

パリ第 13 大学 Aude GREZKA 氏講演の記録¹

オード・グレッツカ

GREZKA Aude

高田 晴夫

TAKADA Hareo

木島 愛

KIJIMA Ai

本稿は、高田晴夫（新潟大学名誉教授）と木島愛（現千葉工業大学准教授、当時新潟大学非常勤講師）の協力により実現した講演の記録である。講演の日時、会場、発表者、講演タイトル、通訳は以下の通りである。当日フランス語で講演が行われたので、以下では、フランス語と日本語の対訳の形で講演内容を公表する。また講演はパワーポイントを使用して行われたが、スライド数の多さ、本稿に割り当てられた紙数制限などの理由から、スライドの掲載は割愛せざるをえなかった。

日時：2017年3月6日（月） 13:00 - 15:00

会場：新潟大学 総合教育研究棟 B351 教室

発表者：Aude GREZKA (パリ第 13 大学 現 LIPN 研究所研究エンジニア、当時 LDI 研究所 語彙部門統括責任者・研究エンジニア、フランス国立科学研究所センター研究員)

講演タイトル：“Élaboration d’outils méthodologiques pour décrire la langue”

「言語記述のための方法論的ツールの開発」

通訳：高田晴夫

D1.² Élaboration d’outils méthodologiques pour décrire la langue	言語記述のための方法論的ツールの開発
D2. La méthode de travail que nous développons au laboratoire part du lexique pour élaborer ou analyser des dictionnaires en utilisant l’informatique et en vue du traitement automatique. Le lexique est le matériau de base qui sert d’entrée à la description des langues	私たちが研究所において展開している研究は、情報科学を利用して言語の自動処理を目指すために、辞書の開発・分析を行うもので、語彙から出発するという方法です。語彙は、言語記述のための入り口としての役に立つものなのです。

D3. Les dictionnaires sont à la fois une source de données et un objet construit	辞書はデータの源であると同時に、新たに構築されるものです。
D4. L'informatique est un outil mis à la disposition des linguistes pour : • analyser/élaborer des dictionnaires • concevoir des méthodologies propres au TAL	情報科学は言語学者にとっても以下の点で利用価値のあるツールです。 • 辞書の分析と開発 • 自動言語処理の方法論の考案
D5. L'objectif est de fabriquer de la ressource linguistique automatisable qui va pouvoir être utilisée dans des applications informatiques : analyse automatique de textes ; aide à la rédaction ; aide à la traduction automatique ; analyse de discours ; etc.	研究所の目的は情報科学の応用として利用できる自動化可能な言語リソースの作成で、これは、テキスト自動分析、作文支援、自動翻訳機械、談話分析などにおいて利用できます。
D6. DIFFICULTÉS	問題点
D7. Cependant le traitement de la langue est difficile car en langue naturelle, il y a un 3 grandes difficultés à gérer : l'inférence ; le figement ; la polysémie	しかしながら、言語処理は難しいものです。それは、自然言語には克服しなければならない、推論、凝結、多義性という3つの大きな問題があるからです。
D8. Le 1 ^{er} problème : l'inférence. C'est-à-dire que l'Information est activée bien qu'elle ne soit pas mentionnée explicitement. Dans l'exemple : <i>Léa gronde son enfant</i>	第1番目の問題は、推論です。 というのも、推論による情報は、明示的に述べられていなくても、得ることができます。例えば「レアは彼女の子供を叱る」という文を見てみましょう。
D9. Nous pouvons déduire 3 informations	この文から3つの情報を引き出すことができます。
D10. <i>Léa a un enfant</i>	まず、「レアは子供がいる」という情報です。
D11. <i>Léa est une mère</i>	次に「レアは母親である」という情報です。
D12. <i>Son enfant a fait une bêtise</i>	そして最後に「レアの子供が悪さをした」という情報です。
D13. Le 2 ^e problème est le figement. Le figement consiste en un certain degré de « fossilisation » d'une formation syntagmatique. Nous retrouvons le figement entre autres dans les	第2番目の問題は凝結です。凝結は(句や文などの)連辞的表現がある程度《化石化》するところにその本質があります。とりわ

<p>expressions, comme on le voit ici avec les exemples suivants :</p> <ul style="list-style-type: none"> • poser un lapin = ne pas venir au rendez-vous 	<p>け、凝結は次のような例において認めることができます。</p> <ul style="list-style-type: none"> • 待ちぼうけを食わせる³ (←ウサギを置く)
<p>D14. Enfin, le 3^e problème est la polysémie. Prenons l'exemple du verbe « descendre » qui est très polysémique en français</p>	<p>最後に、3番目の問題は多義性です。非常に多義的なフランス語の動詞 descendre を例として取り上げて見ましょう。</p>
<p>D15. Voici quelques exemples d'emploi de ce verbe</p>	<p>この動詞の幾つかの用法です。</p>
<p>D16. Luc a descendu un criminel [= <i>tuer</i>] Luc descend les escaliers [<i>≠ monter : aller de haut en bas</i>] Luc a descendu cinq whisky [= <i>boire</i>] Luc descend d'une grande famille [= <i>filiation : être issu</i>]</p>	<p>リュックは犯罪者を処刑した。[=<i>tuer</i>]殺す リュックは階段を降りた。[<i>≠ monter : aller de haut en bas</i>]上から下へ行く リュックはウイスキーを5杯飲んだ。[=<i>boire</i>]飲む リュックは大家族の出である。[=<i>filiation : être issu</i>]出である</p>
<p>D17. Rien que pour le verbe « descendre » nous avons 23 sens possibles en français</p>	<p>フランス語では、動詞 descendre だけでも23の意味があります。</p>
<p>D18. MÉTHODOLOGIE Nous allons maintenant voir la méthodologie pour faire face à ces difficultés</p>	<p>方法論 ここで、これらの問題を解決するための方法論を考えてみましょう。</p>
<p>D19. Mes travaux portent essentiellement sur la réflexion et sur la conception de méthodologies appropriées au traitement automatique des langues Le traitement automatique des langues exige une reconnaissance précise des formes. Cela implique donc :</p> <ul style="list-style-type: none"> • recensement lexical aussi rigoureux et complet que possible 	<p>私の研究は、主として言語自動処理に合った方法論の考察と構想に基づいています。</p> <p>言語の自動処理は正確に言語形態を認識することが求められます。したがって、これは以下のことを前提としています。</p> <ul style="list-style-type: none"> • 可能な限り厳密で完全な調査

<p>• degré suffisant de couverture, que seul peut apporter un <u>dictionnaire électronique</u> construit à cet effet.</p>	<p>• 自動処理の目的のために構築されたコンピュータのための辞書のみが到達できる十分な語彙カヴァー率</p>
<p>D20. L'objectif est donc de construire des dictionnaires électroniques. Mais qu'est-ce qu'un dictionnaire électronique ?</p>	<p>したがって目的はコンピュータのための辞書を構築することです。しかし、コンピュータのための辞書とは何でしょうか？</p>
<p>D21. IL est tout d'abord important de faire la différence entre 2 types de dictionnaires qui ont recours à l'informatique : dictionnaire informatisé/ dictionnaire électronique Un dictionnaire informatisé est une version électronique du dictionnaire papier. L'avantage est que la recherche de l'information y est plus rapide. Mais les informations mises en évidence ne sont pas de <i>nature différente</i> de celles que l'on trouve dans un dictionnaire papier</p>	<p>まず、重要なことは、いずれも情報科学を利用している人間のための電子辞書とコンピュータのための辞書との2種類の辞書の違いを区別することです。 人間のための電子辞書とは単なる紙媒体の辞書の電子版のことです。人間のための電子辞書の利点は、情報検索がより早いことです。 しかし、そこから取り出される情報は紙媒体の辞書から取り出される情報と異なる性質のものではありません。</p>
<p>D22. Je vais illustrer le dictionnaire informatisé par le dictionnaire Le Petit Robert que l'on peut télécharger.</p>	<p>私は、ネット上でダウンロードできる電子辞書『プチ・ロベール』によって、人間のための電子辞書について例をあげて説明したいと思います。</p>
<p>D23. Je prends l'exemple du verbe <i>voir</i> (verbe très polysémique) pour vous montrer comment il est décrit.</p>	<p>私は、非常に多義的な動詞 <i>voir</i><見る>を例にとって、それがどのように記述されているかを皆さんに示したいと思います。</p>
<p>D24. ⁴</p>	<p>赤い下線が引かれている部分に注目してください。これは語の定義の部分です。</p>
<p>D25. Nous voyons au niveau des définitions du verbe que nous avons :</p> <ul style="list-style-type: none"> - très peu d'informations sémantiques, syntaxiques ou morphologiques. - très peu de synonymes - et enfin les emplois sont mélangés : on trouve par exemple les synonymes <i>apercevoir</i> 	<p>この動詞の定義を見ると、次の点がわかります。</p> <ul style="list-style-type: none"> - 意味・統辞・形態に関する情報が非常に少ない。 - 類義語が非常に少ない。 - 最後に、用法の混同が見られる。同じ定義の中に、類義語として、例えば、

<p>et lire dans la même définition. Et pourtant ils sont sémantiquement très différents.</p>	<p>apercevoir<気づく>と lire<読む>がありますが、これらは用法的に大変異なるものです。</p>
<p>D26.</p>	<p>これは D23.と同じ画面です。</p>
<p>D27.</p>	<p>赤い丸で囲まれている部分に注目してください。これは成句です。</p>
<p>D28. Même chose avec les expressions. On constate qu'elles sont peu nombreuses. Nous n'avons également aucune information sur leurs emplois, sur leur mode d'utilisation (peut-on les mettre au féminin ou au pluriel, par exemples). Enfin, on constate qu'on a très peu d'informations sémantiques</p>	<p>成句についても同じであります。成句が少ないことが分かります。また、それらの用法や、女性形や複数形が作れるのか・主語は動物なのか人間なのか・という使用法について、いかなる情報もありません。さらに意味的情報が非常に少ないです。</p>
<p>D29. En revanche, un dictionnaire électronique est conçu dès l'origine pour le traitement automatique → Reconnaissance → Génération</p>	<p>逆に、コンピュータのための辞書は初めからテキストの認識・生成の自動処理のために構想されているのです。 → (テキストの) 認識 → (テキストの) 生成</p>
<p>D30. Voyons les caractéristiques du dictionnaire électronique. Premièrement, les informations sont explicites et reproductibles : Destiné au traitement automatique, ne spéculer pas sur des "rectifications analogiques" qu'un lecteur humain est en mesure de faire quand une information est insuffisante</p>	<p>コンピュータのための辞書の特徴を見てみましょう。 第一点目に、それは情報が明示的で再現可能であるということです。 また、自動処理の用途に充てられているので、情報が不足している時人間がするようないわば「類推的修正」に基づいて思索することはできません。</p>
<p>D31. Deuxièmement, il ne présuppose pas une connaissance préalable de la langue : seules les informations qui lui sont fournies et leurs structurations permettent de reconnaître ou de construire les phrases d'une langue naturelle</p>	<p>第二点目に、言語についての事前の知識を前提としないことがあげられます。与えられた情報とその情報の構造のみから、自然言語の文を認識したり、作ったりすることができます。</p>

<p>D32. Enfin, troisièmement, le dictionnaire électronique est une <i>base de données</i> cohérente</p>	<p>最後、第三点目として、コンピュータのための辞書は、首尾一貫性のあるデータベースなのです。</p>
<p>D33. Ce sont ces dictionnaires électroniques que nous développons. Ils sont essentiels pour traiter les 3 problèmes de la langue naturelle auxquels nous sommes confrontés (la <i>polysémie</i>, la <i>figement</i>, l'<i>inférence</i>)</p>	<p>私たちが開発しているのはこのような特徴を持つコンピュータのための辞書であります。そのような辞書は、私たちが直面している自然言語の3つの問題、すなわち多義性・凝結・推論を解決するために本質的に重要なものであります。</p>
<p>D34. Le point de vue théorique adopté est celui de G. GROSS (modèle des classes d'objets) Cette méthode fusionne le lexique, la syntaxe et la sémantique dans des emplois Un mot ne peut pas être défini hors contexte, mais seulement dans un environnement syntaxique donné :</p> <p>→ le lexique ne peut donc pas être séparée (= la combinatoire des mots).</p> <p>→ La sémantique n'est pas autonome non plus : elle est le résultat des éléments lexicaux organisés d'une façon déterminée (distribution).</p>	<p>さて言語の記述のために私たちが採用している理論的視点は、ガストン・グロスの対象部類⁵のモデルであります。この方法は、語彙・統辞・意味の全てを一つの用法として統合するものであります。</p> <p>語は文脈の外では定義することはできません。ある与えられた統辞的環境の中でしか定義できないのです。</p> <p>→ したがって語彙はそれだけを切り離すことはできません。(それは語の結合の中にあるからです)。</p> <p>→ 意味もまた自律的ではなくそれだけを切り離すことはできません。それは語をある決まった方法で配列した結果なのです。それは分布の表れです。</p>
<p>D35. Dans ce cadre théorique La 1^e étape est la description de la distribution (c'est-à-dire le contexte gauche ou droit) Exemple avec le verbe Descendre :</p> <ul style="list-style-type: none"> • N0<hum> descendre N1<loc:voie> <i>Il a descendu le fleuve</i> • N0<hum> descendre N1<hum> <i>Il a descendu le malfaiteur</i> • N0<hum> descendre N1<inc:boisson> 	<p>このような理論的枠組みにおいて、第一段階は、分布すなわち左右の文脈の記述です。</p> <p>以下は動詞 <i>descendre</i> の記述の例です。</p> <ul style="list-style-type: none"> • N0<人間> descendre N1<場所:道路> <i>Il a descendu le fleuve</i> • N0<人間> descendre N1<人間> <i>Il a descendu le malfaiteur</i> • N0<人間> descendre N1<具体無生物>

<p><i>Il a descendu trois whiskies</i></p> <p>• N0<hum> descendre de N1<inc:mtr></p> <p><i>Il est descendu de l'autobus</i></p>	<p>飲み物></p> <p><i>Il a descendu trois whiskies</i></p> <p>• N0<人間> descendre de N1<具体無生物:移動手段></p> <p><i>Il est descendu de l'autobus</i></p>
<p>D36. La 2^e étape est la description de toutes les variations possibles du mot (c'est-à-dire l'étiquetage des mots simples). L'étiquetage permet de donner les traits flexionnels des mots fléchis (genre, nombre, temps, personne) et leur forme de base (par ex. le masculin sing. pour les adj).</p>	<p>第二段階は、テキストの中で取る語の全ての形の記述、すなわち、語のタグ付けです。語のタグ付けにより、語の全ての形には、性・数・時制、人称などの屈折語の屈折特徴やその基本形、例えば屈折語が形容詞ならば形容詞の男性単数基本形などの特徴が付与されます。</p>
<p>D37. TRAVAUX</p>	<p>業績</p>
<p>D38. Dans ce cadre théorique, mes 1^{ers} travaux ont porté sur la description des prédicats de perception</p>	<p>このような理論的枠組みの中において、私が行った最初の研究は知覚の述語の記述に関するものでした。</p>
<p>D39. Pourquoi ces prédicats ? Les prédicats de perception sont un sujet d'analyse privilégié dans le cadre d'une recherche sur les rapports entre syntaxe et sémantique</p> <p>Au niveau de la syntaxe : Variété de la complémentation</p> <p>Au niveau de la sémantique : Diversité des interprétations</p> <p>Au niveau du lexique : Verbes sémantiquement proches mais syntaxiquement différents</p>	<p>なぜ、知覚の述語の研究を選んだかと申しますと、知覚の述語の研究は、統辞と意味の関係についての最適な研究テーマであるからです。</p> <p>統辞面では、補語の種類が豊かです。</p> <p>意味面では、解釈に多様性があります。</p> <p>語彙面では、意味的に類似しているけれど統辞的に異なる動詞が存在しております。</p>
<p>D40. Les objectifs de ce travail étaient :</p> <p>i. Le traitement de la polysémie (4721 prédicats)</p> <p>ii. La constitution de sous-classes sémantiques de prédicats</p> <p>iii. La constitution d'une taxinomie des prédicats de <i>perception</i></p>	<p>また、この研究の目的は、以下の通りでした。</p> <p>i. 4721 の知覚の述語の多義性の処理</p> <p>ii. 知覚の述語の意味的下位クラス的确立</p> <p>iii. 知覚の述語の分類学的确立</p>

<p>D41. PRESENTATION DE 3 RESSOURCES</p> <p>Maintenant, je vais vous présenter 3 ressources que nous avons développées avec mon équipe :</p> <ul style="list-style-type: none"> • FixISS • Néoveille • Morfetik 	<p style="text-align: center;">3つのリソースの紹介</p> <p>今度は、皆さんに、私が研究チームとともに開発した、3つのリソースを紹介します。</p> <ul style="list-style-type: none"> • FixISS フィクシス • Néoveille ネオヴェュー • Morfetik モルフェティック
<p>D42. Tout d’abord, la ressource FIXISS</p>	<p>まず最初にフィクシスです。</p>
<p>D43. Comme nous l’avons vu, le figement est une des grandes difficultés pour le traitement automatique et donc pour la constitution de ressources.</p> <p>Les ressources sont plus faciles à constituer pour les mots simples</p> <p>La difficulté réside au niveau de la reconnaissance de l’unité globale et de sa génération puisqu’il n’y a pas d’autonomie entre les éléments constitutifs de la suite</p>	<p>すでに私たちが見てきたように、凝結は自動分析にとって、したがって、リソースの構築にとって、困難な点の一つであります。</p> <p>単純語に関しては、リソースの構築は容易です。</p> <p>困難は、一つの塊として捉えられる単位としての凝結表現の認識とその生成にあります。なぜなら、凝結表現は、構成要素の間には自律的關係がないからです。</p>
<p>D44. Une des premières difficultés pour l’ordinateur, pour le traitement automatique va être le repérage des expressions dans un texte.</p>	<p>コンピュータにとっての最初の困難の一つは、すなわち自動処理にとっての最初の困難の一つは、テキストにおける凝結表現のタグ付けです。</p>
<p>D45. Prenons pour exemple l’expression française : <i>enterrer la hache de guerre</i> qui veut dire <i>suspendre les hostilités</i>.</p>	<p>次の例を見てください。これは<敵対関係を中断する>を意味する「<i>enterrer la hache de guerre</i> (字義通りの意味：矛を埋める)」というフランス語の凝結表現です。</p>
<p>D46. Si on fait des recherches dans la presse quotidienne sur internet, on va constater que l’expression va subir de nombreuses transformations.</p>	<p>もし、インターネットの日刊紙の中でこの表現を検索すると、そのままの形ではなく変形を受けた形で表れることに気がつきます。</p>
<p>D47. Dans l’exemple suivant, on voit qu’il y a l’insertion d’une locution (« <u>pour l’heure</u> ») dans l’expression elle-même : l’UMP enterre</p>	<p>次の「民衆運動連合は矛を、さしあたり、収める」の例では、凝結表現の中にくさしあたり>を意味する成句 <u>pour l’heure</u> が挿</p>

<p>pour l'heure la hache de guerre. « UMP » en France est un parti politique.</p>	<p>入されていることがわかります。ちなみに民衆運動連合はフランスの政党です。</p>
<p>D48-49. Dans ce nouvel exemple, « La hache de guerre est enterré par Deschamps », nous avons une inversion. Le complément « hache de guerre » est devenu sujet</p>	<p>また、次の「矛はデシャン氏によって収められた」の別の例では、倒置が起きています。そのため目的補語の <i>hache de guerre</i> が主語になっております。</p>
<p>D50-51. Dans ce 3^e exemple, nous avons un défigement. Le verbe « enterrer » a été remplacé par son antonyme « déterrer ».</p>	<p>さらに次の第3番目の例では、凝結解消⁶が起きています。 それにより動詞の <i>enterrer</i> <埋める> が反対語の <i>déterrer</i> <地中から取り出す> によって置き換えられています。</p>
<p>D52. On voit donc que les expressions sont des types d'élément lexical délicats à représenter car il y a divers degrés de figement. Certaines expressions sont complètement figées et n'acceptent donc aucune variation (par exemple : « <i>au fur et à mesure</i> »). Ces expressions sont peu nombreuses. On considère qu'il y en a à peu près 10% dans la langue française. Toutes les autres expressions vont subir des variations.</p>	<p>したがって、凝結表現は、凝結度が様々で表示するのが難しい語彙的単位のタイプであることがわかります。ある凝結表現は完全に凝結していて、いかなる変異⁷も許しません。例えば、「それに応じて」を意味するフランス語の <i>au fur et à mesure</i> がそうです。このような完全に凝結した表現はフランス語では数多くはなく、全体の約10パーセント程度であると考えられています。その他の凝結表現は、全て変異を許す表現です。</p>
<p>D53. Ces variations sont très diversifiées. En voici quelques exemples :</p> <ul style="list-style-type: none"> • Des changements de détermination : • Des variantes lexicales : • Des changements de position : • Des variations morphologiques : • Des variations graphiques : 	<p>変異はとても多様です。以下はその幾つかの例です。初めは、</p> <ul style="list-style-type: none"> • 限定詞の変異 • 語彙的変異 • 語順の変異 • 形態的変異 • 正書法の変異
<p>D54. L'autre difficulté relative au traitement automatique des expressions et donc du figement va être la traduction.</p>	<p>凝結の自動処理に関係している他の困難な点といえば、翻訳でしょう。</p>

<p>D55. Nous allons illustrer cette difficulté avec les traducteurs automatique que l'on trouve en ligne gratuitement.</p> <p>Nous allons voir comment est reconnue et traduite l'expression idiomatique française : <i>être bête comme ses pieds</i>, qui veut dire être très bête.</p> <p>Nous allons faire le test en traduisant cette expression en espagnol à l'aide des 3 traducteurs : Reverso, Systran et Google traduction</p>	<p>オンラインで、無料で使用出来る自動翻訳機械を使って、この困難を、例をあげて説明します。</p> <p>< (自分の足のように馬鹿→) とても馬鹿である > を意味するフランス語の慣用的表現 <i>être bête comme ses pieds</i> がどのように認識され翻訳されるのかを見てみましょう。</p> <p>この表現を、Reverso, Systran, Google Traduction の3つの翻訳機械を使ってスペイン語への翻訳を試してみたいと思います。</p>
<p>D56. Pour ce 1^{er} test avec le traducteur Reverso, le résultat est satisfaisant. L'expression a été identifiée. Nous n'avons pas l'expression équivalente en espagnol mais nous avons le paronyme de l'expression « ser muy tonto ». C'est-à-dire <i>être très bête</i>. L'intensité est bien présente puisque nous avons l'intensifieur « muy » (= beaucoup)</p>	<p>Reverso を使ったこの最初のテストでは、結果は満足がいくものであります。表現は認識されています。スペイン語にはフランス語と完全に同じである等価表現はありませんが、< ser muy tonto > という準類義的表現があります。つまりこれは< とても馬鹿である > という意味です。そこにおいて強調はしっかりと存在しております。なぜなら、muy は強調を表す言葉だからです。</p>
<p>D57. Avec Systran, le résultat n'est pas bon puisque l'expression n'a pas été identifiée. Nous avons en fait une traduction de l'expression mot à mot donnant lieu à une incohérence sémantique tout en gardant une cohérence syntaxique. Cette expression en espagnol n'a aucun sens pour un hispanophone. Nous avons même une erreur sur le verbe (<i>ser</i> et non <i>estar</i>)</p>	<p>Systran を使うと、今度は、結果は良くありません。なぜなら、凝結表現が認識されなかったからです。それは逐語訳なため、統辞的には一貫性があるけれど、意味的に支離滅裂です。このスペイン語の表現はスペイン語の話者には何の意味も伝えません。動詞に誤りさえあります。estar の代わりに ser を使わなければなりません⁸。</p>
<p>D58. Enfin, avec Google traduction, le résultat est « intéressant ». L'expression espagnole « ser duro como un ladrillo » signifie littéralement « être dur comme une brique ».</p>	<p>最後に、Google Traduction を使うと、結果は興味深いものです。スペイン語の表現 <i>ser duro como un ladrillo</i> は、文字通りの意味で</p>

	<p>は、＜レンガのように堅い＞を意味します。</p>
<p>D59. Le problème est que l'expression « ser duro como un ladrillo » n'existe absolument pas en espagnol. Un espagnol ne peut pas la comprendre.</p>	<p>問題は、ser duro como un ladrillo＜レンガのように堅い＞という表現はスペイン語には全く存在しない表現だということです。スペイン人はこの表現を理解できません。</p>
<p>D60. Du coup se pose la question de savoir comment Google a construit cette expression qui n'existe pas en espagnol.</p>	<p>ふと、Google Traduction はスペイン語に存在しないこの表現をどのようにして作ったのだろうかという疑問が浮かんできます。</p>
<p>D61. L'expression espagnole a été construite sur celle anglaise « to be as sick as a brick », qui elle, existe (et qui veut dire littéralement « être dur comme une brique »).</p> <p>Google Traduction est basé, en grande partie, sur une méthode de traduction automatique statistique. Sa performance dépend considérablement de la qualité du corpus parallèle utilisé. La plupart des corpus parallèles proviennent du domaine politique ou juridique. Il est fort probable que Google Traduction passe par l'anglais pour combler son corpus.</p> <p>Google Traduction fonctionne avec le même principe de « langue pivot », utilisé notamment au parlement européen. Si l'on veut traduire des langues communes (anglais, allemand, français), la traduction ne sera pas parfaite, mais globalement acceptable. Par contre, si l'on veut traduire des langues moins riches en termes de corpus parallèle pour l'apprentissage automatique, Google Traduction passe par une</p>	<p>Google Traduction のスペイン語の表現は、実際に存在する英語の＜to be as thick as a brick ひどく馬鹿である（←レンガと同じぐらい頭が堅い）＞に基づいて翻訳されたものです。</p> <p>Google Traduction の翻訳は大部分が統計的自動翻訳の方法に基づいています。その性能は、並行して使われている並行コーパス⁹の質に大きく依存しています。（Google Traduction の場合は）大部分の並行コーパスは政治的あるいは法律的分野に（のみ）由来しています。</p> <p>Google Traduction はそのコーパスの欠陥を補うため英語を経て翻訳している可能性が高いです。特に欧州議会の翻訳方法と同様に＜ピボット言語¹⁰＞の原理で機能しています。もし、（英語、ドイツ語、フランス語）のような公用語を翻訳したいなら、翻訳は完全ではありませんが、しかし全体的に容認できるものになるでしょう。これに対して、もし、機械学習のための並行コ</p>

<p>langue intermédiaire, l'anglais, avant de traduire de nouveau dans la langue finale (ici l'espagnol), etc.</p> <p>Cela montre que la polylexicalité n'est pas prise en compte de façon stable dans ses modèles.</p>	<p>一パスのより貧（とぼ）しい言語を翻訳したいなら、Google Traduction は、対象言語へ再び翻訳する前に（ここでは対象言語はスペイン語ですが）、英語という中間言語を経ることになるわけです。</p> <p>このことから、Google Traduction はそのモデルにおいて、凝結表現のポリレキシカリティすなわち複合的語彙的単位であるという性質を、考慮していないことがわかります。</p>
<p>D62. Pourquoi m'intéresse-je donc au figement ?</p> <p>Comme on peut le voir avec ces exemples, le figement occupe une place importante dans les préoccupations de la recherche actuelle :</p> <ul style="list-style-type: none"> • rôle déterminant dans les systèmes linguistiques de reconnaissance : - traduction/correction automatique - extraction d'information - enseignement des langues, etc. <p>• donnée de base incontournable dans la description des langues</p>	<p>それではなぜ私は、凝結に関心を持つのでしょうか？</p> <p>翻訳機械で試した先ほどの例からわかるように、凝結は現在の研究の関心事において重要な席を占めているからです。</p> <ul style="list-style-type: none"> • (凝結は) 次のような応用分野の言語認識システムに対して決定的な役割を果たします。 <ul style="list-style-type: none"> - 自動翻訳/自動校正 - 情報抽出 - 言語教育 • (凝結は) 言語記述において不可欠なデータベースです。
<p>D63. Mais on constate que le figement reste un fait problématique pour les descriptions et les méthodologies actuelles.</p> <p>En linguistique :</p> <p>Pour le français : beaucoup de théories, peu d'applications</p> <p>Difficultés à trouver véritablement des base de données de séquences figées accessibles, structurées, formalisées, exploitables</p>	<p>しかし、凝結は、現在の記述や方法論において依然として問題点が多い現状を確認できます。</p> <p>言語学で :</p> <p>フランス語に関しては、たくさんの理論があるのに、応用研究がほとんどありません。アクセス可能で・構造化・形式化された・利用可能な凝結表現のデータベースを見つけることは本当に難しいのです。</p>

<p>En informatique : Même les étiqueteurs comme <i>Unitex, Freeling</i> n'utilisent pas de dictionnaires de séquences figées</p>	<p>情報科学で : <i>Unitex, Freeling</i> のようなテキストに対して語彙情報タグ付けることが可能なソフトでさえ凝結表現辞書を使っていないのです。</p>
<p>D64. Pour les outils de traduction en ligne : Le problème est différent. Les logiciels de traduction répondent à un besoin croissant et pressant : traduire vite et pour pas cher. Les résultats sont relativement corrects pour les mots/phrases simples mais pour les expressions figées les erreurs et les contresens sont nombreux Dans les dictionnaires traditionnels : nous avons très peu d'informations sur les expressions</p>	<p>オンライン翻訳機械に関しては : 問題は違います。翻訳ソフトは、差し迫りかつますます増大しつつある必要すなわち早く安くという必要に応えています。単純な語や単文の翻訳に関しては、結果は比較的正しいのですが、凝結表現に関しては、多数の誤りや誤解が見つかります。 伝統的辞書に関しては : 凝結表現についての情報はごくわずかです。</p>
<p>D65. A partir de ce constat, nous avons donc réalisé notre propre ressource.</p>	<p>したがって、このような現状確認から出発して、私たちは、私たち独自のリソースを作りました。</p>
<p>D66. L'objectif est de proposer un modèle capable (i) d'exposer toutes les formes possibles d'une expression donnée ; (ii) de fournir une ressource capable d'être projetée sur un document écrit pour l'analyse automatique. Une expression figée sera reconnue automatiquement que si le système d'analyse contient une base de données donnant sa forme <i>in extenso</i> et ses propriétés</p>	<p>目的は、(i)まず、与えられた凝結表現のすべての可能な形を示すことができるモデルを提案し、(ii)次に、自動分析のために、文書¹¹へ投影できるリソースを提供することです。 凝結表現は、分析システムが、その形と特性を網羅的に与えてくれるデータベースを含んでいる場合にのみ、自動的に認識できるのです。</p>
<p>D67. Nous avons commencé à travailler sur les locutions adverbiales en français et en espagnol. Nous avons actuellement plus de 6000 entrées</p>	<p>まず第一段階として、私たちはフランス語とスペイン語の副詞的凝結表現から作業を始めました。現在6000以上の見出語(す</p>

<p>(locutions adverbiales). Ces expressions ont été décrites en français et en espagnol. Le projet a d'ailleurs intéressé l'entreprise Systran, qui développe des technologies de traduction automatique.</p>	<p>なわち副詞的成句) があります。これらの表現はフランス語とスペイン語で記述されました。さらに、プロジェクトは、自動翻訳のテクノロジーを開発している Systran社の関心を引きました。</p>
<p>D68-69. Je vous présenter rapidement la ressource</p>	<p>次に急いで私たちのリソースを紹介します。フィクシスの画面です。</p>
<p>D70. On peut actuellement faire des recherches par mot vedette ou par construction</p>	<p>現在、「見出語」か「構文」で検索できます。</p>
<p>D71. La 2^e étape a été de travailler sur les locutions comparatives en « comme »</p>	<p>第二段階は、「なにになにのような」を意味する <i>comme</i> を使った直喩的成句について作業することでした。</p>
<p>D72. Ce travail a débuté en 2013. Nous avons travaillé sur les locutions construites sur un adjectif ou sur un verbe comme :</p> <p><i>Manger comme un cochon</i> <i>Soûl comme un polonais</i></p> <p>Elles expriment toute l'intensité. Nous avons pour le moment travaillé sur 4 langues :</p> <ul style="list-style-type: none"> • Locution en <i>comme</i> (française) : 800 entrées • Locution en <i>comme</i> (polonaise) : 1300 entrées • Locution en <i>comme</i> (espagnol) : 1100 entrées • Locution en <i>comme</i> (portugais du Brésil) : 900 entrées 	<p>この研究は 2013 年に始まりました。私たちは形容詞か動詞に基づいて作られた成句について作業をしました。例えば、次のような成句です。</p> <p>とてもよく食べる とても酔っ払っている</p> <p>これらはすべて強調を表します。 我々は差し当り 4 言語について作業しました。</p> <ul style="list-style-type: none"> • (フランス語)<～のような>型成句: 見出語数 800 • (ポーランド語)<～のような>型成句: 見出語数 1300 • (スペイン語)<～のような>型成句: 見出語数 1100 • (ブラジルのポルトガル語)<～のような>型成句: 見出語数 900
<p>D73. A la suite de ce travail, nous avons réfléchi à l'exploitation possible de cette ressource en imaginant 2 scénario.</p>	<p>この作業が終わった時点で、私たちはこのリソースの活用について、2つのシナリオを考えました。</p>

<p>D74. Le 1^{er} scénario imaginé est de créer une ressource pour l'aide à la traduction.</p> <p>En effet, lorsque 2 langues expriment une même propriété, les comparants (images employées) des locutions correspondantes ne sont pas forcément identiques comme on le voit dans le tableau suivant. Pour décrire la laideur, des langues comme le français, l'italien ou le polonais n'auront pas le même comparant. Pour la France, la laideur est comparée à un pou ; pour l'Italie, à un pêché ou à la faim ; et pour la Pologne à la nuit.</p>	<p>想像した第一のシナリオは、翻訳支援のためのリソースを作ることです。</p> <p>実際に、二つの言語がある一つの同じ性質を表現するとき、対応する成句の「たとえる辞項」（すなわち使用された表象を表す語で、「～のような」を意味する語の直後の語）は、次の一覧表からわかるように、必ずしも同一ではありません。「醜さ」を表現する際に、フランス語、イタリア語、ポーランド語のような言語は同じ「たとえる辞項」をもたないのです。フランスならば、「醜さ」はノミにたとえられ、イタリアなら罪あるいは飢え・飢餓にたとえられ、ポーランドなら夜にたとえられます。</p>
<p>D75. La ressource va ainsi fournir une liste d'expressions correspondantes, classées par rapport à un degré de synonymie.</p> <p>Par exemple, l'expression française <i>manger comme un ogre</i> se verra proposer pour l'espagnol la liste suivante :</p> <p style="text-align: center;"><i>Comer como un ogro</i> <i>Comer como un regimiento</i> <i>Comer como un tudesco...</i></p> <p>A la suite de cette proposition, l'utilisateur pourra effectuer son choix en fonction de critères choisis : <i>classe sémantique du comparant, niveau de langue, ...</i></p>	<p>リソースはこのように同義性の度合い¹²との関連で分類された対応表現リストを提供します。</p> <p>例えば、<i>manger comme un ogre</i> というフランス語の表現はスペイン語に関しては、次のようなリストが提供されるでしょう。</p> <p style="text-align: center;">[食人鬼のように食べる] [軍隊のように食べる] [ドイツ人のように食べる]</p> <p>このようなリストに基づいて利用者は選択した基準に応じて、「たとえる辞項」すなわち「～のような」を意味する語の直後に現れる語の意味的クラスはどうかとか、言語位相はどうかといった風に自らの研究を選択することができるでしょう。</p>
<p>D76. L'autre scénario imaginé est celui d'une aide à la rédaction. Ce type de situation peut se</p>	<p>想像した別のシナリオは作文支援です。この種の状況は次のように述べることで</p>

<p>décliner de la manière suivante : le rédacteur ne connaît que le sens qu'il veut exprimer et cherche à employer l'expression la plus appropriée au contexte. Par exemple, quelles sont les expressions françaises pour dire <i>manger beaucoup</i> ?</p>	<p>きます。作文している人は表現したい意味しか知らずコンテキストに最も適合した表現を用いようとしています。例えば、<i>manger beaucoup</i> <たくさん食べる> という意味を表すフランス語の表現にはどんなものがあるだろうか？と考えるでしょう。</p>
<p>D77. A partir de ces scénarios et de notre ressource FixISS, nous avons créé quelques outils (qui sont encore des prototypes pour le moment).</p>	<p>このようなシナリオと私たちのリソース FixISS から出発して、私たちはまだ試作品ではありますが、幾つかのツールを作りました。</p>
<p>D78. Je vous présente le 1^{er} outil. A partir de la base de données, nous associons un graphe de relation (<i>outil développé au LDI</i>) qui permet d'illustrer les relations entre des unités lexicales de nature diverse.</p> <p>Le 1^{er} graphe montre les adjectifs français qui sont les plus employés dans les locutions comparatives et quels sont les animaux qui illustrent ces adjectifs.</p> <p>Ainsi, on peut être gros comme une baleine, un éléphant, une vache ou un cochon</p> <p>Le 2^e graphe illustre les relations entre les expressions françaises et polonaises. Nous avons pris comme exemple toutes les expressions françaises et polonaises qui vont comparer certaines caractéristiques du lion. En rose, nous avons les expressions françaises et en bleu les expressions polonaises. Avec ce graphe, nous voyons facilement si pour une expression française un équivalent polonais existe.</p>	<p>私は、皆さんに最初のツールを紹介します。それは、多様な語彙的単位の間の関係を例示的に説明できる関係グラフ（これは LDI 研究所が開発したツール）というツールで、これをデータベースに連結します。左側のグラフは、直喩的成句の中で最もよく使われるフランス語の形容詞と、この形容詞を例示的に説明する動物が何であるかを示しています。</p> <p>例えば、人はクジラ、象、牛、豚のように太っている、とすることができることを教えてくれます。</p> <p>右側のグラフは、フランス語の表現とポーランド語の表現の関係を例示的に説明しています。例として人間をライオンの特徴と比較しているフランス語とポーランド語のすべての凝結表現です。ピンクがフランス語の表現で青がポーランド語の表現です。このグラフによって、あるフランス語の表現に対して、ポーランド語に等価表現があるのか容易にわかります。</p>
<p>D79-83. Autre outil que nous avons développé un éditeur assistant. Cet éditeur offre un retour</p>	<p>私たちが開発した別のツールは、補助エディターと言うツールです。このエディター</p>

<p>dynamique, c'est-à-dire au fur et à mesure de la frappe. A la fin de la frappe, nous avons une précision sémantique sur l'expression (en vert)</p>	<p>は、学習者にキー打ちに応じた（作成物への）ダイナミックなインターフェースを可能にすることで、学習者の作文を支援してくれます。キー打ちが終わると、表現に関する意味的な詳細が（緑色で）表示されます。</p>
<p>D84. La ressource Morfetik Autre ressource, la ressource Morfetik</p>	<p>モルフェティック 今回はモルフェティックについて説明します。</p>
<p>D85. Morfetik est une ressource linguistique ayant les caractéristiques suivantes :</p> <ul style="list-style-type: none"> ➤ large couverture ➤ précision et fiabilité des informations ➤ évolutivité <p>Cette ressource est le résultat du travail d'une vingtaine d'années de collecte et de description.</p>	<p>モルフェティックは次のような特徴を持つ言語リソースです。</p> <ul style="list-style-type: none"> ➤ 語彙のカヴァー率が高い ➤ 情報が正確で信頼性がある ➤ 拡張性がある <p>このリソースは約 20 年のデータ収集と記述の成果です。</p>
<p>D86. Morfetik est en fait un dictionnaire morphologique des mots simples du français. Le système permet de générer automatiquement l'ensemble des formes fléchies simples du français</p> <p>Quelques chiffres :</p> <p>758035 formes fléchies, pour 103312 lemmes/graphies</p> <ul style="list-style-type: none"> • 66393 noms • 10223 verbes • 24391 adjectifs • 2305 autres 	<p>さて、モルフェティックとは何でしょうか？モルフェティックとは、事実上、フランス語の単純語の形態論的辞書です。このシステムによりフランス語の単純語の屈折形の全てを生成することができます。数の面から言うと、103312 のレンマ¹³あるいは書記¹⁴に対して 758035 の屈折形が存在します。その内訳は以下の通りである。</p> <ul style="list-style-type: none"> • 名詞 66393 • 動詞 10223 • 形容詞 24391 • その他 2305
<p>D87-89. La ressource peut soit être consultée via un moteur de recherche, soit téléchargée dans un format XML</p>	<p>モルフェティックは検索エンジンを通して参照することもできるし、XML¹⁵の形式でダウンロードすることもできます。</p>
<p>D90-92. La ressource permet pour n'importe quel mot français :</p> <p>- d'obtenir l'ensemble de ses formes :</p>	<p>モルフェティックはフランス語のどんな語に対しても :</p> <p>-その語の屈折形のすべてを手に入れることができます。例えば、</p>

<ul style="list-style-type: none"> - pluriel des N - féminin et pluriel des Adj - formes conjuguées des V <p>etc.</p> <p>- d'identifier le mot (= lemme) : correspondant à n'importe quelle forme fléchie</p>	<ul style="list-style-type: none"> - 名詞の複数形 - 形容詞の女性形と複数形 - 動詞の活用形 <p>など</p> <p>—どんな屈折形にも対応して存在する 語 (=レンマ) も教えてくれます。</p>
<p>D93. Un ensemble d'outils de navigation et de consultation de la ressource sont actuellement en finalisation de développement. Nous avons un LEMMATISEUR DE FORMES et un GENERATEUR DE FLEXIONS</p>	<p>リソースの閲覧と参照のツール全体が現在開発の最後の段階にあります。これには レンマ生成ツールと 屈折形生成ツールがあります。</p>
<p>D94. A partir d'une forme quelconque, l'outil fournit :</p> <ul style="list-style-type: none"> ➤ les différents lemmes possibles (= forme de base) ➤ les différentes formes liées à chaque lemme 	<p>任意の形式に対して、ツールは以下のことを示してくれます :</p> <ul style="list-style-type: none"> ➤ 基本形であるレンマ ➤ レンマに結びついている屈折形
<p>D95-96. Prenons l'exemple du mot « marche »</p>	<p>marche を例にとって説明します。</p>
<p>D97.-D98. Nous obtenons le résultat suivant. Le moteur a identifié les différents lemmes possibles de ce mot : un verbe « marcher », un nom féminin singulier « une marche » et un nom masculin singulier « marché ».</p>	<p>以下のような結果が得られます。検索エンジンはこの単語の様々なレンマを見つけしてくれました。動詞の marcher<歩く>、女性単数名詞の une marche<歩行>、男性単数名詞の marché<市場></p>
<p>D99-101. L'utilisateur peut cliquer sur l'une des formes pour obtenir les formes correspondantes</p>	<p>利用者はこれらのレンマの一つをクリックすると、これに対応する形式を手にすることができます。</p>
<p>D102. Par exemple en cliquant sur le verbe « marcher », nous obtenons tout le tableau de conjugaison de ce verbe.</p>	<p>例えば marcher をクリックするとこの動詞の活用表を手にすることができます。</p>
<p>D103. La ressource Neoveille La dernière ressource que je vais vous présenter est Neoveille</p>	<p style="text-align: center;">リソース・ネオヴェューユ</p> <p>最後のリソース・ネオヴェューユを紹介します。</p>

<p>D104. Toute langue vivante est en constante évolution et une des marques les plus visibles de cette instabilité est le renouvellement de son lexique, avec la disparition de mots obsolètes (peu étudiés) et l'apparition de nouvelles unités : les néologismes.</p> <p>On estime à environ 10% le nombre moyen de mots inconnus dans les textes de langue générale. Une grande partie de ces mots sont des noms propres et des erreurs typographiques, mais également des néologismes proprement dits.</p> <p>L'étude des néologismes est incontournable dans l'étude d'une langue.</p>	<p>すべての生きた言語は常に進化しており、この不安定さの最も目に見える特徴の一つは、すたれた語（ちなみにこれは研究されていないが）の消失と新しい語の出現が引き起こす語彙の更新です。</p> <p>一般的な言語のテキストの中における知らない語の平均的な数は約10パーセントと見積もられます。このような語の大部分は、固有名詞、印刷上の誤り、また同様にいわゆる新語です。</p> <p>新語の研究は言語の研究において避けられません。</p>
<p>D105. Pluridisciplinaire, le projet associe des compétences en informatique (TAL) en linguistique (lexicologie, terminologie, sémantique) et en sociolinguistique. Collaborations entre 4 laboratoires USPC et un partenaire international (univ. de Sao Paulo).</p> <p>Neoveille une plateforme de veille multilingue destinée au repérage et au suivi des néologismes en 7 langues à partir de flux de documents en provenance du web et de corpus diachroniques préexistants.</p> <p>Ce projet est innovant : aucune plateforme de repérage ni de suivi des néologismes à partir de flux de corpus n'existe actuellement, sauf pour l'anglais.</p>	<p>この学際的なプロジェクトには言語自動処理、言語学（語彙論、術語論、意味論）、社会言語学の専門家を結集しています。これはソルボンヌ・パリ・シテと呼ばれるUSPC 大学連合の中の4研究所¹⁶とサンパウロ大学の計5機関の間の共同研究です。ネオヴェューユは7つの言語¹⁷の新語の認識と追跡を目的とした多言語監視プラットフォームで、ウェブと既存の通時的コーパスに由来する多量の資料¹⁸からデータを集めています。</p> <p>このプロジェクトは革新的です。現在、このようなあふれんばかりの資料からデータ収集が可能な（新語）認識・追跡プラットフォームは、英語を除けば存在していません。</p>

<p>D106. L'une des fonctionnalités de Neoveille consiste en une « absorption » quotidienne des articles de presse.</p> <p>Cette étape est suivie d'un repérage automatique des formes potentiellement néologiques (néologismes candidats) : grâce un corpus d'exclusion, créé à l'aide des dictionnaires électroniques du laboratoire (Morfetik, FixISS, ...)</p> <p>Notre tâche consiste ensuite à :</p> <p>(i) valider des néologismes (ii) procéder à une étude du type de néologisme</p>	<p>ネオヴェューの機能の一つは日刊紙の社説をいわばく吸い取るように取り込める点にあります。</p> <p>次に潜在的に新語的な形式（つまり新語候補）の自動認識が続きます。この認識機能は（Morfetik や FixISS のような）研究所が構築したコンピュータのための辞書に基づいて作られる排除コーパスのおかげなのです。</p> <p>今後、私たちの研究は以下のような発展を考えています：</p> <p>(i) 新語の妥当性を検証する (ii) 新語のタイプの研究を行う</p>
<p>D107. Nous avons 3 objectifs :</p> <ul style="list-style-type: none"> • Mettre en place une plateforme multilingue de veille et de suivi des néologismes à partir de corpus diachroniques de très grande taille ; • Utiliser cette plateforme pour mener une étude des emprunts (notamment anglicismes) dans différentes langues (français, portugais du Brésil, chinois, russe, tchèque, polonais et grec) ; • Utiliser cette plateforme pour étudier la notion d'innovation sémantique et pour proposer de nouvelles mesures pour identifier les nouveaux emplois. 	<p>私たちには3つの目的があります。</p> <ul style="list-style-type: none"> • 一つは、大変大きな通時的コーパスに対して、新語を監視・追跡する多言語プラットフォームを設置することです。 • もう一つは、プラットフォームを利用して、様々な言語（フランス語、ブラジルのポルトガル語、中国語、ロシア語、チェコ語、ポーランド語、ギリシャ語）における借用語（とりわけ英語特有語法）の研究をすることです。 • 最後に、プラットフォームを利用して、意味革新の概念を研究し、語の新しい用法を認定するための新しい手段を提案することです。
<p>D108. Nous voyons ici des exemples de néologisme que la ressource a trouvé. Ce sont des néologismes qui ont été construits à partir du nom du président des Etats-Unis TRUMP. Le</p>	<p>ここにネオヴェューが見つけた新語の例があります。それはアメリカ合衆国の大統領トランプの名前から作られた新語です。この名詞は大変生産的です。Trumpmania, trumpiser, trumpiste など。</p>

nom est très productif : Trumpmania ; trumpiser ; trumpiste, etc.	
D109. Dans la 1ere colonne, on trouve le néologisme	この画面を説明します。左の列から新語です。
D110. Le type de néologisme	次の列は新語のタイプです。
D111. La reconnaissance automatique de la catégorisation	次は、新語の範疇化の 自動認識 です。
D112. La fréquence d'utilisation du néologisme	次は、新語の頻度です。
D113. La date de la 1ere apparition du néologisme.	次は新語が最初に現れた日付です。
D114-115. En cliquant sur la croix verte, on obtient le contexte d'apparition du néologisme.	緑の十字のボタンをクリックすると、新語の出現したテキストの抜粋が入手できます。
D116-117. Enfin, on peut avoir diverses informations statistiques sur le néologisme : avec par exemple la répartition par domaine (ici domaine général ou domaine de la presse), la répartition par journaux.	最後に、新語に関する様々な統計的情報が得られます。例えば、分野ごとの統計（ここでは一般的な分野か出版物の分野か）、新聞ごとの統計など。
D118-119. En conclusion <ul style="list-style-type: none"> • Nous avons présenté des ressources basées sur des données linguistiques fines et des outils permettant la manipulation de ces données • Les outils proposés répondent à des tâches de traduction ou de rédaction (aide à la traduction automatique, recherche d'informations sur le Web, aide à la rédaction...) • Caractéristiques des travaux : volonté d'élaborer des dictionnaires couvrant la totalité du lexique d'une langue, au regard de propriétés explicites et reproductibles qui puissent résister à l'exigence d'un traitement informatique. 	結論 <ul style="list-style-type: none"> • 私たちは厳密で・正確さを備えた・言語的データに基づくリソースとこれらのデータを取り扱うことができるツールを紹介しました。 • 紹介したツールは翻訳または作文の仕事の要望に応えるものです（翻訳支援、ウェブ上での情報検索、作文支援など）。 • 仕事の特徴は、情報処理の要求に応えることができる明示的かつ再現可能な特性という観点から、言語の語彙全体をカバーする辞書を、強い意志をもって、開発している点です。

<p>• Ces ressources ont reçu une aide financière de la Délégation Générale à la langue française et aux langues de France (DGLFLF), qui est, en France, un service rattaché au Ministère de la Culture et de la communication.</p>	<p>• これらのリソースは、フランスにおいて文化コミュニケーション省に所管されている「フランス語とフランスで話されている諸言語のための一般委員会」から研究支援を受けています。</p>
<p>D120. Je vous remercie pour votre attention et j'espère que ces travaux vont ouvrir la voie à de nouvelles collaborations avec le Japon.</p>	<p>この発表が日本との新しい共同研究への道を開いてくれることを切に願ってやみません。 ご静聴ありがとうございました。</p>

注

- 1 この講演会は、科学研究費助成基金 (JSPS Kakenhi) 若手 (B) 課題番号 16K16822「フランス語の知覚動詞に関する凝結表現の研究」(研究代表者: 木島愛) の助成を受け実現したものである。
- 2 以下の講演記録の中の左上の D1、D2・・・はパワーポイントのスライドに対応している。本稿では、スライドは割愛するので、形式的な記号にすぎないが、全体の構成がつかみやすくなるのでそのままにしておいた。
- 3 現在この表現は<待ちぼうけを食わす>(字義通りの意味は「ウサギを置く」という意味であるがどうしてこのような成句が誕生したのか。これを理解するためには、18世紀のフランスではウサギといえば狭い小屋に飼われる生き物というイメージがあったことを想起しなければならない。このイメージは<定員オーバーの乗客>(1783)、<馬車の御者が、乗車登録されていない客からお金を受け取って自分の懐にしまう>(1876)という悪いイメージと結びつく。こうしたイメージが廃れると、今度は「女にウサギをプレゼントする」<喜びを与えてくれた女に代金を支払わない>という表現が現れる。この最後の表現が「ウサギを置く」という表現に置き換えられ、現代の意味に発展して行ったようである(1888)。これは『*Dictionnaire historique de la langue française*』(1995)の説明であるが、こうした権威ある辞書をもってしても、現代の<待ちぼうけを食わす>という意味の誕生は未解明のままだと言っても間違いではないだろう。だからこそ、この成句は多くのフランス人が関心を持ち、議論の対象となり続けているのだろう。
- 4 このスライドは、直前のスライドにアニメーションをつけたものなので空欄になっている(以下同様)。
- 5 対象部類とは述語を定義するために設定される語の下位集合のこと。例えば、日本語で「星空を見る」と言った時、「見る」は「目で事物の存在を捉える」という意味だが、「答案を見る」と言った時は「採点する」という意味である。これは日本人なら皆知っているが、機械にこの違いを教えるためには、「見る」の後ろの目的語の位置に「自然の類」を表す単語が来たら前者の意味、「答案の類」が来たら後者の意味と決めておかなければならない。ただ単に「見る」の後ろにこの二つを包摂する上位の集合、例えば「物の類」を設定しておくだけでは区別できない。対象部類は意味的に同類の単語の集合であり、意味的に同類の物体の集合ではないことに注意しなければならない。この概念は、90年代の初めに、ミシェル・マティユ=コラ(Michel MATHIEU-COLAS)によって提唱された(マティユ=コラ 2007 参照)。
- 6 凝結解消とは、凝結表現が一時的に解消され、普通の表現としての解釈も可能になる現象のこと。凝結表現は一塊(かたまり)で一つの特異な意味を持つが、普通の表現としては、個々の要素が意

味を持つ。「油を売るな！」を普通の会社の社長が社員に言う場合と、ガソリンスタンの社長が社員に言う場合で違う。前者は凝結解消が起きやすい例、後者は凝結解消が起きない例である。凝結解消は、凝結表現が一時的に解消する現象であり、言葉遊びや文体的効果を高めるために利用される。

- ⁷ 変異とは、基本形からの、形の上でへだたりないしはずれが生じているものを指す。
- ⁸ *estar* は一般に既知の人・動物・事物についてそれがどこにあるのかを示す。*ser* は主語の性質や属性を表す。
- ⁹ 並行コーパスとは同じ内容のものが2言語以上で、対訳の形で存在しているコーパスのことで、その規模が膨大になればなるほど、精度の高い機械翻訳が可能になる。
- ¹⁰ ピボット言語とは、話者の人口が少ない言語の短所を補うために使われる仲介的な言語のこと。
- ¹¹ 例えばツイートのようなオンライン上に書かれた文書。
- ¹² 意味、語の使用レベル、対象部類 (*classe d'objets*) のレベルで判断。
- ¹³ 語の基本形。語が文の中で実現する様々な形をまとめる抽象的でかつ集約的な単位。実際にテキストに実現した形である実現形に対立する概念。
- ¹⁴ 語を書き表した基本形の全て。屈折形は数えない。レンマとほぼ同じ。
- ¹⁵ 拡張可能マークアップ言語。Extensible Markup Language の略。コンピュータ間でデータ通信を行うためのデータ形式の一種。
- ¹⁶ USPC 大学連合の中の4研究所とは、具体的以下の機関を指す。
- LIPN (=Université Sorbonne Paris Cité: Laboratoire d'Informatique Paris Nord「北パリ情報科学研究所」)
 - LDI(UP)
 - CLILLAC-ARP(UP7)(= Centre de Linguistique Inter-langues, de Lexicologie, de Linguistique Anglaise et de Corpus-Atelier de Recherche sur la Parole 「言語間・語彙論・英語学・話言葉研究アトリエ・コーパスセンター」)
 - ERTIM (INALCO) (ERTIM = Equipe de Recherche Textes, Informatique, Multilinguisme 「テキスト・情報科学・多言語主義研究グループ」, INALCO = Institut National des Langues et Civilisations Orientales)。
- ¹⁷ フランス語、スペイン語、ポーランド語、チェコ語、ロシア語、ポルトガル語、中国語を指す。
- ¹⁸ 例えば fluxRSS などのウェブサイト。

参考文献 (著書)

- Gross, G., 1996, *Les expressions figées en français*, Ophrys.
- Gross, G., 2012, *Manuel d'analyse linguistique*, Presses Universitaires du Septentrion, Université de Lille 3, Lille.
- Mathieu-Colas, M., 2007, «Domaines et classes sémantiques», *Verbum*, t. XXIX, n° 1-2, Presses universitaires de Nancy.

参考文献 (辞書)

- Rey, A., 1995, *Dictionnaire historique de la langue française*, Le Robert.
- Rey, A., Chantereau, S., 2007, *Dictionnaire des expressions et locutions*, Le Robert.