

国立国語研究所学術情報リポジトリ

Disfluency of Japanese : Quantity Survey of Unfilled Pause and Prolongation

メタデータ	言語: jpn 出版者: 公開日: 2019-02-14 キーワード (Ja): キーワード (En): 作成者: 佐々木, 藍子, 砂川, 有里子, 浅原, 正幸 メールアドレス: 所属:
URL	https://doi.org/10.15084/00001681

日本語の非流ちょう性 —とぎれと延伸の数量調査から—

佐々木 藍子（国立国語研究所日本語教育研究領域
/東京学芸大学大学院連合学校教育学研究科）[†]
砂川 有里子（国立国語研究所日本語教育研究領域客員教授）
浅原 正幸（国立国語研究所コーパス開発センター）

Disfluency of Japanese —Quantity survey of unfilled pause and prolongation—

Aiko Sasaki (National Institute for Japanese Language and Linguistics/
Doctoral Course The United Graduate School of Education Tokyo Gakugei University)
Yuriko Sunakawa (National Institute for Japanese Language and Linguistics)
Masayuki Asahara (National Institute for Japanese Language and Linguistics)

要旨

これまで日本語教育では、非流ちょうな発話の指導はほとんど行われてきていない。しかし、実際には日本語母語話者であってもよどみのない流ちょうな発話を行うことはまれであるし、非流ちょうな発話が話し手のストラテジーとして用いられることや、聞き手の理解の促進につながることもある。そこで本研究では、非流ちょう性の要因となる「とぎれ」と「延伸」を取りあげ、「多言語母語の日本語学習者横断コーパス (I-JAS)」に収録されている日本語母語話者データの数量調査を行った。その結果、ストーリーテリング (ST1・ST2) とロールプレイ (RP1・RP2) において、とぎれと延伸ではとぎれのほうが多いが、頻度に男女差がないこと、ST1 と ST2 の間、RP1 と RP2 の間のとぎれと延伸の生起の仕方に差がないこと、および、ストーリーテリング (ST1・ST2) とロールプレイ (RP1・RP2) のタスク間においてとぎれと延伸の生起の仕方に大きな違いがあることが分かった。

1. はじめに

近年、日本語教育の現場でもコミュニケーション能力の向上が重視され、指導法も変わりつつある。しかし、学習者の日本語発話は聞き取りにくいことが往々にしてある。聞き取りにくさの要因には、文法の間違いや発音の悪さなども関係するが、それだけではないと考えられる。

一般的に流ちょうだと感じられる日本語母語話者の発話の中にも非流ちょうな要素が出現することも少なくない。また、よどみなくすらすらと話されると、かえって聞き取りにくさを感じることもさへある。そのため、ある程度の非流ちょうさは、何らかの機能を持っているのではないと思われる。以下の例は、同じタスクを行った日本語母語話者および日本語

[†] a.sasaki@ninjal.ac.jp

学習者の発話をできる限り忠実に文字化したものである¹。例1は日本語母語話者の発話で、例2はオーストラリア人日本語学習者の発話である。

(例1) 日本語母語話者

JJ02: あのアルバイトのシフトなんですけどー、〈はい〉今週三日、〈うんうん〉あの一入れさしてもらってるんですけど、でもできれば、〈ええ、ええ〉それをこれから〈ええ〉週二日に一、変更していただきたいんですけど

(例2) オーストラリア人日本語学習者

EAU18: あの一、実はねあの一の〈うん〉、最近一なんか、大学が一すごく一忙しくなりましたけど〈うん〉、あの一今は一、たっみつか、三日間の一バイトをしていますが一〈うーん〉、それは一あの一、二日間に一、あの一変えてもよろしいですか？

日本語母語話者の発話にも、読点部分が表すように頻繁に発話がとぎれていたり、長音記号で表されるように母音が伸びていたりというような現象が見られるが、日本語学習者の発話では日本語母語話者より、とぎれている部分や引き伸ばしている部分が多いことが分かる。これらのとぎれている部分や母音を引き伸ばしているものは、非流ちょう性のつかえの一部であり、「とぎれ」、「延伸」と呼ばれている(定延; 2016)。日本語母語話者の非流ちょう性のパターンに日本語学習者の非流ちょう性を近づけることができれば、聞き取りにくさが減少し、聞き手に負担をかけない話し方ができるようになるのではないかと思われる。

これまでの非流ちょう性の研究では、とぎれや延伸に関する研究はあまり多くない。その中で、定延(近刊)では、非流ちょうな発話の許容可能性が当該言語の膠着性の高さという文法的な概念に影響される可能性を示している。また、砂川・佐々木(2016)では、I-JASの日本語学習者と日本語母語話者の発話データを分析し、以下のような結果を提示している。ハンガリー語を母語とする学習者はとぎれが多く、韓国語を母語とする学習者は延伸が多い。また、スペイン語母語の学習者のとぎれと延伸の使用頻度は日本語母語話者とほぼ同程度で、ハンガリー語、韓国語を母語とする学習者より少ないが、延伸の方がとぎれより多いことが、日本語母語話者と異なっている。これらの結果からはそれぞれの母語におけるとぎれと延伸の使い方が影響していることが推察され、言語によってとぎれと延伸に異なる選考性が存在している可能性が考えられる。しかし、日本語母語話者の非流ちょう性については、まだ明らかになっていない部分も多い。

そこで、本研究では日本語学習者のコミュニケーション能力養成のための基礎研究として、日本語母語話者の非流ちょう性の実態を探ることを目的とする。具体的には日本語母語話者の発話のデータを用い、(1) とぎれと延伸のどちらが多いか、(2) とぎれと延伸に男女差があるか、(3) タスクの違いによってとぎれと延伸の頻度に差があるか、という3つの観点について分析を行い、日本語母語話者の非流ちょう性の一端を明らかにする。

¹ 例文の〈 〉内の発話は対話者のあいづちである。

2. 分析

2.1 対象データ

分析対象のデータは、国立国語研究所で公開されている『多言語母語の日本語学習者横断コーパス International corpus of Japanese as a second language』（以下、I-JAS）を使用した。I-JAS は主に 12 の言語それぞれを母語とする日本語学習者の話し言葉と書き言葉のデータが収録されたコーパスであるが、比較コーパスとして日本語母語話者のデータも収録されている。I-JAS は現在も増補中で、完成すると国内外の日本語学習者 1,000 名分、日本語母語話者 50 名分のデータが収録された大規模な学習者コーパスとなる。現在は第三次公開データにより、日本語学習者 610 名分、日本語母語話者 50 名分のデータが公開されている。I-JAS の特徴の 1 つとして多様なタスクのデータが収録されている点が挙げられるが、本研究では第一次および第二次公開データの日本語母語話者 50 名のストーリーテリング（以下、ST）とロールプレイ（以下、RP）を分析に使用した。ST は 4 コマと 5 コマの 2 種類（ST1、ST2）のコマ割漫画を見ながら、調査者に対してそのストーリーを語る独話である。そして、RP は指定された役を提示された内容に沿って演じながら、調査者と対話するタスクである。RP も 2 種類（RP1、RP2）あり、1 つは日本料理店の店長に週 3 日のアルバイトを週 2 に変えてもらうよう依頼するタスク、もう 1 つは日本料理店の店長から、ホールの仕事から調理の仕事に変更することを依頼されるが、それを断るタスクである²。

2.2 分析方法

I-JAS の言語データは、コーパス検索アプリケーション『中納言』が使用できるよう、自動の形態素解析が行われ、一部の誤解析やミスについては人手での修正も行われている。形態素解析器は正しい書き言葉に対して処理を行うよう設計されているため、言い淀み、フィラー、無意味語などを含む学習者の発話データでは特に誤解析が生じやすい。そこで、それらの誤解析を防ぐため、I-JAS の言語データには 9 種類のタグが付与されており、『中納言』ではタグによる検索も可能である³。

本研究では、タグの 1 つである T タグを分析対象の 1 つとした。T タグとは、自動の形態素解析を行う上で解析の妨げとなる語中の読点や長音記号を削除したり、逆に読点がないため、誤解析を誘発するような箇所を読点を付与したりするための形態素解析用のタグである。この T タグに加え、T タグの付与されていない読点と長音記号を分析対象とし、『I-JAS 中納言』バージョン 2.2.3 短単位データ 20170519 版を使用して検索した。検索したデータのうち、読点がないため誤解析を誘発しそうな箇所に読点を付与する場合の T タグは、検索したデータから除外した。また、あいづちの後にある読点、文末だと思われる箇所や応答詞の後の読点、フィラーの長音も除外し、とぎれの読点と延伸の長音記号に分類し、数量調査を行った。

日本語母語話者の ST と RP について、とぎれと延伸の出現数を集計し、それぞれのデータの総語数に対してとぎれと延伸の出現した割合を算出し、その割合をもとに t 検定および一般化線形モデルを用いた回帰分析を行った。

回帰式は以下の通りである。

$$\text{rate} \sim \text{age} + \text{gender} + \text{task} + \text{feature} + (1|\text{subj})$$

² I-JAS について詳しくは迫田編（2016）を参照されたい。

³ タグについて詳しくは迫田他（2016）を参照されたい。

ここで **rate** はとぎれや延伸の生起確率で推定する値である。**age** は年齢、**gender** は性別（男、女）、**task** はタスク（RP1, RP2, ST1, ST2）、**feature** はとぎれか延伸かで、以上を固定要因とした。**subj** は実験協力者で、ランダム要因とした。

3. 結果

3.1 とぎれと延伸の頻度

タスク別および男女別にとぎれと延伸の出現頻度の割合を産出し、図1の箱ひげ図を作成した。また、統計的な有意差の有無を確認するため、一般化線形モデルでも確認した。以下図1の箱ひげ図は、データのばらつきを反映したものであるが、STにおいてもRPにおいても、とぎれが延伸に比べてかなり多いことがわかる。また、表1の一般化線形モデルの結果においても、とぎれが延伸より、6.9%多い ($p<0.01$) ことが分かった。

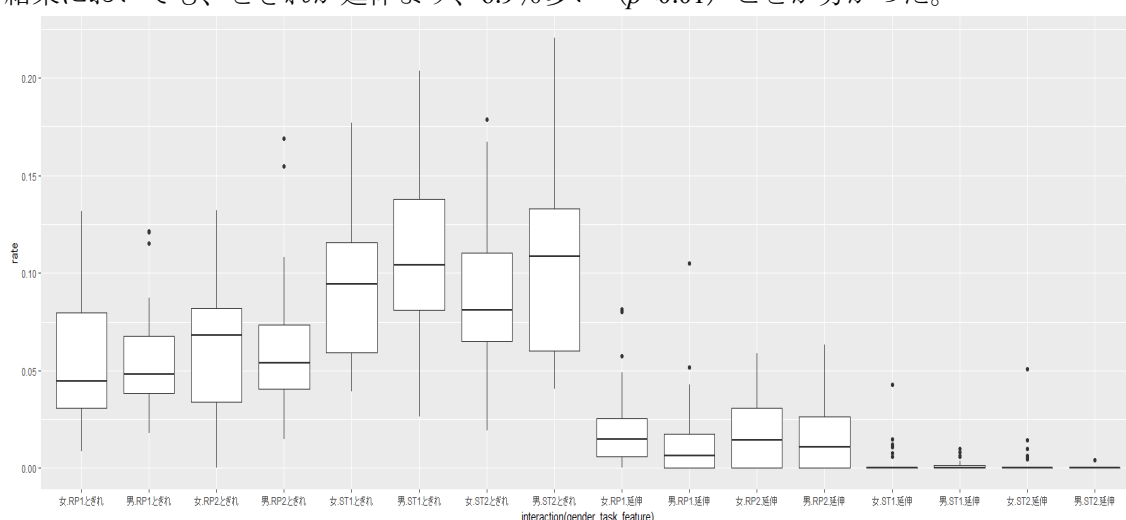


図1. タスク別のとぎれと延伸の箱ひげ図

表1. とぎれと延伸の割合による一般化線形モデル

変数名			推定値	標準誤差
年齢	1歳ごと	age	0.0004**	0.0002
性別：男	vs 女	gender 男	0.003	0.004
タスク:RP2	vs RP1	taskRP2	0.002	0.004
タスク:ST1	vs RP1	taskST1	0.014***	0.004
タスク:ST2	vs RP1	taskST2	0.011**	0.004
延伸	vs とぎれ	feature 延伸	-0.069***	0.003
切片			0.056***	0.008

観測数	400 事例
Log Likelihood	766.802
AIC	-1,515.603

Note:

* $p<0.1$

** $p<0.05$

*** $p<0.01$

3.2 とぎれと延伸の男女差

次にとぎれと延伸の使用頻度に男女の差が統計的に有意かを確かめるため、有意水準 5% で Welch の t 検定を行ったところ、とぎれにおいては、ST1 が $t(43)=1.37, p=0.18$ 、ST2 が $t(43)=1.23, p=0.22$ 、RP1 が $t(48)=0.04, p=0.97$ 、RP2 が $t(42)=0.25, p=0.81$ であり、有意な差は見られなかった。また、延伸においては、ST1 が $t(33)=-1.03, p=0.31$ 、ST2 が $t(26)=-1.66, p=0.11$ 、RP1 が $t(45)=-0.84, p=0.41$ 、RP2 が $t(47)=-0.40, p=0.69$ であり、こちらでも有意差は見られなかった。

3.3 タスクの違いによるとぎれと延伸の頻度差

最後に、タスクによるとぎれと延伸の頻度差の違いを明らかにする前に、同じ種類のタスクである ST1 と ST2、RP1 と RP2 の間で有意差がないか確認した。これは同じ種類のタスクでも、内容によって違いがないかを確認するものであった。有意水準 5% で Welch の t 検定を行ったところ、ST1 と ST2 におけるとぎれは $t(97)=0.51, p=0.61$ 、延伸は $t(97)=0.49, p=0.62$ で、RP1 と RP2 におけるとぎれは $t(97)=0.78, p=0.43$ 、延伸は $t(91)=0.41, p=0.68$ であり、有意な差は見られなかった。そのため、ST1 と ST2、RP1 と RP2 においては、とぎれ、延伸共に同じタスク間では頻度がないことが明らかとなった。そこで、タスクの違いによって、とぎれと延伸に有意な差があるかを確認したところ、差があることが分かった。表 1 を見ると、全体としてはとぎれが 6.9% ($p<0.01$) 多いことが分かるが、図 1 から分かるように、ST にはとぎれが多く、RP には延伸が多いことが分かった。

4. 考察

本研究では、日本語母語話者の非流ちょう性の実態を探ることを目的に 3 つの観点で分析を行った。以下、それぞれの観点ごとに考察する。

(1) とぎれと延伸はどちらが多いか

とぎれは延伸に比べ、有意に多かった。とぎれは非流ちょうなものばかりでなく、発話を流ちょうに進めるためにも必要である。延伸も同様に、非流ちょうな延伸とそうでない延伸が考えられる。本来ならこれらを峻別して、非流ちょうなとぎれと延伸のみを対象とした調査を行う必要があるが、その区別は容易ではない。そのため、今回の調査ではその区別を行わず、流ちょうなとぎれや延伸も含んだ結果を示している。今後はこれらの扱い方も検討していきたい。

(2) とぎれと延伸に男女差があるか

とぎれと延伸については、性別の違いでは有意な差は見られなかった。分析対象の男女の内訳は男性が 23 名、女性が 27 名であった。人数に差はあるが、統計処理を行う際、データごとに発話の総形態素数に対するとぎれと延伸の割合を産出しているため、人数の違いは問題にならない。そのため、とぎれと延伸の使用傾向は男女ともに同じように使用していると考えられる。

(3) タスクの違いによってとぎれと延伸の頻度に差があるか

同じ種類のタスクである ST1 と ST2、RP1 と RP2 のとぎれと延伸の頻度については、有意な差が見られなかった。つまり、ST1 と ST2 のとぎれと延伸の頻度差は見られず、

RP1 と RP2 のとぎれと延伸の頻度差も見られなかった。そのため、ST1 と ST2、RP1 と RP2 をタスクごとにまとめ、ST と RP で比較してみたところ、図 1 から分かるように、ST にはとぎれが多く、RP には延伸が多かった。これは、タスクのタイプの違いによるものだと考えられる。ST は独話であり、RP は対話である。独話は完成された文が使用されることが多いのに対し、対話は中途終了文など、文としての完成を見ない形式が多く使用される。そのため、タスク別のとぎれと延伸の頻度の差は、この影響を受けている可能性が高い。しかし、その他にも、次に述べるようにとぎれと延伸の発話機能の異なりも影響している可能性があるのではないかと考える。

上記 3 つの観点からの分析結果をもとに考察する。ST は、提示された 4 コマと 5 コマのコマ割り漫画のストーリーを調査者に対して話すものである。このタスクはただイラストを見て話せばよいと、一見簡単そうに見える。接続助詞の「て」を多用すれば、1 文で話すことも可能であるが、それだと冗長になり聞き手にとってわかりにくいものになってしまう。このタスクを行う際には、どのような語彙を使用し、どこで文を区切り、きちんと順を追ってストーリー展開を組み立て、最後には結末をどうつけるなど、さまざまなことを考えながら話す必要があるため、とぎれが多くなったのではないかと考える。

一方、RP は依頼や断りをする場面であるが、自身の要望や意向などを相手との関係性や自身の立場を考慮して話さなければならない。さらに、自身の発話に対する相手の反応を伺いつつ、次の発話を考え、自分の要望や意向を伝えていく必要がある。そのため、話し相手に言いにくいことを言う場合や話し相手からの要望を断る場合に延伸を使うことで、ポライトネスを維持しているのではないかと考える。このように非流ちょうな発話はときとしてコミュニケーション方略として用いられていることが考えられる (Lickley (2001)、Bortfeld, et. al. (2001)、伝・渡邊 (2009) など)。

5. おわりに

本研究は、日本語学習者のコミュニケーション能力養成のための第一歩として、日本語母語話者の非流ちょう性の一端を探った。そして、ST と RP において、とぎれと延伸ではとぎれのほうが多いが、頻度に男女差がないこと、ST にはとぎれが、RP には延伸が多く使用され、生起の仕方に大きな違いがあることが分かった。

今後はこの結果をどのように指導に活かすかという方法を検討していく必要がある。そのため、まず日本語学習者の非流ちょう性の習得状況を確認したい。その際、日本語能力や母語の違いによって違いがあるか、違いがあるとすればどのような違いがあるかという点にも着目し、分析を進めていきたい。

謝 辞

本研究は国立国語研究所のプロジェクト「多文化共生社会における日本語教育研究」および科研費基盤(A)「海外連携による日本語学習者コーパスの構築－研究と構築の有機的な繋がりに基づいて－」による成果『I-JAS』を利用して行われたものである。

文 献

迫田久美子編 (2016) 『海外連携による日本語学習者コーパスの構築－研究と構築の有機的な繋がりに基づいて－I-JAS 構築に関する最終報告書 (International Corpus of Japanese as

- a Second Language)』科研費研究報告書 24251010 代表者迫田久美子
迫田久美子・小西円・佐々木藍子・須賀和香子・細井陽子 (2016) 「NINJAL-多言語母語の日本語学習者横断コーパス International Corpus of Japanese as a Second Language」『国語研究プロジェクトレビュー』第6巻3号, pp. 93-110.
定延利之 (2016) 「4つの発話モード」庵功雄・佐藤琢三・中俣尚己 (編)『日本語文法研究のフロンティア』くろしお出版, pp. 205-223.
定延利之 (近刊) 「言語類型からみた非流ちょう性—膠着語と延伸型続行方式のつかえ」『社会言語科学会』21-1
砂川有里子・佐々木藍子 (2016) 「I-JAS を使った非流ちょう性の研究」日本語音声コミュニケーション教育研究会青山研での配付資料
伝康晴・渡邊美知子 (2009) 「音声コミュニケーションにおける非流ちょう性の機能」『音声研究』13-1, pp. 53-64.
Bortfeld, H., S. D. Leon, J. E. Bloom, M. F. Schober, & S. E. Brennan (2001) Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44-2, pp.123-147.
Lickley, Robin J. (2001) Dialogue moves and disfluency rates. In *Proceedings of Disfluency in Spontaneous Speech (DISS)*, pp. 93-96.

関連 URL

コーパス検索アプリケーション『中納言』

<https://chunagon.ninjal.ac.jp/>