

語彙概念構造で表現できる言い換え*

藤田篤[†] 乾健太郎[†] 松本裕治[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科
{atsush-f,inui,matsu}@is.naist.jp

1 はじめに

言語表現の多義性を解消することが重要なように、同じ意味を表す異なる表現を捉えること、すなわち言い換えを生成したり認識したりする処理も重要である。

言い換えの実現方法は、テキスト中のある単語列を直接別の単語列に書き換える、あるいは統語構造の一部に変換を施す手法が主流である。一方、言い換えが「同じ意味を伝える表現」であることを考えると、異言語間機械翻訳における意味レベルのトランスファ方式 [1, 12] やピボット方式 [2, 5] と同様に、意味表現を明示的に使うアプローチは無視できない。たとえば、近年盛んに研究されている語彙意味論の体系は、統語的振る舞いを意味の特徴で表現するため、言い換への処理に有効な語彙的制約を提供すると考えられる。

語彙意味論の一つに、Jackendoff[8] によって提案された語彙概念構造 (Lexical Conceptual Structure; LCS) がある。英語では大規模な LCS 辞書が作られ、中間言語として機械翻訳や言語生成に用いられている [2, 5]。日本語に関しても、大石ら [15]、影山 [9] の包括的な分析に加え、数々の言語学的分析 [11, 16, 6] がある。さらに、竹内ら [17, 18] が影山の分析に基づいて LCS 辞書を整備・公開している。

このような背景から、近い将来実用に耐える規模の LCS 辞書が構築されると期待できる。そこで本稿では、(i) LCS によって表現できる言い換えにはどのようなものがあるか、(ii) LCS のどのような特徴がどのような場面で寄与するのかを論じる。

2 動詞の意味表現—語彙概念構造

本稿での議論は、基本的に竹内らが公開している日本語語彙概念構造 (LCS) 辞書 [17, 18]¹の体系に沿って進める。この体系は影山 [9] が提案する意味述語体系、それを識別するための言語テストに基づいて作成されたものである。所与の動詞に対する LCS はいくつかの言語テストを組み合わせた決定木に基づいて付与される。なお、他の体系については必要に応じて取り上げる。

2.1 語彙分解と意味役割

語彙概念構造は、種々の動詞をいくつかのタイプに分類したときに、それぞれのタイプが持つ基本的意味、概念的意味の記述の枠組みである。具体的には、動詞句の意味構造を、CONTROL, BECOME, BE AT などの抽象的な意味述語を用いて表現する (語彙分解)。たとえば、例 (1) の文 a~c は (2) の S_{a-c} で表される。

- (1) a. 家具屋の店員が僕の家大きな棚を届ける。
b. 僕の家大きな棚が届く。
c. 僕の家大きな棚がある。
- (2) S_a [[店員] x CONTROL [BECOME [[棚] y BE AT [家] z]]]
 S_b [BECOME [[棚] y BE AT [家] z]]
 S_c [[棚] y BE AT [家] z]

項 x, y, z は、意味述語によって意味役割を与えられる。たとえば、CONTROL はその左側の項を行為者 (Agent) と、BE AT はその右側の項を状態あるいは場所 (Goal) とみなす。上の S_{a-c} のいずれにおいても、意味述語と「店員」、「棚」、「家」の位置関係は等しく、それぞれ Agent, Theme, Goal という意味役割が与えられる。これに従うと、(2) の全体の解釈は「棚 (Theme) の状態または場所が家 (Goal) という状態に変わるように店員 (Agent) がはたらきかける」となる。

2.2 アスペクトに関する情報

LCS は語のアスペクトに関する情報を含んでいる。たとえば (2) の S_a は、(1a) の『達成』という語彙的アスペクトを示している。「店員」に関する陳述を除くと (2) の S_b の LCS が得られ、これは (1b) の『到達』のアスペクトを示している。さらに、BECOME が表す状態変化を捨象すると、ある『状態』を表す (2) の S_c の LCS ((1c) に対応) が得られる。このほか、竹内らの体系では、例 (3) のような『活動』、例 (4) のような『継続』を区別している。これらが未完了 (atelic) を表すためである。

- (3) a. 監督が選手を褒める。
 S_a [[監督] x ACT ON [選手] y]
- (4) a. ブログが社会に浸透する。
 S_a [[ブログ] y MOVE TO [社会] z]

2.3 ヴォイスに関する情報

LCS は陳述の視点に関する情報も表す。たとえば、(1c) のような「そこにもものがある」という存在文は、(5c') のような所有文に書き換えることができる [9]。(2) の S_c は (5c') のように表現できるが、ここで、 $[z$ BE WITH ...] という部分構造は視点に関する情報を表していると見なせる。

- (5) c'. 僕は (僕の家) 大きな棚を持っている。
 $S_{c'}$ [[僕] z BE WITH [[棚] y BE AT [僕の家] z]]

このように、単語 (動詞) に含まれる視点の情報は、 $[x$ CONTROL ...] や $[z$ BE WITH ...] のような構造で表現される。ある語の LCS がこれらの構造を含む場合、その語は語彙的使役、語彙的受身を表すことになる。

2.4 項と格のリンキング

LCS は、働きかけや状態変化などのイベントに関わる要素が意味構造上でどのように関係しているかだけ

*Exploiting Lexical Conceptual Structure for Paraphrase Generation. Atsushi Fujita[†], Kentaro Inui[†], and Yuji Matsumoto[†]

[†]Graduate School of Information Science, Nara Institute of Science and Technology

¹<http://cl.it.okayama-u.ac.jp/rsc/lcs/>

でなく、意味構造と統語構造を結び付ける情報（リンクルール）も備えている。

竹内ら [17, 18] が提案する日本語動詞の LCS 体系では、3 種の意味役割 Agent（行為者）、Theme（対象）、Goal（状態や空間上の着点）で表現できる LCS に限定して 16 種類の LCS を定義し、約 1,300 語を記述している。この体系では外項、内項の区別に基づいて項と格が結び付けられる。項 x は、 $(2S_b)$ や $(2S_c)$ のように状態変化や存在を表す LCS には現れない。すなわち Agent、外項である。したがって、 x を含む LCS では格格に対応付けられる。ただし、 $(5S_c)$ のように視点が移る場合もあるため、LCS の左端の項を格格に対応付けるとしている。 y は Theme を、 z は Goal を表すために用いられる。LCS が x を含む場合、 y はヲ格に、 z は二格に対応付けられる ($(2S_a) \Leftrightarrow (1a)$, $(3S_a) \Leftrightarrow (3a)$)。 x を含まない場合は y はガ格に、 z は二格に対応付けられる ($(2S_b) \Leftrightarrow (1b)$, $(2S_c) \Leftrightarrow (1c)$, $(4S_a) \Leftrightarrow (4a)$)

他方、大石ら [15]、Dorr ら [2] は、動詞の統語構造および各種動詞交替の可能性に基づいて LCS 体系を設計し、LCS の種類ごとに統語構造との対応付けの情報を記述している。大石ら [15] は約 700 語の日本語動詞の分析に基づいて 30 種以上の LCS を設計している。一方、Dorr ら [2] の英語の LCS 体系²では、Levin [13] の動詞分類に基づいて 450 種以上の多様な LCS を定義し、約 4,000 語の動詞を記述している。

3 言い換えの統語的スキーマ

まず、次の 2 つの言い換いの例を見てみよう。

- (6) s. 風が勢いを弱めている。
t. 風の勢いが弱まっている。
- (7) s. 2 位の選手が先頭との距離を縮めている。
t. 2 位の選手と先頭との距離が縮まっている。

これらは他動詞文から自動詞文への言い換え（自他交替）の例である。格要素が連体化したり異なる助詞に割り当てられたりするため、自他交替に関しては (8) のような言い換いのスキーマが導かれる。

- (8) V_i (他動詞) $\Rightarrow V_i$ (V_i に対応する自動詞)

格要素をどう変形するべきかは、LCS のリンクルールによって決定できる。次に、例 (9) を見てみよう。

- (9) s. 子供がストーブで身体を暖めている。
t. ~~ストーブで~~子供の身体が暖まっている。
- (9s) が変化過程に焦点を当てているのに対し、(9t) は「すでに暖まっている」という結果状態しか表さないため、(9t) は言い換えとしては適切でない。どのような条件を満たせば適切な言い換えであるかは、(8) のスキーマでは言及されていないが、LCS のアスペクト情報を用いれば計算できる。

例をもう 1 組見てみよう。機能動詞結合の言い換えを取り上げる。例 (10s)、(11s) の「刺激を受ける」「刺激を与える」のような構文を機能動詞結合と呼ぶ [14]。この構文では、動詞「受ける」「与える」は文法的な機能のみを果たし（機能動詞）、動詞の格要素となっている動詞派生の名詞（動作性名詞、e.g. 「刺激」）が実

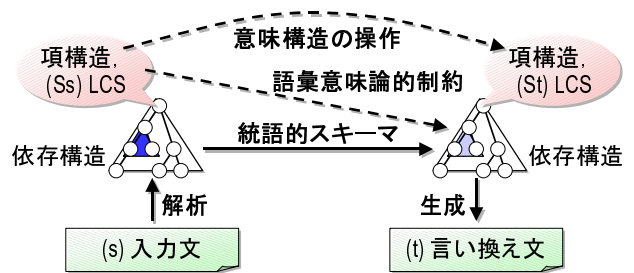


図 1: 言い換いのどの場面で LCS を用いるか

質的な意味を担っている。機能動詞結合は動作性名詞を動詞化して (10t)、(11t) のように言い換えられる。

- (10) s. 息子が友人の活躍に刺激を受ける。
t. 息子が友人の活躍に刺激される。
- (11) s. 友人の活躍が息子に刺激を与える。
t. 友人の活躍が息子を刺激する。

これらの例では、「刺激される」「刺激する」のように言い換え後の動詞のヴォイスが異なる。また、(11t) のみ「息子」の格助詞が変化している。格助詞やヴォイスは対象となる動作性名詞や機能動詞によって異なるし、格格、二格の動作性名詞も機能動詞結合を形成しうる [14]。これをふまえ、(12) に示すような機能動詞結合の言い換いのスキーマが考えられる。

- (12) N_v (動作性名詞) を V (機能動詞)
 $\Rightarrow N_v$ (動詞形) $_$ (ヴォイス)

格助詞やヴォイスは、動作性名詞と機能動詞のあらゆる組み合わせについて記述しなくても、LCS が提供する視点の情報やリンクルールを用いることで決定できる可能性がある [4]。

このように、言い換えは特定の統語構造あるいは語を別のそれと置き換える言い換いのスキーマによって動機付けられ、実現される。一方、LCS は異形同質の動詞を同じ意味構造で、同形異質の動詞は別の意味構造で表現できるため、統語構造レベルであらゆる入力に対する処理を記述するのに比べて、様々な選択をコンパクトに記述できる可能性がある。我々は、LCS を語彙意味論的制約として用いたり、それに基づく意味構造の操作を考える。そして、図 1 に示すように、依存構造上の言い換いのスキーマを組み合わせることで個々の言い換いの実現を目指す。

4 語彙概念構造に基づく言い換え

LCS の特徴は少なくとも次のような場面で言い換いのスキーマに情報を補ってうまく作用する。

- 生成 ヴォイス、アスペクト、リンクルールの点で、
 - (i) 言い換え後の構造が言い換え前と等価か否かを計算し、
 - (ii) 言い換え後の構造のうちどこまでが一つの語の意味の範囲で、どこからが統語的に表すべき部分かを同定する。
- 交替・派生 様々な交替・派生関係を LCS の操作で表現する。ヴォイス、アスペクトに関するものであれば、交替・派生関係は LCS の組（意味構造の操作の結果）で表現できる。
- 意味的構造変換 LCS を参照して構造変換を制御する。

²<http://clipdemos.umiacs.umd.edu/englcslx/>

4.1 生成時の制約

2 節で示したヴォイスやアスペクトを示す情報は、言い換えを生成する際の表現の適否や言い換えにおける等価性を判定する制約となりうる。3 節の例を再考する。

- (13) s. 風が勢いを弱めている (変化過程)
 S_s [[風] x CONTROL
 [BECOME [(風の 勢い) y BE AT [弱い] z]]]
 t. 風の勢いが弱まっている (変化過程/結果状態)
 S_t [[[(風の 勢い) $x=y$ CONTROL
 [BECOME [(風の 勢い) y BE AT [弱い] z]]]]
- (14) s. 子供がストーブで身体を暖めている (変化過程)
 S_s [[子供] x CONTROL
 [BECOME [[身体] y BE AT [暖かい] z]]]
 t. ストーブで子供の身体が暖まっている (結果状態)
 S_t [BECOME [[身体] y BE AT [暖かい] z]]]

影山 [9] は、(13 S_s) から (13 S_t) への派生を反使役化、(14 S_s) から (14 S_t) への派生を脱使役化と呼んでいる。(13) と (14) の違いは、他動詞の LCS から自動詞の LCS を派生する際に生じるアスペクト性の違いによって説明されている [16]。

4.2 意味構造の着脱による交替・派生関係の表示

2.3 項では、ヴォイスに関する情報を持つ部分構造 [x CONTROL ...], [z BE WITH ...] を示した。この部分構造を付加あるいは削除することで、語彙的なヴォイスに関わる言い換えを表現できる。例 (15) (文献 [10] より引用) は、異なる視点から行為を陳述する「教える」と「教わる」の間の言い換えを示している³。(15 S_s), (15 S_t) はそれぞれ (15s), (15t) を表す LCS である。

- (15) s. 先輩が後輩に合格の秘訣を教える。
 S_s [[先輩] x CONTROL
 [[秘訣] y MOVE FROM [先輩] x TO [後輩] z]]]
 S_s' [[先輩] x CONTROL [BECOME [[後輩] z BE WITH
 [[秘訣] y MOVE FROM [先輩] x TO [後輩] z]]]]]
 S_t [BECOME [[後輩] z BE WITH
 [[秘訣] y MOVE FROM [先輩] x TO [後輩] z]]]
 t. 後輩が先輩に合格の秘訣を教わる。

「教える」では「秘訣が先輩から後輩に移動する」という命題を「先輩が使役している」ことが [x CONTROL ...] によって、「教わる」では同じ命題を「後輩が受けている」ことが [z BE WITH ...] によって表されている。これより、この授受の動詞交替については、(16) のような言い換えのスキーマ、および (17) のような LCS の操作が考えられる。

- (16) $V \Rightarrow V'$ (V に対応する授受動詞)
 (17) [x CONTROL ...] \Leftrightarrow [z BE WITH ...]

(15 S_t) は、(2 S_c) から (5 S_c') への変形と同様に (15 S_s) を所有文の構造 (15 S_s') に変換し、例 (14) と同様の脱使役化を施すことで得られる。

4.3 意味的構造変換

LCS が備えている情報を駆使すれば、交替や派生の関係にない動詞間の言い換えにおいても、言い換え前

³文献 [10] では、文献 [9] と異なり、存在文を所有文にする際に BECOME も付与されている。例 (15) は文献 [10] に従う。

の動詞の格要素にどの格を割り当てるか、アスペクトに関する情報、言い換え後の動詞のヴォイスなどを決定できる [4]。

機能動詞結合の動作性名詞を動詞化して (18t1)、あるいは (18t2) のように言い換えることを考える。

- (18) s. 息子が友人の活躍に刺激を受ける。
 S_s [BECOME [[息子] z BE WITH
 [[刺激] y MOVE FROM [活躍] x TO [息子] z]]]
 S_{t_0} [x ACT ON y]
 S_t [BECOME [[息子] z BE WITH ...] \oplus
 [[活躍] x' ACT ON [息子] y']]]
 t1. («息子」の視点) 活躍が息子を刺激する。
 t2. 息子が友人の活躍に刺激される。

まず、(18s) に対して (18 S_s)、および「刺激」の動詞形である「刺激する」の LCS (18 S_{t_0}) が与えられる。次に、MOVE FROM TO と ACT ON の視点および意味役割の共通性に基づいて (18 S_s) の項がそれぞれ対応付けられる。(18 S_s) のうち、対応付けられなかった部分構造 [BECOME [太郎] z BE WITH ...] がそのまま引き継がれ (18 S_t) を為す。「刺激する」がこの部分構造を語彙的に表せないため、記号「 \oplus 」を用いて示してある。(18 S_t) から (18t1) への投射はリンクルールに基づいて実現され、最後に視点の情報を考慮して受動化が施され (18t2) を得る。

ここで、[z BE WITH ...] という視点を表す部分構造の役割に注意されたい。この構造は、(15) では語彙的受身として「教わる」の意味に含まれていたが、(18 S_t) では「刺激する」という語で表せる意味の範囲に入らなかった。このように、LCS によって語の意味の範囲が与えられているので、[x CONTROL ...] や [z BE WITH ...] が語の意味の範囲からはみ出る場合には、統語的使役「~させる」、統語的受身「~られる」として語彙化すべきことが分かる。

この言い換えは例 (13) や (15) と異なり、変換の際に多くの仮説に依っている。たとえば、異なる意味述語の対応付けなどである。語彙意味論的制約に基づいてこのような LCS の操作を実現する際には、意味述語あるいは意味役割の共通性にも着目する必要がある。

4.4 語彙概念構造で表現できない言い換え

複数の節に関わる言い換えや主題化などの談話要素を扱う言い換えは、語彙意味論で形式化できる範疇を超えるため、LCS が有用な場面は今のところ見つからない。また、一つの節の中の言い換えの場合でも、慣用句のような非構成的な表現の言い換えは表現できない。たとえば、例 (19) の実現には、「脚光を浴びる」 \Leftrightarrow 「注目される」のような対が必要である。

- (19) s. 「ひかり都市」として脚光を浴びました。
 t. 「ひかり都市」として注目されました。

5 議論

5.1 関連研究

本稿では LCS とその操作による言い換えの表現方法を提案したが、(i) 項構造と格を結び付ける知識、(ii) アスペクト情報、(iii) ヴォイス情報などを個別に蓄えてお

けば実現できそうに見えるかもしれない。実際、Dorrら [2] は、LCS に基づいて辞書項目を記述しているが、単に同じ語の複数の辞書項目を結びつけるだけで交替の可能性を表現している。また「売る」⇔「買う」のような交替は、より高い抽象度の意味構造を新たに設計して表現するとしている [3]。

個々の知識をすべて組み合わせる用いるならば、古典的な語彙意味論の一つである Meaning-Text Theory [19, 12] の立場に近くなる。この理論では、個々の語に対する交替や派生は項構造を引数/返り値とする語彙関数として体系化されるが、個々の語に対して経験的に作るしかなく、異なる単語の同じ振る舞いを実現するのに多くの冗長な記述を要する。これに対し LCS は、表現の多様性を吸収するための一貫性のあるコンパクトな記述体系である。ゆえに、LCS に基づいて言い換えの操作を実現するのは現実的なアプローチといえる。

5.2 必要な語彙資源

LCS の体系化に関する進行中のプロジェクト [18, 6] があることから、本稿では LCS 辞書を仮定して言い換える表現方法について議論した。一方、個々の種類の言い換えを実現するためには、LCS 辞書とは別の語彙資源が必要になる。

例 (13), (15) のような言い換えの実現には、それぞれ「弱まる」⇔「弱める」、「教える」⇔「教わる」のような語の対を蓄える必要がある。ただし、これらを必ずしも「自動詞と他動詞の対応表」や「授受動詞の対応表」のような別々の辞書として持っておく必要はない。所与の語の組の各々の LCS が得られた時点で、LCS 間の関係（意味構造の操作の結果）が分かるからである。言い換えれば、「弱まる」⇔「弱める」という語の対に対して、 $(13S_s)$, $(13S_t)$ という LCS の組が得られると、これらが反使役化によって結ばれる他動詞と自動詞の関係であることがおのずと分かる。

5.3 項構造選択

本稿では LCS を用いて言い換えを表現することを目的とし、所与の統語構造に対して LCS が一意に付与できると仮定していた。しかし、LCS の項から統語構造の格への投射と異なり、格から項への対応付け（意味解析）は一筋縄ではいかない。統語構造においてヲ格や二格が多様な役割を担うためである。仮に LCS の違いが語義の違いによるならば、まず所与の統語構造に対する項構造解析が必要であろう。そこで我々は、格要素および動詞の類似性に基づいて格フレーム用例をクラスタリングし、クラスと項構造を効率的に結び付けることによって大規模な用例付き項構造辞書を構築する手法を検討している [7]。

6 おわりに

本稿では語彙概念構造 (LCS) の理論を取り入れた言い換える表現方法を提案した。そして、動詞の語彙意味論的特性に従って規則的に実現可能と考えられるいくつかの種類 of 言い換えを取り上げ、LCS による説明を試みた。LCS は生成、交替・派生、意味的構造変換

の各場面で統語的スキーマに情報を補う。

現状では、LCS を介して表現できる言い換える種類を網羅的には列挙できていない。ゆえに、今後はまず、より広い範囲の言い換えるの調査を考えている。そして、交替・派生や意味的構造変換の操作を実装し、その妥当性を経験的に確かめる。実際、LCS 辞書や体系を仮定して言い換えるのモデルを論じたが、辞書の編纂にはまだ多くの課題が残っている [18, 6]。だからこそ、言い換えというアプリケーションが分析の材料を提示し、体系そのものの形式化に関する議論の発展に寄与すると考える。

参考文献

- [1] M. Dorna, A. Frank, J. van Genabith, and M. C. Emele. Syntactic and semantic transfer with F-structures. In *Proc. of COLING-ACL*, pp. 341–347, 1998.
- [2] B. J. Dorr. Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Translation*, Vol. 12, No. 4, pp. 271–322, 1997.
- [3] B. J. Dorr, R. Green, L. Levin, O. Rambow, D. Farwell, N. Habash, S. Helmreich, E. Hovy, K. J. Miller, T. Mitamura, F. Reeder, and A. Siddharthan. Semantic annotation and lexico-syntactic paraphrase. In *Proc. of LREC Workshop on Building Lexical Resources from Semantically Annotated Corpora*, pp. 47–52, 2004.
- [4] A. Fujita, K. Furihata, K. Inui, Y. Matsumoto, and K. Takeuchi. Paraphrasing of Japanese light-verb constructions based on Lexical Conceptual Structure. In *Proc. of ACL Workshop on Multiword Expressions: Integrating Processing*, pp. 9–16, 2004.
- [5] N. Habash, B. J. Dorr, and D. Traum. Hybrid natural language generation from Lexical Conceptual Structures. *Machine Translation*, Vol. 18, No. 2, pp. 81–127, 2003.
- [6] 畠山真一, 坂本浩, 加藤恒昭, 伊藤たかね. 日本語動詞の LCS 推定に関して—他動詞を中心に—. 情報処理学会研究報告, NL-165-1, pp. 1–8, 2005.
- [7] 平野徹, 飯田龍, 藤田篤, 乾健太郎, 松本裕治. 動詞項構造辞書への大規模用例付与. 2005. (in this proceedings).
- [8] R. Jackendoff. *Semantic structures*. The MIT Press, 1990.
- [9] 影山太郎. 動詞意味論—言語と認知の接点. くろしお出版, 1996.
- [10] 影山太郎. 非対格構造の他動詞—意味と統語のインターフェイス. 伊藤たかね (編), 文法理論:レキシコンと統語, pp. 119–145. 東京大学出版会, 2002.
- [11] 岸本秀樹. 壁塗り構文. 影山太郎 (編), 日英対照 動詞の意味と構文, pp. 100–126. 大修館書店, 2001.
- [12] B. Lavoie, R. Kittredge, T. Korelsky, and O. Rambow. A framework for MT and multilingual NLG systems based on uniform lexico-structural processing. In *Proc. of ANLP-NAACL*, pp. 60–67, 2000.
- [13] B. Levin. *English verb classes and alternations: a preliminary investigation*. Chicago Press, 1993.
- [14] 村木新次郎. 日本語動詞の諸相. ひつじ書房, 1991.
- [15] 大石亨, 松本裕治. 格パターン分析に基づく動詞の語彙知識獲得. 情報処理学会論文誌, Vol. 36, No. 11, pp. 167–180, 1995.
- [16] 杉岡洋子. 形容詞から派生する動詞の自他交替をめぐる. 伊藤たかね (編), 文法理論:レキシコンと統語, pp. 91–116. 東京大学出版会, 2002.
- [17] 竹内孔一, 内山清子, 吉岡真治, 影浦峯, 小山照夫. 語彙概念構造を利用した複合名詞内の係り関係の解析. 情報処理学会論文誌, Vol. 43, No. 5, pp. 1446–1456, 2002.
- [18] 竹内孔一. 語彙概念構造による動詞辞書の作成. 言語処理学会第 10 回年次大会発表論文集, pp. 576–579, 2004.
- [19] L. Wanner. *Current issues in Meaning-Text Theory*. Pinter Publishers, 1994.