言語のダイナミクスへの構成論モデル

北陸先端科学技術大学院大学・知識科学研究科
橋本 敬

hash@jaist.ac.jp

Constructive Modeling toward the Dynamics of Language

Japan Advanced Institute of Science and Technology
School of Knowledge Science
Takashi Hashimoto

Abstract

Significance and implication in the constructive study of dynamics of language are discussed. The constructive studies admit three important viewpoints in study of language. These are dynamics, emergence and subjectivity. However, showing global dynamics in language structure based on internal dynamics of subjects has not been succeeded. An attempt to synthesize language dynamics with the developing word-web model is introduced.

Keywords: Dynamics of Language, Developing Word-web Model

1 構成的言語研究に見る言語観

言語を研究するにあたり重要なことは、「言語の本質」がどこにあるのかということを常に念頭におき、それについての考え方をつきつめるということである[1]。以下ではまず、進化言語学、特に、 その構成論的な研究における言語の本質観はどのようなものなのかということを見る。

1.1 言語のダイナミクスと創発

人類学、歴史言語学、構成論的モデル研究といった幅広い領域を含みながら立上りつつある進化言語学では、時間を止めた言語の構造自体ではなく、言語が変化する点に着目している。その点で、進化言語学においては、多くの場合、言語は本質的にダイナミックなものであると考えられていると言えよう。

では、言語のダイナミクスとして想定されているものはどんなものだろうか。よく研究対象とされるものとしては、言語進化や言語発達がある。

例えば、言語の起源と進化の研究においては、比較的長期的な言語構造の変化が着目点となる[2]。 言語の起源とは、言語のないところからあるところへという変化がいかにして起きたのか、あるいは、いかにして可能なのかという問題であり、言語進化とは、一旦成立した言語の文法や意味の構造的な変化がいかなるものかという問題となる。このような、具体的な証拠としては残されにくく、観察や実験を行うことが困難な問題については人工生命研究で採られるような構成論的なモデル化を行う研究手法の特色が生かされてる。すなわち、コンピュータの中に作られたモデルは一つの実験室としてみなされ (Casti, 1999)、そこでいろいろな言語現象の創発を構成することを試み、 その情報構造を明らかにすることを目指す (Steels, 1997)。

また、 意味や文法などの社会的に共有される言語的

[1] 時枝 (1941) は、言語本質観と実際的な言語研究の関係について、「言語研究の使命は、個々の言語的事実を法則的に整理し、組織することにあるというよりも、先ず対象としての言語の輪郭を明らかにする所になければならない」と主張している。

[2] 言語獲得は言語進化に比べると短期的な時間スケールのものである。そこでは、 幼児が既存の言語 (第二言語習得においてはある言語の話者が別の言語) の構造をいかに学習するか、 ということが論点となる。

ルールのようなものはアプリオリに存在するのではなく、言語使用者の個別の活動から中央のコントロールといった特権的な外部構造なしに創発すると考えられている。その点でも、言語のダイナミクスの研究は、人工生命研究における創発の考え方との親和性が強い。

言語の本質を以上のように捉えること、すなわち、それ自身のダイナミクスを持ち、創発的なものであるという言語観は、人工生命の分野における様々な言語研究、例えば、 単語の指し示すもの (Steels, 1996) やその多様性 (Arita and Koyama, 1998)、 あるいはネットワークで共有される文法 (Hashimoto and Ikegami, 1996) などの創発や進化についての研究に特徴的に見てとられるであろう。

1.2 構成論的モデルにおける「主体」の地位

構成論的言語研究に見られる言語観をさらに分析すると、もう一つ重要な論点が見えて来る。それは、言語には主体的活動が不可欠だと考えているという点である。記述的に言語を扱おうとするならば、「言語的ルールのようなもの」を、あたかも言語使用者としての主体を離れて実体化しているかのようにみなしてしまう。すなわち、主体が対象から消し去られてしまうのである。これに対して構成論的言語研究においては、言語を話し、聴き、理解する主体の活動を基礎におき、「言語的ルールのようなもの」はその主体の活動の結果として見い出されるのである。

そこでは主体とは、単に自律的に振る舞うのではなく、系全体の性質を主体内になんらかの形で埋め込み、それに基づいた行動を行う存在である (金子邦彦・池上高志, 1998)。このことは、創発システムである言語においては、主体が社会的に共有されうる「言語的ルール」をつくり出すとともに、それによって制約を受ける存在であることを意味する。この制約によって、主体はあらゆることが言語的に表出可能なのではなく制限を受ける。そして、また同時に、 この制約によって言語がコミュニケーションの媒体足り得るのである。

経済や社会システム、言語の研究において、できるだけ取り払われて来た主体性を系のなかに埋め込み、数理的な研究の対象に含むことができるというのが、構成論的研究手法の最大の利点だと考える。

2 言語を構成するとはいかなることか

前節において、主体とは、言語における (社会的) 制約をつくり出し、かつ、それにさらされる存在であると述べた。成立した言語構造が単に主体を制約するだけならば、言語構造は静的なものになってしまうが、実際には言語は常に変化しつづけている。また、その制約としての言語構造があたかも客観的に存在するかのように扱うならば、主体は言語的ルールをただ追認する存在になってしまうが、主体は、言語を話すとき文法や語用のルールに常に従っているばかりではなく、時にはそのルールを破ることも可能な存在である。そういう自由な言語活動は、結果として言語構造の変化に現れる場合もある。

例えば、創造的なメタファーを考えてみよう。一般によく使われてるメタファーは既に言語体系の一部と化しているが、われわれは新たなメタフォリックな表現をつくり出すことができる。そういった新しいメタファーは、リジッドな言語のルールがあるという立場から見るならば「間違った文」であり、ルールを破っていることになる。しかし、われわれはそのような表現を読んだり聴いたりしたとき、ある程度自分なりに理解することができる。この時なにが起きていると理解すべきか。

言語に主体が不可欠なものとして捉える立場に立ってみると、そのような新しいメタファーの理解において、主体は意味を創造していると見ることができる。ここではわかりやすい例として創造的メタファーを取り上げたが、言語を話す、理解するということは、多かれ少なかれこのような主体による意味創造が関わって来る。すなわち、ごくふつうの叙述文においてさえ、それは客観的現象の表明ではなく、主体がそれをいかに理解したかを表し、その文を了解する主体は、そこで叙述される現象について、場面、文脈、経験、言語体系に制約されながら概念を形成しており、一つの文が誰にとっても常に同じ意味を持つことはない。

著者は、 こうした、 言語を用いる主体が意味を創造する動的な過程として言語を捉え、主体の内的なダイナミクス自体を意味作用とする立場を、(橋本敬, 1999; Hashimoto, 1998) において「動的言語観」として表明した。

主体の概念作用を基礎として言語を見るという考え

方は、時枝の言語過程説 (時枝誠記, 1941) の主張に通じる。 まさに時枝はこの説を「言語の本質を心的過程と見る言語本質観の理論的構成」として位置付けている。また、バフチンは，個人主義的主観論の見方においては，言語とは「活動そのものであり，個人の発話行為によって実現される絶え間ない創造の過程」として理解されるべきものであるとしている (Vološinov, 1986)。

前節で述べたように、構成論的手法の利点は、主体を対象系に作り込むことにある。よって、われわれは、時枝の言語過程説やバフチンの個人主義的主観論で考察されたことを、広げ深めるための概念装置を手にしていると言えるのである。われわれは言語を構成論的に研究するにあたり、主体の意味創造活動を基礎にして、そこから現れる言語構造や、言語の絶えざるダイナミクスを構成し理解することを目指すべきである。

人工生命の言語研究は、あるシンボルとある対象の結び付きが集団内で共有されることや、その結び付きが多様なシンボルと対象の対において実現されたことをもって言語が創発したとみなすことを越えて行くべきである。また、言語によるコミュニケーションは、自分の言いたいことを相手の中にコピーするというものではない。むしろ、そのようなシンボルの結び付きを解き放ち、新たな対象をつくり出すという主体的な活動を基礎とし、互いに影響しあい内部構造が発展する動的なプロセスであるコミュニケーションの結果として、言語構造自体が変化しつづけるような系を構成し、そのダイナミクスを明らかにすることを目指したい。

3　語間関係網のダイナミクス

本節では、上で表明した構成論的言語研究への試みとして、単語間の関係の網目が会話とともに発展するというモデル (Developing Word-web model) を用いたカテゴリー形成についての研究を簡単に紹介する[3]。

3.1　主体の内部構造のモデル化

本モデルでは、主体の内部構造を単語間関係で表現する。単語間の関係は、それが指示あるいは参照する対象の関係を写し取ったようなものではなく、会話の中で各単語がどのような使われ方をしているかによって発展し形成されてくるという使用法ベースの考え (Hashimoto, 1997) を採っている。

主体が言語を使用するときの意味創造活動は、単語を語間関係網の中に位置付ける過程としてモデル化される。今の場合、単語を関係網の中に位置付けるとは、ある文を受け取ったとき、そこで使われている単語と他の全ての語との関係を計算し更新するということである。よって、主体は会話を通じてその内部構造を動的に変化させて行く。 これを模式的に図 1 に示した。

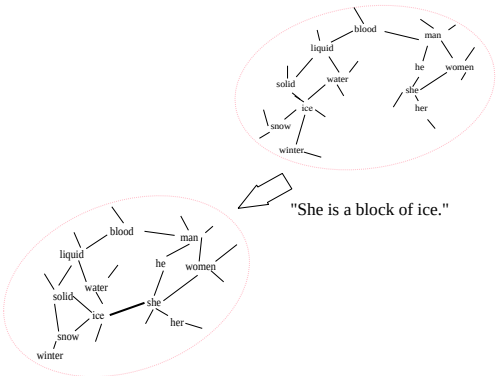

図 1: 単語間関係網が一つの文により変化することの模式図。 上の状態 ('she' と'ice' の間にはつながりがない) にあるときに、“She is a block of ice.” という文を受理し、下の状態 ('she' と' ice' の間につながりができ、 それにともなって全体の形も変わる) へと変化する。

語間関係を更新するアルゴリズムは、 Karov and Edelman (1998) によるコーパスの文中の単語関係を計算する手法に二点の変更を加えたものである。一つは、会話とともに動的に計算を行うようにするという変更である。これは、語間関係の最終的な構造よりも、会話により内部構造が変化し、カテゴリーが形成されるダイナミクスに興味の中心があるからである。 もう一方は、 文の上位階層として「テキスト」というものを考え、テキスト中の単語の出現頻度の相関も考慮して単語間の関係を決めるという点である。今の場合テキストとは、会話の中で発話し受理される一連の文のことを指す。

テキスト t 中の n 番目の文を受理したときの二つの

[3]詳しい内容については (Hashimoto, 1998; 橋本敬, 1999) を参照してほしい

単語 w_i と w_j の間の関係 $R(w_i, w_j)$ は、語間類似度と語間相関を次のように線形結合したものとして定義する。

$$R_{t,n}(w_i, w_j) = \alpha^w(\text{語間類似度}) + (1-\alpha^w)(\text{語間相関})$$

α^w は結合パラメータである。語間類似度の項は、一つの文中での単語の現れ方を考慮するものであり、文中で同時に使われた単語は関係を強めるようになる。第二項は、テキスト中での単語の出現パターンの類似性を考慮するものであり、ある二つの単語の出現パターンが同じであれば正に、異なっていれば負になる。これはテキスト中の単語の出現確率の相関を計算することによって実現している。 これらの値は、 直前の文の時点での値 ($R_{t,n-1}$、$n=1$ の場合は直前のテキスト $t-1$ の最後の文の状態) から計算されるものである。よって、文を連続的に発話・受理していくと、このアルゴリズムを繰り返し適用することになるので、一つの文中ではなく異なる文中で使われる単語間の関係も変化していく。

3.2 会話のモデル化

上で定義した語間関係を内部構造に持つ主体の間の「会話」を、 単純に「文のやりとり」としてモデル化する。主体の集団を考え、そこから選ばれた二つの主体に対しあるトピックが提示される。 会話は、一方の主体 (発話者) がそのトピックについての文を発話することから始まる。 次にもう一方 (聴取者) が返答をするが、その後の会話の行方は提示されたトピックについてのものに制限されず、直前の文に含まれる単語のうちの一つを用いた文を発話する。また、主体が生成した文は発話の前に変更される可能性があり、その確率を creativity rate(c) とする。これにより新しい文が生成され、このとき新しい単語を作る場合もある。

会話の手順を次にまとめる。

1. 発話者がトピックについての文を生成する
2. 文をある確率に従って変更し、聴取者に対して発話する
3. 発話した文により、発話者の語間関係を更新する
4. 聴取者はその文を受理するかどうかを決める。その基準は、今の場合、文中に知らない単語が二つ以上あれば受け取らない、 とする[4]。文が受理されなかったら、発話者は別の文を作る (1. へ)
5. 文を受理した場合、聴取者の語間関係を受け入れた文により更新する
6. 発話者と聴取者の役割を入れ換え、新たな発話者は返答する文を生成する (2. へ)

このプロセスを、受理された文がある一定量に達するか、受け入れられなかった文がある数を越えるまで続ける。一つの会話を終えると、別の主体の組を集団中から選び、別のトピックを提示して会話を行う。

語間関係は行列で表現される。 単語 w が含まれた文を受け取ったとき、あるいは、発話したとき、既知の全ての w' について $R(w, w')$ を計算し、行列中の値を更新する。文中に未知単語が一つだけある場合はその文を受理し、 未知単語を含めるように行列を拡大して他の語との関係を計算する (図 2)。

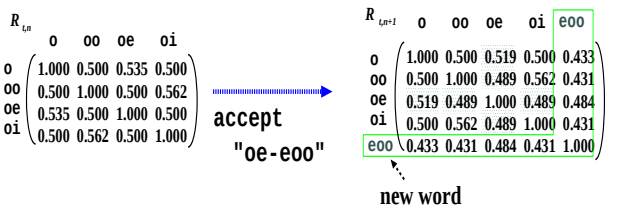

図 2: 行列として表された語間関係が文に応じて更新される例。左の行列で表されるような四つの単語を知っている状態で “oe-eoo” という文を受理した場合、単語’oe’ に関する部分が更新されるとともに、新たな単語’eoo’ を組み込んで行列が大きくなる。

3.3 結果のまとめ

5 主体の集団で上記の会話を行ったシミュレーションの結果をまとめる。初期には、各主体は単語に関してなにも知識を持たないとする。単語に使える文字数は 5 で、単語の長さ、文の長さの上限は 8 とした。語間関係の結合パラメータを $\alpha^w = 0.4$、 文を修正する確率を $c = 0.1$ とした。シミュレーションにより得られた結果は以下のようなものである。

1) 主体の持つ語間関係の中で、関係の強弱に応じて

[4] これは、主体が単語を自分の内部に位置付ける能力、言い替えるならば、知らない単語に意味付けをする能力の限界を設定している。

単語がクラスタを形成する。

2) 新しい単語が使われた時、あるいは、 単語の使われ方が新しい時、語間関係が大きな変化を見せる。主体の内部状態の変化を図 3 に例示した。これは、ある単語と、他の全ての単語との関係が会話ごとにどう変化したかを重ね書きしたものである。21 回目の会話の前後で、いくつかの強い関係をもつ単語と弱い関係を持つ単語の関係が逆転している。この変化は 1 回の会話中にゆっくりと起きるのではなく、単語が新しい用法で使われている 1 文を受理することで引き起こされる。

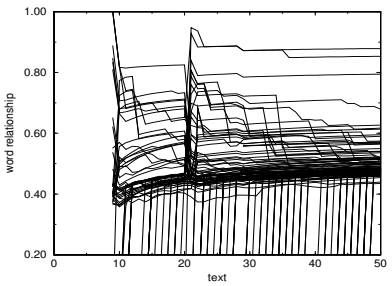

図 3: 会話による単語間関係の変化

3) 図 4 に図 3 で示した時点での単語間関係の変化を、クラスタ構造の変化がわかりやすいように示した。変化の前後での対応する単語を矢印で結んでいる。 新しい使われ方の単語 (破線で結ばれている) は、属しているクラスタが変わっている。 他の単語は、同じクラスタ内の語と同じ方向に動いており、全体の構造は大きくは変化していない。すなわち、単語のクラスタ構造は全体的な安定性と、局所的に新しい用法によって変わって行くという適応可能性をもつ。

4) 主体の内部構造のうち主体間で共有される部分が、会話とともにできてくる。

5) 会話によって主体間で共有されない部分もまた発展する。これは、 各主体は異なる会話の経験を積んで行くので、同じ文を受理しても異なったように位置付けを行うためである。

3.4 考察

主体はその内部に単語の関係の強弱に応じたクラスタを形成するが、これは一種のカテゴリーを作っていると解釈される。主体の集団は、そのカテゴリー構造のうち、共有する部分とそれぞれに個別な部分の両方を、会話を通じて発展させることができた。共有の構造が得られるという点は、社会性の萌芽と考えれる。その共有構造は、会話によって必ずしも大きくなるとは限らず、会話によって共有される程度が低くなる場合もある。これが主体の個別性の発展へとつながっている。集団を構成する全ての主体が同じになってしまうのではなく、主体が個別性を維持することにより、主体間の関係が常に変化し続け、社会性としての共有構造が変化していくことも考えられる。

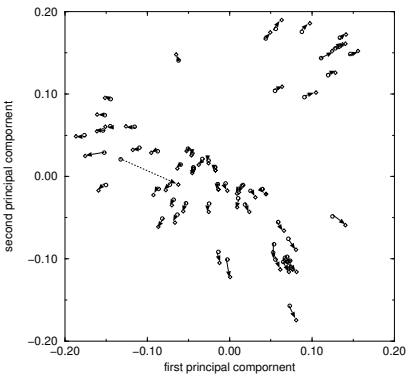

図 4: クラスタ構造の変化 (単語間関係の計算に使われる行列を主成分分析したもの)

本研究においては、単語を新しい用法で使うということが、めだった主体内部のダイナミクスを引き起こす。 新しい用法の単語は、 図 4 に見るように、 二つのクラスタを結び付けるような働きがあり、 あたかもメタファーのようなものである。§2 で議論したように、こういった文が集団の構造のレベルでのダイナミクスを導く可能性もあるが、現在この点については確定的な結果は得られていない。

4 結語–複雑系としての言語研究

前節で紹介した研究は、 主体の内部ダイナミクスに基づいて、カテゴリー形成や社会性と個別性が発展するという振る舞いを見せた。 しかし 2 節で述べた、絶えざる変化を示すシステムとしての言語を構成するという目標に対しては、まだまだ稚拙な段階にあると言わねばならない。新しい語が使われたときの主体内の

ダイナミクスがいかに全体へ影響するかという点の分析、および、主体の内部ダイナミクスをどのようにつくるかということの研究を進め、われわれの目指すべきものである、言語構造のダイナミクスの生成という創発現象を示したい。

人工生命の研究において「創発」とは、上位レベルにおける機能や大域的な構造が下位レベルにおける (または局所的な) 相互作用から自発的に現れる現象と言われる[5]。しかし、実際の研究では、創発した大域的な構造は局所的なダイナミクスを制限し、大域的な構造自体は静的なものになりがちである。 これでは、非線形・非平衡系の研究において、 散逸構造 (Nicolis and Prigogine, 1977; Nicolis and Prigogine, 1989) やシナジェティクス (Haken, 1978) が示したこととそれほど変わらなくなってしまう。

上位レベルの構造自体がダイナミックに変化しつづけるような動的な現象は、言語だけではなく、経済や社会システムなどにもよく見られる。そこで、複雑系としての構成論的な言語研究では、言語現象だけにとどまらない動的で複雑な現象の一般性を見出し、そのような現象についての理解を深めるためのテスト・ベッドとして言語を考え、 動的なシステムの (数理的) 構造の理解へ至ることが課題である[6]。

References

T. Arita and Y. Koyama. 1998. Evolution of linguistic diversity in a simple communication system. *Artificial Life*, 4:109–124.

J. L. Casti. 1999. The computer as a laboratory. *Complexity*, 4(5):12–14.

H. Haken. 1978. *Synergetics*. Springer-Verlag, Berlin.

T. Hashimoto and T. Ikegami. 1996. Emergence of net-grammar in communicating agents. *BioSystems*, 38:1–14.

T. Hashimoto. 1997. Usage-based structuralization of relationships between words. In P. Husbands and I. Harvey, editors, *Fourth European Conference on Artificial Life*, pages 483–492, Cambridge, MA. MIT Press.

T. Hashimoto. 1998. Dynamics of internal and global structure through linguistic interactions. In *The Proceedings of Multi-agent systems and Agent-Based Simulation*, pages 124–139, Berlin. Springer-Verlag.

Y. Karov and S. Edelman. 1998. Similarity-based word sense disambiguation. *Journal of Computational Linguistics*, 24:41–59.

G. Nicolis and I. Prigogine. 1977. *Self-organization in nonequilibrium systems*. John Wiley & Sons, New York.

G. Nicolis and I. Prigogine. 1989. *Exploring complexity*. R. Piper GmbH & Co., München.

L. Steels. 1996. Emergent adaptive lexicons. In P. Maes *et al.*, editor, *From animals to animats 4: Proceedings of the fourth international conference on simulation of adaptive behavior*, Cambridge, MA. MIT Press.

L. Steels. 1997. The synthetic modeling of language origin. *Evolution of Communication*, 1(1):1–34.

V. N. Vološinov. 1986. *Marxism and Philosophy of Language*. Harvard University Press, Cambridge, MA.

橋本敬. 1999. 動的言語観に基づいた単語間関係のダイナミクス. 認知科学, 6(1):55–65.

金子邦彦・池上高志. 1998. 複雑系の進化的シナリオ. 朝倉書店, 東京.

金子邦彦・津田一郎. 1997. 複雑系のカオス的シナリオ. 朝倉書店, 東京.

時枝誠記. 1941. 国語学原論. 岩波書店, 東京.

有田隆也. 2000. 人工生命. 科学技術出版, 東京.

[5] 人工生命の研究における「創発」についての議論は (有田隆也, 2000) によくまとめられている。

[6] 現在知られている、 数理モデルにおけるそういった現象の候補としては、 高次元カオス系に見られる「カオス的遍歴」(金子邦彦・津田一郎, 1997) がある。しかし、これが言語や経済といった系にそのまま当てはまるかどうかは不明である