様に「面白い」と評価された割合は高くなかった。大人と異なっていたのは1人も「面白い」と評価しなかったダジャレが存在していたことである。その一つが2.2.4で意図的に選出していた重畳型ダジャレ「い列車ーい!(いらっしゃーい)」である。筆者が普段子供と会話している際に感じていた「重畳型ダジャレは子供からダジャレとして認識されていない」ことがアンケートで誰からも「面白い」と評価されず、「どこがダジャレになっているか分からない」を選択した被験者が多かったことから考えられた。

4.2. SVM

4.2.1. 学習データに適した指標

3.3.1の結果から、学習データに適している指標は、TPS全てのモーラ数、語の類似度においてはTPlevenSim、TPVowelCons、Distance、そして0/1で測る指標の中ではMeaningfulnessとMotonetaが学習に適していることが分かった。ここで、モーラ数に関しては、TPSの中でTPが一番正答率を上げていることが明らかになった。つまり、ダジャレ文の表層に現れているTPのモーラ数の方が表層に現れていないSよりもダジャレの面白さに影響を与えているということである。Sはダジャレを生成する際には必要不可欠な存在であるが、ダジャレを受け取る側にとっては表層の語の方が重要であるということが分かった。また、この結果から「T-P間の類似度がP-S間、S-T間よりも与える影響が大きい」という仮説iは支持された。

SVMに教え込んだ語尾一致度(=Distance)は「T-P間で語尾が一致しているほうがより面白いと感じる」という仮説iiの元で作成された数値だったが、3.3.1の語の正答率の計算から分かるように、Distanceは他の類似度の指標よりも比較的高い正答率となっている。同じT-P間でもアクセントやバウンダリーの違いも組み込まれているレーベンシュタイン距離よりも正答率が高いことから、ダジャレにおいてTとPの語尾が一致しているとより面白いダジャレを生成できることが分かる。よって、仮説iiは支持された。

3.3.1の子音の類似度と母音の類似度でSVMの正答率に与える影響を読み取ると、子音の類似度と母音の類似度に大きな差は見られなかったため、仮説iiiは支持されなかったといえる。しかし、子音の類似度と母音の類似度を同時にSVMに教え込んだ際にそれぞれ単体で教え込んだ時よりも正答率が上がったため、どちらも正答率を上昇させることに貢献していることが明らかになった。

学習データに適したデータを下に作成したSVMの中で最も正答率が高かったのは

- ①TPS全てのモーラ数 (moraT+moraP+moraS)
- ②T-P間のレーベンシュタイン距離 (TPlevenSim)
- ③文章として意味が成立しているか (Meaningfulness)
- ④ダジャレ内に事前に必要な情報があるか (Motoneta)

の4つを学習させたデータだった。ここに0/1で判断している音韻的情報とはまったく関係ない要素が2つ含まれていた。つまり、ダジャレの面白さは音韻的情報以外からも影響を受けていることである。これを以って「ダジャレの面白さは音韻的類似度以外にも文章としての意味の有無や事前に必要な知識の有無などからの影響も受けている。」という仮説ivが支持された。また、このSVM

の正答率は 91.935%であり、人間何を面白いと思うかをシステムによってかなり高い精度で理解できるようになっていたことが分かった。

4.2.2. 面白いと評価されやすいダジャレ

3.1.8 で明らかにした音韻的でない指標ごとの面白さの差の中央値を出した指標のなかで 0 の時と 1 の時で 5 以上の大きく差が見られたのは指標⑦、指標⑩、指標⑫である。3.1.2 で述べた通り、本研究ではアンケートにおいて「面白い」の合計値が 36 を超えたものを「面白い」ダジャレとし、36 以下のものを「面白くない」ダジャレにしている。指標⑦の結果より、T-P 間でアクセントずれが無いダジャレの方が面白くなる傾向にあり、ずれが有るダジャレは面白くない傾向にあることが分かる。指標⑩の結果より、必ずしも文章として意味が成立しているダジャレが面白いとは限らないが、文章として意味が成立していないダジャレは面白くない傾向にあることが分かる。指標⑫の結果より、必ずしもダジャレに感嘆詞が追加されていなければ面白いというわけではないが、感嘆詞が追加されてダジャレになっているダジャレは面白くない傾向にあるということが分かる。

4.3. 反省点

アンケート作成時にいくつかの反省点が挙げられた。1 つ目に挙げられるのは「面白い」と答えにくい選択肢だったということである。「面白くない」に関しては選択肢を 3 つに分けて答えやすくしていたのにも関わらず、面白いは 1 項目のみだった。これは 2.2.5.1 で述べたとおり、後の教師データに教え込む際に 0/1 で区別しやすくするために「やや面白い」というような項目を入れなかった。しかし、被験者からアンケート回答後に「面白いとは認めたいが、面白くないわけでもないダジャレが多く、回答が難しかった」という声が多数寄せられた。アンケート結果も「面白い」が 50%を超えるダジャレはほとんど無かった。

2 つ目に挙げられるのは面白いと評価されやすいダジャレが少なかったということである。元データの中に含まれている自作のダジャレは面白いと評価された実績があるものだったが数は非常に少なく、全体の 93%である本から選出したものはその実績があるかは不確かなものであった。先行研究でブログサイトからダジャレを選出したために重複データが生じ、苦戦していたことから本を利用したが、2001 年に発行された本だったため、今の時代に合った生きたダジャレが少なかったことが考えられる。「面白い」と答えにくい選択肢であっただけでなく、いわゆる「寒い」、「面白くない」と評価されるダジャレを入れすぎたことが被験者を退屈させる原因にもなっていた可能性があると考えた。

また、普段からダジャレに対して抱いている印象を聞く項目があれば被験者のダジャレに対する意識を明らかにすることができたと考えた。4.1 で全体的に「面白い」と評価されなかったことが「ダジャレは寒い」という表現を後押ししていると考察したが、実際にダジャレに被験者がダジャレに対して抱いている印象を明らかにすることで面白さの評価傾向と比較できたと考えた。

4.1.1 で述べたように、一番長く住んでいる地域が日本国外の人が「どこがダジャレになっているのかが分からない」を選択していることが多いという結果から、日本国内に長く居住し、日本語に触れる機会が多い人により面白いと評価されやすい、と考察したが、一番長く住んでいる地域だけでなく第一言語を聞く項目を設けることでこの考察がより強固なものになっただろうと考えた。

4.4. 今後の展望

今回はダジャレの分類方法として川原・荒木(2016)の併置型/重畳型分類法と本研究で提案する「TPS 分類法」を併用した分類方法であったが、この方法は複雑であったため、これらを包括するより単純明快な分類方法を作ることでより簡潔で明確な分類ができるだろう。

また、今回は被験者の負担を考慮し、音声データを用いずにダジャレ文を読んで面白さの判断を行ってもらった。しかし、音韻論的分析をするためにも音声データと共にアンケートを行うことで、より精度の高い聞き手側のダジャレの面白さの判断を明らかにすることができると考えられる。そしてその際にダジャレ文として読んだ場合と一般文として読んだ場合のイントネーションやアクセントの違いを調べて違いを明らかにすることはダジャレを発話するにあたってより面白く聞こえる方法を明らかにすることに繋がると考えた。

本研究では Shadow という Paronomastic の裏に存在する語を発見したが、それはダジャレの聞き手側が面白さを評価する際には大きな影響を与えてはいなかった。しかし、Shadow はダジャレを生成する話し手側にとっては重要なプロパティであることが予想できる。ダジャレを生成する側の思考のプロセスをより深く考察することは、本研究でアンケートをとることで行った聞き手側の思考のプロセスを考察することと同等に重要であると考えられる。

5. 結論

人は何を「面白い」と思い、何を「面白くない」と思うのか。言葉では説明できない「面白さ」をシステムに教え込むことで機械は人間の面白さを判断できるのか。先行研究ではダジャレの子音の類似度から一般文とダジャレ文を判別するシステムやダジャレを発話するシステムが開発されてきていたが、学習データに限りがあり、課題のあるシステムであった。そしてその原因の一つとしてダジャレの表層には表れていない変形表現の裏である「Shadow」について考慮されていないダジャレの分類法だったからだと分かった。また、音韻的情報以外にも面白さに影響を与えているものがあるという仮説を設定した。

そこで、Shadow についても考慮した「TPS 分類法」とそれまでの併置型/重畳型分類法を併用した上でダジャレを分類した上で、被験者にアンケートをとり人間の思う「面白さ」のデータを取った。また、語の類似度やモーラ数を含む音韻的情報、文章として意味が成立しているかなどの音韻的でない情報を組み込み、SVM に学習させた。

アンケートの集計結果からも面白さの評価に性差は無いこと、年代が上がるとより多くのダジャレを「面白い」と評価することを明らかにすることができた。

SVM の分析を行っている際に T-P の語尾が一致している方が語頭や語中が一致している時より面白く感じることも明らかになった。しかし、仮説では子音の類似度と母音の類似度で面白さに与える影響は異なると設定したが、母音の類似度の方が僅かに高いだけで大きな差は見られなかった。

SVM の正答率を上げることに貢献する指標は Paronomastic のモーラ数、T-P 間のレーベンシュタイン距離、文章として意味が成立しているか、そして固有名詞など、事前に知っておくべき情報があるか、の 4 つであることが判明した。つまり、ダジャレの表層に現れる Target と Paronomastic の

情報が重要であり、ダジャレの面白さは音韻的情報以外からの影響も受けていることが明らかになった。

作成された SVM の正答率は 91.935% であり、類似した方法論で作成された先行研究の SVM の正答率よりも 1% 高い結果となった。

言語遊戯について語る綿谷(1964)は洒落について以下のように述べていた。

「われわれは、日常生活の会話に急激なショックを導入するために、時として即興的な洒落をいう習慣がある。あまり頻繁な場合にはショックにならないで、かえってわずらわしい感情を生起するが、ともかく洒落が人間間の言語のやり取りにとって一種の潤滑油であることは明白だ。」(p.343)

ユーモアは人間の心理的ストレスを軽減させる。そしてユーモアは人間と人間の間だけでなく、人間とロボットとの間の会話でも潤滑油になるはずだろう。

本研究は人間の面白さを体系的に明らかにしただけでなく、未解明であるユーモアを理解するシステムの開発に貢献する実りある研究となった。

謝辞

本論文は上智大学外国語学部 2019 年度卒業研究をまとめ直したものである。本研究の実験に参加した参加者の方々 174 名の皆さまに感謝申し上げます。ならびに、本研究の遂行、また本研究の執筆につき、多大な時間をかけ指導して下さった北原真冬教授に感謝申し上げます。

参考文献

- 大深悦子(2003)「促音/t/の知覚: アクセント型と促音・非促音語の音響的特長による違い」『音声研究』, 7, 70-76.
- 金子冨(2018)「【技術解説】似ている文字列がわかる! レーベンシュタイン距離とジャロ・ウィンクラー距離の計算方法とは」https://mieruca-ai.com/ai/levenshtein_jaro-winkler_distance/#toc_3 (2019 年 12 月 10 日最終閲覧).
- 川原繁人・篠原和子(2009)「ダジャレから見る母音の近似性」『音声研究』, 13, 101-110.
- 菊池佑太(2012)「R 言語で SVM(Support Vector Machine)による分類学習」
<http://yut.hatenablog.com/entry/20120827/1346024147>(2019 年 12 月 10 日最終閲覧).
- 多治家礼(2001)『これをダジャレで言えますか?—ダジャレ練習帳〈2の巻〉』角川春樹事務所, 東京.
- 谷津元樹・荒木健治(2016)「子音の音韻類似性及び SVM を用いた駄洒落検出方法」『知能と情報(日本知能情報ファジィ学会誌)』, 28, 875-888.
- 馬場信哉(2018)「サポートベクトルマシンの考え方」
<https://logics-of-blue.com/svm-concept/>(2019 年 12 月 10 日最終閲覧).
- 深川英雄・相沢秀一・伊藤 徳三(2005)『時代を映したキャッチフレーズ辞典』電通, 東京.
- へボン式変換君(2019) <http://hebonshiki-henkan.info/>(2019 年 12 月 9 日最終参照).

文部科学省(2019)「各資格・検定試験と CEFR との対照表」[PDF]

http://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afieldfile/2019/01/15/1402610_1.pdf (2019年12月10日最終参照).

吉田裕介・萩原将文(2014)「複数の言語資源を用いたユーモアを含む対話システム」

『知能と情報(日本知能情報ファジィ学会誌)』, 26, 627-636.

Kawahara, S., & Shinohara, K. (2009). The role of psychoacoustic similarity in Japanese puns: A corpus study. *Journal of Linguistics*, 45, 111-138.

バイリンガル話者の音声単語認識における音韻的近傍密度効果

松井 さなえ

上智大学大学院 言語科学研究科 言語学専攻

E-mail: sanaematsui1107@gmail.com

1. はじめに

本研究の目的は、母語の音韻的近傍密度が第二言語の単語の認識に影響を与えるかどうか、影響を与えるのであればどのような影響かを調査することである。

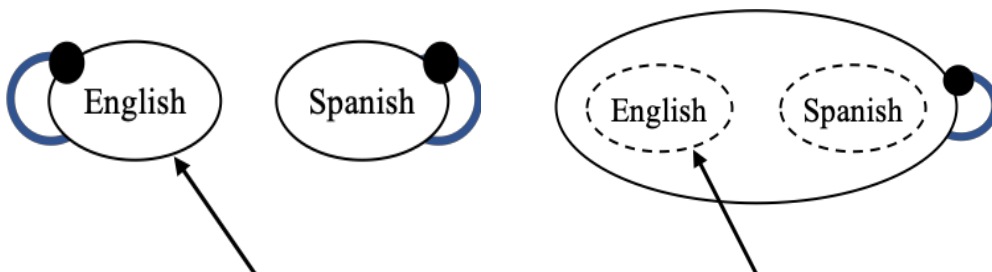
1.1. 先行研究

バイリンガル話者の語彙認識に関する研究では、従来、以下の二つの問題が活発に取り組みられてきた。

- (1) “integrated lexica or separate lexicon” 統合的/独立的心内辞書
バイリンガル話者は二言語の語彙を、一つの統合された心内辞書に貯蔵しているのか、言語ごとに独立した心内辞書に貯蔵しているのか。
- (2) “selective or non-selective access” 選択的アクセス/非選択的アクセス
バイリンガル話者の心内辞書では、一方の言語の単語が与えられた時に、その言語の語群のみが単語候補として活性化するか、もう一方の言語の語群も単語候補として活性化するか。

問題(1)(2)は、図1のように図式化することができる。

A: 独立的心内辞書・選択的アクセス B: 統合的心内辞書・選択的アクセス



C: 独立的心内辞書・非選択的アクセス D: 統合的心内辞書・非選択的アクセス

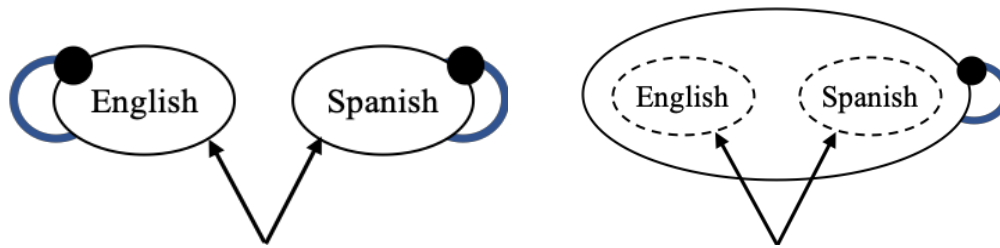


図 1: 例示として英語・スペイン語バイリンガル話者の心内辞書に関する 4 つの図式化(黒丸は、語彙項目間での作用があることを示す。(Van heuven & Dijkstra, 1998))

本研究では、バイリンガル話者の心内辞書が統合的で且つ語彙アクセスが非選択的なのかどうかを明らかにする。図 1 に基づいて言い換えると、D なのかそれ以外なのか、を明らかにすることを目的とする。

モノリンガルの語彙認識研究の中で相互活性モデルに基づく枠組みでは、心内辞書内の単語と単語の間には、抑制的な結合が存在すると仮定されている (McClelland & Rumelhart, 1981)。ある単語が与えられた時に、心内辞書ではその単語に関連する複数の単語が単語候補として活性化し、その語群がターゲット語に対して抑制的に働くと考えられている。

それでは、バイリンガル話者の場合はどうなるのだろうか。図 1 の A や C のような独立的心内辞書では、そのような単語間での相互作用が言語内でのみ起こると仮定される。一方で、B や D のような統合的心内辞書では、単語間での相互作用が言語をまたいで起こると考えられる。これが、先に挙げた問題(1)に相当する。

問題(2)は語彙アクセスの選択性についてである。図 1 の A や B のような選択的アクセスでは、ターゲット語が属する言語の語彙のみにアクセスする。例えば、英語とスペイン語のバイリンガル話者が “dog” という単語を見た場合、英語の語彙のみにアクセスする。一方で、C や D のような非選択的アクセスでは、ターゲットではない言語の語彙にもアクセスすることになる。例えば、英語とスペイン語のバイリンガル話者が “dog” という単語を見た時には英語の語彙だけではなくスペイン語の語彙にもアクセスする。

(1)(2)の問題は、これまでに様々な方法で研究されてきた。表 1 にその概略を示す。

表 1: 問題(1)(2)に対する従来のアプローチの概略

	(1)心内辞書の統合/ 独立	(2)選択的/非選択的アクセス (同時活性の有無)
視覚	言語をまたいだ正書法近傍密度効果 (Van heuven & Dijkstra, 1998 等)	<ul style="list-style-type: none"> 同系語や言語間同形異義語のプライミング効果 (Beauvillain & Grainger, 1987 等) ストループ課題 (Dyer, 1971 等)
聴覚	-	・アイトラッキング (Spivey & Marian, 1999 等)

これまでの研究で、問題(2)選択的/非選択的アクセスについては、視覚・聴覚モダリティー共に、非選択的アクセスを支持する結果が優位である。問題(1)については、正書法近傍密度効果を利用した研究では心内辞書は統合的であるという仮説が支持されている。具体的には、オランダ語と英語のバイリンガル話者が英単語を認識する場合、オランダ語の正書法近傍語を多く持つ英単語と、オランダ語の正書法近傍語が少ない英単語は、有意に異なる速度で認識されることを示した。それをもとに、彼らはバイリンガル話者の心内辞書は統合的であると結論づけた。

しかしながら、表 1 に示されているように、聴覚モダリティーを用いて心内辞書の独立性を検証した研究は筆者の知る限り見当たらない。そこで本研究では、聴覚モダリティーを用いて心内辞書の独立性を検証する。

1.2. 目的

本研究では一方の言語の単語を認識する時に、もう一方の言語の単語候補が、ターゲット語に対して作用するかどうかを検証する。検証方法としては、Van heuven & Dijkstra (1998)の実験 4 に対応する実験を行う。表 2 に、Van heuven & Dijkstra (1998)の実験 4 と本研究の実験の概要を示す。

表 2: Van heuven & Dijkstra (1998)の実験 4 および本研究の実験の概要

	実験 4	本研究の実験
実験課題	語彙性判断課題	
モード	視覚	聴覚
参加者	オランダ語と英語のバイリンガル話者	日本語とスペイン語のバイリンガル話者
条件	言語をまたいだ正書法近傍密度が高い VS. 低い	言語をまたいだ音韻的近傍密度が高い VS. 低い
刺激語	実在語(英語)と無意味語	実在語(スペイン語)と無意味語

表 2 のように、本研究ではスペイン語と日本語のバイリンガル話者に対して、

聴覚による語彙性判断課題を行う。刺激語には、スペイン語の实在語と無意味語を用いる。实在語には二つの水準を作り、一つは言語をまたいだ音韻的近傍密度が高い (日本語の音韻的近傍を多く持つスペイン語の単語)刺激群、もう一つは言語をまたいだ音韻的近傍密度が低い刺激群とする。例えば、“queso”/keso/は「基礎」/kiso/や「過疎」/kaso/など日本語の音韻的近傍語を多く持つため、言語をまたいだ音韻的近傍密度が高い。それに対して、“palma”/palma/や“fecha”/fetja/は、日本語の音韻的近傍語が少ないため、言語をまたいだ音韻的近傍密度が低い。

語彙性判断課題での反応時間と正確さを測定し、従属変数とする。

2. 方法

2.1. 参加者

実験の参加者は、スペイン語と日本語のバイリンガル話者 16 名だった。年齢の平均は 26.8 歳、標準偏差は 8.55 だった。参加者のほとんどは、母語が日本語であり、大学入学からスペイン語の学習を開始していた。一律にスペイン語の単語の知識を測るため、ランカスター大学の語彙力テスト(Lancaster University, 2014)を行なった。その結果、参加者のレベルはヨーロッパ言語共通参照枠の B1 から C2 にあった。

2.2. 刺激

語彙性判断課題を行うため、スペイン語の实在語 100 語と、スペイン語らしい無意味語 100 語からなる 200 語を刺激語とした。实在語を選定する際、(1)二音節であること、(2)名詞であること、(3)親密度が高いこと、(4)言語内での音韻的近傍密度が 20 から 50 であることを条件とし、統制した。さらに言語をまたいだ音韻的近傍密度を操作するため、实在語のうち 50 語は、言語をまたいだ音韻的近傍密度が高く、残りの 50 語が低くなるようにした。選定した实在語の特性は表 3 の通りである。

表 3: 選定した実在語の語彙特性

統制した剰余変数				独立変数
(1)音節数	(2)品詞	(3)親密度	(4)言語内での音韻的 近傍密度	言語をまたいだ音韻的 近傍密度
2	名詞	5~7 mean=5.71 SD=0.69	20~50 mean=34.61 SD=8.43	15~50 mean=26.97 SD=9.94
2	名詞	5~7 mean=5.79 SD=0.61,	20~50 mean=23.81 SD=8.01	0~5 mean=1.14 SD=1.399

音韻的近傍密度は、ある単語から一つの音素を入替、削除、添加してできる単語の数と定義した(Greenberg & Jenkins, 1964)。言語をまたいだ音韻的近傍語を定義するにあたって、日本語とスペイン語の間で、筆者が知覚的に類似していると考えた音素は、同じ音素として扱った。例えば、日本語の音素/b/とスペイン語の音素/b/では、音声的には異なると考えられるが(Celdrán, 2013)、知覚的には類似していると考え、同じ音素として扱った。

2.3. 録音

メキシコ出身のスペイン語母語話者が、上智大学音声学研究室の防音室で刺激語の読み上げを行った。ラップトップのコンピュータ(AppleのMacBook Air)にランダムに刺激語を提示し、読み上げるように指示した。読み上げは各単語について二回ずつ行なった。マイクはSONYのECM-M975を用い、録音はSONYのPCM-D50で行なった。サンプリング周波数は44kHz、量子化ビット数は16bit、録音レベルは6で行われた。刺激音声の音の大きさを一致させるため、ソフトウェアPraatのversion 6.0.40(Boersma & Weenink, 2018)で、各刺激の最大振幅が一致するように編集した。

2.4. 手続き

実験は次の三つの過程で構成されていた。まず(1)語彙性判断課題を行い、次に参加者の単語の知識を測るために(2)スペイン語の語彙力テストを行い、最後に(3)語彙性判断課題で提示した実在語100語を紙で提示し、刺激語をすでに知っていたかどうかを確認した。実験は、上智大学音声学研究室の防音室、神田外語大学の静かな研究室、立教大学の静かな教室で行なった。参加者への指示は日本語母語話者が日本語で行った。実験は約20分で終了した。

語彙性判断課題ではE-Prime 3.0を用い、ラップトップのコンピュータ(EPSON

Endeavor)にヘッドフォン(Pioneer SE-M531)とレスポンスパッド(Cedrus RB-840)を繋げて行なった。参加者には、スペイン語の単語が聞こえたら“Yes”ボタンを、スペイン語の単語ではない音声聞こえたら“No”ボタンを押すように指示した。本番を開始する前に、本番の刺激語に含まれない実在語 5 語、無意味語 5 語の計 10 語を刺激語として、練習を行った。

刺激語に対する反応時間を測定した。反応時間は、刺激音声の開始時点から参加者がボタンを押すまでの時間である。

2.5. 分析

実在語に対する反応時間と誤答率を分析した。誤答率が 40%以上の参加者の結果は、分析対象外とした。それにより、二人の参加者の結果が分析対象外となったため、分析されたのは 14 人分である。

実在語に対して“No”(スペイン語の単語ではない)と回答された回答の反応時間は分析対象外とした。さらに、参加者が知らないと答えた単語に対する反応時間も、分析対象外とした。また、各参加者の試行において、反応時間が、標準偏差の 2.5 倍以上または以下のものは、分析対象外とした。

3. 結果

反応時間と誤答率の分析結果を以下に示す。

表 4: 反応時間と誤答率

	反応時間 (ms)		誤答率 (%)	
	平均	標準偏差	平均	標準偏差
高 PhND	974.0	260.5	13.15	11.25
低 PhND	970.7	254.3	13.0	11.13

3.1. 反応時間

線形混合効果モデルによる分析を行なった。反応時間の分布は正規分布には見られなかったため、対数変換を行ない、線形混合効果モデルに当てはめた。言語をまたいだ音韻的近傍密度が反応時間に影響を与えるかどうかについて、尤度比検定を行った。検定には統計ソフトウェア R(R Core Team, 2017)および R の追加パッケージ lme4(Bates et al., 2015)を用いた。従属変数は反応時間、予測変数には、固定因子として言語をまたいだ音韻的近傍密度、言語内音韻的近傍密度、親密度を含めた。また、実験参加者と刺激語を、ランダム切片として含めた。

言語をまたいだ音韻的近傍密度は、反応速度に有意には影響を与えていなかった(Estimate= 0.00082, s.e=0.00082, $\chi^2 = 1.01, p = 0.31$)。言語内音韻的近傍密度は、反応速度に有意には影響を与えていなかった(Estimate = -0.0015, s.e = 0.0012, $\chi^2 = 1.66, p = 0.19$)。有意水準を 5%とすると、単語の親密度は、反応速度に有

意に影響を与えた(Estimate = -0.037, s.e = 0.017, $\chi^2 = 4.71$, $p = 0.029$)。親密度が高いほど、反応速度は早かった。三つの予測変数の交互作用に有意な効果はみられなかった。

3.2. 正確さ

回答の正誤データに混合ロジスティック回帰モデルを当てはめ、言語をまたいだ音韻的近傍密度が回答の正確さに影響を与えるかどうかについて前節と同様に尤度比検定した。従属変数は、正答または誤答の二値データであり、正答が 1、誤答が 0 のダミー変数とした。予測変数には、固定因子として言語をまたいだ音韻的近傍密度、言語内音韻的近傍密度、親密度を含めた。また、実験参加者と刺激語を、ランダム切片として含めた。

言語をまたいだ音韻的近傍密度は、回答の正確さに有意には影響を与えていなかった(Estimate = -0.0027, s.e = 0.011, $\chi^2 = 0.0475$, $p = 0.82$)。言語内音韻的近傍密度は、回答の正確さに有意には影響を与えていなかった(Estimate = -0.028, s.e = 0.017, $\chi^2 = 2.44$, $p = 0.11$)。単語の親密度は、回答の正確さに有意には影響を与えていなかった(Estimate = 0.31, s.e = 0.24, $\chi^2 = 1.53$, $p = 0.216$)。三つの予測変数の交互作用にも有意な効果はみられなかった。

4. 考察と今後の課題

本研究で、日本語の音韻的近傍密度が、スペイン語の単語への反応時間や誤答率に影響を与えるかどうかを検証した。結果として、統計的に有意な効果は得られなかった。そのため、バイリンガル話者の単語認識において、ターゲットではない言語の単語が活性化し、ターゲット語の認識に影響を与える可能性は低いと考えられる。あるいは、他の要因によって、本実験では有意な結果が得られなかったと考えられる。

考えられる他の要因としては、例えば、刺激語の長さや方言が挙げられる。本実験では、独立変数である音韻的近傍密度をカテゴリカル変数として扱った。そのため、音韻的近傍密度が高い刺激群では、日本語に多い音節構造(e.g. CVCV)の単語が多く、低い刺激群では、日本語に珍しい音節構造の単語(e.g. CVCVC や CCVCV)が多かった。それに従い、二つの群では刺激語の長さが異なっており、それが結果に影響を与えていた可能性がある。

また、スペイン語の方言が要因の可能性もある。今回の実験で用いた刺激語は、メキシコ出身の話者によって発話されたが、実験の参加者の多くは、スペインに長く滞在経験があった。方言や外国語なまりの音声的な特徴は、音韻的近傍密度効果に影響を与えることがわかっている(Imai et al., 2005)。本研究でも、刺激語の音声的特徴と、参加者の心内辞書の音韻表象との差異が、影響を与えた可能性も考えられる。今後の実験では、日本に住むバイリンガル話者が多く耳にしている可能性がある方言を刺激語として使用することで、解決できる可能性がある。

謝辞

本論文は上智大学大学院 2019 年度修士論文に加筆修正をし、まとめ直したものである。本研究の実験に参加して下さった皆様、実験場所を提供して下さった皆様に、深くお礼申し上げます。

参考文献

- Amano, S. & Kondo, T. (1999). *Nihongo-no goitokusei*, Sanseido, Tokyo.
- Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of memory and language*, 26(6), 658-672.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. *Speech perception and linguistic experience*, 171-206.
- Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer Program]. Version 6.0. 40 retrieved in 2018.
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Mit Press.
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, 45(4), 1246-1258.
- Dyer, Frederick N.(1973). The Stroop phenomenon and its use in the study of perceptual, cognitive, and response processes. *Memory & Cognition*, 1.2, 106-120.
- Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, 20(2), 157-177.
- Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *The Journal of the Acoustical Society of America*, 117(2), 896-907.
- Kroll, J. F., & De Groot, A. M. (Eds.). (2009). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.
- Lancaster University (2014). DIALANG. Disponible en : <https://dialangweb.lancaster.ac.uk/>
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, 19(1), 1.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, 6(2), 97-115.
- Martínez Celdrán, E. (2013). Caracterización acústica de las aproximantes espirantes en español. *Estudios de Fonética Experimental*, XXII, 11 - 35.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context

- effects in letter perception: I. An account of basic findings. *Psychological review*, 88(5), 375.
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological science*, 10(3), 281-284.
- Team, R. C. (2013). R: A language and environment for statistical computing.
- Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, 2, 75-94.
- Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50(1), 1-25.
- Yoneyama, K. (2002). Phonological neighborhoods and phonetic similarity in Japanese word recognition (Doctoral dissertation, The Ohio State University).

COVID-19 と教育・研究

北原 真冬

上智大学 外国語学部

E-mail: mafuyu@sophia.ac.jp

1. はじめに

上智大学音声学研究室だけでなく、上智大学および大学院、さらには教育行政全体が、この原稿を執筆している 2020 年 3 月、いわゆる新型コロナウイルス 感染症により、未曾有の危機と呼んでも誇張ではない非常事態を迎えています。この音声研 Working paper の第一義的な目的は、音声学・音韻論の発展に資する研究成果を発表することにあります。この機会に、教育研究機関に人が集まるという基本的な営みにはどんな構造と意義があり、人が集まること自体が阻害されてしまった時にどのような代替手段があるのか、その時考えなくてはならない教育・研究の本質とは何か、について雑感をまとめておきます。

もとより、筆者は教育学・教育工学を専門とするわけではありません。しかし、音声学・音韻論という学問分野の視点を通して、非常時における教育・研究のあり方について今見えることを書き留めておきたいと思っています。

2. 平時の研究と教育

筆者が担当している授業を例にとります。どのクラスにも専用のウィキページをまず作り、パスワードを知る者のみが閲覧可能にしています。書き込みは原則として教員と TA のみが可能ですが、コメント欄を設けることで、授業内の発言をその場で拾い、参加者全員で検討することができます。例えば、英語学科の 2 年生の必修科目である英作文(Composition)の授業では、その場その場での小さな課題として、パラグラフを読んでキーワードを抜き出したり、そのキーワードを検索して関連事項を調べたり、という作業は、コメント欄に書き込むように指示しています。一方、教員側も授業のための資料を紙で配ることはなく、全てウィキに貼っておくか、その場で書き込んでいきます。つまり、授業内での一対一、一対多、多対多、のリアルタイムな細かな教育内容のやりとりはほぼウィキの上で行われていて、音声は「今、こういう作業をしてください」という指示や、「今、すべきことのやり方がわからない」という学生からの質問として主に発せられています。

より講義という形に近い授業の場合も、ウィキに教科書の該当範囲の要点を書いておき、それに沿って説明をしていきます。要点と言うのはテキストで書ける範囲の情報であり、例えば音韻論の授業では、音声記号や樹形図などはその場で板書しながら単語例を書き込んでいきます。音声として筆者が

発するのは、脱線しつづける余談や、インパクトを強めるための比喻、堅い説明をなるべく柔らかい言葉で言い換えること、などであり、時にはマイクを持って学生の間を歩き回り、具体例を聞き出したりもします。

最後に、いわゆるゼミ形式の授業では、学生の側がテキストの担当範囲、あるいは自分の研究成果を発表することに時間の多くを費やし、教員側はその発表内容に対して質問をしたり、少し大きなコンテキストに置き直したまとめを述べて、学生に考えさせたり、という、その場でのやりとりはかなり依存した時間が残りを占めます。ここでもウィキに学生の発表スライドを集約し、授業中および授業終了後も自由に見ることができるようにしています。ゼミ形式の特徴は受講者の側が発信者であること、すなわち講義形式で教員側が用意している作業を部分的にせよ受講者が務めなければならず、それに対して教員は濃いフィードバックを与えなければなりません。

ここまで、いわゆる教育活動における授業時間内での情報のやりとりを描写しました。ウィキをよく使うという以外はそれほど特別なことをしているわけではありませんが、いずれもテキストあるいは図表を含むスライドとそれに重なる音声という2層のコミュニケーションが根幹を成しています。

研究活動においては、はるかにインタラクティブなやりとりが中心となります。例えば卒論や修論のアイデアを練ったり、学会発表のためのポスターを作りあげたりするにあたっては、ホワイトボードの上で即興のアイデアを殴り書きしては、それをお互いにあちこち消したり継ぎ足したりしていく作業が欠かせません。特に音声学の領域においては、全くの単独で行う研究は少なく、大抵は複数人の共同作業になることが多いため、アイデアが創発される場としての対面でのやりとりは非常に重要です。

もう一つ、教育と研究のどちらにおいても、実習的要素が入ってくることがあります。音声学・音韻論の分野では、スペクトログラムを見てセグメンテーションを行ったり、知覚実験のために刺激を提示するプログラムを確認したりするには、マイクの前で録音したり、PCの前でソフトウェアを実際に操作して見せる必要があります。これは必ずしも入門レベルの授業で操作法を手ほどきする時に限ったことではなく、例えば、ある音声コーパスの分析方法について研究グループ内で統一基準を立てる場合にも必要な手順です。このような実習において最も重要なのはリアルタイムで作業が行われることです。参加者が実際に試し、失敗し、すぐにそこでアドバイスを与えて、もう一度試してみる、そのような過程が即時に行われることが必要でしょう。

以上の記述から教育と研究の平時における活動において普段行われていることをキーワードとして取り出すと、[1] 2層のコミュニケーション、[2] インタラクティブ性、[3] リアルタイム性、の3つが挙げられます。次節ではこの3つを非対面でどこまで実現できるかを考えます。

3. 非常時の研究と教育

2020年2月末あたりからCOVID-19に対して日本国内の行政的な急展開が始

まり、多くの大学で卒業式、入学式、ガイダンスなどの行事がまず中止と決まり、続いて授業開始時期を遅らせる判断が出始めました。一方でオンライン授業という形で教育の空白を埋められないかという検討も始まり、技術に明るい一部の教員を除いて、多くは戦々恐々としているのが3月末時点での現状です。前節でまとめたような情報のやりとりを非対面で行うことがどこまでできるか、何を準備すればよいのか、予想できる問題点は何か、ということについての私見を述べます。

[1] 2層のコミュニケーション、すなわち、テキストあるいはスライドとそれに重なる音声というコミュニケーションが根幹であるとまず述べました。このうち教員側から受講者側への単一方向の情報伝達については、いわゆるスライドショーにナレーションを入れるという機能が標準で PowerPoint にも Keynote にも付属していて、すぐにでも作成ができます。しかしながら、まず教員側にとって、その方法で 100 分も喋り続けるのは至難と言えます。30 分ほどの学会発表であっても、スライドを 20-30 枚は用意することから敷衍すると、スライドが 80-100 枚も必要になります。ぶっつけ本番ではつかえつかえでとても聞くに耐えない音声になるため、原稿を用意して練習もしておかなければなりません。

仮に 100 分の授業一コマのために多大な労力と時間をかけてナレーション入りのスライドショーを用意したとして、より大きな問題は、受講する側にとってはほぼ確実に猛烈に退屈で、時間の無駄にしか感じられないものになりやすいことです。結果として、しごく当たり前のことですが、対面式の授業は、その場で実時間耐え忍ぶしかないのに比べて、スライドショーは早送りもスキップも自由自在。スライドにある視覚情報だけを自分のペースでサッと見てしまえば、受講者としては、それで済みとしてしまいがちです。

少しでも受講者の覚醒を保つ（と信じられている）ために取り入れられる方法の一つは、話者の顔を画面に入れることです。スライドの片隅に小さなウィンドウを重ねて、実際に生身の人間が話かけているという風を装うわけですが、これも作る側の苦労は、単なる録音よりもさらに大きくなる割りに、受講者側にもたらす覚醒効果は極めて小さいものです。他人の作ったそのような動画を見たらひどく眠くなるというのに、いったん眠くなったり関心を失ってしまった受講者を引き止める効果が自分の顔にあると信じられるほど筆者は楽観的ではありません。

もう一段覚醒レベルを上げる（と信じられている）ために取り入れられる方法は、話者の半身あるいは全身を画面に入れ、大きなスクリーンやボードの前でダイナミックに話す場面を収録することです。多数のスタッフと多額の予算をかけて作られるテレビ番組としてのニュースは、そのような収録の一つの完成形と言えます。しかし、これを教員が単独で作ることはほぼ不可能です。より個人レベルの製作物である、いわゆる YouTuber のビデオ作品にもかなうはずはありません。自分の授業内容が、日々収録するネタを骨身を削って考え、照明や機材に投資し、編集に膨大な時間をかけている人気 YouTuber と肩を並べる

とは到底思えません。

受講者の覚醒レベルに焦点を合わせると、一見地味ではありますが、意外と効果があるのは Khan Academy 形式の動画ではないかと考えます。画面には黒板のみが映っているような状況で、電子ペンを何色か使って手書きで板書しながらナレーションが入るといった形式です。もちろんナレーションの語り口や書く内容はよく工夫されている必要があり、準備はそれなりに必要です。しかし、iPad やタブレットを使い話すリズムで書いていくのは、作る側にとってはなかなか楽しく作業ができ、普段の授業のようにちょっと軽口を言ったり脱線したりもしやすい印象があります。スライドショーに合わせてのナレーションはなぜか声が重苦しかったり早口になりがちで、脱線もしにくいものです。

[2] インタラクティブ性については、前節の冒頭で描写したウィキのコメント欄という形でもある程度実現はできます。受講者はただビデオを見るだけではなく、見たものについてのコメントを投げ、教員はそれに対してフィードバックを与えるというチャンネルが用意されていることが最低限必要になります。受講人数が少なければ、ひとつひとつのコメントについて丁寧に答えることも可能ではありますが、20 人を超えてくると難しくなります。また受講者側から教員への情報伝達についても[1]の 2 層のコミュニケーションを保証しようとする一気に難度が上がります。[3]のリアルタイム性とも密接に関わってきますが、テキストを打ち込むのに比べて音声はより即時的であり、しかも音声学においては、まさにそれこそが対象物なので、音声についてのインタラクティブ性を確保したい場面は往々にして存在します。

例えば、IPA 表記の記号列をいくつか声に出して読む、という課題を与え、受講者は自分の発音を録音して、オンラインの指定場所に指定時間までに投稿する、その後、正解を発表する、という形のインタラクティブなやりとりは不可能ではありません。ただし、提出物の時間管理、提出ファイルの容量を受け止めるだけのサーバー、そして受講者側の録音品質管理、という種々の技術的な課題は解決しておかなければなりません。

[3] リアルタイム性については、音声と動画が双方向につながっていることが求められます。ソフトウェアの使い方の実習を例にとると、教員側の操作を見ながら説明を聞き、続いて受講者が同じ操作をしているところを教員側がモニターできること、さらにそこでの失敗について即座に質問ができ、回答が得られること、これらを実現するためには zoom や skype のようないわゆるビデオ会議システムの導入が必要であり、しかも初回からいきなり滑らかに双方向のやりとりができるとは限らず、教員側にも受講者側にもテストや慣れが必要となります。

また、受講者側の機材、通信環境を一定水準以上に揃えなければ成立しないというより根本的な問題もあります。ソフトウェアの使い方であれば PC がなければ成り立たないのはもちろんですが、例えば zoom を使用すると通信量が 1 時間で 500MB にも達することもあります。一人暮らしで下宿に光回線などを引いておらず、スマホのモバイル回線のみであれば、あつと言う間にいわゆる「ギ

ガが足りない」状況になってしまいます。Free-WiFi のサービスがあるところに出ていくことになってしまえば、せつかく教員とは非対面を確保しても、不特定多数と長時間近い距離で過ごすハメになります。COVID-19 に対しては特に避けなければならないまずい事態であることは自明です。

4. まとめ

ここまで「非常時」に向けていろいろな方法を紹介してきましたが、あらためて列挙します。

- ウィキサイトにテキスト等を提示しコメント欄を活用
- スライドショーにナレーションを録音
- スライドショーにナレーションを録音し、顔も出す
- テレビ番組あるいは YouTuber 風にダイナミックな講義を収録
- Khan Academy 形式で板書にナレーション
- 受講者は音声録音してアップロード
- zoom などビデオ会議システム利用

「平時」に近づけるためには[1] 2層のコミュニケーション、[2] インタラクティブ性、[3] リアルタイム性、の確保が必要であるとしましたが、上記の方法にはそれぞれ問題があり、「非常時」に「平時」をそのまま持つてくることは容易ではありません。筆者が提案したいのは、ここに列挙した方法を 10-20 分単位でいくつもつなぎ合わせていく手法です。前節で述べた覚醒レベルの問題も、比較的短い時間で目先が変わっていくことによってかなり防げます。例えばスライドショーにナレーションという形で原理的な部分の講義を行い、板書形式で例題を解き、コメント、チャット、あるいは投稿で問題演習を行う。答え合わせと質疑応答には zoom を用いる、というような複合的な形式であれば何とか目を覚ましていられるのではないのでしょうか。

「非常時」がいつまで続くのか、これが教育の根幹を揺るがせてしまうのではないかという危惧は非常に強いものがありますが、非対面で行える方法として、現時点で筆者が考える手段を本稿では提示しました。しかしながら、2 節で触れた「研究」におけるアイデアの創発と言う部分については、残念ながらどの手段でも補いきれないと思います。逆に、溢れるほどの才能があれば手段は何であっても構わないのかもしれないかもしれません。粋なアイデアをどんどん持ち込んでくれる才能を待ち望みます。そして、それこそが「非常時」を乗り越えて生き延びるための最大の希望です。



Sophia University
Phonetics Lab
上智大学音声学研究室

Sophia University Working Papers in
Phonetics

Vol.26 2020

2020年4月30日発行

編集発行 上智大学音声学研究室
代表者 北原 真冬

〒102 - 8554 東京都千代田区紀尾井町7 - 1
電話 03 - 3238 - 3492