

In France, corpus is a collection of entire texts on a particular research subject. The major tasks of linguists of French corpus studies are the construction of the corpus and the interpretation of it. In this paper, two characteristic corpus studies in French tradition will be examined: lexicology, especially lexicostatistics on the one hand, and political discourse analysis on the other. Corpus linguistics is a study of the “norms” of various genres by analyzing appropriate corpus. This approach contributes to general linguistics by narrowing the gap between a linguistic system and usage.

## 0. はじめに

フランス語の Corpus は、「資料体」という意味で、人文科学、社会科学、哲学、法学などの分野で用いられてきた伝統的な用語である。現代の言語科学においても、コーパスには必ずしも電子化された巨大なデータというイメージが伴うわけでないこと、またコーパスの構築・選択、およびコーパスから得られたデータの解釈に重大な関心が払われることは、この伝統に由来している。本稿では、英語のコーパス研究と対照させる意図のもとに、フランスにおけるコーパス研究の特徴と、代表的な二つのコーパス研究である語彙研究と政治ディスコース研究を、その歴史を踏まえて紹介し、最後にフランスにおけるようなコーパスの扱い方が言語学一般にもたらす貢献についてまとめる。研究史を一覧するために、末尾には年表を掲げた。

## 1. フランスにおけるコーパス研究の特徴

フランスにおけるコーパス研究の特徴を、英語圏で行われているものと対比して強調するならば、次の4点を挙げることができる。第一に「コーパス」の定義の違い、第二に意味の問題への傾斜、第三に語彙の計量研究が伝統的に盛んであること、第四にテキストやディスコース研究との親和性である。フランス



的に記述する」という考え方には懐疑的である。

### 1.2. 意味論への傾倒

コーパスを用いた研究の主目的が意味の生成への関心であるということも、フランスの言語科学におけるコーパス研究の特徴である。フランスの言語学一般の伝統として、統語論や形態論など言語の表現に関する研究に比べて、言語の内容である意味への関心が高いということがある。語の *signification* (=辞書的意味) は *langue* において規定されるが、*sens* (=意味) と *valeur* (=価値) はテキストとコンテキストから生成されると考えられているので、意味の研究の対象はテキストの集合、すなわちコーパスということになる。テキストは言語学の最小単位であり、一つずつのテキストに対してコーパスはコンテキストとなる。コーパスはテキストの単なる寄せ集めであってはならず、テキスト間の関係が重視される。

### 1.3. 語彙計量学 (*lexicométrie*): 情報工学との近接性

フランスは、語彙の計量的研究が盛んである。それは、1950-60年代に欧米・ロシア・日本において始まった機械翻訳の研究にさかのぼり、言語学と情報工学の親密な関係が当初から今日に至るまで続いている。この点は、とくにアメリカにおける状況と異なっていると Léon(2001)はいう。1950-60年代のアメリカには機械翻訳に興味を持つ言語学者は存在せず、経験的・統計学的研究方法は好まれなかった。また、当時のアメリカの言語学では二言語間の翻訳研究という観点が希薄であったため、言語分析の単位として、単一の言語の中で形式的に確定可能な「形態素」が選ばれ、2言語間の辞書の編纂において必須の単位となる「語」には注意が払われなかった。一方フランスでは、当初より多くの著名な言語学者 (Cohen、Culioli、Benveniste、Gougenheim など) が言語の自動処理の研究に多かれ少なかれ参加した。フランスで盛んだったのは、英国やアメリカとは異なり、語彙表の自動生成であり、語彙研究の機械化であった。それに伴い、分析の単位としての「語」が議論され、イディオムなど、語の認定に関する研究が始まった。同時に、統計学を文体論、文献学、方言学、言語教育に応用する研究も行われた。

#### 1.4. ディスクール分析との関係

フランスのコーパス研究はさらにディスクール分析 (*Analyse du discours*)との繋がりがある。上に述べたように「コーパス」はテキストの集合と定義されるのでこの連携に不思議はないが、フランスのディスクール分析がイデオロギーとコンピュータ解析とに強い親和性がある点は注目すべきである。フランスの「ディスクール分析」の創始者といわれるペシュュー (M. Pêcheux)は 1969 年に *Analyse automatique du discours* [ディスクールの自動分析]を公表した。彼のディスクール論は、ディスクールはイデオロギーの具体化であるという、マルクス主義者であったアルチュセールの流れを引いている。ディスクールの分析にあたり、ペシュューは、分析方法自体がもつイデオロギー性を排除するためには経験論的・実証的方法が必要であるとして、情報工学・数学の専門家の協力のもとにコンピュータを使った分析方法を提案した<sup>1</sup>。それが「ディスクールの自動分析」である。今も、フランスのコーパス研究には、ペシュューの創始したディスクール分析の方法論が影響を及ぼしている。また、コーパス研究と政治との強いつながりもここに起源がある。

## 2. 二つの代表的な研究

以下では、フランスで行われている代表的なコーパス研究である語彙研究と政治ディスクール研究を、その歴史とともに紹介する。この二つの研究のジャンルは対照的ではあるが、意味に関心の中心がおかれたものであるという点は共通している。

### 2.1. 語彙論・語彙意味論

フランスの語彙研究の最も重要な成果は、辞書 *Trésor de la langue française* (フランス語宝典<sup>2</sup>、以下 TLF と略記)と、その編纂のためのデータベースの *Frantext* である。また、重要な役割を果たしてきたのは、語彙計量学 (*lexicométrie*)である。既述のとおり、フランスの言語学の関心の中心は常に意味の問題であったが、語彙論はかつて意味論と同義であった。すなわち、意味論とは語彙の研究であり、語彙論とは語の意味の研究であった。

### 2.1.1 Frantext と TLF

TLF 編纂のプロジェクトが始まったのは 1957 年のことである。Lexicologie et lexicographie française et romanes : orientations et exigences actuelles [フランス語とロマンス諸語の語彙論と語彙記述：現代の動向と課題] という名の会議がストラスブールで開かれ、Oxford English Dictionary に匹敵するような辞書を、新しいテクノロジーを使った国家プロジェクトとして編纂することが決議された (cf. Antoine, G. & B. Cerquiglini, 2000 : 969)。この時期はコーパス研究の歴史においてきわめて重要な時期である。年表に掲げたように、1950-60 年代には、今も積極的な活動を続ける研究所、雑誌、学会が続々と設立された。TLF のための研究所は 1960 年にナンシーに作られ、1963 年には大型コンピュータが稼動を始めて、のちに Frantext となるテキストデータベースが執筆者のために整備された。紙媒体による出版は 1971 年からアルファベット順に始まり、1994 年に完成した。その後、電子版の *Trésor de la langue française informatisé* (TLFi) が 2001 年にインターネット上に全面無料公開された。Frantext は 1992 年以來、大学や研究機関を対象に有料で公開されており、言語研究をはじめとするさまざまな研究のために利用されている。

#### 2.1.1.1. Frantext <http://atilf.atilf.fr/frantext.htm>

Frantext は、16 世紀から 20 世紀の間の 4000 件以上の作品からなる巨大なデータベースである<sup>3</sup>。テキストは適宜追加され、規模は拡大し続けている。総語数は 2 億 1000 万語を越える。コーパスは使用者が研究目的に合わせて自分で構成することになっているので、コーパス言語学でいう「代表性」は考慮されていない。文学作品が 80%、学術書が 20% という編成からわかるように、フランスを「代表する」評価の高い書かれたテキストのみからなり、メディアのテキストや商業的印刷物、話し言葉は含まれていない。Frantext には二つのバージョンがある。一つは品詞解析が施されたテキストの集合で、19 世紀から 20 世紀の 1940 作品からなる。もう一つは品詞解析されていないもので、Frantext の全体に対応する。アクセスはブラウザ経由で行い、検索は Frantext 独自の検索プログラムによる。検索結果は文脈 (400 字) を含めて一括してブラウザからダ

ウンロード可能である。

#### 2.1.1.2. TLFi

TLF は 19-20 世紀のフランス語の辞書であり、16 巻と補遺 1 巻からなる。見出し語は 10 万語、27 万件の語義、43 万の例文 (Frantext からの引用がほとんど) からなる。単語の語義分析、語源、歴史、Frantext での頻度などが充実していて、それぞれの単語に関して通時的にも共時的にも十分な情報を得ることができる。語の意味の構造的部分(signification)とそこから派生する多様性(sens)のバランスが考慮されているといわれる。TLFi は TLF の電子版であり、360 万個の XML タグが施され、辞書に記載された情報に関してさまざまなタイプの検索が可能のように工夫されている。素晴らしい点は、ハイパーリンクによって、Frantext やアカデミーフランセーズの辞書などの外部のデータに、直接リンクしていることである。ハイパーリンクをたどることによって、多くの例とそのコンテキストにシームレスでアクセスが可能である。

#### 2.1.1.3. 語彙研究のための組織

フランスの研究所の多くは文系・理系を問わず全ての分野にわたって、CNRS(=フランス国立科学研究センター)が束ねる国立の組織である。研究所間の関係は競争的ではなく協力的である。

- Analyse et Traitement Informatique de la Langue Française – Computer Processing and analysis of the French Language, <http://www.atilf.fr>

「TLF のための研究センター」は、「国立フランス語研究所」を経て「フランス語の分析と情報処理研究所」と名前を変えたが、フランス語の語彙研究の中心的存在として、種々の研究を行っている。

- Centre National de Ressources Textuelles et Lexicales, <http://www.cnrtl.fr/>

Frantext のうち著作権が消滅している 500 のテキスト (18-20 世紀) をこのサイトからダウンロード可能である。また、それ以外のリンクも充実している。

- Lexique 3, <http://lexique.org/>

135000 語に関して Frantext における頻度が調査されている。

#### 2.1.1.4. Frantext を使った言語研究

Frantext は現代フランス語の共時的研究を行うための理想的なコーパスではないのは明らかである。直近の時代区分のテキストが少ないこと、ジャンルが文学に偏っていることは欠点である。しかし、通辞的な研究のためのコーパスとしては理想的といってよい。以下に二つの研究をあげる。

- Schenedecker C. (2007) “fichu, foutu et maudit, dammé : de drôles d’adjectifs !” *Colloque « Adjectifs »* 悪態語が形容詞として用いられていく通時的過程 (19-20 世紀) と、現代における語の意味および形態統語的共時的価値の対応を検証したもの。Frantext は、第 1 に用例検索のためのアーカイブ、第 2 に 19・20 世紀の変化を見るためのコーパスとして用いられている。
- Fujimura, I, M.Uchida, & H.Nakao (2004) “De vs des devant les noms précédés d’épithète en français : le problème de *petit*” *Le Poids des mots* vol 1, Presses Universitaires de Louvain, p.456-467. 複数形容詞 + 複数名詞の前の複数不定冠詞の *des* と *de* の交代の問題を扱う。Frantext から、ジャンル別 (学術書、小説)、時代別のコーパスを構成し、ジャンルと時代に関わらず、同一の原理による選択が行われていることを計量的に示した。

#### 2.1.2. Lexicométrie 語彙計量学

TLF のプロジェクトが開始した 1950-60 年代は、フランスの語彙計量学の開始の時代である。この当時に考案され、今なお世界のコーパス研究において用いられている二つの計量的方法を以下に挙げる。

まず、Guiraud Index<sup>4</sup>である。Guiraud Index はテキストの語彙の多様性を計る指標の一つであるが、TTR と比べて、テキストのサイズによるぶれが少ないという点が評価され、今、注目されている。

Guiraud Index :  $\text{Type} \div (\text{Token の平方根})$  (cf. TTR :  $\text{Type} \div \text{Token}$ )

ギローはフランスの著名な言語学者であり、その著作は、クセジュ文庫のシリーズの中に『意味論』、『文法』、『フランス詩法』、『記号学：意味作用とコミュニケーション』などとして日本に紹介されているが、最も重要な二つの業績は知られていない。一つは上記の言語統計学の業績であり、Guiraud (1954)において発表された。もうひとつは、1978 年に初版が発行された *Dictionnaire érotique*

[エロティックな辞書]である。これは中世から現代までの 7000 の性的な語と表現の辞書であり、600 ページを超える大著である。まさに語彙研究の成果であるといえる。

フランスに起源のある語彙計量学の貢献として忘れてならないもう一つの業績は、Benzécri の対応分析（コレスポンド分析）である。対応分析も Guiraud Index と同時代の 1962 年ごろに生まれた。フランスではよく使われてきた方法である。

## 2.2. 政治的ディスコース研究：言語学の一分野

フランスに特徴的な 2 番目のコーパス研究として、政治ディスコースの研究を挙げる<sup>5</sup>。フランスのディスコース分析の創始者のペシュは 1967 に研究所 *Lexicométrie et textes politiques*[語彙計量学と政治的テキスト]を ENS Fontenay/Saint-Cloud（高等師範学校フォントネー・サンクルー校）に設立し、以来、ここがフランスの政治的ディスコース分析の中心となる。計量的方法と政治性がフランスのディスコース分析一般の特徴であるといわれている。

第一人者であるマンガノー（Mangueneau）は次のように、コーパスとコンピュータを使った計量的方法の、フランスにおける意味を語っている。

In France, discourse analysis has always been in a symbiotic relationship with information technology. Today, discourse analysts dispose of a whole array of programs in order to scan corpuses, test hypotheses or construct new ones. The constant growth of all sorts (data, bibliographies, software) on the Internet considerably changes the way research is done. Those working on archives can claim exhaustiveness. Thus, notions like “representative sample” get profoundly transformed, and discourse analysis participates in the remarkable development of “corpus linguistics”, which investigates enormous masses of data. (Mangueneau, D. & J. Angermüller (2007))

政治的ディスコースの研究は典型的なコーパス研究である。なぜなら、政治的ディスコース分析の研究対象は完結したテキストの集合としてのコーパスだからである。しかし、ディスコースは実際の社会の中に存在して初めて意味をも

つので、現実の社会もコンテキストして分析の対象に入る。したがって、政治ディスカールの分析は、言語学と社会科学の交差する場所に位置する応用科学である。

政治と日常の乖離が少ないフランスの社会においては、一般の人々は政治的ディスコースの分析に関心がある。インターネット上には、政治家のディスカールを論じたサイトが山のようにあり<sup>6</sup>、2007年5月の大統領選挙中には、サルコジ候補とロワイヤル候補のディスカールが種々の分析の対象となった。たとえば、サルコジは *identité* (=identity), ロワイヤルは *sécurité* (=security) がキーワードであり、3-gram としては、サルコジは “je veux que (I want that)”, “je veux être (I want to be)”, “parce que je (because I)”, “je veux dire (I want to say)” が多く、ロワイヤルは “une France qui (France which)”, “la lutte contre (the fight against)”, “je vous propose (I propose you)” が多かった<sup>7</sup>。自己アピールの強さにおいて、二人の間に差があったことがわかる。

以下ではこの分野の研究活動や成果を挙げる。

- 学術誌 *Mots: les langages du politique* [語：政治的な種々の言語活動]は、1980年に政治的語彙の計量的研究のために創刊された雑誌。
- 雑誌の特集号 *Corpus No4, 2005*, 特集 *Les corpus politiques : objet, méthode et contenu*, [政治コーパス：対象、方法、内容]  
<http://corpus.revues.org/document292.html>。
- 書籍 *Le discours gouvernemental. Canada, Québec, France (1945-2000)* [政府のディスカール。カナダ、ケベック、フランス (1945-2000)]. Labbé, D.& D. Monière. (2003). Paris :Champion.  
カナダ政府、ケベック州政府、フランス政府による所信表明演説の計量的比較分析。55年間の演説が対象となっている。
- 書籍 *Paroles de président. Jacques Chirac (1995-2003) et le discours présidentiel sous la V<sup>e</sup> République* [大統領のことば：ジャック・シラク (1995-2003)と第五共和制大統領のディスカール], Mayaffre, D. 2004. Paris : Champion.
- ウェブサイト PoliText : Base de données de discours politiques française

(1789-2002) [政治テキスト：フランスの政治ディスコースデータベース (1789-2002)] <http://www.unice.fr/ILF-CNRS/politext/>. 政治家のディスコースのデータベース。約 200 万語、600 件からなり、議会での演説、テレビやラジオのインタビュー、講演、雑誌でのインタビュー記事などを含む。

● 書籍 Encrevé, P., *La liaison avec et sans enchaînement*, Seuil, 1988.

Encrevé はシラクやミッテランなどの政治家の演説を音声面から研究して、アンシェヌマン抜きのリエゾンという現象の政治的意味を分析した<sup>8</sup>。

3. ラングの言語学とパロールの言語学を結ぶものとしてのコーパスの価値  
フランスにおけるコーパス研究の特徴を再度まとめると以下ようになる。

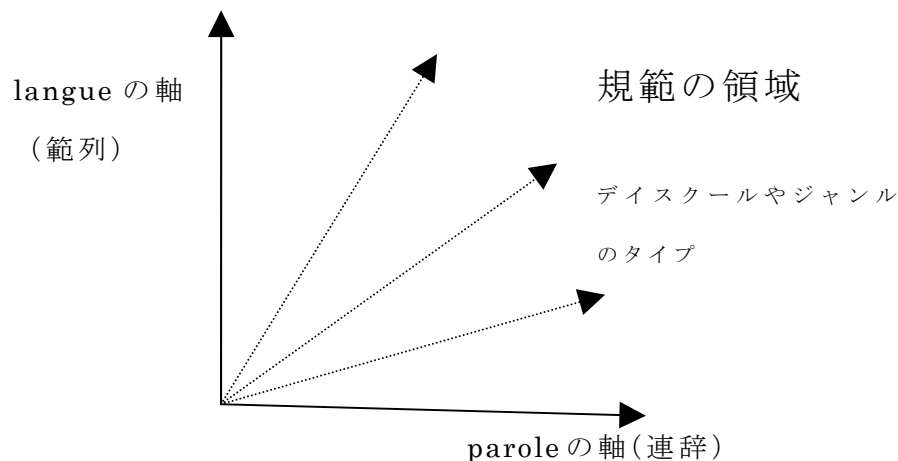
1. コーパス言語学 (linguistique de corpus) とは、テキストの集合としてのコーパスの言語的研究であるということ。「言語的」という語は広くとらえられ、ディスコース分析や文学研究も含まれる。コーパスはツールではなく、コーパスそれ自体が研究対象である。
2. 意味論研究が中心にあること。テキストやディスコースの意味がどのようにして生成するかという問題が言語学の関心の中心であり、テキストの構成単位である語の意味とテキストにおける分布が研究されてきた。
3. 計量的・機械的研究方法に対するアレルギーの不在とフランス的難解さの融和。人文系の研究者に自然科学的方法に対する不信感が少ないことは注目すべきである。ペシュエの『ディスコースの自動分析』において、言語学者の仕事は、第一にコーパスの構築、第二に結果の解釈であるとされている。この二つの間にデータの処理と解析があるが、それは言語学者の仕事ではなく、自然科学者の仕事である。言語学者自身が情報工学や統計学の知識を持つことは少なく、分業体制になっている。データの解析のみには価値が認められず、準備と解釈を入念を行わねばならないという点は見習うべきであろう。一方で、解釈が飛躍につながる場合もあり、それは問題点である。

最後に、これまでに述べてきたフランス的文脈において、コーパス研究は一般

言語学に対してどのような貢献ができると考えられるかを Rastier (2002)から引用しつつ検討する。

Rastier によると、コーパスはラングとパロールをつなぐものとしての規範(Norms)の観察を可能にする<sup>9</sup>。コーパス研究は実際の言語使用をデータとするので、パロールが研究の対象とする。しかし、パロールは無限で雑多なものであり、パロールをいくら多量に集めて分析しても、ラングとはどのようなものか、また、個別の言語はどのような言語かということに解答を与えることはできない。一方で、ラングとパロールの区別は、理論的には十分に認知されているにせよ、その二つがどのように接合するのかという問題は理論的にも実証的にも未解決のままである。ソシュールのラングとパロールの対立は、アリストテレスの可能態と現実態、フンボルトのエネルギアとエルゴン、チョムスキーの *competence* と *performance* に対応すると考えられるが、どのように言い換えたところで未解決である点に変わりはない。

コーパス言語学は、体系（ラングに相当）とパロールの間にコセリウ (E. Coseriu) が設定した、規範を対象にする言語学であると Rastier はいう。規範の言語学は、ラングの言語学とパロールの言語学の間であり、規範の領域は、以下の図のようにラングとパロールの軸の狭間にある (Rastier 2002)。規範はラングによって定められるのではなく、テキストの属する分野やジャンルごとに定められている。したがって、規範は社会的な存在である。たとえば、文学テキストと法学のテキストの規範は異なるし、同じ文学テキストでも、小説と演劇の規範は同じでない。また、紀行文と推理小説の規範も同じではない。しかし、ラングは規範と無関係であるわけではない。ラングの規則は規範が凝結したものである。また、パロールの実践に規範性があることは明らかである。パロールは、ラングの規則と、社会的規範に順応して実践される。



図：「規範」の言語学の領域<sup>10</sup>

Rastier は、書かれたテキストだけを取り上げているが、テキストジャンルに代えて地理的方言や社会的方言、時間軸などを規範の領域に置き、それぞれの規範を研究対象とする学問としてコーパス言語学をとらえるなら、コーパス言語学を一般言語学の全体的な枠組みの中に位置づけるのが容易になる。そのとき、「規範」は「規則性」ということばで言い換えることも可能と思われる。

#### 【参考文献】

- コセリウ、エウジェニオ (1981)『うつりゆくこそことばなれ：サンクロニー・ディアクロニー・ヒストリア』クロノス社。
- 藤村逸子・内田充美(2003)「情報ファイル：オンラインデータベース-Frantext, WordbanksOnline, TLFi-」『フランス語学研究』37: pp.87-93.  
(URL : <http://www.gsid.nagoya-u.ac.jp/fujimura/gyoseki/johofile.pdf>)
- Antoine, G. & B. Cerquiglini, (2000) *Histoire de la langue française, : 1945-2000*, CNRS Éditions.
- Benzécri, J. & F. Benzécri (1981) *Pratique de l'analyse des données*. Paris : Dunod.
- Condamines, A. (dir.) (2005) *Sémantique et corpus*. Paris : Lavoisier.
- Fujimura, I, M.Uchida & H. Nakao (2004) “De vs des devant les noms précédés d'épithète en français : le problème de *petit*” *Le Poids des mots* vol 1, Louvain : Presses Universitaires de Louvain. pp. 456-467.

- Guiraud, P. (1954) *Les caractères statistiques du vocabulaire*. Paris : P.U.F.
- Guiraud, P.(1978) *Dictionnaire érotique (Dictionnaire historique, stylistique, rhétorique, étymologique, de la littérature érotique)*. Paris : Payot.
- Habert, B., A. Nazarenko & A. Salem (1997) *Les linguistiques de corpus*. Paris : Almand Colin.
- Hak, T. & N. Helsen (eds) (1995) *Michel Pêcheux. Automatic Discourse Analysis*. Amsterdam : Rodopi.
- Léon, J. (2001) “Conceptions du ‘mot’ et débuts de la traduction automatique,” *Histoire Épistémologie. Langage* 23-1; 81-105.
- Mangueneau, D. & J. Angermüller (2007) "Discourse Analysis in France. A Conversation." *Forum qualitative Sozialforschung/ Forum : Qualitative Social Research*, 8(2)Art.21, [Online].  
URL: <http://www.qualitative-research.net/fqs-texte/2-07/07-2-21-e.htm>.
- Rastier, F. (2002) Enjeux épistémologiques de la linguistique de corpus, in G. Williams (éd.) (2005) *La linguistique de corpus*. Rennes : PUR, pp. 31-45. [En ligne sur *Texto !* (URL : [http://www.revue-texto.net/Inedits/Rastier/Rastier\\_Enjeux.html](http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html))].

#### 【年表】

- 1954 : 機械を使った語彙研究、自動翻訳研究の開始。アメリカ、イギリスより 10 年ほど遅れる。
- 1954 : Guiraud P. *Les caractères statistiques du vocabulaire* [ギロー、『語彙の統計的特徴』]の出版。
- 1957 : 会議 *Lexicologie et lexicographie française et romanes : orientations et exigences actuelles* [フランス語とロマンス諸語の語彙論と語彙記述 : 現代の動向と課題] の開催。
- 1959 : 雑誌 *Cahiers de lexicologie* [語彙論ノート]の創刊。
- 1959 : 研究所 *Laboratoire d'analyse lexicologique* [語彙分析研究所] がブザンソン大学に創設され、語彙の計量研究の中心地となる。
- 1959 : 学会 *Association pour l'étude et le développement de la Traduction*

Automatique et de la Linguistique Appliquée (ATALA) [自動翻訳と応用言語学の研究と開発のための学会]の設立。

1960 : 研究所 Centre de recherche pour un Trésor de la Langue Française [TLF のための研究センター] がナンシーに設立。

1960 : 雑誌 *TAL* (*Traitement automatique des langues*) [言語の自動処理]の創刊。

1962 : 雑誌 *Etudes de Linguistique Appliquée* [応用言語学研究]創刊。

1963 : ナンシーの研究所においてコンピュータ Gamma 60 Bull の使用開始。

Trésor de la langue française 編纂のためのデータベースの構築と処理の開始。

1964 : 会議 Statistique et analyse linguistique [統計学と言語分析]。多数の重要な言語学者が参加。

1967 : 研究所 Lexicométrie et textes politiques [語彙計量学と政治的テキスト]の創設。

1969 : Pêcheux, M., *Analyse automatique du discours*, Dunod [ペシュュー、『ディスクールの自動分析』]の出版。

1969 : Faucault, M., *L'Archéologie du savoir*, [フーコー『知の考古学』]の出版。

1971-1993 : *Trésor de la langue française* の紙媒体による出版。

1977 : Nancy の研究所が改名 Institut national de la langue française [国立フランス語研究所]。

1996 : データベース Frantext がインターネット上で有料公開開始。

2001 : オンライン辞書 *Trésor de la langue française informatisé* の完成。インターネット上で無料公開。

2001 : Nancy の研究所が再び改名 : *Analyse et traitement informatique de la langue française* [フランス語の分析と情報処理]

2005 : 研究所 Centre National de Ressources Textuelles et Lexicales [テキストと語彙データ国立研究所]の創設。Frantext のうち著作権が消滅したテキスト 500 を無料公開。

現在 : 中世フランス語の電子辞書のプロジェクトなど

---

<sup>1</sup> Harris, Z. (1952) “Discourse analysis”, *Language*, 28, 1-30 から大きな影響を受けたといわれている。Hak, T. & N. Hulsloot (eds) (1995)において英訳を読むことができる。

<sup>2</sup> この名称は 17 世紀のフランス語辞書 : *Trésor de la langue françoise, tant ancienne que moderne*, Jean Nicot (1606) が起源である。

<sup>3</sup> Frantext の詳細については、藤村、内田 (2003)を参照されたい。最近では、個人単位のパスワード方式の利用も可能である。

<sup>4</sup> Guiraud Index というのは後になって付けられた名称であり、本人がそのように呼んだわけではない。

<sup>5</sup> ディスクール (=discours)という語のフランス語における基本的意味は「演説、スピーチ」である。ディスクールには、英語の discourse に対応する言語学の用語としての意味も存在し、ディスクール分析は「演説分析」に限られるというわけでは決してないが、「ディスクール」と「政治」の深い関係は語源にさかのぼると思われる。

<sup>6</sup> たとえばパリ第 3 大学の TAL のサイトを参照のこと。

<http://tal.univ-paris3.fr/blogtal/>

<sup>7</sup> <http://members.unine.ch/jacques.savy/Papers/Royal-Sarkozy.html>

<sup>8</sup> アンシェヌマンのないリエゾンとは、例えば Il faut en être (そうでなければならぬ) の場合に、faut の語末子音を発音するにも関わらず次の母音との間にアンシェヌマンが起こらず [ilfotʔə̃nɛtʁ] のように声門閉鎖が挿入されることをいう。Encrevé は Chomsky & Halle の *The Sound Pattern of English* を仏訳した音韻論学者である。

<sup>9</sup> コセリウ (1981)参照。

<sup>10</sup> Rastier (2002)より。