

CABRÉ, M. T.; DOMÈNECH, M.; ESTOPA, R.; FREIXA, J.; SOLÉ, E. (2003) «L'observatoire de néologie: conception, méthodologie, résultats et nouveaux travaux». En: Sablayrolles, J. F. (ed.) *L'innovation lexicale*. Paris: Honoré Champion Éditeur, p. 125-147. ISBN: 2-7453-0800-9 / ISSN: 1279-8207 (CL).

L'OBSERVATOIRE DE NÉOLOGIE: CONCEPTION, MÉTHODOLOGIE, RÉSULTATS ET NOUVEAUX TRAVAUX¹

1. Introduction

Objectifs

Il y a plusieurs objectifs à cet exposé. Dans la première partie, nous présentons, d'abord, l'Observatoire de Néologie², un projet de travail en néologie ayant douze ans d'expérience; ensuite, nous réfléchissons sur quelques problèmes théoriques et méthodologiques parmi lesquels nous soulignons la notion même de néologisme et la sélection des matériaux pour la documentation des nouvelles unités; et, finalement, nous présentons les applications et les résultats de la recherche en néologie. Dans la deuxième partie, nous présentons une piste de travail que nous venons d'entreprendre: la néologie dans des textes oraux. D'abord, nous présentons les conséquences méthodologiques et

¹ Ce document a été traduit du catalan par Helena Pàmpols, boursière de recherche à l'Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra, et supervisé par Lydia Fernández, maître de conférence de la Facultat de Traducció i Interpretació de la Universitat Pompeu Fabra.

² Pour plus d'information sur l'Observatoire de Néologie (membres, histoire, publications, etc.) veuillez consulter <http://www.iula.upf.es/obneo>

quelques résultats de cette nouvelle piste et, finalement, nous finirons notre exposé par la présentation de possibles applications de la néologie orale.

L'Observatoire de Néologie, qu'est-ce que c'est?

L'Observatoire de Néologie est un groupe actuellement rattaché à l'Institut Universitaire de Linguistique Appliquée (Université Pompeu Fabra) qui analyse depuis 1988 le phénomène de l'apparition de mots nouveaux ou néologismes dans la presse à large diffusion, en catalan et en espagnol.

L'Observatoire, dirigé par M. Teresa Cabré, est constitué par une quarantaine de collaborateurs provenant de différentes universités catalanes (mais, fondamentalement de l'Université Pompeu Fabra), licenciés ou étudiants en dernière année de Philologie ou de Traduction et Interprétation. Les principaux objectifs de ce projet sont les suivants:

- Disposer d'un corpus de néologie lexicale à large diffusion en catalan et en espagnol
- Décrire et analyser les ressources qu'utilisent le catalan et l'espagnol pour actualiser leur lexique
- Diffuser périodiquement les nouvelles créations lexicales
- Contribuer à l'actualisation du lexique contenu dans des dictionnaires de langue générale
- Etablir une plate-forme qui permette l'échange avec d'autres groupes de néologie

2. La néologie dans les textes écrits

Comme nous avons mentionné, depuis 1998 l'Observatoire travaille avec des textes de la presse à large diffusion. Toutefois, plus récemment, l'Observatoire a commencé à travailler avec des textes oraux. On abordera le travail concernant le dépouillement des textes

oraux dans la deuxième partie de cette présentation. Dans celle-ci, nous n'abordons que le dépouillement de la presse écrite.

2.1. Quelques aspects méthodologiques à relever:

La notion de néologisme

Le critère avec lequel l'Observatoire de Néologie travaille pour établir le caractère néologique des mots est le critère lexicographique: on considère comme néologique n'importe quel mot qui n'apparaît pas dans le corpus lexicographique d'exclusion. Ce critère, avec lequel on a travaillé depuis le début pour les textes écrits, a été aussi adopté pour les textes oraux.

Compte tenu des caractéristiques et des objectifs du projet, on a écarté le critère psychologique et le critère temporel et on a choisi le critère lexicographique, mais on a établi quelques restrictions afin que les données résultantes puissent s'adapter plus largement à nos nécessités. Ainsi, a-t-on considéré comme des néologismes, les mots qui sont signalés comme tels par un astérisque ou toute marque équivalente, même s'ils sont répertoriés dans notre corpus d'exclusion. En revanche, on n'a pas considéré comme des néologismes les augmentatifs, les diminutifs et les superlatifs, même s'ils n'apparaissent pas dans les ouvrages d'exclusion, à cause de leur potentialité dérivative quasi illimitée, de même pour les adverbes en *-ment*, les mots formés par le préfixe *ex-* quand il s'ajoute à des radicaux simples ou dérivés qui font référence à des professions ou des relations personnelles, entre d'autres.

En dehors de ces aspects que nous venons d'exposer, il est évident que le critère lexicographique permet de recueillir, par an, comme néologismes de nouvelles unités qui ne représentent aucune nouveauté parce qu'elles sont parfaitement intégrées dans la langue, mais pour des raisons diverses, elles n'apparaissent pas dans les dictionnaires.

Ainsi, *western* est, depuis longtemps, un des néologismes les plus fréquents et stables pour le catalan.

Le corpus à dépouiller

Comme l'on a dit, les sources à dépouiller sont des journaux à large diffusion et l'Observatoire en dépouille à peu près cent cinquante pour les deux langues. Il est important de remarquer que ces journaux sont d'information générale et, donc, que les néologismes que nous recueillons ont un caractère général et non pas spécialisé.

Pour le catalan, on dépouille l'*Avui* et *El Periódico* et pour l'espagnol *La Vanguardia* et *El País*, mais, étant donné que les articles passent par des services de révision et de correction, et qu'ils peuvent, donc, avoir perdu une partie de leur potentiel néologique, on a commencé, depuis un certain temps, à recueillir et sélectionner d'autres textes écrits, de caractère général, qui n'ont pas été vérifiés par des services linguistiques.

Le corpus d'exclusion

L'établissement du corpus d'exclusion détermine les néologismes obtenus après le dépouillement des textes; si les dictionnaires sélectionnés sont des ouvrages normatifs, la quantité de néologismes qu'on obtiendra du dépouillement sera plus élevée que si on prend des ouvrages descriptifs, avec une nomenclature plus large. Depuis le début, l'Observatoire travaille avec un corpus qui combine des ouvrages normatifs et descriptifs et ce corpus a été actualisé avec la parution de nouveaux dictionnaires.

À ce jour, les dictionnaires qui constituent le corpus d'exclusion pour le catalan, aussi bien pour les textes écrits que pour les textes oraux, sont les suivants:

- *Diccionari de la llengua catalana* (1995, Institut d'Estudis Catalans)
- *Gran diccionari de la llengua catalana* (1998, Enciclopèdia Catalana)

Et pour l'espagnol, on travaille avec les dictionnaires suivants:

- *Diccionario de la lengua española* (1992, Real Academia Española)
- *Diccionario General Ilustrado de la Lengua Española* (1987, Alvar Ezquerro)³

2.2. Voies d'automatisation des données

Pendant les premières années, le dépouillement des journaux a été fait manuellement au moyen des fiches et l'automatisation de celles-ci a été faite par un boursier du groupe de recherche. Toutefois, dès le début, on a commencé à développer des outils nous permettant de faire le dépouillement d'une façon plus assistée.

La détection automatique

Le Sextan est un outil logiciel qui a pour objectif de faciliter au maximum la tâche du néologue et qui a été créée à partir de la réutilisation des modules de traitement du Corpus Technique de l'IULA. Le Sextan permet la détection automatique de la néologie formelle monolexématique.

Dans la première phase du traitement, on récupère massivement les journaux en format électronique et on fait le marquage structurel HTML. Le Sextan considère un néologisme potentiel toute unité

³ Dans la détection automatique on utilise une version électronique avec l'ensemble des lemmes actualisés, qui contient plus de 4.000 nouvelles entrées.

lexicale que l'analyseur ne reconnaît pas dans le dictionnaire électronique; le résultat est une liste de mots non reconnus (complétés par des informations, comme les contextes d'apparition de chaque unité, la section du journal, etc.) que le néologue doit vérifier.

Finalement, le Sextan propose au néologue une interface graphique qui permet la vérification et la sélection des néologismes, la sélection du contexte d'apparition le plus remarquable et la fiche est complétée par l'inclusion de celle-ci dans la base de données.

Le dépouillement assisté

Indépendamment du fait que la plupart du dépouillement est déjà fait par voie automatique comme on vient de l'expliquer, les collaborateurs de l'Observatoire continuent à dépouiller de façon manuelle, mais plus assistée, car c'est la seule manière, jusqu'à présent, de couvrir les principales lacunes de l'autre méthode: la néologie sémantique et la néologie formelle syntagmatique.

On dit qu'actuellement le dépouillement est "assisté" car on ne travaille plus avec des fiches en papier remplies manuellement, puis introduites dans la banque de données; des collaborateurs ont accès par Internet à une fiche-modèle qu'ils remplissent avec chaque néologisme et qui va directement à la banque de données.

2.3. Quelques résultats

Nous présentons, ci-dessous, quelques résultats que l'on a obtenus, conformément aux objectifs fixés par l'Observatoire : ceux-ci concernent aussi bien la néologie écrite que la néologie orale, même si les résultats présentés font référence aux douze premières années du travail sur des textes écrits.

Les banques de données

On a dit que le premier objectif du projet est celui de disposer d'un corpus de néologie lexicale à large diffusion en catalan et en espagnol. En effet, les résultats les plus importants de ces douze années de travail sans interruption sont les banques de données de néologismes que l'on a créées. Actuellement on dispose des quantités suivantes:

	Catalan	Espagnol ⁴
Occurrences	78.055	50.523
Unités	28.789	21.990

Chaque néologisme est recueilli une seule fois dans chaque journal (et on laisse de côté les autres occurrences de ce même néologisme dans ce même journal), mais on le recueille si le néologisme apparaît dans un nouveau journal dépouillé. Cela explique que le nombre d'unités, qui fait référence au nombre réel de néologismes différents pour chaque langue, soit inférieur au nombre d'occurrences, qui répond aux répétitions d'un même néologisme dans des journaux différents. Le fait d'avoir travaillé avec ce critère nous a permis d'obtenir une information sur la fréquence relative des différentes unités au cours d'une même année; et ce critère nous a permis d'observer le mouvement de fréquence au cours des années, ce qui nous semble le plus important.

Le tableau suivant montre quelles ont été les unités ayant une fréquence supérieure à 100 pour le catalan et pour l'espagnol dans la période 1998-1999 :

Catalan	Espagnol
---------	----------

⁴ Pour l'espagnol, les quantités sont plus basses parce que, bien qu'on travaille dans les deux langues, le catalan (la seule langue pour laquelle on a obtenu une aide financière) a toujours été prioritaire et on a fait plus de dépouillement dans cette langue.

<i>thriller</i> , 182 <i>miscampista</i> , 157 <i>western</i> , 148 <i>rànquing</i> , 102	<i>antiterrorista</i> , 132 <i>thriller</i> , 120
--	--

On peut observer que *thriller* est la seule unité commune deux langues, et il est intéressant de remarquer que dans les autres cas, les unités (qui n'en sont pas, pour autant, moins fréquentes) ne sont pas des néologismes puisque ces mots apparaissent dans le corpus d'exclusion. Dans le cas de *rànquing*, par exemple, l'unité n'est plus néologique car elle a été introduite dans les dictionnaires.

Même si les unités du tableau sont les plus fréquentes au cours de la période mentionnée, elles ne l'ont pas été chaque année; l'an 2000, par exemple, a comme unité la plus fréquente *lehendakari* pour les deux langues, suivie par *kale borroka* et *play-off* pour le catalan, et par *centroizquierda* et *web* pour l'espagnol.

L'actualisation lexicographique

La contribution à l'actualisation du lexique contenu dans les dictionnaires de langue générale est le deuxième objectif visé, objectif qu'on a atteint d'une manière plus nette pour le catalan, bien que l'on doive encore continuer dans le même sens.

Les néologismes de l'Observatoire de Néologie ont servi à la sélection de nouvelles entrées et sous-entrées, et de nouvelles acceptions d'entrées déjà existantes. Ainsi, les données du projet ont-elles été utilisées pour l'élaboration du *Diccionari de la llengua catalana* (1995, Institut d'Estudis Catalans) et pour l'élaboration du *Gran diccionari de la llengua catalana* (1998, Enciclopèdia Catalana).

Le nombre de sous-entrées provenant de l'Observatoire et incorporées dans un de ces ouvrages, est très bas et il s'agit d'unités sélectionnées à partir de critères de nécessité interne du dictionnaire

(fondamentalement l'incorporation de néologismes formés par des syntagmes qui sont incorporés comme sous-entrées d'entrées lexicographiques déjà existantes) et non pas à partir du critère de fréquence, comme on peut voir dans les exemples ci-dessous:

príncep blau (3) *cotxe bomba* (24) *quart món* (1)

Quant aux nouvelles acceptions incorporées dans les dictionnaires actualisés avec les données de l'Observatoire, elles correspondent à des néologismes formés par des processus sémantiques à partir d'entrées lexicographiques déjà existantes, soit par homonymie et polysémie, soit par nominalisation d'adjectifs:

- *marató* (3): Programme de radio ou télévision, de longue durée, qui recueille de l'argent pour des causes bénéfiques grâce aux dons des auditeurs ou des téléspectateurs.
- *torejar* (1): Se moquer de quelqu'un, railler quelqu'un.
- *disjuntiva* (4): Alternative entre deux possibilités, dont une doit être nécessairement choisie.

Toutefois, la plupart des dictionnaires ont utilisé les données de notre corpus afin d'incorporer de nouvelles entrées; fondamentalement on a incorporé des noms, mais aussi des adjectifs, quelques verbes et très peu d'adverbes ou de locutions.

Du point de vue du type de néologisme, on a incorporé quelques emprunts de diverses langues, mais fondamentalement on a incorporé des néologismes formels et, en plus, on peut remarquer la prédominance des néologismes suffixés (plus de la moitié des exemples).

Divers sont les critères d'incorporation utilisés par les dictionnaires pour l'inclusion de nouvelles entrées. Tout d'abord, le besoin dénomiatif de l'unité, surtout pour les emprunts et néologismes de langue; puis, les synonymes et les variantes très productives; enfin,

l'ajout de familles de mots déjà existants dans le dictionnaire ou l'inclusion de familles entières.

La description et l'actualisation grammaticale

Quant au troisième objectif, c'est-à-dire, la description et l'analyse des ressources qu'utilisent le catalan et l'espagnol pour actualiser leur lexique, il faut souligner deux lignes de travail. D'un côté, les travaux des différents membres du projet portant sur l'évaluation de la vitalité de la langue à partir de l'analyse des ressources d'où sont créées de nouvelles unités lexicales. Et, de l'autre, les travaux portant sur le contraste des processus de formation de mots que décrivent les grammaires et les processus qu'on observe dans les données néologiques, travaux élaborés par la plupart des membres de l'Observatoire qui ont travaillé ces deux dernières années sur ce sujet.

L'Observatoire de Néologie va recevoir une importante reconnaissance de la tâche de description grammaticale qui est en cours car, dans un acte public, l'Institut d'Estudis Catalans a annoncé que cet organisme tiendra compte de ces travaux pour la grammaire normative qui est en cours d'élaboration.

La diffusion des données

L'accomplissement du quatrième objectif, c'est-à-dire, la diffusion périodique des nouvelles créations lexicales, se fait de différentes manières, mise à part celle que nous venons de commenter avec l'actualisation des dictionnaires. La plus exhaustive est la publication annuelle de la liste complète des néologismes documentés chaque année.

En outre, en 1998, le projet a élaboré le *Diccionari de paraules noves*, un dictionnaire qui contient les mille unités néologiques les plus fréquentes recueillies entre 1988 et 1996 et qui contient différentes

informations linguistiques, pragmatiques et encyclopédiques et ce, pour chaque unité.

On est en train de travailler, actuellement, sur une nouvelle voie de diffusion des donnés, qui permettra de consulter en ligne les unités documentées.

L'échange avec d'autres groupes

Etablir une plate-forme qui favorise l'échange avec d'autres groupes de néologie est le dernier des principaux objectifs prévus par l'Observatoire de Néologie. Même s'il est vrai que le projet a maintenu des contacts avec d'autres groupes de néologie dès pratiquement son début (avec le Centre de Terminologie et Néologie de l'Université de Paris XIII, avec des groupes de travail de néologie de l'Universidade Nova de Lisboa, de l'Université de Tessalonique et de l'Université de Paris-Sorbonne —Paris VIII—, par exemple), il est aussi vrai qu'au fur et à mesure que le projet est devenu plus important et que de nouveaux groupes de travail en néologie sont apparus, ces contacts ont augmenté.

De cette façon, l'Observatoire collabore étroitement avec l'Observatorio de Neoloxía pour la langue galicienne (Université de Vigo) et avec l'observatoire qui est en train de se créer pour l'italien (Universita La Sapienza, Rome).

Depuis un certain temps, l'Observatoire gère un projet de suivi et de contraste de la néologie économique dans la presse, dans le cadre de Realiter (Réseau panlatine terminologie), un réseau formé par des groupes de toutes les langues romanes.

3. La néologie dans les textes oraux

Comme nous venons de dire, pendant les dix premières années de fonctionnement, l'Observatoire de Néologie a analysé la néologie lexicale en catalan et en espagnol, apparue dans la presse à large diffusion, de façon que les données obtenues informen sur un usage réel et particulier de la langue qui reflète des usages néologiques non spontanés, puisque la plupart des textes des journaux sont révisés par les services de correction. Dans le but d'enrichir et de diversifier les bases néologiques et de permettre l'observation de la création spontanée, depuis 1999 l'Observatoire a commencé à analyser la néologie provenant des textes oraux et des textes écrits les plus informels. Quant aux textes écrits les plus spontanés, le dépouillement se trouve encore dans une phase très embryonnaire, mais, pour la langue orale, nous avons déjà une méthodologie de travail propre et nous disposons des premiers résultats qui nous permettent de définir ses spécificités et de prévoir ses différentes applications dans l'avenir.⁵

La méthodologie générale de travail (concept de néologisme, corpus d'exclusion, procédé et diffusion des données) est partagée par tous les groupes de travail de l'Observatoire, mais les documents internes du projet, à partir desquels ont été établis les processus d'identification, enregistrement et analyse des données, s'adaptent aux spécificités de chaque type de dépouillement. Dans ce sens, le dépouillement de textes oraux que mène l'Observatoire peut être décrit, d'une manière comparable au dépouillement de la presse écrite, comme un procédé qui consiste à identifier les néologismes employés habituellement à la radio et à la télévision, conformément aux critères de sélection fixés auparavant, et à les enregistrer sur des fiches manuelles selon des normes de rédaction établies. L'objectif fondamental du dépouillement de textes oraux, dans cette première étape de fonctionnement correspondant à la période 1999-2000, a été l'élaboration d'un protocole propre d'élaboration de la fiche de dépouillement à partir des données extraites

⁵ Les personnes qui ont fait partie du Groupe de Néologie Orale pendant cette première étape de fonctionnement sont les suivantes : Meritxell Domènech, Rosa Estopà, Maite Ferré, Cristina Gómez, Cristina Mayoral, Noelia Palacios, Cristina Peña, Sara París et Mireia Raymi.

de programmes de radio à large diffusion (informations, débats, documentaires).

3.1. Spécificités du dépouillement des textes oraux

Le fait que la méthodologie générale de travail soit partagée par tout l'Observatoire, implique que le critère de sélection néologique commun à tous les groupes de dépouillement soit le critère lexicographique. Celui-ci, comme on a dit, consiste, *grosso modo*, à considérer comme néologiques toutes les unités lexicales identifiées dans le corpus de dépouillement et qui n'apparaissent pas dans un corpus lexicographique d'exclusion, formé par des ouvrages normatifs et descriptifs.

La conséquence fondamentale de l'application de ce critère sur le dépouillement de textes oraux a été de s'attacher exclusivement à l'observation de la néologie lexicale et d'écarter, par conséquent, les formes néologiques qui peuvent être expliquées par des questions strictement phonétiques (sons adventifs ou de support, propres au registre oral) ou dialectales (ex.: *pa*i*ella* au lieu de *paella*; *així*ns** au lieu de *així*)⁶.

D'après ce que nous venons de dire, pour l'élaboration de la fiche de dépouillement de textes oraux nous sommes partis de la fiche de dépouillement de la presse écrite et nous avons supprimé ou adapté les champs qui ne sont pas nécessaires ou pertinents pour la néologie orale. Quant à l'information que nous obtenons à propos des néologismes relevés, on peut distinguer, donc, deux grands groupes: celui qui concerne l'information recueillie pour tous les néologismes (oraux et écrits) et celui qui concerne l'information spécifique de la néologie orale.

⁶ Il faut souligner le fait que, bien que la prévision soit de faire le dépouillement des textes oraux en catalan et en espagnol, conformément aux objectifs généraux de l'Observatoire, dans cette première phase de travail de néologie orale, nous n'avons travaillé qu'avec des textes en catalan. C'est pour cela, que les exemples que nous exposerons tout le long de cette communication seront toujours en catalan.

Les informations recueillies pour tous les néologismes, oraux ou écrits, relevés à l'Observatoire sont les suivants:

l'entrée lématisée, la catégorie grammaticale, le type de néologisme (formel, syntagmatique, sémantique ou emprunt), les dictionnaires consultés, le contexte d'apparition, la référence complète de la source et, s'il est nécessaire, une note avec des informations sur un certain aspect du néologisme identifié comme tel (orthographique, sémantique, contextuel, rémissions, etc.).

Les informations spécifiques de la néologie orale qui présentent, d'après nous, un intérêt pour l'étude de la néologie lexicale et que nous expliquerons maintenant plus en détail, sont au nombre de trois :

la transcription phonétique, les marques métalinguistiques et les caractéristiques de l'émetteur.

Quant à **la transcription phonétique** des néologismes repérés, il est intéressant de la recueillir fondamentalement dans le cas des emprunts d'autres langues, afin de voir s'il existe des prononciations ou des hésitations chez un même émetteur ou chez différents émetteurs et afin de savoir jusqu'à quel point, l'emprunt s'adapte à la prononciation de la langue propre (ex.: *revival* /ri'baɪbal/, *e-mail* /íméil/, *crêpe* /krép/); et aussi il est intéressant de la recueillir dans le cas des néologismes de composition savante, car la transcription phonétique nous permet de savoir si certains sons (plus concrètement les voyelles de liaison) se neutralisent ou pas (ex.: *albanokosovar*, *teleaddicte*, *internet*). Cette information peut être utilisée pour l'actualisation lexicographique et grammaticale de la langue, puisqu'elle peut nous donner des pistes sur l'usage réel des emprunts et leur niveau d'introduction dans la langue propre ou sur la catégorisation de certains préfixes et suffixes comme toniques ou atones.

L'information concernant les **marques métalinguistiques** contient des commentaires ou des observations qui nous apportent des données sur la perception que l'émetteur a du degré de fixation dans la langue du néologisme documenté. Cette information est, en quelque sorte, la corrélation orale de l'information donnée par les marques typographiques qui accompagnent parfois les néologismes relevés lors du dépouillement des textes écrits (guillemets, ou caractères gras ou italiques).

Dans le cas des néologismes oraux, cette information métalinguistique peut être explicite (ex.: "Acabo d'arribar de Boston i estic desencaixat... Deu ser això del *jet-lack*, com diuen els anglesos."); "...diguéssim, que no té gaires miraments a l'hora d'encolomar-te feina. Vaja, que és un *jeta*, com diu el jovent d'avui dia.") ou implicite (un changement de voix ou, même, de gestes quand la source de dépouillement est un média comme la télévision).

Finalement, il faut souligner qu'il est important, pour la néologie orale, d'explicitier les **données relatives à l'émetteur**. Dans le cas de la néologie écrite identifiée dans la presse, la fiche indique seulement si le néologisme se trouve dans un texte signé (articles d'opinion, lettres au directeur, etc.) ou non signé (chroniques de correspondants, nouvelles des rédacteurs ou d'agence, etc.).

Dans le cas de la néologie orale, on a beaucoup d'information sur la personne qui émet un néologisme, ce qui permet d'obtenir des données pragmatiques sur les néologismes relevés. Et ces données sont, d'après nous, très intéressantes pour les futures études à propos des origines et des motivations de la création néologique des locuteurs d'une certaine langue. Actuellement, nous constatons cinq caractéristiques quant à l'émetteur de chaque néologisme:

- le rôle de l'émetteur: on tient compte fondamentalement s'il s'agit d'un expert dans la matière et si son intervention est spontanée ou préparée.

- l'âge de l'émetteur.
- le sexe de l'émetteur.
- le dialecte de l'émetteur: on distingue, pour le catalan, trois variétés: la centrale, la baléare et l'occidentale et, pour l'espagnol, on a initialement prévu de distinguer essentiellement entre la variété péninsulaire et l'américaine.
- la langue maternelle de l'émetteur: nous pensons qu'il est intéressant d'en tenir compte quand la langue maternelle n'est ni le catalan ni l'espagnol, surtout à cause de l'incidence d'emprunts et de calques qu'utilisent les émetteurs.

3.2. Corpus à dépouiller et premiers résultats

Le dépouillement des textes oraux réalisé au cours de cette première étape de fonctionnement a été fait de façon manuelle et cela nous a permis de détecter les spécificités propres de ce type de dépouillement par rapport à l'écrit et de rédiger une première version du document "Protocole d'usage de la fiche de dépouillement de textes oraux". Ensuite nous avons semi-automatisé ce dépouillement en créant une fiche de dépouillement électroniquement sur Internet qui peut être introduite, une fois remplie avec le néologisme repéré, dans la banque de données de l'Observatoire.

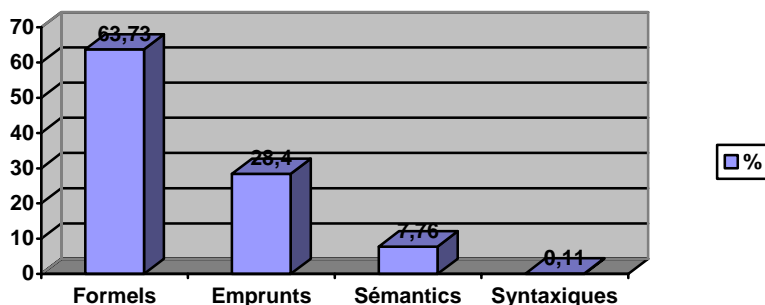
Les sources de dépouillement utilisées dans cette première étape de fonctionnement ont été des programmes radiophoniques émis par Catalunya Ràdio —le poste émetteur officiel du gouvernement de la Catalogne, qui émet totalement en catalan. Les programmes dépouillés ont été sélectionnés parce qu'ils contenaient des espaces radiophoniques spontanés (débat, interviews, etc.) mais aussi d'autres programmes plus contrôlés (nouvelles, chroniques, etc.).

La distribution globale des différentes sources utilisées dans cette première étape du dépouillement de textes oraux est la suivante :

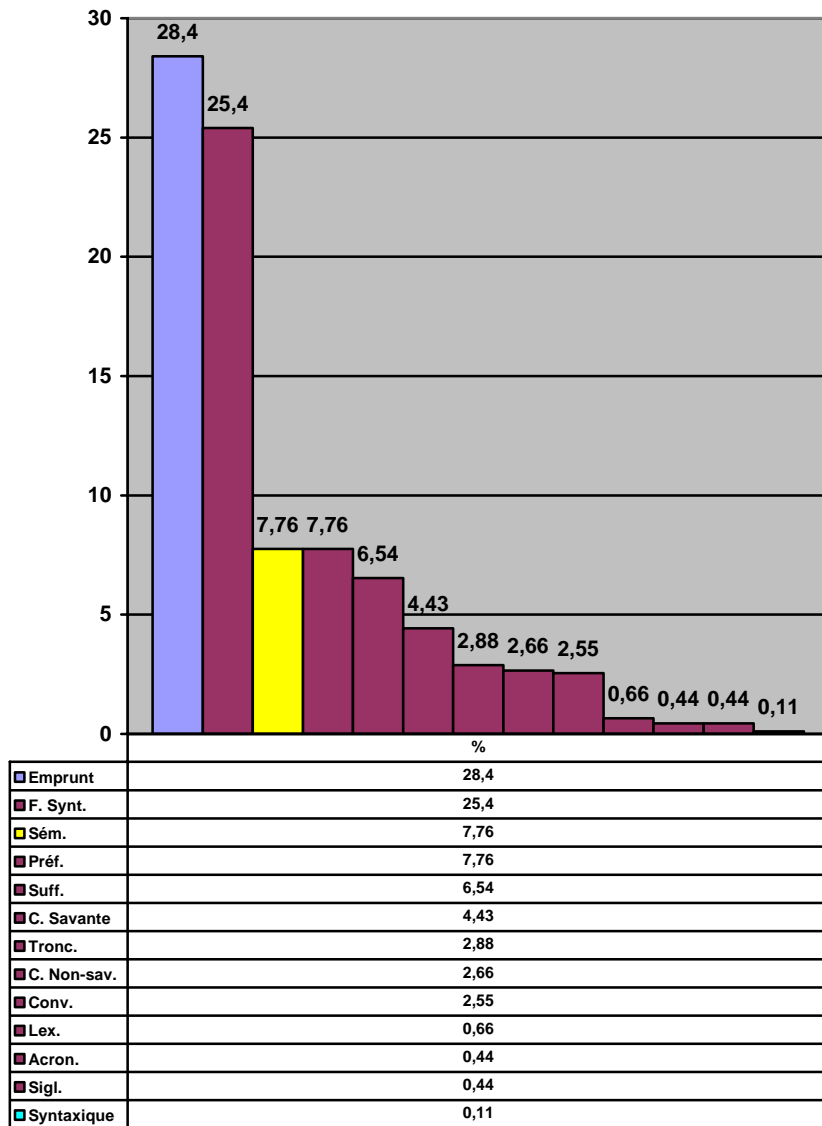
- 23 informations (60%), qui présentent la particularité d'être des programmes assez fermés, avec des sections, qui se répètent chaque jour.
- 12 débats (32%); ce sont de programmes de discussion assez ouverts et dynamiques, qui traitent des thèmes d'actualité où interviennent, par conséquent, des émetteurs avec des caractéristiques très variés.
- 3 émissions du programme "Mites del s. XX" (8%); un programme d'actualité caractérisé par la co-présence d'un documentaire et de commentaires spontanés de personnages invités ayant vécu le thème traité.

Quant aux résultats obtenus, il faut tenir compte du fait que le dépouillement réalisé jusqu'à maintenant n'a pas encore été automatisé, de sorte que les données que nous allons présenter ci-dessous, obtenues manuellement, correspondent à des pourcentages généraux. De même, il faut souligner le fait que pour les néologismes répétés, nous ne recueillons qu'une seule occurrence pour chaque type générique d'émetteur

Aujourd'hui nous disposons d'environ 1000 unités néologiques extraites de textes oraux (sans tenir compte des occurrences répétées), qui se distribuent selon les groupes suivants:



À partir de ce graphique nous pouvons affirmer que la néologie formelle prédomine à l'oral, résultat convergeant avec ceux obtenus à partir du dépouillement de la presse écrite. Mais, si on tient compte de la diversité des processus néologiques groupés sous la dénomination «néologie formelle», les différences entre le registre oral et l'écrit se manifestent plus nettement. Voir ci-dessous un graphique plus détaillé:



Les données du graphique nous permettent de constater que les deux types de formation néologique prédominants sont l'emprunt (28,4%) et le figement syntagmatique (ou néologismes créés par une structure syntaxique lexicalisée) (25,4 %), suivis de loin par les néologismes sémantiques, les préfixés (7,76 %) et les suffixés (6,54%).

Si nous comparons ces données avec celles provenant du dépouillement des textes écrits, en particulier avec les données de 1995,⁷ nous pouvons observer dans le tableau suivant que, dans la néologie écrite, le processus de formation prédominant est la préfixation, suivie par les emprunts et par les formes suffixées, tandis que le figement syntagmatique n'occupe pas un rang prédominant.

TYPES DE NEOLOGISME	ÉCRIT (1995)	ORAL (2000)
EMPRUNT	21,24 %	28,4 %
FIGEMENT SYNTAGMATIQUE	5,53 %	25,4 %
N. SÉMANTIQUE	3,64 %	7,76 %
PRÉFIXATION	23,27 %	7,76 %
SUFFIXATION	17,94 %	6,54 %
COMPOSITION NON-SAVANTE	8,05 %	2,66 %
COMPOSITION SAVANTE	13,41%	4,43 %
N. SYNTAXIQUE	0,45%	0,11 %
TRONCATION	0,99%	3,76%
CONVERSION SYNTAXIQUE	2,89%	2,55%
AUTRES	2,55 %	10,65 %

De ces premiers résultats, nous remarquerons fondamentalement les deux aspects suivants:

- Premièrement, le fait qu'il semble que la spontanéité propre des moyens oraux favorisent des processus de formation néologique déterminants, comme l'emprunt à d'autres langues et le figement

⁷ Données extraites de : OBSERVATORI DE NEOLOGIA (1998).

syntagmatique (ou néologismes créés par une structure syntaxique lexicalisée), différents quant à la fréquence de ceux qui sont utilisés dans les textes oraux. Voici quelques exemples des deux types:

Emprunts: *bueno, lograr, regalo (m), hòlding*.

Néologismes syntagmatiques: *acord de pau, alt càrrec, lluita armada, neteja ètnica*.

Dans le cas des emprunts, il faut tenir compte du fait que les formes provenant de l'espagnol –qui se prononcent généralement adaptées à la phonétique catalane– sont ceux qui prédominent. Ce fait s'explique essentiellement par la situation sociolinguistique de la Catalogne, où la coexistence de deux langues officielles favorise énormément les interférences linguistiques constantes entre les deux langues. Mais, en outre, il faut souligner que beaucoup de ces emprunts relevés ne se trouvent pas dans les textes écrits, car normalement ceux-ci sont filtrés par le service de correction des médias qui les considèrent comme des barbarismes ou des espagnolismes inutiles.

- Deuxièmement, il faut remarquer que la néologie sémantique a un rôle aussi important que la néologie formée par préfixation et suffixation –c'est-à-dire, celles qu'on considère comme les processus néologiques par excellence pour n'importe quelle langue. Voici quelques exemples de néologismes sémantique provenant de textes oraux: *aliat* (en politique), *bloquejar* (en économie), *cervell*, *contingent* (en politique), *cúpula* (en politique), *traspasar* (en sport).

En définitive, ces premiers résultats obtenus nous permettent de saisir déjà des différences avec les données provenant de la presse écrite et, en même temps, de continuer à approfondir notre connaissance sur les

facteurs qui conditionnent la création néologique dans les sociétés actuelles. C'est pourquoi nous avons commencé, cette année, le dépouillement de textes provenant des médias (programmes de télévision), qui, bien sûr, vont nous donner une quantité et une diversité de formes néologiques très intéressantes à recueillir.

3.3. Applications de la recherche en néologie orale

Les applications possibles de la recherche en néologie orale que mène l'Observatoire coïncident, *grosso modo*, avec celles qui ont été faites, jusqu'à présent, avec les données provenant de la presse écrite et, en plus, elles les complètent et les diversifient.

Par conséquent, les prochains projets du groupe de dépouillement des textes oraux sont conformes aux objectifs généraux de l'Observatoire qu'on a déjà vus; ces projets sont les suivants:

1. Constituer une banque de données de néologismes oraux.
2. Fournir l'information obtenue avec le dépouillement des textes oraux à des institutions officielles et à des entreprises qui élaborent des ouvrages lexicographiques.
3. Rendre possible les travaux de recherche concernant la création néologique spontanée.
4. Publier périodiquement les néologismes oraux documentés.

A part ces applications, qui complètent et enrichissent celles que l'Observatoire a réalisées, l'information fournie par le dépouillement de textes oraux élargit les horizons du projet car elle permet d'envisager de nouveaux domaines d'application, comme:

- l'application des nouvelles technologies au traitement des langues naturelles. L'Observatoire pourra collaborer à l'élaboration de programmes de correction orale automatisés, à l'usage des énonces des moyens audiovisuels. Un tel programme est en train

d'être mis au point par l'Unité de Recherche en Génie Linguistique de l'IULA, que dirige le professeur Lluís De Yzaguirre, en collaboration avec Catalunya Ràdio;

- l'enseignement des langues; les données extraites du dépouillement des textes oraux peuvent contribuer, ici, à l'élaboration de matériaux visant à atteindre la maîtrise orale de la langue.

En définitive, le fait que l'Observatoire ait commencé ce nouveau domaine de recherche de la néologie orale contribue à enrichir considérablement le volume et la diversité du matériau recueilli et contribue aussi à élargir et à diversifier les objectifs du projet. Comme l'a dit le professeur Badia i Margarit "il y a, en effet, plusieurs façons de créer des néologismes, et donc, il doit y avoir plusieurs manières de les chercher »⁸.

4. Conclusions

Pour conclure, nous voulons remarquer encore une fois, que le travail accompli par l'Observatoire de Néologie pendant ses douze années d'existence nous a permis de constater que la recherche en néologie, outre son intérêt indiscutable pour approfondir l'étude et la compréhension du phénomène de l'évolution des langues, présente de multiples possibilités d'application dans différents domaines linguistiques. En ce sens, nous avons vu comment les données de l'Observatoire ont été utiles pour des différentes applications: l'enseignement et la recherche par des membres du groupe et d'autres personnes qui les ont sollicitées; l'actualisation des dictionnaires et des grammaires, et la création de nouveaux outils lexicographiques.

⁸ Badia i Margarit, A. M. (2000). A: Cabré, M. T.; Freixa, J.; Solé, E. (ed).

En outre, ces objectifs ont été élargis actuellement grâce à la création du groupe de dépouillement de textes oraux, lequel permet de compléter et d'enrichir les données extraites de la presse écrite, et de commencer, donc, un nouveau domaine d'étude concernant l'observation de la création néologique spontanée, qui nous aidera, sans doute, à approfondir l'étude des motivations et du fonctionnement de la création néologique dans les langues.

M. T. Cabré, M. Domènech, R. Estopà, J. Freixa et E. Solé
Institut Universitari de Lingüística Aplicada
Universitat Pompeu Fabra

Bibliographie

Cabré, M.T.; Freixa, J.; Solé, E. (1997) "À la limite des mots construits possibles". Corbin, D. et al. (ed.) *Mots possibles et mots existants. Forum de Morphologie (1ères rencontres). Actes du colloque de Villeneuve d'Ascq (28-29 avril 1997)*. Lille: Université de Lille III, p. 65-78. (Silexicales, 1).

Cabré, M.T.; Estopà, R.; Freixa, J.; Solé, E. et al. (sous presse) "Evaluación de la vitalidad de una lengua a través de la neología". *La Néologie Iberique*. (Paris, 25-26 mai 1998). Paris: Université de Paris IV.

Cabré, M. T.; (2000) «La neologia com a mesura de la vitalitat interna de les llengües». Cabré, M. T. ; Freixa, J. ; Solé, E. (ed.). *La neologia en el tombant de segle. Actes del I Simposi sobre Neologia (2000) i el I Seminari de Neologia (2000)*. Barcelone: Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, p. 85-108. (Sèrie Activitats, 5).

Cabré, M. T.; Freixa, J.; Solé, E. (ed.). (2000) *La neologia en el tombant de segle. Actes del I Simposi sobre Neologia (2000) i el I*

Seminari de Neologia (2000). Barcelone: Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, p. 85-108. (Sèrie Activitats, 5).

Freixa, J.; Solé, E.; Cabré, M. T. (1998) "Observació de la variació i el contacte de llengües en els neologismes". Boix, E. et al. (ed.) *El contacte i la variació lingüístics: descripció i metodologia (CLUB-4)*. Barcelone: PPU: Secció de Lingüística Catalana, Universitat de Barcelona.

Freixa, J.; Solé, E. (sous presse) "Neologismos para la actualización de gramáticas y diccionarios". *II Jornadas sobre Neología*. Facultade de Filoloxía e Traducción. Universidade de Vigo. 11 mai 2000.

Observatori de Neologia. (1998) *Descripció quantitativa dels neologismes documentats durant l'any 1995 a la premsa en català*. Barcelone: Institut Universitari de Lingüística Aplicada. (Papers de l'IULA. Sèrie Informes, 23).

Observatori de Neologia. (2000) *Neologismes documentats a la premsa en català (1999)*. Barcelone: Institut Universitari de Lingüística Aplicada. (Papers de l'IULA. Sèrie Informes, 28).

Observatori de neologia. (2000) *Neologismes documentats a la premsa en castellà (1999)*. Barcelone: Institut Universitari de Lingüística Aplicada. (Papers de l'IULA. Sèrie Informes, 29).