

文法形成の計算論的研究： 非論理的推論による発話の不一致と多義発話の発生

金野 武司, 橋本 敬

北陸先端科学技術大学院大学 知識科学研究科

key words : 合成的言語, 分析的プロセス, 計算機実験

はじめに

ヒトの言語が持っている文法的特徴の1つに合成性がある。合成性を持った言語とは、単語の意味の組み合わせによって文の意味が決まるような形式を持つ言語のことである。このような合成的な言語が文化進化の過程で形成される仕組みを計算モデルを用いて検討した研究がある (Kirby, 2002)。この計算モデルでは、状況に対する発話として最初に文を形成し、文どうしの類似性を使って単語を切り出していくプロセスによって合成的な言語が形成される¹。この合成的な言語の形成によって、計算モデルにおいて言語を学習する個体 (エージェント) は、複数の状況に対する発話を少ない規則から生成できるようになる。初期言語が形成された頃の人類にとっては、少ない規則で多くの状況を表現できる仕組みを持つことは、記憶容量への負担を軽減できるという意味で重要だったのではないかと考えられる。

ところが、この計算モデルは合成度を高めることができる反面、状況に対する発話が他者と一致しない言語 (発話の不一致) や、異なる状況に対して同じ発話 (多義発話) を生成する言語が形成されることがある。これらは、状況に対する他者の認識をその発話から推論するという、言語が持つ基本的な機能を損なう可能性があることを意味している。本論ではこの問題が、単語を切り出す過程で発生する多義単語を抑制する仕組みによって回避できることを報告する。この取り組みを通じて、本研究はヒトの言語学習に関わる認知能力的な仕組みを考察するための知見を得ることを目的とする。

計算モデル

エージェントどうしの相互作用には繰り返し学習モデルが用いられる (Kirby, 2002)。このモデルでは、親と子が同じ状況を見ながら親が子に向かって一方的に発話する。子はある一定回数状況に対する発話を聞くと、今度は親になりまた別の子に向かって発話する²。このモデルには100個の状況が用意され、エージェントはその状況に対して発話 (アルファベットの文字列) を割り当てる。エージェントは状況と発話の対応関係をルール形式で保持する (これを文ルールと呼ぶ)。エージェントはこの文ルールどうしの類似性を使って発話の区切りを見つけ、単語を切り出す (chunk 操作)。そして、切り出された単語ルールを他の文ルールに拡大適用することで文ルールを一般化していく (replace 操作)。金野・橋本 (2009) は、この replace 操作で非論理的な推論が行なわれることによって、状況に対する発話が親子間で一致しない言語が形成されることを見出している。

実験とその結果

全ての状況に対する発話を文ルールとして形成するなら、エージェントは100個のルールを持つことになる。ところが、chunk と replace の操作が実施されると、十数個程度のルールによって100個の状況全てを表現できる言語 (表現度 100%の言語) が形成されるようになる。表1は、親から子への言語伝達1万世代を100回実施したときに得られるルール数と表現度、そして多義発話比率の平均値を示したものである (親から聞く発話の回数

¹このような過程は分析的プロセスと呼ばれ、逆に単語から文を構成していく過程は合成的プロセスと呼ばれる (Smith, 2006)。言語の形成過程には両者のプロセスが混在していると考えられるが、本論では分析的プロセスに着目して議論を進める。

²計算モデルの具体的な構成は橋本・中塚 (2007) に基づく。

表 1: ルール数, 表現度, 多義発話比率の平均値

	ルール数	表現度	多義発話
何もしない場合	16.5	99.8 %	42.2 %
多義単語の抑止	15.5	99.9 %	0.1 %

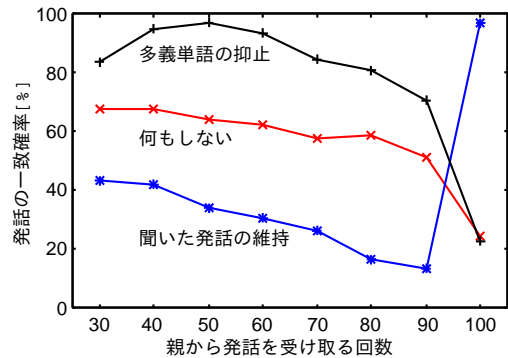


図 1: 状況に対する発話の親子間での一致確率

は50回に設定)。また表1の下段には、chunk 操作において生成される多義単語を抑止した場合の結果を示した。この結果から、多義単語の抑止によって多義発話を生成する言語の形成が抑制できることが分かる。

続く図1は、親から聞く発話の回数を変化させたときに、形成される言語において親子間での発話が一致する確率を示したものである。ここには、先ほどの多義単語を抑止した場合に加えて、親から聞いた発話を維持する仕組みをエージェントに導入した場合の結果を示した。この結果から、状況に対する発話の一致確率は多義単語の発生を抑止すれば上がるが、聞いた発話を維持した場合、もしくは全ての状況に対する発話を聞いた場合には下がることが分かる³。

議論

非論理的な推論によって発話可能な状況を増やしつつ合成的な言語を形成する方法は、そのままでは多義発話や状況に対する発話の不一致を発生させる。しかしこれは、多義単語の抑止によって大きく回避することが可能である。また聞いた発話を維持せずに、なおかつ他者から聞く発話の数が少ない方が、同じ状況には同じ発話を用いる言語を形成できる確率が高くなる。

翻ってヒトの言語は、本質的に多義的であり、状況に対する発話も変化し続けている。しかし、ヒトはこういった言語を生み出す仕組みを持っていても、他者の認識を推論する道具としてその言語を用いることができる。本論の結果は、ヒトの言語がこういった特徴を持つようになる仕組みを考察するための足掛かりにできるのではないかと考えられる。

文献

- 橋本敬・中塚雅也 (2007) “文化化の構成的モデル化—進化言語学からの考察—”, 日本認知言語学会論文集, vol.7, 33–43.
- Kirby, S. (2002) “Learning, bottlenecks and the evolution of recursive syntax,” in E.J. Briscoe ed. *Linguistic evolution through language acquisition*: Cambridge University Press, 173–203.
- 金野武司・橋本敬 (2009) “文法形成における非論理的推論と多義的発話の役割”, 情報処理学会研究報告 MPS-73, vol.2009, no.19, 101–104.
- Smith, A. D. M. (2006) “Semantic reconstructibility and the complexification of language,” in *Proceedings of the 6th International Conference on the Evolution of Language*, 307–314.

³100個の状況に対する発話を聞いてそれを維持すれば一致確率は100%に達するが、そのときにはルール数が100個以下にならない。