

IRINA PRODANOF

À LA RECHERCHE D'UN MODÈLE DE DÉRIVATION
EN ITALIEN

Les travaux de réorganisation du Dictionnaire Machine de l'Italien se divisent en plusieurs branches dont l'une, la plus indépendante du point de vue des résultats scientifiques, est la recherche d'un modèle de dérivation. Nous nous limitons dans cet exposé aux mots suffixés.

1. Notre point de départ est le dictionnaire de l'Italien et le Dictionnaire Machine à notre disposition, et la recherche se base sur une interprétation du métalangage de la définition du dictionnaire. Le dictionnaire nous fournit pour les mots suffixés une définition formée d'éléments que nous divisons conventionnellement en fonction et base. Pratiquement la définition se compose d'un mot qui s'identifie généralement avec la base d'une série dérivative et d'une séquence qui peut contenir les éléments suivants:

a) des mots syntaxiques liés directement à la base:

Ex. (1): *ABBAIATORE*: *chi abbaia*

b) des mots pleins liés à la base d'une façon plus ou moins étroite:

Ex. (2): *CADUTA*: *ATTO del cadere*

Ex. (3): *CIMICIAIO*: *LUOGO pieno di cimici*

Le fait que ce deuxième type d'énoncé et les éléments qui le composent soient employés avec constance dans tout le dictionnaire nous a porté à considérer les mots pleins de la même manière que les mots syntaxiques, c'est-à-dire vides, comme de simples indicateurs de fonction dans le métalangage.

2.1. La première conséquence opérative d'une telle régularité est évidente. En effet, si nous pensons à placer fonctions et bases dans des

positions fixes, la manipulation mécanique de telles données devient une simple opération de sélections et de comparaisons.

La seconde conséquence est également évidente parce que la position fixe de la base dans l'espace réservé à la définition nous permet de sélectionner facilement les séries dérivatives.

2.2. La troisième conséquence d'ordre supérieur est qu'un métalangage comme celui que nous examinons et duquel nous avons déjà identifié le caractère formel, peut être traduit facilement dans un autre métalangage plus synthétique, des codes rapidement manipulables par l'ordinateur.

3. En partant de telles considérations nous pouvons projeter toute la recherche, à condition que l'on fasse, au préalable, recours au postulat de l'identité substantielle entre procès dérivatif et la forme de la définition donnée par le dictionnaire. Si, en effet, il est vrai que la définition peut être exprimée par la formule $f(x)$ dans laquelle x est la base, il est également vrai que la même formule peut décrire un dérivé, en tant que ce dernier résulte composé d'une base et d'un suffixe qui a les mêmes caractéristiques de régularité que nous avons reconnues dans la fonction.

Donc, si les deux parties de l'article du dictionnaire (lemme et définition) sont décrites par la même formule $f(x)$, étant donnée l'identité de base, et présumée la réelle adhérence de l'une à l'autre, les f coïncideront, de manière que la fonction devient définition du suffixe.

4.1. Le premier pas à faire sera la sélection des lemmes suffixés par suffixes formellement identiques. En interrogeant en même temps la définition et le code grammatical nous nous prémunirons de ne pas entraîner dans les listes les mots qui ne sont pas réellement des dérivés:

- | | | | | |
|----------|------------|------|----|-----------------------|
| Ex. (4): | ABBAIATORE | : | SN | chi abbaia |
| (5): | TRATTORE | (1): | SN | chi trae |
| (6): | TRATTORE | (2): | SM | veicolo che... |
| (7): | TRATTORE | (3): | SN | gestore di trattoria. |

Si nous sélectionnons par exemple les dérivés personnels en *-TORE* nous excluerons de la liste l'Ex. (6) puisque le code grammatical ne répond pas à la demande (*SM* au lieu de *SN*) et en même temps la définition ne commence pas par *chi*.

L'Ex. (7) ne pourra pas être sélectionné dans cette liste à cause de

sa définition; pourtant il sera introduit dans une liste à part de pseudo-dérivés pour être réexaminé.

Cette sélection nous fournira une vision exhaustive de la fonction ou des fonctions exprimées par chaque suffixe et nous permettra de grouper les fonctions qui se trouvent régulièrement ensemble dans une catégorie unique,

Ex. (8): *-MENTO*: atto, effetto di...
-ALE : di, relativo a...

De plus, l'analyse des fonctions nous permet d'établir une première distinction entre suffixes formellement identiques et nous fournit en même temps la liste complète des bases compatibles avec chaque suffixe.

Ex. (10): *BARC-ATA* (1): quantità di carico portata da...
GOMIT-ATA (2): colpo di.....
CAMMIN-ATA (3): atto, effetto del...

etc.

4.2. Ensuite nous sélectionnerons les matériaux en interrogeant la définition, mettant en évidence ainsi les suffixes doués de fonctions identiques. En effet, en triant les définitions, celles qui commencent par le même mot ou groupe de mots (fonctions), apparaîtront l'une après l'autre.

Une fois obtenues ces listes, nous regrouperons les suffixes qui ont les mêmes fonctions et en même temps une catégorie grammaticale de base identique, en leur attribuant un code d'homo-fonctionnalité.

Cette classification devrait nous permettre, dans une dernière phase du projet, d'établir une compatibilité réelle entre classes homologues de bases et classes homofonctionnelles de suffixes, réduisant ainsi le nombre des suffixes à examiner comme indépendants.

En d'autres termes, une fois prouvée la réelle compatibilité d'une série de bases de la même classe avec un suffixe appartenant à un certain groupe, sa compatibilité avec tous les autres suffixes homofonctionnels sera implicite.

Ex. (11): *MONT-A-TURA* **GUAR-I-TURA*¹
MONT-A-MENTO *GUAR-I-MENTO*

¹ La base verbale *MONT-* choisit les deux suffixes homofonctionnels pour former deux mots presque synonymes attestés dans la langue, lorsque la base *GUAR-* a choisit sans raison évidente seulement le suffixe *-MENTO*. Nous pouvons pourtant étendre à l'autre la compatibilité de *GUAR-*.

4.3. Dans une phase ultérieure nous regrouperons automatiquement les mots dérivés en tenant compte du code grammatical de la base et, à l'intérieur, de l'identité formelle des bases mêmes. Nous obtiendrons, ainsi, en séquence, les séries dérivatives de chaque base, groupées par catégorie grammaticale.

Cela mettra en évidence l'extension complète de la série dérivative de chaque catégorie grammaticale.

4.4. Ces matériaux nous mettent dans la condition de délimiter à l'intérieur d'une catégorie grammaticale de base, des sous-ensembles qui sélectionnent un certain groupe de suffixes ou bien un autre.

Ex. (12): base = *V* dérivé = *S*

<i>MONT-ARE</i>	<i>REGISTR-ARE</i>	<i>GUARD-ARE</i>
<i>MONT-A-MENTO</i>	—	<i>GUARD-A-MENTO</i>
<i>MONT-A-TURA</i>	<i>REGISTR-A-TURA</i>	<i>GUARD-A-TURA</i>
<i>MONT-A-TA</i>	—	<i>GUARD-A-TA</i>
<i>MONT-A-TOIO</i>	—	—
<i>MONT-A-TORE(1)</i>	—	<i>GUARD-A-TORE(1)</i>
—	<i>REGISTR-A-TORE(2)</i>	—
—	<i>REGISTR-A-ZIONE</i>	—

L'hypothèse la plus optimiste serait celle de la parfaite coïncidence entre classes de bases et séries de suffixes. Nous nous attendons pourtant à ce que le rapport réel soit exprimé par une intersection ou bien par une inclusion, desquelles nous devons rendre compte. Cela signifie que nous devons expliquer pourquoi certains sous-ensembles de bases homologues sélectionnent des séries non coïncidentes de suffixes. Suivant notre opinion la réponse est à rechercher exclusivement dans la base.

4.5. Il ne reste, à ce moment-là, qu'à faire recours au dictionnaire en cherchant les traits classificatoires communs des bases considérées comme lemmes et d'établir la réelle compétence dans le procès de dérivation. Pour les classes de bases identiques nous pourrions étendre la compatibilité avec tous les suffixes acceptés par chaque base de la classe.

4.6. A la fin d'un travail de recherche, le suffixe sera identifié par les codes suivants:

a) un code grammatical se rapportant au mot dérivé résultant;

b) une indication de fonction tirée du type général de définition attribuée aux dérivés;

c) un code indiquant la où les classes de bases avec lesquelles il est compatible.

Tandis que pour le lemme-base, sa propre classification lexicale sera suffisante à exprimer la compatibilité avec une certaine série de suffixes. Evidemment, pour obtenir un résultat correct, il sera nécessaire d'attacher au lemme des codes phonétiques.

5. L'intérêt d'une telle recherche porte, selon nous, sur deux points:

5.1. Il va de soi que cette étude peut nous fournir des données très intéressantes sur le processus de dérivation en acte aussi bien qu'en puissance, de façon que l'on aura le moyen de former un modèle du fonctionnement de la dérivation.

On s'attend aussi, sur le plan lexicographique, à constituer une première base pour la réduction du nombre des entrées du dictionnaire même. Ce résultat sera évident si l'on pense qu'au lieu d'une série dérivative on pourrait avoir une base et un algorithme général capable d'en produire les dérivés.

5.2. D'autre part, l'équivalence entre code et définition nous met en condition de transformer, s'il est nécessaire, le mot dérivé en la phrase type, ampliable par peu de règles.

De plus, comme cette recherche nous amène à manipuler les définitions structurellement les plus constantes, un bon résultat dans cette direction nous confirmerait l'hypothèse générale que certains traits classificatoires peuvent être déduits des définitions mêmes.

