数理的研究

中野洋

はじめに

数理的研究の分野がいつも問題になる。他の展望の多くが対象を規定しているのに対し、数理的研究は方法を規定しているためである。研究が数理的であるためには方法が次の条件を満足していなければならないと考えられる。すなわち、本来、多面的である言語現象を、客観的・再現可能なたった一つの切り方（あるいは、その組み合わせ）でとらえている必要がある。それは質的相違を量的相違に変換することに通じる。その結果、処理を電子計算機（以下計算機と略記）にゆだねることも可能となる。

この分野の研究は、統計的分析によって言語現象を明らかにしようとするものと、計算機によって言語の情報処理機構を明らかにしようとするものとに大別することができる。したがって、以下統計的分析研究・言語情報処理の研究のそれぞれについて、各研究の力点あるいは筆者から見た価値がその理論・分析方法・処理手順・結果のどこにあるかによって分類し紹介する。

概観

国家予算（文部省科学研究費：特定研究「言語」、通産省工業技術院大型プロジェクト「パターン情報処理システム」）が投入され、言語情報処理の研究が盛んであった。

言語情報処理のみならず、言語の統計的研究においても、道具として計算機が用いられるようになった。言語研究ソフトウェアの開発、安価で高い機能を持ったハードウェアの開発、言語データの蓄積、および、それらの公開と共同利用の推進、啓蒙活動等によるものである。

道具の開発は、研究のスピードを高めるものではあるが、新しい研究を生むものではない。この時期に二つの新しい研究の流れが確認できた。一つは、従来も行なわれて来た文学作品の語彙・表現論的研究の一つに統計的手法を用いることが一般的になった。他の一つは文章に関する研究である。この動きは国語学界だけでなく、文学・心理学・工学の分野でも同時におこっている現象である。

理論的研究およびモデルの構築

新しい理論的研究は多くない。言語モデルの構築に、樺島忠夫に「文章の構造」（『日本語のスタイルブック』第二章、大修館、昭54・4）等の一連の研究がある。これは文章の構造を、意図の構造・意味内

容の構造・文脈の切れ続きの構造によってとらえ、これらによって文章を生成するための規則を提案したものである。この研究は文章作法や国語教育に役立つだけでなく、言語情報処理における文章解析・生成にも有益な情報を与えるだろう。

初めから言語情報処理をねらったものに、石綿敏雄に「『デラシネの旗』の文章解析」(『計量国語学』11—6、昭53・9)等がある。これは一つ一つの文を受け取って、文章をどのように理解するか、内容を理解できる手がかりがどこに存在し、どのようにとらえることができるのかを現時点で既得の知識、技術によって解析しようとしている。

モンテギュー文法の紹介・適用が前期に続いてあった。仁科弘之「モンテギュー文法と特定性」(『計量国語学』11—4、昭53・3)、石本新「モンテギュー文法をめぐって——自然言語の分析」(『数理科学』193、昭54・7)等がそれである。

論理学の観点からの言語の記述・分析について、次の二点があった。小野寛晰「名前と名指し関係」(『計量国語学』11—4、昭53・3)は、ものとものの名前を同じレベルで取り扱うことができる一階の述語論理上の一つの形式体系を提案したもの、草薙裕「時制の論理と自然言語表現Ⅰ、Ⅱ」(『計量国語学』12—1、2、昭54・6、9)は、自然言語の時に関する表現およびアスペクトの意味記述に時制の論理を応用しようとするものである。

漢字についてのモデル論が二つあった。山中光一「日本語文字システムのモデル的考察」(特定研究「言語」研究報告資料、昭54)、海保博之「漢字情報処理機制をめぐって」(『計量国語学』11—8、昭54・3)である。計算機で文字・漢字をどう扱うか、人間は漢字をどう処理しているかについての論究である。

統計的研究のための分析方法

どのような単位で現象をはかり、予測するかは統計的研究の出発点である。語彙調査においては同じ語か異なる語かの判定は調査の基本となる。土屋信一「高校教科書の同語異語判別システム」(国研報告61『電子計算機による国語研究Ⅸ』以下この書名は省略する、昭53・3)は計算機による語彙調査でのこの問題を論じている。計算機にデータを入れれば何らかの結果は出るが、それが意味を持つためにはこの問題を避けて通れない。

数理的研究では主観的分析はさけなければならない。野元菊雄「ていねいさの順位」(『国語学』115、昭53・12)は、主観的判断によるていねいさの順位がどの程度一致するか、そしてそれは何によるかについて論究している。

文学作品の語彙・表現論的研究に統計的分析法が用いられた。浅見徹「万葉集の語彙構造」(『国語と国文学』55—5、昭53・5)は、延べ語数六三、二六三を各巻別、および分類語彙表の分類別に集計し、各巻の類似度の計算、集中度×使用率の値による特徴語彙の抽出等の分析をおこなった。集計だけでも大変な作業量であるのにもかかわらず、さらに統計的な新しい試みを加えている。この種の研究方法の一つの水準を示している。西田直敏「平家物語語彙論序説」(『国語国文研究』60、昭53・7)、進藤義治『源氏物語形容詞類語彙の研究』(笠間書院、昭53・11)、森下金二郎「新古今集の表現形態について——統計的にみた一つの推論——」(『宮城学院女子大学研究論文集』50、昭54・6)は、それぞれの作品と他の数作品の語彙の比較をおこなうことにより、その作品の語彙・表現論的考察を加え、

成果を上げている。この種の方法に重要なことは、比較すべき対象が同じ規準で単位切りされ、同語異語判別されていることである。共通の規準で調査を進め、結果を公表することは、単にその研究だけでなく学界の水準をも高めることにもなる。そのような注意が払われることを期待したい。

これまでにない研究対象には、これまでにない分析方法が必要である。談話行動の研究と文章構造の分析にそれが見られる。

この分野の計量的研究として、米田正人「談話行動の計量的研究について」・江川清「談話行動の実験社会言語学的研究――目標と資料収集方法について――」・「身ぶりの記述について」、杉戸清樹「身振りを記録する――「変化」の記録表試案――」（以上四点、国研報告62『研究報告集1』〈以下この書名は省略〉、昭53・3）がある。これらは、非言語行動を含めた談話行動をいかに記述するかについての考察であって、今後が期待される。南不二男編『講座言語第三巻　言語と行動』（大修館書店、昭54・11）はこの分野の現時点での研究水準を示す好著である。

文章構造の研究は語彙の面からの分析と内容の面からの分析の二つがある。共に時系列的解析によっている。

語彙の面からの分析に、寺川みち子「段落・談話の構造」（『東海学園国語国文』16、昭54・9）は文章論上の単位の具体的な切れ目を求めるために、語・形態素の時間的な現われを利用し、いくつかのタイプの段落・談話を分析している。この種の分析では、分析対象が長くなると、それだけ語彙量も多くなり、分析し切れなくなるのが普通だ。これは計算機を用いても同様だ。村井護晏「理科教科書の分析――その内容のスペクトル分析――」（『大分大学教育学部研究紀要』5―4、昭54・10）では、因子分析法によって因子を求め、それぞれの教科書の特性項目（単語で代表）の自己関連度の強いものから順に10になるまで拾い出し、それらの特性項目についてのみ分析するという方法を用いてこの問題を解決している。杉戸清樹・沢木幹栄「言語行動の記述――買い物行動における話しことばの諸側面――」（『言語と行動』大修館書店、昭54・11）は一つの談話資料に対し語彙内容、発話時間、発話回数、文長等各種の分析をほどこし文章構造を総合的に解析しようとする研究である。前述、進藤では源氏物語の文章を自立語二〇〇で区切り、各区間ごとに調査することにより、文章の流れと語彙（形容詞）の変化をとらえている点に新しい分析法がみられる。

内容の面からの分析には、前述、樺島、石綿がある。これらの言語学的な多種多様な文章解析はもっとなされなければならない。これらの他に、田中穂積「談話理解の構造」（『情報処理』20―10、昭54・10）等の工学的研究、高木和子・丸野俊一「幼児の物語理解に及ぼす先行情報の質的効果」（『教育心理学研究』27・4、昭54・12）等の心理学的研究、ジョルフコスキーおよびメルチューク等の研究を紹介した北岡誠司「生成詩学の試み――解説にかえて――」（『文学』47―6、昭54・6）、文結合関係について述べた、下川浩「文結合関係――命題相互関係――全体テキストの階層的構成――」（『エネルゲイヤ』5、昭53・6）等、国語学界以外の分野でも盛んになっている研究分野である。

言語データにおける多変量を解析する手段として、数量化理論や因子分析法が用意されている。これらは、計算機を用いることによって容易になった統計的分析法である。水谷静夫「数量化理論とそ

の応用」(『情報管理』21—9、昭53・12)は言語データへの適用法について述べ、「計算機が十数秒で解いた結果の検討に、四日も五日も掛かるのは、珍しくない。(中略)この手間を惜しんでは、数量化理論の真価は発揮されまい」と述べている。この手法は決して「魔法の杖」ではないのである。水谷静夫の以下の研究、すなわち、「昭和前期流行歌と相の語」(『国語学』115、昭53・12)、「続形容動詞辨」(『計量国語学』11—7、昭53・12)、「用語による梅・桜の歌の弁別」(計量国語学』12—1、昭54・6)、「二つの流行歌——言語の統計的処理法——」(『国文学解釈と鑑賞』572、昭54・10)および、井上史雄「地域差の実態——北海道浜言葉——」(『東京外国語大学論集AREA AND CULTURE STUDIES』28、昭53・3)は数量化理論第Ⅲ・Ⅳ類による分析をおこなっている。

大城宜武「チェック・リスト法による語意味の分析——計量意味論からの接近——」(『計量国語学』11—8、昭54・3)は、60の形容詞からなるチェックリストによって動詞の意味特性について因子分析法を用いて分析した、新しい意味の計量的研究である。この他、因子分析法を用いた分析に、中野洋「擬声語・擬態語のイメージ——意味微分法による分析——」(『計量国語学』11—7、8、昭53・12、昭54・3)、石川清治・本永守靖・大城宜武・東江平之「沖縄の児童の言語能力の分析的研究(2)(3)」(『琉球大学教育学部紀要』22、昭53・12)、関根正俊・服部桂・前東彬・田中末雄「因子分析法による万葉集の音形上の特徴」(『電気通信学会オートマトンと言語研究会資料』78—19、昭53・6)がある。

横山晶一・板橋秀一「弁別的特徴と2次系モデルに基づいた日本語単語間の距離」(『計量国語学』11—4、昭53・3)は、単語の連続パラメータの上での比較について述べている。

言語研究のための電子計算機

計算機は言語情報処理のためだけでなく、言語研究のための道具として用いることができる。実際、そのような使われ方が国語研究所ではなされてきたし、故岩淵悦太郎先生もそのような使われ方を期待されて計算機を導入されたのである。この経過は石綿敏雄「日本のコンピュータ言語学と岩淵悦太郎」(『言語生活』324、昭53・12)にくわしい。荻野綱男「コンピュータと言語」(『言語』8—4、昭54・4)では、計算機を言語研究のための道具として使うための第一段階は文科系の人達の認識、第二段階は計算機処理の普及、第三段階は文学部に計算機を設置することといっている。すでに、個人でマイクロコンピュータを持って研究に活用している人がいる。マイコンは大型機に接続できるから、かなりのことが可能になるだろう。いまや、国語学界全体に言語研究の道具として計算機を普及させる環境はととのったといえる。

水谷静夫「連糸操作プログラミングの誘ひ」(『東京女子大学論集』28—2、昭53)、同「俳句を作る計算機」(『東京女子大学日本文学』52、昭54・9)で、ソフトウェアの設計そのものが俳句理論や言語理論の達成水準を客観的に評価できる道具となり得ること、また、同「国語学の授業と計算機」(国語学会昭53年春季大会)で、実験的要素の持ち込み、算法の考え方の教育等にも計算機は利用できることを示し、石田敏子「コンピュータを利用した漢字の個別学習」(『日本語教育』36、昭54・2)では教育機器としての使い方が示され、研究そのものや教育にも有用であることが述べられた。

大型計算機センターには統計処理プログラムパッケージがある。

それらを利用する前に言語データの整理・集計プログラムが必要である。荻野綱男「コンピュータ言語地理学――言語分析の新手法――」(『言語研究』74、昭53・10)はプログラムGLAPSの説明とその分析例を示している。GLAPSはすでに各種の方言研究に利用され成果を上げている。長尾真・池田義雄「特殊文字・言語地図の入出力システム」(『情報処理』19―6、昭53・6)も言語地理学のための道具となるプログラムシステムである。中野洋「言語処理における一貫処理法の研究」(『国研報告』61、昭53・3)は分かち書きされていない漢字かな混り文に単位切り・よみがなつけ・品詞つけ・分類語彙表の番号つけ・文脈つき索引作成を自動的におこない語彙調査用データを作成するプログラムであり、谷垣正道「小型計算機による語彙調査プログラム」(『釧路論集』11、昭54)ではプログラムリストが提示されている。田中卓史「表記変容のシミュレーションシステム」(『計量国語学』11―5、昭53・6)は漢字かな混り文における漢字の割合をいくつかのパラメータによって自動的に変化させるプログラムで、漢字使用の実験研究に用いることができる。

国文学の世界にも計算機利用がはかられつつある(田嶋一夫「国文学研究におけるコンピュータの活用――コンピュータ国文学提要への覚書――」(『文学・語学』80・81合併号、昭53・2))。

以上のプログラムの分析目的に合っていないデータを計算機と対話しながら作成・加工するためのプログラムが、荻野綱男「日本語テキストエディタ」(『情報処理』20―10、昭54・10)によって報告されている。各研究機関で開発された日本語研究の道具としてのプログラムおよび資料は、長尾真・辻井潤一「日本語分析資料およびツールの調査」(『情報処理』20―10、昭54・10)に紹介があり、利用を呼びかけている。実際、同じプログラムやデータを別々のところで作るのは無駄なことである。

計算機利用の一つの利点は、データのコピーが容易であることにある。この利点により、各種の分析、フォーマット変更、データ変換が可能になる。国語研究所で作成された新聞の語彙調査データ約三百万語のうち、約百万語が共同研究用にいくつかの研究機関に渡され、田中穂積・荻野孝野・荻野綱男『新編日本語品詞列集成右順篇(上・下)左順篇(上・下)』(電子技術総合研究所、昭55・3、1)などが作られた。電子技術総合研究所では三省堂新明解国語辞典を入力し、長尾真・辻井潤一・山上明・建部周二「国語辞書の記憶と日本語文の自動分割」(『情報処理』19―6、昭53・6)の成果を生んでいる。逆引き辞典の出力などは、このデータを用いれば簡単にできるだろう。

言語現象の統計的分析

はじめに、田中章夫『国語語彙論』(明治書院、昭53・2)が刊行された。現代日本語語彙についての統計的研究の成果が手際よくまとめられており、語彙の計量的研究を始める人のよき出発点となろう。

語彙の分析では、現代俳句の語彙について二つの報告があった。寿岳章子「現代俳句の語彙」(『論集日本文学 日本語』5、昭53・8)、佐伯哲夫「現代俳句の語種」(『関西大学国文学』56、昭54・12)である。前者には、頻度分布、高頻度語の分析、基礎語を抽出、後者には、語種の分布、俳句の文体差等の考察がある。

児童読物の語彙についても、二編の報告があった。野村雅昭・柳瀬智子「児童読物の語彙構造」(『計量国語学』12―2、昭54・9)

は、語種・品詞・度数分布等の語彙量の分析、低中高学年、フィクション・ノンフィクションの層間の類似度の計算、数量化理論Ⅳ類による分析を行なっている。江連隆「児童文学における色彩語」(『弘前大学教育学部紀要』39・昭53・2)は色彩語の使われ方と効果について論究している。

度数分布についての報告は、前述、寿岳、野村・柳瀬にもあるが、水谷静夫「昭和初期流行歌での度数分布――昭和初期流行歌の調査から2――」(『計量国語学』11―5、昭53・6)がある。これは水谷(昭50)を承けてデータを増し、調査範囲を拡げたものだ。語彙の計量的性質についての調査、分析はもっとなされてよい。

日本語教育のための語彙調査に、専門家(20人)判定方式により「分類語彙表」から基本語彙を抽出した、国立国語研究所日本語教育センター第一研究室『日本語教育基本語彙第一次集計資料(1)(2)』(国立国語研究所、昭53・3、8)、計算機を使用して初級日本語教科書五冊と『現代雑誌九十種』の調査の上位語とを比較した、中村妙子・堀口秀嗣・今田滋子「初級日本語教科書の語彙に関する一考察(Annual Reports (ICU) 3、昭53・3)がある。また、語彙の対照研究として、根本今朝男・浜本純逸・森田信義「日米ソ教科書語いの比較研究Ⅲ――入門期教科書における動詞・その他――」(『福岡教育大紀要』28、昭54)があった。

言語生活関係のものとしては、杉戸清樹「職場敬語の一実態――日立製作所での調査から――」(『言語生活』328、昭54・4)、藤田克彦「住宅地での敬語行動」(『言語生活』328、昭54・4)、森下喜一「十代のことば――特に敬語を中心に――」(『野州国文学』23、昭54・2)、語彙の位相に関して、飛田良文「明治初期東京人の階層と語種との関係――『安愚楽鍋』を中心として――」(『国研報告』62、昭53・3)があった。

文字・表記では、宮島達夫・高木翠「雑誌九十種資料の漢語表記」(『国研報告』62、昭53・3)、宮島達夫「新字体の画数」(『計量国語学』11―7、昭53・12)、伊藤菊子「形声文字と漢字指導」(『言語生活』326、昭54・2)、野村雅昭・伊藤菊子「漢字の表音度」(『計量国語学』11―7、昭53・12)があった。

野村雅昭「接辞性字音語基の性格」(『国研報告』61、昭53・3)では、接辞性字音語基全般を対象として、その種類・機能についてのべている。語彙の中で重要な役割を占める漢語の構成を明らかにするとともに、言語情報処理における複合語の分割にも重要な情報をもたらす論究である。

計算機の処理結果である文脈つき用語索引の分析で、「～な」と「～の」のゆれを扱った、鶴岡昭夫「『～な』と『～の』について――漱石と鷗外の場合――」(『国研報告』61、昭54・3)、あまり使い分けを意識されていない「へ」と「に」について、著者の出生地と使い分けとの関連についてのべた、鶴岡昭夫「近代口語文章における「へ」と「に」の地域差――「行く」と「来る」について――」(『中田祝夫博士功績記念国語論集』、昭54・2)、同じ意味の二つの言語の刺激に対する連想について分析した、芳賀純・サトー＝アメリア＝ルイサ・萩原裕子「一言語使用者の2言語連想について」(『計量国語学』12―3、昭54・12)があった。

最後に、統計的方法によった安本美典・本田正久「日本語の誕生」(大修館書店、昭53・11)についての、R・A・ミラーの書評(『言語生活』332、昭54・8)に発した各二回の論争(『言語生活』333～335、昭54・9～11)があった。はげしい論争から生まれたもの

は少なかった。もっと続ければ、あるいは争点がはっきりしたかも知れない。書評者に統計的処理に対する基本的な不信があったのではなかろうか。

言語情報処理の研究

計算機による日本語文の出力は高速漢字プリンタの開発によってほぼ解決した。入力法については、音声による入力法、印刷漢字の光学読取り、手書き文字の入力法、カタカナによる入力と漢字変換・編集についての研究が盛んである。高橋延匡「日本語入出力装置」(『情報処理』20—10、昭54・10)、中川聖一「機械による会話音声の認識・理解研究の動向」(『情報処理』19—7、昭53・7)、安田道夫・中野康明・藤田好司「手書き文字の認識技術」(『電気通信学会誌』61—2、昭53・2)、斎藤秀紀「多目的漢字入力システムの試案」(『国研報告』61、昭53・3)に詳しい。

西村恕彦・水谷静夫・尾上圭介・田中幸子「日本語基本文法　単文篇・複文篇」(『電総研報告』783、784、昭53・3)は、計算機による構文解析システム《やまと》上で働くための文法規則の組を提案している。日本語についての機械処理用の完全な文法規則集としては最初のものであろう。機械上で動くということは、前述したが、その組自体に重なり・矛盾・抜け落ちを含んでいないことを示す。著者の努力に深い敬意を表するものである。桃内佳雄「手続記述文章における動作の制御に対する言語表現」(『計量国語学』12—1、昭54・6)は、料理の文章を資料に、動作の制御構造と言語表現の対応関係について考察している。

日本語を機械処理するためには、処理法の開発とともに、日本語そのものの研究が進められなければならない。もちろん、石綿敏雄「コンピュータと言語」(『岩波講座日本語』別巻、昭53・3)のいうように、「言語情報処理の研究は工学者の責任だけではなく、それに利用できる、より精密化、有機化した言語構造の記述は言語研究者の本来の任務である。」しかし、ここでは主に工学者の日本語研究の代表的なものだけについて紹介する。ただし、*印をつけた人は言語学側の人である。これにより何がわれわれに求められているかがわかるだろう。

カナと漢字の変換については、*石綿敏雄「〝カナ漢字情報処理〟の問題点——自然語処理とカナ漢字変換——」(『情報科学』141、昭53・8)、森健一・河田勉「かな漢字変換」(『情報処理』20—10、昭54・10)、荒木啓介・板山和彦「JICSTの実用的全自動漢字—カナ変換システム、K—KACSについて、(『情報処理』20—10、昭54・10)坂本義行「日本語の点字情報処理におけるカナ漢字変換」(『計量国語学』12—3、昭54・12)、単語分割については、坂本義行「文節単位の自動分割法——字種と平仮名連系による——」(『計量国語学』11—6、昭53・9)、日本語文の述部の分析について、石原好宏・田町常夫「日本語終端表現の生成のための言語情報の分析と分類」(『電子通信学会論文誌』J61D—9、昭53・9)、構文について、*柴谷方良「日本文法の仕組み」、佐藤泰介・田中穂積「日本語の構文解析」(以上二点、『情報処理』20—10、昭54・10)、意味について、田中穂積『計算機による自然言語の意味処理に関する研究』(『電総研報告』797、昭54・7)、*山梨正明「言語理論からみた意味表現」辻井潤一「計算機からみた意味表現」(以下二点、『情報処理』20—10、昭54・10)、知識の蓄積と利用について、田中穂積「言語理解システムと知識表現」(『数理科学』193、昭54・7)、田中卓史「人工知能のため

の言語分析――基礎的考察――」(『国研報告』61、昭53・3)があった。この時期の特色としては、実用からの要請によってカナ・漢字変換の研究が盛んだったこと、文章、談話構造、知識体系、意味処理等、これまでになく広く、深い範囲での研究・分析が活発だった。

以上の研究成果を総合的に利用したシステムとして、文章処理・情報検索・文献検索・機械翻訳等がある。文章の処理実験としては、岡本哲也・佐藤雅之「計算機による法令の一部改正の試み」(『計量国語学』12―2、昭54・9)があり、これは一定の形式と用語によって文書化されている法令のテキスト構造に注目し、その一部改正と改正経過の蓄積を実験した報告である。文献検索についてはドキュメンテーション関係の雑誌での報告が多い。高松忍・藤田米春・西田富士夫「係り受け関係に基づく文献の検索」(『情報処理』19―12、昭53・12)は構文情報を検索に利用している。情報検索では、藤崎哲之助・間下浩之・渋谷政昭「計算機の日本語理解：データベース照会システム『ヤチマタ』」(『計量国語学』11―8、昭54・3)、絹川博之「情報検索のための日本語解析」(『情報処理』20―10、昭54・10)がある。これらは語彙、構文、意味の処理を行ない目的をはたそうとしている。機械翻訳の将来については、「やや肯定的」(*石綿敏雄「機械翻訳の再評価」(『言語』8―1、昭54・1)に変わっているらしく、日本でも長尾真・辻井潤一・建部周二「技術論文表題の英和自動翻訳の試み」(『情報処理学会CL研究会資料』19―2、昭54・9)の報告がある。ここでも形態素・構文・意味の解析と合成をおこなっている。

これらの成果は、徐々に、しかも確実に、我々の生活・研究に影響を与えるだろう。よりよい日本語処理システムができるようにするために、我々はこの分野に対してもいろいろなレベルでの、いろいろな発言をしなければならない。

与えられた頁数はすでに尽きている。取り上げなかった論文は数多い。また、当然読んでいるべくして見落した論文も多いかと思う。著者の真意を理解していないこともあるかもしれない。お許しいただきたい。

〈補足〉NHK『文研月報』中の文献の紹介をおとしていたので補足する。お許し頂きたい。

〈統計的研究のための分析方法〉にはいるものとして、菅野謙「テレビニュースの主要表現連鎖500種」(343、昭54・12)がある。これは「新聞」と「NC9」の文章を「表現連鎖」の使用量という点で比較し、テレビニュースの文体を明らかにしようとする研究である。ここで「表現連鎖」とは、学校文法でいう「単語」とほぼ等しい単位を最小の表現単位として扱い、二個以上の表現単位が結合したものをいう。「こと　に」「し　て　いる」などがそれである。このような「連鎖」による研究は、これまでにない調査単位や分析方法を生むことが多い。菅野は文体研究の新方法を提出している。

〈言語現象の統計的分析〉に、石野博史「テレビニュースの用語10年間の変化(1)―外来語」・菅野謙「同(2)―サ変動詞」・「同(3)―文末表現」(以上三点、338、339、340、昭54・7、8、9)、菅野謙「テレビニュースのタイトル表現」(330、昭53・11)、同「漢音・呉音・慣用音の発音のゆれ(1)、(2)」(341、342、昭54・10、11)がある。

――国立国語研究所員――