

潜在学習を利用した映像・音声の伝送遅延の心理的影響の測定

Application of Implicit Learning to Measurement of Psychological Effect of Transmission Delay
in Audiovisual Communication

大西 仁^{1,3)}, 平賀 健²⁾, 望月 要^{1,3)}, 結城暁曠^{1,4)}

Hitoshi OHNISHI, Ken HIRAGA, Kaname MOCHIZUKI, and Kiyohiro YUKI

1) メディア教育開発センター

(〒261-0014 千葉市美浜区若葉2-12, {ohnishi,moc,yuki}@nime.ac.jp)

2) 千葉工業大学工学研究科情報工学専攻

(〒275-0016 習志野市津田沼2-17-1, hiraga@nao.net.it-chiba.ac.jp)

3) 総合研究大学院大学文化科学研究科, 4) 図書館情報大学情報メディア研究科

Abstract : We propose an objective method to measure a psychological effect caused by transmission delay of the image and sound in audiovisual communication. The proposed method uses participant's performance of learning of an artificial grammar (finite state automaton), which is a paradigm in the study area of implicit learning, as an index of psychological effect of transmission delay. This article explains an overview of implicit learning used as the theoretical background of the measurement, and discusses advantages of the proposed method over the existing methods. Subsequently, it reports the result of preliminary experiments, and discusses the problems of measurement of psychological effect caused by transmission delay.

Key Words: 伝送遅延, 心理的影響, 測定, 潜在学習

1. はじめに

マルチメディア通信における映像・音声の伝送遅延、映像と音声のずれは、『コミュニケーションに違和感がある』、『コミュニケーションが妨害されている』、『相手の話がわかりにくい』、『会話が盛り上がらない』等、コミュニケーションの円滑さを妨げる原因になると言われている。この問題は、早くから認識されており、主に通信・放送の品質評価の観点から研究されてきた [2]。さらに、最近の社会心理学研究では、映像・音声の伝送遅延やずれによりコミュニケーションの円滑さが低下させられた時、その原因が伝送経路ではなく、対話相手に帰属されることにより、『相手は話が下手』、『相手は頭が悪い』というように対人印象を低下させる可能性も指摘されている [7]。したがって、映像・音声の伝送遅延が与える心理的影響を正しく理解すること、正確に評価することの重要性は高まっていると考えられる。

映像・音声の伝送遅延の心理的影響に関する研究の主流は、伝送遅延の検知限・許容限、あるいは相手の話のわかりやすさ・面白さ等の印象を、主観評価法で測定するものである。しかし、これらの方法では、映像・音声の伝送遅延が認知にどのような影響を与えているかを特定することができない。また、被験者の評定に頼る方法では被験者の意識に上らない心理的な影響を測定することができない。通

信システムの物理的特性から通信品質を客観的に評価する研究も見られるが、これらは主観評価における評価者の負担やバラツキを減ずることを目的としており、物理的特性から主観評価値を推定するというものである。したがって、このような客観的評価の方法では、主観評価に関する問題は解決できない。

主観評価法に代わる方法として、皮膚電導反射 (SCR) や脳波 (EEG) 等の生理的指標を用いることも考えられる。適切な生理的指標を用いれば、被験者が意識化、あるいは言語化できないような微妙な影響をリアルタイムで測定できる可能性がある。しかし、現時点では、これらの生理的指標と認知・行動のメカニズムとの関係は明らかではない。したがって、生理的手法では何を測定しているのかわからないという問題がある。

本研究では、主観評価法に代わる、映像・音声の伝送遅延の心理的影響の評価法として、潜在学習の一種である人工文法学習のパフォーマンスを指標とする方法を提案する。本稿では、測定法の理論的背景となる潜在学習について説明した後、潜在学習を利用した映像・音声の伝送遅延が与える心理的影響の測定法について述べる。ついで、提案した測定手法による予備実験の結果について述べ、効率的な測定のための課題について議論する。

2. 人工文法の潜在学習

人間が行なう学習は、教科学習に代表されるような意識的な学習と、刺激を生成する環境の構造や手続きの無意識的な学習に分類され、後者は潜在学習と呼ばれている。実際、人間は無意識のうちに多くのことを学習し、それを行動に反映させていることが知られている。例えば、母国語の学習や、人間の顔のような複雑なパタンの分類はその代表例である。潜在学習では、学習された知識は後の課題に利用できるが、その知識を完全には言語化できなかったり、意図的な利用ができないとされている。

人工文法の学習は、潜在学習研究の最も代表的な課題であり、その典型的な概観は以下の通りである。有限状態機械に基づく文法(図1)によって生成された文字列が課題として利用される。生成された文字列のうちのいくつかは学習時に提示され、被験者にはその記銘(記憶すること)が教示される。記銘後、文字の並びを規定する規則(文法)の存在が教示され、転移課題では、提示される新しい文字列が文法的か否かを判断することが要求される。被験者は、チャンスレベル以上に正しい判断ができるが、判断の理由を明確に述べることができない。この実験で重要なのは、被験者は単に記銘時に提示された文字列を記憶しただけでなく、非明示的にはあるが、文法構造を学習していることである。通常の記憶テストとはこの部分が異なる。

潜在学習で用いられる課題に共通しているのは、規則によって規定された刺激を繰り返し経験することで、被験者がその規則を知っているかのように振る舞うこと、それにも関わらずそれを明示できないことである[4]。また、学習後のパフォーマンスは、チャンスレベル以上で正答するようにはなるが、100%近い正答には至らない場合が多く、学習の成立が微妙であることも挙げられる。そのため、たとえ微小でも学習の阻害要因があると、学習が成立しなくなったり、学習のパフォーマンスが低下する可能性がある。

3. 潜在学習を利用した映像・音声の伝送遅延の心理的影響の測定

本研究では、前節で述べた潜在学習の特徴を利用し、人工文法の潜在学習を課題として、映像・音声の伝送遅延や映像と音声のずれの心理的影響を測定する方法を提案する。潜在学習の研究者は、潜在学習は単に提示された刺激を機械的に記憶するのではなく、刺激の構造を学習していること、学習が無意識的に行なわれることを強調している。しかし、本研究は映像・音声の伝送遅延が与える心理的影響の測定を目的としているので、これらの主張には特にこだわらない。

映像・音声の伝送遅延の影響の測定という目的に合うように、次のような手続きをとる。

1. 文法の存在の教示 提示される刺激音声を、ある規則に従って生成されたものとそうでないものに分類することが、課題であることが教示される。
2. 刺激音声の提示 刺激は音声(必要なら映像と共に)で提示される。刺激音声はある有限状態機械に受理される

もの(文法的)及び受理されないもの(非文法的)が提示される。

3. 文法判断 被験者は、刺激音声は1つ提示されると、提示された刺激音声が文法的であるか否かを判断する。
4. 正誤フィードバック 被験者が判断を行うと、設定した遅延時間の後、判断の正否がフィードバックされる。
5. 2~4を繰り返す。

被験者の判断に対する正誤フィードバックに遅延を入れることが、通信における伝送遅延に相当する。また、刺激音声の提示において、音声と映像を併せて提示し、音声と映像の同期を外せば、これが映像・音声のずれに相当する。

フィードバック遅延が大きくなると、学習のパフォーマンス(e.g. 正答率)が低下することが予測される。認知心理学分野の潜在学習研究ではその例を知らないが、人間を含む動物のオペラント学習においては、強化(フィードバック)に遅延が入ると、学習のパフォーマンスが低下することが知られている。また、機械学習分野の強化学習においても、即時強化に比べて、遅延強化時の学習は難しい。

したがって、学習のパフォーマンスを指標とすることにより、映像・音声の伝送遅延が与える心理的影響を客観的・定量的に測定することができると考えられる。

4. 実験

提案した手法により、実際に映像・音声の伝送遅延が与える心理的影響を測定できることを確認するために、音声刺激を用いた潜在学習時において、正誤フィードバックに遅延を加えた場合に生じる学習成績の変化を検討した。

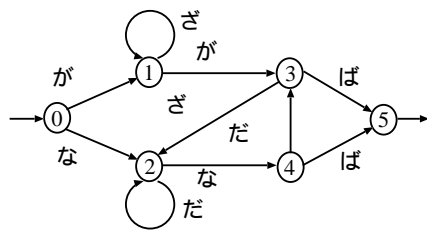
4.1 方法

被験者 大学学部生 32 人を、フィードバック遅延の大きさ(150, 300, 450, 600ms の 4 種類)、文法(全員が文法 1, 2 の 2 種類を学習。いずれか一方は遅延有り、他方は遅延無し)、提示順序(文法 1 が先、後)の組合せによる 16 群のいずれかにランダムに割り当てた。

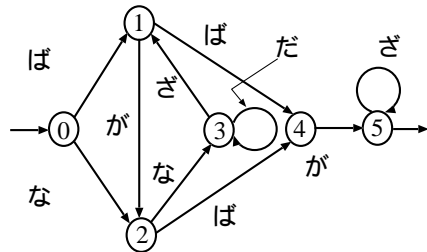
材料 図1に示すような2種類の有限状態機械から生成される文字列を音声合成ソフトウェア(Microsoft's Speech API)で音声ファイルに変換し、刺激として提示した。音声ファイルのフォーマットは、RIFF (little-endian) data, WAVE audio, Microsoft PCM, 16 bit, mono 22050 Hz とし、日本人女性風の音声を使用した。

文法 1, 2 は Reber[6] で用いられたものをそのまま使用した。本研究では、刺激は音声として提示することから、文字列は発音可能なものに変更した。文法 1, 2 の複雑さは情報理論的観点からは同じである。

文法 1, 2 とともに、文字列の長さが 3~8 で文法的な(有限状態機械受理される)ものは全部で 43 個である。文法 1, 2 に関して、各々文法的な文字列を 43 個、非文法的な文字列を 43 個を刺激のセットとした。発音のスピードは 1 文字あたり 250ms とした。



文法 1



文法 2

図 1: 実験で使用した人工文法 .

非文法的な文字列は、文法的な文字列の最初と最後以外の 1 文字をランダムに他のアルファベットに変化させることにより生成した。非文法的な文字列が実際に非文法的であること、86 個の文字列がすべて異なることが確認された。

刺激音声の提示はヘッドフォンを通して行なわれた。86 個の刺激音声の提示順序はランダムとした。

手続き 実験の課題、および手順はインストラクションシートにより教示された。

刺激音声はヘッドフォンを通して提示され、被験者は 1 文字列分の刺激音声提示される毎に、それが文法的であるか否かを判断し、キーボード入力によって回答した (文法的なら “A”、非文法的なら “R” のキーを押す)。被験者の回答入力後、設定された遅延時間の空白を空け、回答の正誤がディスプレイに表示された (図 2)。

各文字列 1 回ずつ計 86 回の刺激提示を 1 セッションとし、1 セッション終了すると、正答率が表示された。実験時間 30 分程度を目安に各文法について 3~4 セッションを実施した。同じ文法については、提示する 86 個の文字列はどのセッションでも同じであるが、提示順序はセッションごとにランダムに変化させた。以上を 1 フェーズとし、被験者全員が、30 分程の間隔を空けて 2 種類の文法を異なる遅延時間の下で学習した。

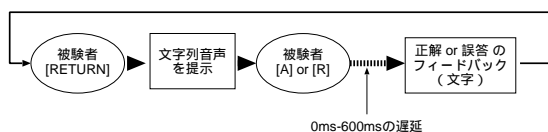


図 2: 測定の流れ

4.2 結果と考察

今回の実験では同一被験者に、2 種類の文法を学習させた。文法 1 と文法 2 は情報理論的観点では同程度の複雑性を持つが、刺激が音声として提示されていることから、音韻的には異なる (e.g. 一方のほう聞き取りやすい) 可能性もある。そこで、実験のデザイン上、データ数が最も多い、遅延無し条件における正答率の推移を調べたところ、文法 1 と文法 2 の正答率には明確な差がなかったため、本節では文法 1 と文法 2 のデータを合わせて議論する。

フィードバック遅延の大きさに関わらず、正答率の個人差が大きく、セッション毎の正答率の平均値を比較しても、有意な差を見出すことはできなかった。そこで、正答率の平均値を比較することに代えて、学習が成立した被験者の比率を比較ことを検討した。各セッション 86 文字列の半数が文法的であったから、完全な当て推量の場合、 n セッション分の正答数は、期待値

$$\mu_n = 86 \times n \times .5 = 43n,$$

標準偏差

$$\sigma_n = \sqrt{n \times 86 \times .5 \times .5} = \sqrt{21.5n},$$

の正規分布に近似できる。 n セッションで正答数が $43n$ を $2\sigma_n$ 以上上回ることを学習の成立とみなすことにすると、その臨界となる正答率 $P_c(n)$ は、

$$P_c(n) = \frac{\mu_n + 2\sigma_n}{86 \times n}$$

となる。表 1 に学習が成立した被験者数を示す。遅延無し条件では 30%以上の被験者が学習成立の基準に達しているのに対し、300ms 以上の遅延がある条件では、600ms 条件の 1 人を除いて、学習成立の基準に達していない。したがって、300ms 以上のフィードバック遅延は学習を阻害したと考えられる。150ms 条件では、第 1~3 セッションの集計で、8 人中 2 人が学習成立の基準に達している。150ms 条件の被験者数は少ないので、現時点では遅延が学習を阻害しているか否かは判断できない。

今回の実験は、遅延有り条件の被験者数が十分でない等の不十分さはあるが、300ms 程度のフィードバック遅延が与える心理的影響を、潜在学習のパフォーマンスを指標にすることで検出可能であることが示唆された。しかし、改善すべき問題がいくつか存在することが明らかになった。

最も重要な問題は個人差の扱いである。実験結果が示している通り、学習のパフォーマンスにはかなり大きな個人差がある。今回は、学習成立の基準に達した被験者数を条件間で比較することにより、フィードバック遅延が学習を阻害することを見出すことができたが、被験者の特性の分布が揃っていることを前提としている。特に被験者数が少ない場合、条件間で分布が異なりやすくなるので、条件間で被験者特性の分布が揃っていることを確認する必要がある。

このためには次のような手続きが考えられる。まず、被験者全員が文法 1 を遅延無し条件で学習し、パフォーマン

表 1: 学習が成立した被験者数。上段は第 1~3 セッションの集計、下段は第 1~4 セッションの集計。分子は学習が成立した被験者数、分母は全被験者数。

$$P_c(3) = .562$$

遅延量	0ms	150ms	300ms	450ms	600ms
第 1 フェーズ	5/15	1/4	0/4	0/4	1/4
第 2 フェーズ	5/16	1/4	0/4	0/4	0/4

$$P_c(4) = .554$$

遅延量	0ms	150ms	300ms	450ms	600ms
第 1 フェーズ	5/15	1/4	0/4	0/4	1/4
第 2 フェーズ	6/9	0/3	0/2	0/3	0/1

スの分布を得る。次に、得られた分布を基に、分布が揃うように被験者を条件数で等分する。この際、極端に逸脱した特性を持つ被験者や等分するのに都合の悪い被験者は排除する。以上の手続きによりスクリーニングされた被験者を用いて、文法 2 の学習を行ない。文法 2 の学習のパフォーマンスを条件間で比較する。この方法により、個人差の扱いに関する問題はある程度解決できると考えられる。

5. 議論

本稿では、映像・音声の伝送遅延が与える心理的影響の評価法として、人工文法学習の潜在学習のパフォーマンスを指標とする方法を提案し、予備実験の結果から、300ms 程度の遅延の影響が検出可能であることを示した。

この結果は、伝送遅延が与える心理的影響の先行研究の結果と比較することにより、その意味が明確になる。栗田ら [2] は TV 電話、および音声電話による会話において、伝送遅延の影響を主観評価により検討した。その結果、自由会話課題における伝送遅延の検知限は、TV 電話において 375ms、音声電話において 432ms であった (各々、往復の遅延量)。したがって、本研究の結果は、自由会話において検知すらされない微小な伝送遅延でも、認知的に影響を与えている可能性を示している。

提案した方法は既存の方法とは異なる特徴を持っている。まず、あらためて言うまでもないことであるが、既存の手法の主流である主観評価と異なり、被験者の言語報告や評定に頼らない客観的手法であることである。より重要な特徴は、伝送遅延がある種の学習あるいは記憶を阻害することが明確に示されることである。すなわち、伝送遅延が与える心理的影響は、情動への影響だけでなく、学習や記憶という認知レベルへの影響もあることが示される。これは、『遅延が妨害になる』、『相手の話がわかりにくい(ような気がする)』、『違和感がある』という類の言語的な評価では示すことができない。

主観評価以外の方法としては、被験者にコミュニケーション課題を課して、そのパフォーマンスを指標に伝送遅延の心理的影響を測定する方法が考えられる。例えば、ITU-T 勧告 P.920[1] では、映像・音声の伝送品質をコミュニケーション課題を通して評価するための指針が示されている。その中では、1~10 の数をペアで交互にできるだけ速く数え上げるという知覚・運動的な課題から、被験者が各々記憶した物語の相違点を探すという高次認知的な課題まで、測定用のコミュニケーション課題が挙げられている。このような方法は、被験者の意識に上らない、あるいは言語化できないような心理的影響を客観的に評価できるという利点がある。しかし、ここで挙げられているような課題では、伝送遅延が認知プロセスのどの部分に影響を与えるかが不明確である。また、挙げられている課題を使用して、映像・音声の伝送遅延の影響をパフォーマンスを指標に測定できるかも定かではない。

一方、提案した方法により測定された心理的影響が、自然なコミュニケーションの状況で伝送遅延により受ける心理的影響の一部であるか、という問題は検討を要する。例えば、ITU-T 勧告 P.920 では、評価に使用する課題は、実生活における映像・音声コミュニケーションに類似したものであることを推奨している。その意味では、提案した手法は、少なくとも表面的には実生活におけるコミュニケーションには類似していない。

今後、検出力の向上と測定の効率化を図るとともに、社会的相互作用を指標にする [3, 5] 等のより自然なコミュニケーションに近い課題による評価、より高次な認知への影響の測定等、順次検討していきたい。

参考文献

- [1] ITU-T, "Interactive test methods for audiovisual communications", ITU-T Recommendation P.920, 1996.
- [2] 栗田 孝昭, 井合 知, 北脇 信彦, "オーディオビジュアル通信における伝搬遅延の影響", 電子情報通信学会論文誌, vol.J76-B-I, no.4, pp.331-339, 1993.
- [3] 望月 要・河野 友香・山岸 直基・大西 仁, "言語条件づけにおける強化遅延の効果", 日本行動分析学会第 17 回年次大会発表論文集, pp.84-85, 1999.
- [4] 村越 真・松井孝雄, "潜在学習", 認知科学, vol.2 no.3, 12-23, 1995.
- [5] 大西 仁・望月 要・河野 友香・山岸 直基, "言語条件づけを利用した通信遅延の心理的効果測定の試み", 電子情報通信学会総合大会, B-11-2, 1999.
- [6] Reber, A.S., Transfer of syntactic structure in synthetic language, *Journal of Experimental Psychology*, 81, 115-119, 1969.
- [7] Reeves, B. and Nass, C., *The media equation*, Cambridge, CA: Cambridge University Press, 1996.