

イントネーションの日英比較

—— まだ AI にできないこと ——

柴 田 知薫子

A Comparative Analysis of English Intonation and Japanese Intonation:

What the AI cannot Generate yet

SHIBATA Chikako

competition に置かれる。一方、(6b)では関係節が主語の名詞句とは別のイントネーション句に属し、音調核は team と competition のそれぞれに置かれる。このことから、(6b)の例文に見られる非制限用法の関係節の内容は主語に関する付随的な情報を示していることがわかる。

このように、文の統語構造はイントネーション句の境界に反映される。Wells(2006)によると、個々のイントネーション句(Intonation Phrase: IP)は一つの情報を提示する。話し手はメッセージをいくつかの情報グループに分割するが、その方法には選択の幅がある。まず、日本語の例を見てみよう。

- (7) a. 私は彼に招待された。
- b. 私は、彼に招待された。
- c. 私は彼に、招待された。

(橋谷 2023 : I-②)

(7a)では、客観的事実が一つの情報として一つのイントネーション句で発話されている。したがって、単一のイントネーション句内で自然下降が発生する。橋谷(2023 : I-②)によると、(7b)では「私は」の後に読点を打つことによって「ほかの人は誰に招待されたか知らないけれど、“私は”彼に招待された」という事実を強調している。このとき「私は」よりも「彼に」の方に強い卓立を置くと、社会的評価の高い男性に招待されたという情報が示唆されることになるだろう。(7c)では、話し手が「招待された身分」だということが強調されている。この場合「招待された」は自然下降せず、独立したイントネーション句の冒頭で焦点化されていることがわかる。これがイントネーションの機能③焦点化機能である。

英語では、焦点化される要素は IP の末尾に配置されるとは限らない。

- (8) a. I 'think you've made good 'progress this year.
- b. I 'think you've made good 'progress | 'this year.
- c. 'I think | you've made good 'progress | this year.

(Wells 2006: 191-192)

(8a)の直示的表現 this year は音調核を担わないため、ここでは progress に卓立が置かれて「進歩したね」という解釈になる。しかし、(8b)のように this year を別の IP に切り離して this に卓立を置くと「今年だね」という意味が付加される。さらに(8c)のように主節の後に IP の境界を置き、通常は音調核を担わない人称代名詞 I に卓立を置くと「私が思うには」という話し手の判断が強調されることになる。つまり、(8a)では焦点の領域外にあった this year や I think といった要素が、IP を分割することによって焦点化されたのである。

橋谷(2023 : I-⑤)は、朗読における間の効果には以下のようなものがあると述べている。

- 1) 意味の切れ目
- 2) 次に出てくる言葉を強調する
- 3) 相手にわざと考えさせる
- 4) 驚きや困惑、大きな喜びなど、言葉を失った時の表現
- 5) 気持ちの変化を表す
- 6) 場面の変化を表す
- 7) 相手が感動した後、余韻を与える

このうち 1) の効果はイントネーションの機能②文法的機能に相当し、2) の効果は③焦点化機能に相当する。

日本語ではイントネーション句の冒頭の語に卓立が置かれるから、強調したい語の前で間をとることによってイントネーション句の境界が生まれ、句頭の語が焦点化されることになる。上記の間の効果を、長田弘の詩「最初の質問」第1連に適用してみよう。

- (9) 今日あなたは／空を見上げましたか。
 空は遠かったですか、近かったですか。
 雲は／どんな形をしていましたか。
 風は／どんなにおいがしましたか。
 あなたにとって、いい一日とはどんな一日ですか。
 「ありがとう」という言葉を／今日口にしましたか。

スラッシュはアクセント句の境界として筆者が挿入したものである。1行目に読点はないが、この連の主題は「空」であるから、その前でイントネーション句を区切れば「空」が焦点化される。これに対して2行目の「空」は、すでに読み手に共有された古い情報として背景化される。英語であれば古い情報は人称代名詞で指示されるはずであるから、音調核を担うことはない。3行目の「雲」と4行目の「風」は主題の「空」に関連する語であるから、焦点化を避けることで1行目から4行目までが結束していることを示し、イントネーションの機能④談話機能が実現する。一方、「雲は」「風は」の後に句境界を置くことによって後続する疑問詞「どんな」を際立たせ、主題が変わる5行目の前に長めの間を置いて読み手に考える時間を保証する。それによって間の効果3)が実現すると同時に、場面の変化を表す間の効果6)が実現される。

以上の観察から、これまで言語学の分析対象となってきた文法・焦点・談話に関するイントネーションの機能は人工知能に学習可能であることが期待される。次節では、イントネーションの機能①態度的機能がAIに学習可能かどうかを考察する。

4. イントネーションの態度的機能

英語では、ある文に通常とは異なるイントネーションが付与されると、話し手が聞き手に伝えようとする何らかの意図や心的態度が付加される。

- | | |
|--|---|
| (10) a. Did you take the money?
L H
お金、取りましたか？ | b. Did you take the money?
H L
お金、取りましたね？ |
|--|---|

Yes/No 疑問文は(10a)のように上昇調イントネーションで発話されるのが普通だが、(10b)のように下降調で発話されると話し手の主張(*insistent fall*)を意味する。他方、日本語ではイントネーションの代わりに文末の終助詞「か」と「ね」によって疑問と主張の意味が区別される。次の例も同様である。

- | | |
|--|---|
| (11) a. Do you have a passport?
L H
パスポートお持ちですか？ | b. Do you have a passport?
H L
パスポートお持ちですね？ |
|--|---|

(11a)は持っているかいないかを尋ねているのに対して、(11b)は下降調で「見せなさい」という命令を意味する。日本語では終助詞という形式で区別される意味が、英語では純粹にイントネーションだけで区別されるのだが、果たして(10b), (11b)に示された意味がAIにも学習できるだろうか。さらに、別の例を見てみよう。

- | | |
|----------------------------|-----------------------|
| (12) a. Take off your cap. | b. Take off your cap. |
| HL | LH |

命令文は下降調で発話するのが普通だが、聞き手が子どもである場合(12b)のように上昇調で発話されることがある。上昇調イントネーションによって、相手を安心させる効果があると言われる。

- | | |
|---------------------------|----------------------|
| (13) a. What's your name? | b. What's your name? |
| HL | LH |

WH 疑問文も下降調で発話するのが普通だが、聞き手に答えを促すために(13b)のように上昇調イントネーションで発話されることがある。(12b), (13b)のように聞き手に対する話し手の心的態度をイントネーションで表現することは、人間の意識と無意識の間に存在する意識下の(subconscious)現象であり、生身の人間でもある程度成熟するまでは習得できない機能である。

Cruttenden (1974)の実験によると、大人はアナウンサーの発話“Forfar 3, Stranraer ...”を途中まで聞いてサッカーの試合結果を正しく推測できたのに対して、10歳以下の子どもはイントネーションからStranraerのスコアがForfarよりも多いか少ないかを推測することが困難だったという。また、松井(2013)によると、日本語話者の子どもは大人の皮肉をイントネーションから理解することができず、事実と異なる発話を「嘘をついている」と解釈するという。Wells(2016)によると、少なくとも英語には皮肉を表す特定のイントネーションはなく、ある発話に通常とは異なるイントネーションを付与することによって伝達される。

- | | |
|------------------------------------|-------------------------------|
| (14) a. She's a talented musician. | b. She's a talented musician. |
| H L | H L M |

(14a)は話し手が事実と認定した内容を聞き手に伝達した発話で、断定を表す下降調イントネーションが最後の要素 musician に付与されている。話し手が「彼女は才能あるミュージシャンだ」ということを認めている場合、(14b)のように文末に下降上昇調イントネーションが付与されると(ただし品性がない／ひどく我儘だ)などの含み(reservation)を持たせた表現になる。本来、下降上昇調は文中において継続を表すイントネーションであるから、文末で用いると言外に何らかの意味を含んでいることを示唆することになる。一方、話し手が「彼女の才能」を認めていないにもかかわらず(14b)のような発話をした場合は、「たいしたミュージシャンだね」という皮肉として聞き手に受け取られるだろう。つまり、イントネーションまで同じ発話であっても、話し手の心的態度によってその意味は大きく異なるということである。日本語では、言外の意味を表すような複雑なイントネーションは「さ、よ、ね」に代表される終助詞が担うことが多い。群馬方言には特に「さ」が多いと言われ、文末に「さ」を使うのは東京方言にはない特徴である。「俺さ、昨日さ、高崎へさ、行ったんさ」と言う場合、文末の「さ」には様々なイントネーションが置かれ、東京方言話者には正確に復元することができない。

以上で観察した通り、イントネーションの態度的機能はきわめて複雑であり、その言語の母語話者がある程度成熟するまで獲得できないものである。以下では星新一の『ポッコちゃん』から、ロボットのポッコちゃんと彼女に熱を上げていた青年との会話を見てみよう。

(15)	青年	ポッコちゃん
	「もう来られないんだ」	「もう来られないの」
	「悲しいかい」	「悲しいわ」
	「本当はそうじゃないんだろう」	「本当はそうじゃないの」
	「きみぐらい冷たい人はいないね」	「あたしぐらい冷たい人はいないの」
	「殺してやろうか」	「殺してちょうだい」

二人とも様々な終助詞を用いているものの、ポッコちゃんはロボットゆえに通常のイントネーションを付与していると想定される。すなわち、1行目の「の」には疑問を表す上昇調、3・4行目の「の」には断定を和らげる下降調イントネーションが付与される。他方、ポッコちゃんとの別れを惜しむ青年の終助詞には、恨み・悲しみ・未練など複雑な感情が宿ることだろう。1行目の「だ」は単なる断定とは違うから、下降調ではなく相手の様子をうかがう上昇調になるかもしれない。そして5行目の「か」は明らかに単なる疑問ではないから、上昇調ではなく相手を威圧する下降調になるだろう。このように、日本語では話し手の心的態度が終助詞に反映されることが多く、同じ終助詞であっても感情によって異なるイントネーションが付与される。他方、終助詞という形式を持たない英語では、イントネーション句内で最も強い強勢を持つ語に通常とは異なるイントネーションが付与されることによって、話し手の心的態度が聞き手に伝達されるのである。²

5. まとめ

『ポッコちゃん』は1958年の作品であるが、2021年に出版されたカズオ・イシグロの『クララとお日さま』ではAIロボットのクララが所有者の少女を観察して行動や思考のパターンを学習している。その結果、少女を含む周囲の人々と摩擦を避けたコミュニケーションをとることに成功しているが、前節で述べたような心的態度を自らの発話に反映させる領域までには到達していないようである。話し手の心の働きである心的態度を表現できるのは、この段階でも生身の人間に限られるようだ。

しかしながら映像の分野では、生体情報を収集・分析した上でデータ化し、「感情を見える化」する試みが始まっている。音声の分野でも、AIで怒りの程度を表現することまではできそうだとされている。一方、1980年代からAI研究に関与してきた西垣通氏は朝日新聞のインタビューで「質問に対し、言葉の意味を理解せずに確率計算を行い、もっともらしく易しい文章で答える」AIが人間のような知性を持つ日は来ないし、「人間と同様のことができる汎用AIも出現しない」という見解を示している。

人間同士のコミュニケーションは、相手の思考過程を探りながら進行する。イントネーションは、その際の心の働きを意識下において音声的に反映するものである。AIの出す答えは多くの場合正しいが、その答えにたどり着くまでの思考過程がわからない。AIは聞き手として相手の感情を計算することはできるかもしれないけれども、西垣氏の言う「人間のような知性」すなわち心の働きを表現することまでは、少なくとも現段階ではできないという結論に至る。

注

1. 日本語は母音の長短で意味を区別するため、アクセントを持続時間の長さ (duration) で表現することができない。現代英語の母音には長さの区別がないので、英語ではそれが可能である。
2. いわゆる部分否定の平叙文では、否定する要素に下降上昇調のイントネーションが付与される。

a. I won't eat anything.

H L

b. I won't eat anything.

H L M

(Wells 2006: 32)

a. は全否定 (何も食べない)、b. は部分否定 (何でも食べるわけではない) の意味を表す。

参考文献

- Cruttenden, A. (1974) "An Experiment Involving Comprehension of Intonation in Children from 7 to 10," *Journal of Child Language* 1: 221-32.
- Davenport, Mike and S. J. Hannahs (2010) *Introducing Phonetics and Phonology*, 3rd ed. Hodder Education, London.
- Haraguchi, Shōsuke (1977) *The Tone Pattern of Japanese: an Autosegmental Theory of Tonology*. Kaitakusha, Tokyo.
- 橋谷能理子 (2023) 「アナウンサーが教える朗読講座 I-①~⑥」よみうりカルチャー荻窪, 東京.
- 星 新一 (1971) 『ポッコちゃん』新潮文庫, 東京.
- イングロ, カズオ (2021) 『クララとお日さま』早川書房, 東京.
- 松井智子 (2013) 『子供のうそ、大人の皮肉—ことばのおモテとウラがわかるには—』岩波書店, 東京.
- 西垣 通 (2023) 「AI と私たち—本質を理解する—」朝日新聞 2023 年 8 月 25 日朝刊 11 面.
- 長田 弘, いせひでこ (2013) 『最初の質問』講談社, 東京.
- Sugahara, Mariko (2003) *Downtrends and Post-focus Intonation in Tokyo Japanese*. Published in 2005 by GLSA, University of Massachusetts at Amherst, Mass.
- Wells, J.C. (2006) *English Intonation: An Introduction*. Cambridge University Press, Cambridge.