

非言語音による発話者の同一性識別

北神慎司・遠藤菜文・池田賢司・高橋知世
(名古屋大学大学院環境学研究科)

キーワード：耳撃証言，同一性識別，非言語音

犯罪場面に遭遇した場合、必ずしも犯人の顔を目撃するとは限らない。マスクを被ったり変装をしたりする犯人や、暗闇にいる犯人、電話を使った犯人については、顔を目撃することはできない (Yarmey, 2006)。また、被害者や証人が視覚障害を持っている場合も同様である。そのような場合では、犯人の顔ではなく、声が重要な証拠となってくる。このような音声の記憶は、目撃記憶に対して、耳撃記憶と呼ばれる。そして、事件や事故といった、何らかの出来事を目撃ではなく耳撃した後に、その記憶を思い出して証言することを、耳撃証言と呼ぶ。例えば、*United States v. Duran* (1993)では、強盗に襲われた銀行の出納係が犯人の声について証言し、その証言が公判での主たる証拠となり、有罪判決が下されている。

さて、実際の犯罪場面においては、被害者が唸り声や咳、溜息などの非言語音しか耳撃していないケースが考えられる。しかしながら、そういった非言語音の同一性識別については、これまで全く検討されてこなかった。そこで、本研究では、耳撃記憶における非言語音の同一性識別が可能であるか否かを検討するために、学習と再認の発話内容の異同を操作して、2つの実験を行った。

実験 1

方法

実験参加者：大学生，大学院生，社会人 44 名
実験計画：学習（言語音・非言語音）×再認（言語音・非言語音）の 2 要因参加者間計画
刺激：音声刺激は、予備調査で選定した発話者 25 名と、言語音 10 種類・非言語音 10 種類の発話内容 20 種類、計 500 個を使用した。学習と再認で発話内容を異なるものにし、再認テストでは、ターゲット 1 人とフォイル 4 人を設定した。

手続き：学習フェーズでは、言語音もしくは非言語音の音声は 2 回繰り返して提示された。その後、干渉フェーズとして 30 秒間計算問題を実施した。最後に、再認フェーズでは、ラインナップを 2 回聞いた後、5 つの音声の中から 1 つを 5 段階の確信度とともに回答するように求めた。

結果と考察

各実験条件の平均再認率を図 1 に示す。再認率について、学習（言語／非言語）×再認（言語／非言語）の 2 要因分散分析を行った結果、学習と再認ともに主効果が有意であった（学習： $F(1, 40) = 4.34, p < .05, \text{partial } \eta^2 = .10$ ；再認： $F(1, 40) = 12.07, p < .01, \text{partial } \eta^2 = .23$ ）。これに対して、学習と再認の交互作用が有意ではなかった（ $F(1, 40) = .02, p = .89, \text{partial } \eta^2 = .00$ ）。また、チャンスレベル（20%）との t 検定を行った結果、言語一言語条件、言語一非言語条件、非言語一言語条件でチャンスレベルよりも有意に高い再認率であった（言語一言語条件： $t(10) = 3.92, p < .01$ ；言語一非言語条件： $t(10) = 2.60, p < .05$ ；非言語一言

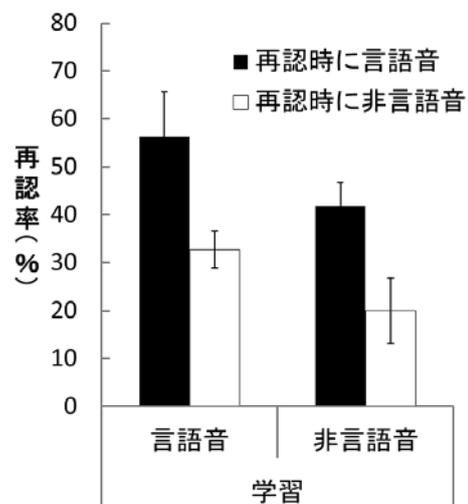


図 1 実験 1 の結果

語条件： $t(10)=3.18, p<.05$ ）。これに対して、非言語—非言語条件では、チャンスレベルと有意な差はなかった ($t(10)=.00, p=1.00$)。

つまり、学習、再認のいずれの条件でも、音声と言語音である場合の方が、再認成績が高かった。この結果は、同一性識別は、学習時、再認時共に、言語音で呈示されていた場合に、より正確になることを示すものである。したがって、同一性識別は、学習時、再認時ともに、言語音で呈示されていた場合に、より正確になることが示唆された。さらに、非言語—非言語条件以外の条件で、チャンスレベルよりも、再認成績は有意に高かったことから、学習、再認のいずれかで、非言語音が呈示されていても、いずれかで言語音が呈示されていれば、ある程度の精度を持って弁別可能であるといえる。

実験 2

方法

実験参加者：大学生，大学院生，社会人 20 名

実験計画：実験 1 と同様

刺激・手続き：実験 1 と同様。ただし、学習と再認で発話内容を同じものとした。

結果と考察

各実験条件の平均再認率を図 2 に示す。各実験条件の再認率の差を見るために、 t 検定を行ったところ、有意な差は見られなかった ($t(18)=-.73, p=.78, d=-.15$)。また、チャンスレベル (20%) との t 検定を行った結果、言語条件、非言語条件ともに、チャンスレベルより有意に高い確率であ

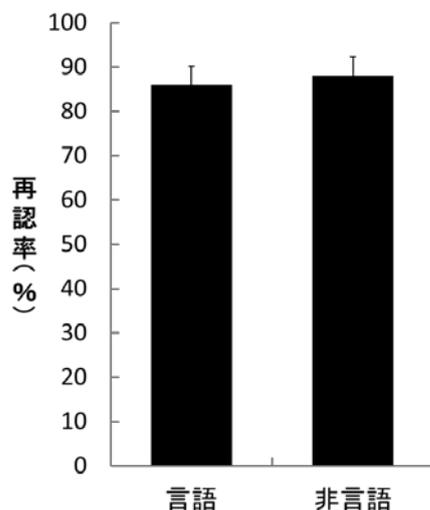


図 2 実験 2 の結果

った (言語条件： $t(9)=15.46, p<.01$ ；非言語条件： $t(9)=15.38, p<.01$)。

つまり、学習時と再認時で、同じ発話内容が呈示される場合は、言語音、非言語音に関わらず、高い精度を持って弁別可能であるといえる。このような結果が得られた背景には、学習と再認で同一の音声を使用したことで、音声の弁別の手がかりが豊富であったためだと考えられる。

総合考察

本研究では、実際の犯罪場面において、犯人の唸り声や笑い声、溜息などの非言語音を耳撃した場合、警察の捜査や裁判において、ラインナップに言語音を使用するか、あるいは、耳撃した音声と同一の発話内容をラインナップとして使用することで、裁判において重要な耳撃証言として採用できる可能性が示唆されたといえる。つまり、非言語音が耳撃証言の対象となることで、警察による捜査手段の幅が広がり、さらに、裁判の円滑な進行につながるであろう。

しかしながら、声のラインナップ手続きにおけるガイドラインには、ラインナップ作成時に、犯人による特定の単語や語句の意図的な歪曲を防ぐため、犯行中に犯人が発話した単語や語句を含めるべきではないと示されている (Yarmey, 2006)。これに対して、耳撃した音声为非言語音の場合、犯人が発話した音声をラインナップ音声に含めることによって、証言者が確信をもって同一性を識別できる可能性があることから、今後、非言語音を耳撃した場合には、耳撃した発話内容と同一の発話内容を、ラインナップ音声として使用可能なように、ガイドラインを改定する必要がある。そのためには、非言語音においても発話者による意図的な歪曲が可能であるのか、また、その歪曲された音声が、同一性識別の正確性に影響を与えるのか否かを、今後の研究で検討することが重要になってくる。

引用文献

United States v. Duran, 4 F.3d 800 (9th Cir. 1993).
Yarmey, A. D. (2006). The Psychology of Speaker Identification and Earwitness Memory. In Lindsay, R. C. L., Ross, D. F., Read, J. D., & Toglia, M. P. (Eds.), *The Handbook of eyewitness psychology: Volume II: Memory for people*. Lawrence Erlbaum Associates, Publishers. New Jersey, pp. 101-136.