第 16 回自立分散システム・シンポジウム資料, 125–130, (2004)
2004 年 1 月 26 日 ~ 27 日、京都、

動的言語観に基づく記号生成の構成論的モデル化

北陸先端科学技術大学院大学・知識科学研究科
橋本 敬

Constructive Modelling of Symbol Formation Based on the Dynamic View of Language

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

Abstract

Dynamics of language including the formation of symbols are studied with a constructive modelling. We discuss the dynamic view of language, which is the basic concept of this study, and the constructive approach. We propose a model, called developing word-web model, that is an implementation of the basic concept by a computational model. Simulation results of the developing word-web model show that the model realizes the coexistence of stability and adaptability and that of commonality with individuality. These features are summarization of the necessities for an established language system that is an essentially dynamic phenomenon.

Keywords: Dynamic View of Language, Constructive Approach, Developing Word-web Model

1 動的言語観

言語は本質的に動的な現象である。そのダイナミクスは、phylogeny、ontogeny、glossogeny[1] の 3 つのレベルに見られる。phylogenetic なレベルとは、生物進化のダイナミクスであり、言語の起源と初期進化にかかわる。ontogenetic なレベルは個体の学習・言語獲得に対応し、glossogenetic なレベルは文化進化のダイナミクスと関連する。本論では、ontogeny と glossogeny のレベルに着目し、言語のダイナミクス、特に、記号生成にからむカテゴリ構造と語彙構造の形成を、構成論的手法により研究する。

記号はなにかを意味する。しかしその意味は記号が指し示す対象ではなく、記号を用いる主体が解釈するものである。よって、記号の体系である言語の動的な側面、とくに、記号の生成過程を考察するにあたっては、言語を用いる主体の意味作用を重要だと考えるべきだろう。言語を用いることは、既に存在する言語構造を再認することではなく、意味を生成する活動として理解されるべきである。そのような生成過程の集積は言語構造自体の変化をもたらす。この考え方は、「動的言語観」として提唱されている (Hashimoto, 1998; 橋本敬, 1999)。

メタファーについて考えてみれば、意味をつくり出す活動としての言語という観点が分かりやすいだろう。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。新しいメタファーは、固定的な言語ルールがあるという立場から見るならば「間違った文」にすぎず、ルールを破っていることになる。しかし、われわれはそういった表現を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だという立場に立ってみると、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みなせる。ここでは典型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。認知言語学では言語におけるメタファーやメトニミーが、単にレトリック的なものではなく、我々の認識を形作る重要な道具としての地位を与えられている (Lakoff and Johnson, 1980; Taylor, 1995)。また、Nogales(1999) はメタファーを reconceptualization という観点から分析している。これらの研究では、普段メタファーとは思わずに使っている言葉や表現が、実はメタファーとして解釈でき、それらが我々の認識の奥深くに根ざしているか社会で

[1] glosso とは、「言語」を意味する接頭辞であり、geny は「発生」「起源」を意味する接尾辞である。すなわち、言語発生・言語起源を意味する。

定着した表現であるがゆえに、メタファーとしては意識されないということが指摘されている。逆にいえば、メタファー的表現と非メタファー的表現の間に明確な差はない。

ならば、メタファーについて言えることが多かれ少なかれ非メタファー的表現についても成り立つ。特に、reconseptualization というメタファーの機能を日常的な言語表現にまで拡張して考えてみてはどうだろうか。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、主体がある状況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。

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、「言語の本質を心的過程と見る言語本質観の理論的構成」を行った時枝の言語過程説 (時枝誠記, 1941) の主張に通じる。 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)。

2 構成論的アプローチ

構成論的アプローチとは、ひとことで言うならば、システムを作って動かすことにより対象を理解する方法である。これは、記号生成のような言語のダイナミクスを研究する手法として有効である。

言語には主体的活動が不可欠である。しかし、記述的に言語を扱おうとするならば、言語的ルールを、あたかも言語使用者としての主体を離れて実体化したものとみなしてしまう。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、言語的ルールはその主体の活動の結果として見い出されるのである。

構成論的手法では、言語活動や認知的振る舞いをする主体と、主体間の相互作用を構成し、そのシステムを分析的対象とする。言語や社会といった主体活動が不可欠な対象を科学的に取り扱う際には、主体性はできるだけ取り払われて来た。その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利点である。

単純なものから複雑化するようなシステムを理解する際に、このアプローチは有効である。通常、複雑な対象をそのまま構成することは難しい。一方、構成的手法では、複雑な状態をそのまま作るのではなく、より単純なシステムとそのダイナミクスを実装し、システムが自律的に複雑化するプロセスと結果を合わせて観測するという手法によって、対象である複雑なシステムの構造とその複雑化のプロセスを理解するのである (Kaneko and Tsuda, 1994)。

構成論的アプローチにより言語システムを研究する際には、意味や文法、記号の使用法、記号と対象の結び付けなどの言語ルールをアプリオリに設定すべきではない。そういった言語ルールは、言語使用者の活動と相互作用から創発すると考えるべきである。また、教育やマスメディアなどの特権的な制御機構も導入すべきではない。また、共有される記号や文法は固定化したものではなく、多様性と可塑性を持ったものでなくてはならない。その性質により、言語システムは常に変化するダイナミクスを持つようになるだろう。

3 Developing Word-web Model

動的言語観の計算モデルとして、単語間の関係の網目が会話とともに発展する Developing Word-web model ついて述べる[2]。

3.1 単語間関係網

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、単語を単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単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。

単語間関係を計算するアルゴリズムは、一文中の単語の用法の類似度と、複数の文にわたる単語の出現パターンの類似度を考慮し、会話とともに逐次的に更新されて行く。すなわち、ある文を受け取った時、その文を受け取る直前の単語間関係の状態から計算を行

[2] モデルの §3.1 と §3.3 については、(Hashimoto, 1998; 橋本敬, 1999) に詳細が述べられている。

う。本論では、会話の中で発話し受理される一連の文を「テキスト」と呼ぶ。出現パターンはこのテキスト中での単語の出現頻度から計算を行う。

テキスト t 中の n 番目の文を受理したときの二つの単語 w_i と w_j の間の関係 $R^{t,n}(w_i, w_j)$ は、用法類似度と出現類似度を次のように線形結合したものとして定義する。

$$R^{t,n}(w_i, w_j) = \alpha^w[\text{用法類似度}] + (1-\alpha^w)[\text{出現類似度}] \quad (1)$$

α^w は結合パラメータである。用法類似度の項は、文中で同時に使われた単語は関係を強めるように働く。第二項は、ある二つの単語の出現パターンがこれまでのテキストと現在のテキストで同じであれば正に、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 語彙

本論で紹介するシミュレーションでは、外界の対象を導入している。そして、主体が持つ単語のうちのあるものは、この対象と結び付いている。すなわち、そのような単語は対象を指し示す機能を持っていると考える。主体は前節で述べた単語間関係網に加えて、このような「語彙」を持っている。

単語 w と対象 o の間に結び付き強度 s_{wo} を考える。複数の単語が一つの対象を指している場合、すなわち、同義語がある場合にこの変数が参照される。ある対象を指す言葉を発する際、同義語の中の一つが

$$\sigma_{wo} = \frac{1}{\exp[-\beta s_{wo}] + 1} \quad (2)$$

に比例した確率で選ばれる[3]。

結び付き強度は、単語の使われ方により

$$s_{wo}^{n+1} = s_{wo}^{n} + \Delta s_{wo} \quad (3)$$

と増減する。増減のしかたは次のように分類される。

- トピック[4]になっている対象 t を指し示す語 w が使われた場合、$\Delta s_{wt} = u^{\mathrm{TU}}$
- トピックである同じ対象 t を指す、使われなかった語 w'、$\Delta s_{w't} = -u^{\mathrm{TN}}$
- トピック以外の対象 o を指す語 w、$\Delta s_{wo} = u^{\mathrm{OU}}$
- トピック以外の同じ対象 o を指す、使われなかった語 w'、$\Delta s_{w'o} = -u^{\mathrm{ON}}$

通常、増減のパラメータは $u^{\mathrm{TU}} > u^{\mathrm{TN}} > 0, u^{\mathrm{OU}} > u^{\mathrm{ON}} > 0$ に設定される。

3.3 会話

上で定義した単語間関係と語彙を内部に持つ主体の間の「会話」を、 単純に「文 (=単語の連なり) 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。

聴取者は聞いた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[5]。文が受理されなかった場合は発話者は別の文を作る。

文を発した際と、受理した際に単語間関係を更新する。単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。

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 $\mathrm{creativity\ rate}(c)$ 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。

最初の発話の際、トピック対象を指す単語を既に語彙中に持つ場合はそれを用いる。ない場合は新しい単語を作ってトピック対象を指すようにする。聴取者は初めの返答では、受け取った文の中にトピックを指す単語があると想定する。もし自分の語彙中にあるトピック対象と結び付いた単語が受け取った文で使われていないならば、文中のランダムに選んだ単語を、新たにトピック対象を指す語として採用する。

[3] σ_{wo} は結び付き強度の値域を $0 \sim 1$ にするために変換したものである。

[4] トピックについては §3.3 で述べる

[5] 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。

以上の文のやりとりのプロセスを、発話者と聴取者の役割を交替しながら、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。これが 1 テキストにあたる。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。

4　シミュレーション結果のまとめ

5 主体の集団で上記の会話のコンピュータ・シミュレーションを行った。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。また、外界の対象の間にはなんら関係を設定しない。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、外界の対象の種類は 10 とした。単語間関係の結合パラメータを $\alpha^w = 0.5$、 文を修正する確率を $c = 0.01$、結び付き強度の初期値を $s_{wo} = -5000$、σ_{wo} に変換する関数のパラメータを $\beta = 0.001$、強度変化の単位をそれぞれ $u^{\mathrm{TU}} = 50, u^{\mathrm{TN}} = 10, u^{\mathrm{OU}} = 10, u^{\mathrm{ON}} = 1$ とした。

シミュレーションにより得られた結果を以下にまとめる。

1) 主体の持つ単語間関係の中で、関係の強弱に応じて単語がクラスタを形成する。

クラスタの形は二つに分けられる。一つは flat 型で、クラスタ内の各単語は互いに強い関係にある。もう一つは gradual 型で、クラスタ内の単語間関係が徐々に変化し、ある二つは関係が強いが、全ての単語間が強い関係にあるわけではない。実際のシミュレーションでは、この二つの型が混ざって存在する。

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、単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 1 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。いくつかの単語間関係がたびたび大きく変化している。 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。

3) 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。

図 2 に単語間関係が大きく変化する時のクラスタ構造の変化を示した。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。

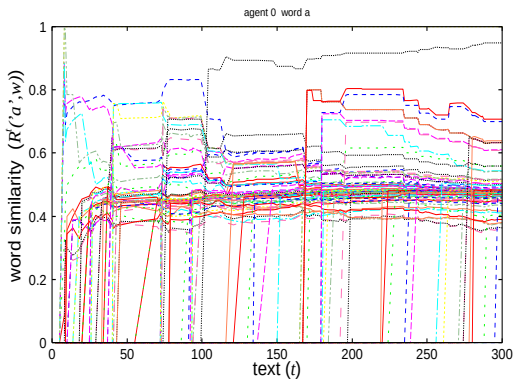

図 1: 会話による単語間関係の変化。ここでは、横軸にテクスト t を、各テクスト t の最後における、単語 $'a'$ と他の全ての語との単語間関係 $R^t('a', w)$ を縦軸に取り、全ての単語間関係を重ね書きしている。

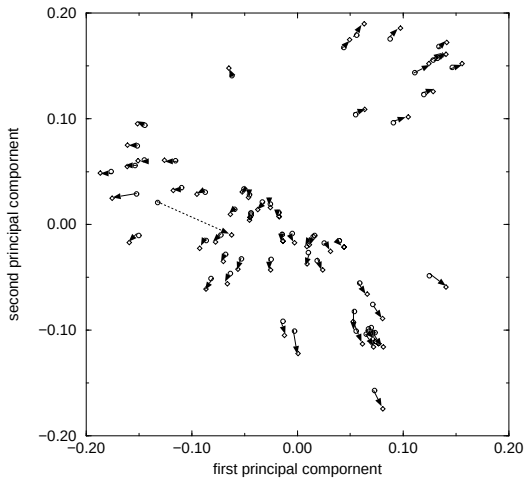

図 2: クラスタ構造の変化 (単語間関係の計算に使われる行列を主成分分析した散布図)

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。

6) 共有語彙の形成、衰退、同義語、多義語の発生、語彙の交替など様々なダイナミクスが見られる (図 3)。

7) 主体は、会話を通じて語彙構造を主体内に自ら組織化する。

図 4 a) はある主体の内部状態を、語彙とともに表している。外界の対象 (図では A∼J で表示) にはあらかじめ決まった構造はないが、主体内ではある構造をもったクラスタに分類されている。

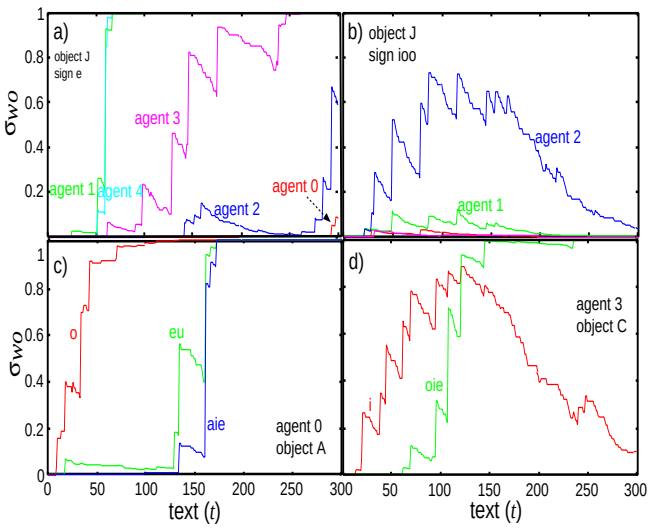

図 3: 単語と対象の結び付き強度の変化。横軸はテクスト t、縦軸は各テクスト t の最後における強度 σ_{wo}。a) 共有された語彙の形成。全主体がひとつの対象に同じ語を割り当てるようになる。b) 語彙の衰退。ある程度強度が強くなった対象と語との結び付きが衰退し、その語が使われなくなる。c) 同義語の発生。ある主体がひとつの対象に複数の語を割り当て、そのどれもが強い強度を持つようになる。すなわち、同じように使用される。d) 語彙の交替。ある対象に割り当てられる語が入れ替わる。

8) 共有される語彙構造は保存され、共有されない語彙構造は変化しやすい。

図 4b) は a) と同じ主体のしばらく後の内部状態である。a) と比べると、図の左側にある単語群の構造 (相互の位置関係) はあまり変化していないが、右側の単語群の構造は大きく変化している。図 4 c) は a,b) とは異なる主体であるが、左側の単語群の構造は図 4a,b) の主体と比較的と共通している。

5 議論

主体が、会話を通じて内部に単語のクラスタを形成し、対象の自律的な構造化を行うという現象は、一種のカテゴリを作っていると解釈される。flat 型クラスタと gradual 型クラスタの組み合わせは、プロトタイプ的カテゴリ構造と類似している。すなわち、カテゴリには典型的な要素 (プロトタイプ) と、非典型的な要素がある。gradual クラスタを介して複数のクラスタがゆるやかにつながるという点は、カテゴリ間の境界はリジッドに決まっているのではないというプロト

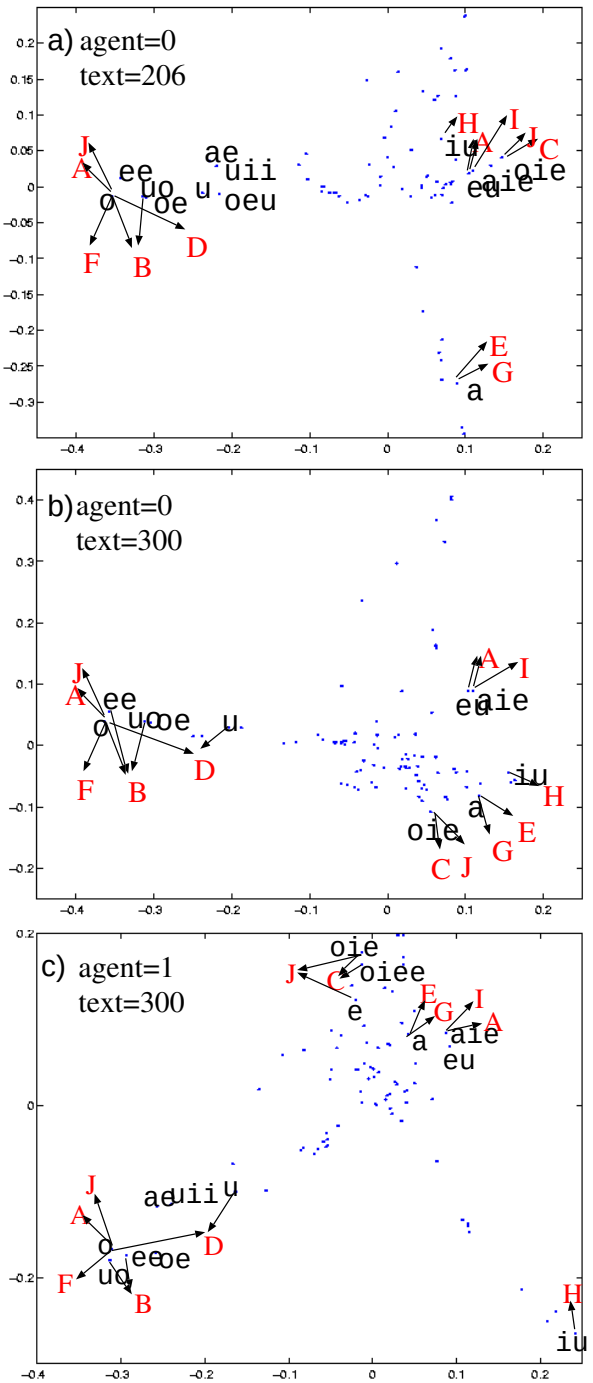

図 4: 語彙構造。単語間関係 $R(w, w')$ を主成分分析した散布図の各点 (単語) のうち、対象を指すものを記入している。a) agent=0,$t = 206$、b)agent=0,$t = 300$、c) agent=1,$t = 300$

タイプ的カテゴリ構造の特徴と共通している。

カテゴリは事物や概念の相互の関係を規定し分類するものであり、記号形成の基礎をなす。言語における記号の体系は外にある構造をただ写し取るものではなく、記号の使用に応じて動的にその体系を変化させながら構造化するものである。そして、上で述べたプロトタイプ的なカテゴリ構造により、記号体系、および、言語システムの安定性と適応可能性という重要な性質が実現されている。

カテゴリー構造のうち、共有される部分と個々の主体に個別な部分の両方が会話を通じて発展する。そして、共有される部分は集団の中で保存されやすく、共有されない部分は変化を蒙りやすい。共有の構造が得られるという点は、記号の社会的流通である。しかし、その共有構造は、会話を通じて常に発展するとは限らず、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。

言語における記号は、ただひとりの内的な体系を持つものではなく、ある程度社会において共有されることにより、コミュニケーションの用をなす。しかし、記号体系は完全に社会で共有され不変な構造を持つわけではない。個々人の多様性があり、多様性による解釈のずれを常にはらんでいる。すなわち、記号は使われることによって不可避的に記号体系自身のダイナミクスを生み出す、共有性と個別性を持ったシステムである。

6 結論 – 動的言語観による記号システムの構成

動的言語観を計算モデルとして実装した Developing Word-Web Model により、プロトタイプ的なカテゴリ構造の形成、新しい表現によるカテゴリ構造の変化、外界対象の自律的な組織化という、個体レベルのダイナミクスが観察された。また、会話を通じて共有されるカテゴリ構造と個別の構造がともに発展し、多義語・同義語の発生や語彙の交替といった現象を伴った語彙の共有が見られ、そして、共有された語彙構造ほど変化しにくいという集団レベルでのダイナミクスも生じた。すなわち、ontogentic なレベルと glossogenetic なレベルの動的な相互作用を通じて、安定性と適応可能性の両立、および、共有性と個別性の両立という、動的な記号生成をともなう言語システムが持つべき性質が、動的言語観に基づいたモデルにより実現されている。

謝辞

本研究は Canon Foundation Europe の Research Fellowship Program、 および、文部科学省科学研究費補助金 (No.12780269) の支援を受けている。ここに謝意を表する。

References

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.

K. Kaneko and I. Tsuda. 1994. Constructive complexity and artificial reality: an introduction. *Physica*, D75:1–10.

G. Lakoff and M. Johnson. 1980. *Metaphors We Live By.* The University of Chicago Press, Chicago.

P. D. Nogales. 1999,. *Metaphorically Speaking.* CSLI Publications, Stanford, CA.

J. R. Taylor. 1995. *Linguistic Categorization – Prototypes in Linguistic Theory.* Oxford University Press, Oxford.

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.

時枝誠記. 1941. 国語学原論. 岩波書店, 東京.