

言語の文化実験



山内 肇 (RIKEN/UC Berkeley)

はじめに

1990年代半ばに出版された“The Major Transitions in Evolution”の中で、著者J.M.SmithとE.Szathmaryは言語の出現を「生物史上の最後の進化的跳躍」として挙げている。しかし、言語の進化・起源は何時の時代においても高い注目を集めつつ、遅々として進まなかった。これには様々な理由が考えられるが、その一つがその進化過程の複雑さにある。言語は目や肺といった単なる物理的器官やホルモン調節機能のような生理的機能の進化とは異なり、脳の認知能力という側面や、コミュニケーションツールといった文化的な側面など多面的な性質を有している。

1990年代以降、コンピュータの発達によりこれらの性質を考慮したシミュレーション研究が可能となり、言語進化研究が一気に盛んとなったが、2000年代初期のまでは、その多くがヒトの脳自体の進化を支配的な説明原理とみなしてきた。つまりヒトの言語がここまで複雑かつユニークなのはそれを受容するヒトの脳の特異性に由来するということである。従って、言語進化の鍵は脳の進化過程にあると考えられた。

言語の文化進化

しかし、仮に脳の進化がヒトの言語の説明原理を司るとすれば、言語がもたらす機能的適応¹というものを考えざるを得なくなる。一方でN.Chomskyのように古くから言語の機能的側面に疑いを持つ者は多く、しばしば論争の種となってきた。このような状況を打破するように、近年では言語自体をヒトの脳に寄生するある種の有機体と見なした上で、言語進化を言語自体の脳へのミーム的な適応、すなわち文化進化の結果であるとの見なし方が主流となってきた。

言語は幼児期に大人から言語入力を受けることによって学習に世代から世代へ言語獲得を通じて受け継がれる際、その学習には強いボトルネック効果²が働くことが言語獲得研究からわかっている。

Kirby[1]はこれが言語自体に対する強い淘汰圧となり、異なった状態を持った言語知識間に一種の生存競争をもたらすと考え「連鎖学習モデル(Iterated Learning Model)」を提唱した。ILMでは文化進化を生物学的な進化をとみなさない³一子相伝型⁴の時間変化プロセスとしてモデル化し(Fig. 1)、ヒトが持つ学習バイアスが言語への淘汰圧となり結果的に言語の構造

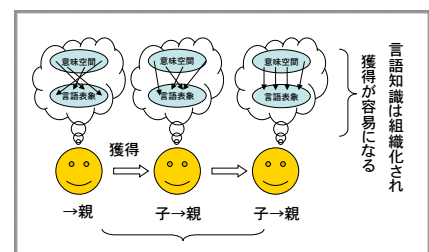


Fig. 1 言語文化進化モデル

被験者 $t+1$ は同じ系列の直前の被験者 t からイベント(意味)と t がそれに対して付与した文字列を入力として受け、それを元に学習を行う。学習者 $t+1$ は、実験の設定によりランダムに選ばれた[イベント:文字列]のペアについて観察することは出来ない(全ペアの約半分)。一方、生成に関しては観察したことのないイベントに対しても文字列を付与しなければならず、ここに観察によって得られた被験者の仮説が利用されることとなる。この連鎖を繰り返すことにより、言語の構造化の文化進化が確認された。

化を引き起こすことを確認した。本研究はこの言語の文化進化モデルを基に、よりヒトの言語使用状況に即した実験を行おうというものである。

Mechanical Turk

ここ数年、アメリカを中心として急速に注目を集めている実験プラットフォームがある。本を始めとするインターネット通販で有名な米Amazon社が運営するMechanical Turk(AMT)と呼ばれるクラウドサービスである。「機械じかけのトルコ人」とはなんと奇妙な名前だが、これはハンガリー人の発明家ケンペレンが作った「トルコ人」と呼ばれるチェス・オートマタに由来する。要は画像・音楽のタグ付、簡単な翻訳、市場調査、データ解析などコンピュータが苦手とする作業を、ネットを介してその向こうにいる実際の人間にやってもらおうという趣旨である。AMTはそのための労働市場や支払スキームを提供する、ある種のクラウド型サービスと考えられる。アメリカの大学を中心として、このサービスを利用して、社会・心理実験を行なおうという動きが急速に盛んになっている。これには以下にあげる利点の存在が大きい。

1.登録者数

AMTには5年ほど前の段階ですでに10万人を超えるワーカが存在しており、特定の実験を考えるにあたって潜在的な被験者の数は無限と考えて良い。また、参加者の地理的・社会的な分布も多様であり、学部生を被験者とするケースが多い対面実験よりもより質のよいサンプルが得られる可能性が高い。

2.謝金

Amazon社を通しての決済システムのため、振込等の複雑な手続きが不要であり、支払いに関して個人情報の授受は一切無い。また、謝金額は概して低く、現在約1時間の実験で\$0.30-\$1.40と、一般的な実験と比べると一桁以上の違いがある。

3.個人情報の秘匿性

各ワーカ（リクエストも含め）には固有のID番号が割当てられており、実験の参加にはこの番号を利用することになる。リクエストにとってIDと実際の被験者の名前、住所等の個人情報を結びつける仕組みは存在しない(無論、そのような質問を行えばその限りではない)。従って、実験の参加から謝金の支払いに至るまで、個人情報に接することがない。その一方で、リクエストとワーカは必要に応じてAmazonのサービスを介し、メール等を取り取りすることができる。倫理委員会が懸念する個人情報の取り扱いを考えると非常に好ましい形態といえる。

もちろん、遠隔かつ匿名の実験ということから、様々な懸念を持つかもしれない。しかし、2011年現在、実験への集中度やその結果等について幾つかの報告が行われているが、対面実験とほぼ差がないことがわかっている[2]。また、参加者数が多くなることから、一般の心理実験では困難な統計的な処理も可能になっている。

AMTを介したタスクは(これをHuman Intelligence Tasks, HITsと呼ぶ)、基本的には同サービスページ内で行われることを前提としているが、Amazon社が提供するAPIを用いることによって外部サイトに誘導することも可能であり、本実験はこの形態を利用している。本実験においては、被験者の集中度や重複参加等を避けるために実験中には様々なパラメータを収集し、実験中のWebブラウザ操作による干渉等を避けるためにJavaScriptやPHPなどの動的な制御を利用した。

実験に際しては理化学研究所の倫理審査委員会の手承を得ることが必要であり、これには相当度の時間を費やした。新しい形態での実験のため、審査委員会が幾つかの点について判断に苦慮したためである。このようなノウハウについてはいずれ機会があればお話ししたい。

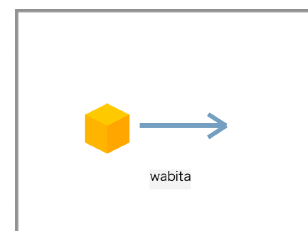


Fig. 2 実験画面の例

被験者は入力として一定数のイベント(色・形・動きを伴った映像)と、前世代の被験者が付与した文字列を一定数観察し、その関連性を学習する。その後、イベントのみを観察し、自ら文字列を付与する(矢印は動きを表しているが実際のイベントでは図形そのものが移動する)。

実験

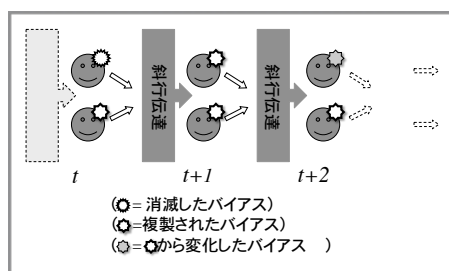


Fig. 3 斜行伝達による文化進化

本実験では、一子相伝型の拡散連鎖ではなく、各世代において各個体の結果がプールされ、その中から次世代の入力がランダムに選ばれる。その結果、各世代内に存在するバイアス(同義語など)の継承が不規則となる。

Kirby et al.[3]はコンピュータを利用した連鎖学習モデルの結果を確かめるために、拡散連鎖モデル(diffusion chain framework)と呼ばれる心理実験フレームワークを用いて言語の文化進化を実験を行った。具体的には被験者は色・形の異なる図形とその動き(ここではこれを「イベント」と呼ぶ)を観察し、同時にそれに名付けられた文字列を記憶する⁵。これを

「学習期」と呼ぶ。学習期の終了後、被験者はイベントのみを観察し、自身で記憶を元に文字列を生成する。これが生成期である。

イベントは3種の色、3種の形、そして3種の動きの組み合わせにより全部で27種類存在する。被験者は学習期にはその内の半分、14種類分のイベントと文字列のペアしか観察することができない一方で、生成期には27種類全部の文字列を生成することが求められる。つまり、被験者は学習期の観察から得られた経験により、ある種の「仮説」を形成しそれを基に未知のイベントについて文字列を生成することが必要になるわけである。

生成期に得られた27のイベントと文字列のペアは、次の被験者の学習期の入力となる(そのうち13ペアは表示されない)。このように第一被験者の実験結果を第二被験者の学習用の入力に用い、さらに第二被験者の実験結果を第三被験者の入力にする、といった個別実験間の関係を操作することで、被験者を仮想的一つの時系列上の連鎖とみなし、イベントと文字列のペア、すなわち言語知識がどのように時系列的に変化するかを見るわけである。

第一被験者への入力には14種類のイベントについて全てランダムな文字列が付与されているが、Kirbyらは学習が連鎖していくことによって、文字列が徐々に構造化していくことを確認している。これは言語文化進化のシミュレーションモデルである連鎖学習モデルの結果とほぼ等しい。しかし、一方で同実験は初期の段階で多くの同音異義語の発生に悩まされており、Kirbyらはある世代

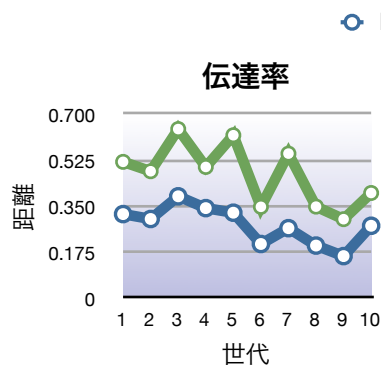


Fig. 4 世代間伝達率

被験者が各イベントに付与した文字列を、その被験者の学習入力となった親世代の同じイベントに対する文字列と比較し、その距離を正規化圧縮距離(NCD)と正規化レーベンシュタイン距離(LEVENSHTTEIN)の2つで計測した。距離が大きいほど、同じイベントに対して、被験者が付与した文字列が異なり、正しく伝達できていないこととなる。

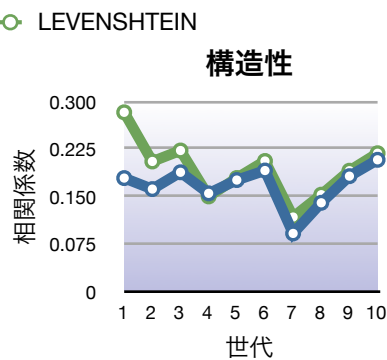


Fig. 5 構造化

本実験におけるイベントは全部で27種類あるが、これは、色、形、動きの3種の特徴領域についてそれぞれ3つの異なるタイプの組み合わせによって得られたものである(例えば色では黄色・灰色・青色など)。従って、各イベントはその近似度を距離として計測する(ハミング距離)ことが可能である。任意のイベントとそれに付与された文字列の距離の相関が高ければ、似たイベントは似た文字列を有していることとなり、言語の構造化を図る指標となる。

で発生した同音異義語を次世代の入力とする際に「間引き」し、ランダムな文字列で置き換えるという少々荒っぽい操作を行うことで、改善を試みている。

Kirbyらの実験では、連鎖学習モデルの成果に関する実証的検証という意味もあり、各被験者系列(論文内では4系列)が独立して10世代分のノードを形成している。つまり、始端となった各系列の第一被験者の入力以外は、全くインタラクションが存在しない、一子相伝型の垂直伝達(vertical transmission)モデル⁶と見ることができる。一方、我々の実験ではより実際の言語獲得プロセスを模すため、各世代における実験結果をランダムに交叉させ複数の被験者の実験結果を入力とする「斜行伝達(oblique transmission)」による実験をAMTを介して行った(Fig. 3)。なお、ここではKirbyらとほぼ同じ状況を模すために、各世代4被験者で10世代行った結果を提示している。

斜行伝達の場合、特定の被験者で発生したバイアス(同音異義語や固有の形態素など)が次世代に伝わらず消滅する可能性が高く、その分情報伝達における不確実性が高まる。その一方、このようなバイアスが適宜取り除かれることで、Kirbyらが人為的に行った「間引き」に似たフィルター現象が発生することが期待される。

実験の結果、Kirbyらのような同音異義語全てを間引いたような強力なフィルター効果は得られなかったが、各世代における同音異義語数や世代間の伝達率(Fig. 4)について、Kirbyらの間引き実験と同じようなプロファイルが確認されている。また、言語の構造的性(Fig. 5)²では、中間の世代までは構造的性が低下するものの、後半において直線的な改善が見られている。これらの結果は、斜行伝達によるバイアスの消滅や部分継承によってフィルター効果が発生していることを示唆している。

まとめ

上述しなかったが、AMTのもう一つの良い点は、実験実施について時間的・空間的制約から解放されることである。そのため、異なるパラメータ下での実験実施が非常に効率良く行え(理論的には24時間実験が可能)、本研究でも異なるパラメータ下においての実験を行い、約350名ほどのデータを収集している。これらの実験についても、今回の結果以外に興味深い結果がいくつか得られており、これらについて論文・学会等を通じて発表をしていきたい。

参考文献

[1]

Kirby, S. (2001) Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity, IEEE Transactions on Evolutionary Computation, 5(2):102-110,

[2]

Paolacci, G., Chandler, J. & Ipeirotis, P.G. (2010) Running experiments on Amazon Mechanical Turk. Judgment and Decision Making 5(5): 411-419.

[3]

Kirby, S., Cornish, H., and Smith, K. (2008) Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. PNAS, 105(31):10681--10686.

¹ 言語が持つコミュニケーションなどの機能に選択圧がかかり、それによって言語が進化したという考え方。一方、他の能力の適応の結果、副産物として言語も機能を持つようになったという外適応(exaptation)という考え方も存在する。

² ここでの「ボトルネック効果」とは、ある情報が伝達される際に入力となるサンプルの数が大きく減ることで、その入力を生み出した知識システムの全容が把握できない様なケースを指す。知識システムを母集団と入力サンプルを標本と考えるとわかりやすいかもしれない。少ないサンプルからシステム全体の性質・構造を想定する場合、そこにバイアスが発生する。たとえば、ある楽曲を完全に演奏するためには、誰かが演奏したその曲全部を聴くか、完全な譜面を手に入れることが必須である。しかし仮にそれらの一部しか入手できない場合に、入力に特定のリフやコード進行が含まれていれば、復元される曲は元の曲よりも強い規則性を持つ可能性が高い。言語の場合、ある特定の言語にしての構造全てを子供が入力として与えることは不可能なため、ここにボトルネック効果が発生すると考えられている。

³ 生物進化をモデル化したシミュレーションでは、個体は世代を経るに従ってその遺伝情報が交叉や突然変異を通して変化していくのが一般的である。自然選択はこの変化に伴う個体の複製率の違いに起因する現象である。一方、ILMの文化進化モデルでは交叉や突然変異は含まれず、どの世代においても遺伝的情報は一定である。もしこのような状況において、学習を通じて伝達される知識システムが変化し、それがシステム自体に何らかの構造性を生み出すのであれば、どこかにその構造性を引き起こす「選択メカニズム」が潜んでいると考えられるだろう。そのシステムを受容するホストである個体の遺伝情報が一定である以上、この選択メカニズムはシステムの伝達そのものにあると考えられる。これが文化進化の基本的な考え方である。上記のボトルネック効果と合わせて考えると、ボトルネック効果がこの選択メカニズムに深く関与していることが予想される。

⁴ 一人の「親」が一人の「子」に対してのみ入力を与える場合をここでは「一子相伝型」と呼ぶことにする。

⁵ 初期入力として27個のランダムな文字列を生成した。具体的には「ekure」「wabita」「kea」などである。なお、これらの文字列は偶然を除き実際の言語との間に対応関係はない。我々の実験ではどの言語を母語とするものでも基本的に発音できるよう、母音(Vowel)と子音(Consonant)の組み合わせの中からCCCVのような子音の連鎖は除外し、VないしはCVを基本構成単位とした。最短の文字列はVCV・CVVであり、最長はCVCVCVである。第一世代以降は、被験者自身が自由に文字列を生成することになるので、このような制約は課していない。

⁶ 親から子へといった世代間での情報伝達を一般的に「垂直伝達」と呼ばれ、一方の同世代内での情報伝達は「水平伝達」と呼ばれる。一般的に「垂直伝達」という言葉はその伝達形態が一子相伝型であるか、多対多であるかを区別しないが、前者と後者を区別する場合は一子相伝型を「垂直伝達」と呼び、後者を特別に「斜行伝達」と呼ぶ。

7 本実験では任意の2つの文字列の近似度を計測するために2つの指標を用いた。一つはコンピュータのスペルチェックなどに広く利用されている「レーベンシュタイン距離」と呼ばれる指標であり、「編集距離」としても知られている。これは、1つ目の文字列を2つ目の文字列に変形するのに要する文字の挿入・削除・置換の回数を距離とする。これを2つの文字列のどちらか長い方で割ったものが正規化編集距離である。例えばMarkとCarならばMをCに置き換え、kを削除すればよいので距離は2(正規化した場合は0.5)である。

もう一方の正規化圧縮距離(Normalized Compression Distance)は、ここ数年で急激に普及した近似度計測方法で、情報理論における複雑性概念で有名なコルモゴロフ複雑性をベースにした手法である。これはbzip2など、コンピュータで日々使用されている可逆圧縮・解凍アルゴリズムを用いて、2つの文字列を別々に圧縮した場合のファイルサイズの合計と、2つを1つの文字列(abcdeadeabc)として圧縮した場合のファイルサイズの差を距離とするものである(正規化には大きい方のファイルサイズで差を割ることになる)。例えば「abcde」と「adeabc」の2つの文字列は互いに部分列「abc」と「de」を共有しているので、それぞれ独立して圧縮する場合と、あわせて圧縮した場合での差が少ない。

NCDの利点は、文字列だけではなくコンピュータ上で扱える情報ならば、画像や音声などその状態を問わないことと、編集距離では難しい、直感的な近似の度合いを比較的スマートに考慮に入れることができる点である(編集距離の場合、abcdeからadeabcへは5回の編集が必要となり、距離としては最大になってしまう)。

異なるイベント間の距離はハミング距離と呼ばれる距離尺度を利用している。例えば色・形は同じだが動きのみが違う2つのイベント間の距離は1であるのに対し、色のみが同じで形と動きが違う2つのイベント間の距離は2となる。本実験では3種の特徴領域が存在するので最大距離は3となる。従って正規化の場合は得られた距離を3で割ることになる)。