

韻律と視線が非難の焦点の解釈に及ぼす影響

The Effect of Prosody and Gaze Direction on the Interpretation of the Focus of Accusation

高木 幸子[†], 坂本 暁彦[‡]
Sachiko Takagi, Akihiko Sakamoto

[†]常磐大学, [‡]東京電機大学
Tokiwa University, Tokyo Denki University
takagi@tokiwa.ac.jp

概要

対格 WH 語を伴う修辭疑問文では発話時に聞き手を非難する解釈が生じうる。非難の解釈には、指示的解釈（非難の焦点が対象物へと向かう場合）と非指示的解釈（行為そのものへと向かう場合）の 2 種類があるとされ、それぞれに対して異なる統語・韻律構造が想定できる。当該統語・韻律構造から、指示的解釈の場合には WH 語に、非指示的解釈の場合には動詞に強調アクセントが置かれると予測され、その妥当性が音声聴取実験から検証されてきた。本研究ではさらに、上記のように予測される音声情報と視線情報を同時に呈示した場合、聞き手による指示性解釈がどのように変化するかについて検証された。

キーワード：修辭疑問文、非難の焦点、韻律構造、視線方向、視聴覚統合

1. はじめに

1.1. 研究の背景と目的

我々の日常生活におけるコミュニケーション活動では、多種多様なモダリティから発信されるメッセージを同時に受信しつつも、他者の意図をある程度正確に読み取ることが可能である。この点について、近年では、複数のモダリティから表出される情報を、我々がどのように用いて他者の感情や意図を解釈しているのかに関する研究が進められている。しかしながら、そうした研究においては、非言語的情報であれば非言語的の情報同士の間でその相互的影響を検討することが多く、言語的情報と非言語的情報の双方を対象にした研究は少ない。

そこで本研究では、言語的情報と非言語的情報の両方に焦点を当てる。言語的情報には、たとえば文法構造から導かれる意味や音声、推論から得られる意味が含まれる。一方、非言語的情報には、表情や声の抑揚、身体動作、視線などが含まれる。本研究では、言語的情報として、発話時に聞き手を非難する解釈が生じ、その解釈は 2 通りに曖昧であるとされるが、アクセント位置によっては解釈が一義的になることが示されている対格 WH 語を伴う修辭疑問文を取り上げた。また、

非言語的情報としては、視線を取り上げた。これらの言語的情報と非言語的情報を聞き手に同時に呈示した際、聞き手の解釈がどのように変化するかを実験によって検証する。

1.2. コミュニケーションにおける視聴覚統合

先行研究では、我々が視覚情報と聴覚情報から得られる感情情報を利用し、これらが統合されて他者感情が理解されることが明らかになっている[1-3]。この視聴覚情報の統合においては、一般的には聴覚情報よりも視覚情報の影響が優位であるというモダリティ優位性が示されてきた。しかしながら、顔と声による感情判断においては、たとえば怒りや悲しみの判断においてモダリティ優位性はみられないが、喜びや驚きは顔優位であるといったように、感情ごとにモダリティ優位性の傾向が異なることが明らかにされている[2]。また、この傾向には文化差があることも指摘されており、Tanaka, Koizumi, Imai, Hiramatsu, Hiramoto., & de Gelder [3]の研究から、オランダ人と比較して日本人が音声を重視して他者の感情を読み取る傾向があることがわかっている。

2. 言語的情報と非言語的情報

上述の内容を踏まえ、本研究では言語的情報として対格 WH 語を伴う修辭疑問文を、また、非言語的情報として視線を用いた検討を行う。

言語的情報として扱う対格 WH 語を伴う修辭疑問文は、発話時に相手を非難するのに用いられうる[4-6]。

- (1) 何を食べているの！[非上昇調]
 - a. 「何を」の指示的解釈：聞き手の食べている対象物に言及する形での行為非難
 - b. 「何を」の非指示的解釈：問題の対象物には無関心状態で行う行為そのものの非難

(1)では、当該発話が 2 通りに解釈可能であることが示

されている。著者らによる一連の研究[7-8]では、これらの解釈を指示的解釈・非指示的解釈と名付け、それぞれに対応する統語・韻律構造を仮定している。当該構造による予測のもと、非難の焦点が対象物へと向かう指示的解釈ではWH語に、非難の焦点が行為そのものへと向かう非指示的解釈では動詞に強調アクセント（以下、emphatic accent (EA)）が置かれることになり、この予測の妥当性が音声聴取実験から検証されている。

非言語的情報として扱う視線は、目を使って表出される非言語的情報のひとつであり、空間的注意の手がかりとなる。たとえば、モニター上に呈示されたターゲットを検出する課題では、左右どちらかに視線を向けた顔画像がモニター中央に呈示されたとき、この視線の方向にターゲットが呈示された場合の方がそうでない場合よりも検出時間が短いことが報告されている[9-11]。つまり、他者の視線方向の変化によって、観察者の注意もその視線方向にシフトすることが示唆された。

対格WH語を伴う修辭疑問文における指示性解釈において、非難の焦点を理解するためには、発話者の注意が向かう先を把握しなければならない。これを音声情報のみから行う場合、手がかりはEAに限定される。一方、他のモダリティ情報から得られる手がかりの影響も同時に検討したいのであれば、視線情報が非常に適していると考えられる。具体的には、発話者の視線が常に自分に向けられていれば非難の焦点が行為そのものにあると解釈され、発話者の視線が行為の対象物に向けられれば、当然のことながら非難の焦点が対象物にあると解釈されると予測できる。

本研究では、EA位置の異なる対格WH語を伴う修辭疑問文という言語的情報と、視線方向の変化という非言語的情報を組み合わせた刺激を用いて、非難の焦点の解釈の変化とモダリティ優位性について検証する。

3. 手続き

3.1 実験参加者

実験参加者は11名（男性6名、女性5名）であり、平均年齢は19.36歳（ $SD = 0.64$ ）であった。

3.2 刺激

音声刺激 音声刺激は、高木・坂本[8]と同一のものを使用した。具体的には、女性発話者から「何を食べてるの」と「何を讀んでるの」の2種類の発話内容について、WH語にEAもしくは動詞にEAを置いて相手を

非難するように発話している4種類の音声を収録した。これらの音声に関して、Praatを用いて発話前後のポーズなどの調整を行った。その結果、4種の音声はいずれも、WH語部分が約0.45秒、動詞部分が約0.75秒であり、発話前後に0.15秒のポーズを加えて全体で約1.50秒であった。

視線動画刺激 視線動画刺激は、Windows Movie Makerを用いて視線静止画を連続的に組み合わせて作成した。発話時の顔画像動画を用いなかった理由は、非難の意図を込めて発話すると、抑制しようとしても感情情報が反映されてしまうことが不可避であり、これが剰余変数となる可能性を避けるためであった。具体的には、女性モデルがカメラのレンズを正視した場合の静止画およびレンズ位置から約30cm下方を正視した場合の静止画を撮影した。この静止画から目元部分のみをスリット状に切り出し、レンズ静止画像のみ（約1.50秒）からなる動画・レンズ静止画像（約0.15秒）と下方正視画像（約1.35秒）からなる動画・レンズ静止画像（0.15秒）と下方静止画像（約0.45秒）およびレンズ静止画像（約0.90秒）からなる動画の全3種の動画を作成した。以下では、これら3種の刺激を順に、人物凝視刺激・対象物凝視刺激・双方注意刺激とする。

視聴覚刺激 先述した音声刺激と視線動画刺激を組み合わせて、視聴覚刺激を作成した。これにより、全12種（発話内容2×EA位置2×視線動画刺激3）の動画が作成された。

3.3 手続き

実験は集団形式で実施され、視聴覚刺激が呈示される第1セッションと、視線刺激および音声刺激が呈示される第2セッションから構成されていた。第1セッションは3つの教示条件（教示なし・視線注意・音声注意）ごとに行われ、各刺激は4回繰り返して呈示された。いずれの条件においても、発話者による非難の焦点が、対象物と行為そのもののどちらに向けられているのかについて2択で判断を求めた。ただし、視線注意条件では出来る限り音声を無視して視線に着目し、音声注意条件では出来る限り視線を無視して音声中に着目して回答することを求めた。第2セッションでは、視線刺激と音声刺激を単独で呈示し、第1セッションと同様の判断を求めた。

4. 結果

本研究の目的は、非難の焦点が異なる音声と視線を

組みあわせて呈示した場合に、非難の焦点に関する指示性解釈がどのように変化するかについて、また、視線変化の違いによってそれらに差がみられるのかについて、心理実験を用いて検討することであった。以下に結果を述べる。

4.1 視線および音声単独呈示実験の結果

視線刺激を単独呈示した場合、非難の焦点が人物凝視刺激では行為そのものにあると選択された割合は1.00 ($SD = 0.00$)、対象物凝視刺激では対象物にあると選択された割合は0.98 ($SD = 0.07$)であった。また、双方注意刺激においては、非難の焦点が対象物にあると選択された割合は0.45 ($SD = 0.46$)であった。

音声刺激を単独呈示した場合、WH語にEAが置かれた刺激において非難の焦点が対象物にあると選択された割合は0.92 ($SD = 0.12$)であり、動詞にEAが置かれた刺激において非難の焦点が行為そのものにあると選択された割合は0.97 ($SD = 0.06$)であった。

ここで、人物凝視刺激と対象物凝視刺激においてそれぞれの選択肢が選択された割合の平均と、2種の音声刺激においてそれぞれの選択肢が選択された割合の平均に差がみられるかどうかを検討するため、対応のある t 検定を実施した。検定の結果、有意差はみられなかった ($t(10) = 1.90, n.s.$)。

4.2 教示なし条件における視聴覚刺激実験の結果

教示なし条件の反応については、非難の焦点が異なる音声を組み合わせた刺激に対する解釈が、視線と音声のどちらに基づいて行われるのかを検討した。そこで、第一に、4.1の結果に基づき、人物凝視刺激および対象物凝視刺激と、EA位置がそれぞれ異なる音声を組み合わせた刺激に対する反応を対象とした分析を実施した。この組み合わせにおいては、非難の焦点に関して、人物凝視刺激と動詞部分にEAが置かれた音声の組み合わせおよび対象物凝視刺激とWH語部分にEAが置かれた音声の組み合わせは一致刺激、それ以外の組み合わせについては不一致刺激となる。この不一致刺激のみを対象に、実験参加者ごとの反応が視線あるいは音声に基づく割合を算出し、 t 検定を実施したところ、有意差はみられなかった ($t(10) = -0.81, n.s.$)。つまり、非難の焦点が異なる音声と視線の組み合わせからなされる指示性解釈は、視線あるいはEA位置のどちらか一方に依拠してなされていない傾向がみられた。

双方注意刺激と、EA位置がそれぞれ異なる音声の組み合わせに関しては、4.1の結果に基づき、実験参加者の視線のみに対する反応に沿って一致刺激と不一致刺

激とに分類を行い、そのうえで実験参加者ごとの不一致刺激のみを対象に同様の分析を実施した。 t 検定の結果、有意差がみられ ($t(10) = 3.88, p < .01$)、視線反応よりも音声反応の割合が高かった。これを図1に示す。つまり、視線が対象物と行為そのものの両方に向けている場合には、視線よりも音声に基づいて非難の焦点が解釈される傾向がみられた。

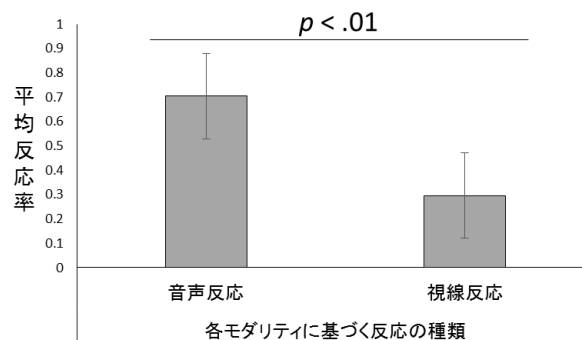


図1 教示なし条件における視線反応と音声反応の平均値 (エラーバーは標準偏差)

4.3 注意あり条件における視聴覚刺激実験の結果

視線注意条件および音声注意条件の反応については、視線と音声というモダリティが異なる情報の影響の強さを比較することを目的として分析を実施した。

第一に、人物凝視刺激および対象物凝視刺激と、EA位置がそれぞれ異なる音声を組み合わせた刺激に対する反応を対象とし、独立変数を注意課題(視線/音声)とする t 検定を実施した。従属変数として一致性効果(=一致刺激における正答率 - 不一致刺激における正答率)を用いた。視線注意課題における一致性効果が高いことは判断の際に無視すべき音声から強く干渉を受けたことを意味し、音声注意課題における一致性効果が高いことは判断の際に無視すべき視線から強く干渉を受けたことを意味する。 t 検定の結果、注意課題間で有意差はみられなかった ($t(10) = -0.34, n.s.$)。つまり、視線が人物あるいは対象物を凝視している場合には、視線注意課題における音声からの干渉と音声注意課題における視線からの干渉に差はみられなかった。

第二に、双方注意刺激とEA位置がそれぞれ異なる音声の組み合わせに関しては4.1と同様に、実験参加者ごとに一致刺激と不一致刺激に分類したうえで一致性効果を算出し、第一の分析と同様の t 検定を実施した。 t 検定の結果、有意傾向がみられ ($t(10) = 1.76, p < .10$)、音声注意課題よりも視線注意課題における一致性効果の方が大きい傾向がうかがえた。これを図2に

示す。つまり、音声注意課題における視線の干渉よりも、視線注意課題における音声の干渉の方が大きい可能性がみられた。

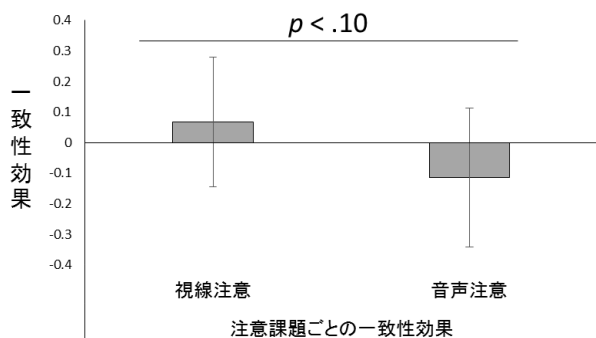


図2 注意課題ごとの一致性効果
(エラーバーは標準偏差)

5. 考察

視線および音声の単独呈示実験の結果から、視線が対象物と人物のどちらに向けられているかに応じて、また音声におけるEAがWH語と動詞のどちらに置かれているかに応じて、非難の焦点の解釈が明確に分かれることが示唆された。具体的には、実験者の予測の通り、視線が対象物に向けられている場合と音声におけるEA位置がWH語にある場合には非難の焦点は対象物にあると解釈され、視線が人物に向けられている場合と音声におけるEA位置が動詞にある場合には非難の焦点は行為そのものにあると解釈されていた。また、2種の視線刺激と2種の音声刺激において、非難の焦点が上述の予測通りに解釈された平均割合に差がみられなかったことから、本実験のために作成された刺激には実験者らの意図が十分に反映されていたと考えられる。また、視線単独呈示実験において、視線がWH語部分では対象物に、動詞部分では人物へと向けられる双方注意刺激では、非難の焦点の解釈は対象物と行為そのものに明確に分かれなかったことが示された。ただし、この傾向には個人内では偏りがみられたため、個人の中では明確な基準があると考えられよう。

次に、視聴覚刺激呈示実験の結果について考察する。視線が対象物あるいは人物を凝視した刺激とEA位置がそれぞれ異なる音声を組み合わせた視聴覚刺激に対する非難の焦点の解釈は、教示なし条件の結果から視線情報と音声情報のどちらか一方に基づいて行われるわけではないことが示された。また、視線もしくは音

声注意条件の結果から、各モダリティ情報からの干渉量には差がないことが示された。しかしながら、視線がWH語部分では対象物に、動詞部分では人物に向けられた刺激の分析においては、教示なし条件では視線情報よりも音声情報に基づいて非難の解釈がなされることが示され、また各モダリティからの干渉量に関しても、音声に注意を向けた場合の視線情報からの干渉よりも、視線に注意を向けた場合の音声情報からの干渉の方が大きい可能性が示唆された。この傾向を感情研究の関連から捉えると、非難の焦点に関する指示性解釈においては、音声重視というモダリティ優位性がみられると解釈できる。

謝辞

本研究は、JSPS 科研費若手研究 (B) (課題番号: 17K13474) の支援を受けた。

文献

- [1] de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, 14, 289-311.
- [2] Takagi, S., Hiramatsu, S., Tabei, K. I., & Tanaka, A. (2015). Multisensory perception of the six basic emotions is modulated by attentional instruction and unattended modality. *Frontiers in integrative neuroscience*, 9, 1.
- [3] Tanaka, A., Koizumi, A., Imai, H., Hiramatsu, S., Hiramoto, E., & de Gelder, B. (2010). I feel your voice: Cultural differences in the multisensory perception of emotion. *Psychological Science*, 21, 1259-1262.
- [4] Konno, H. (2004). The *nani-o* X-o construction. *Tsukuba English Studies*, 23, 1-25.
- [5] 天野みどり (2008). 「拡張他動詞文—「何を文句を言ってるの」—」『日本語文法』8(1), 3-19.
- [6] 高見健一 (2010). 「「何を文句を言ってるの」構文の適格性条件」『日本語文法』10(1), 3-19.
- [7] 坂本暁彦・高木幸子 (2018). 「日本語修辭疑問文における対格WH語の指示性—統語構造に基づく韻律構造の予測—」『電子情報通信学会技術研究報告』117(509), 101-105.
- [8] 高木幸子・坂本暁彦 (2018). 「日本語修辭疑問文における対格WH語の指示性—音声聴取実験による二種類の統語構造の検証—」『電子情報通信学会技術研究報告』118(49), 83-88.
- [9] Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review*, 5(3), 490-495.
- [10] Driver IV, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual cognition*, 6(5), 509-540.
- [11] 吉川左紀子・佐藤弥 (2000). 「社会的メッセージ検出機構としての顔知覚—表情と視線方向による促進 (特集: 日本における表情研究)—(表情研究の現状 (2) 心理的, 社会的観点)」。『心理学評論』43(2), 259-272.