IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

FRONTIÈRES DE LA LEXICOLOGIE ITALIENNE: des unités multilexicales aux collocations

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1556196 since 2018-12-18T15:10:36Z

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

12 October 2024

FRONTIÈRES DE LA LEXICOLOGIE ITALIENNE

Des unités multilexicales aux collocations*

Résumé

Le mot lexicologie n'apparaît pas souvent dans les récentes études des chercheurs italiens concernant le lexique. Il a perdu sa valeur d'hyperonyme fort : celui qui fait des études d'étymologie, d'onomastique, de néologie admet faire des études de lexicologie, mais il préfère les classer sous ces termes.

Les contributions italiennes les plus récentes et les plus intéressantes en matière de lexicologie sont les études consacrées aux unités multilexicales, c'est-à-dire les unités formées de plusieurs mots graphiques, et aux collocations restreintes. La recherche lexicale à partir de corpus bien construits a mis en lumière dans les langues romanes des rapports significativement fréquents sur l'axe des combinaisons qui, auparavant, étaient passés inaperçus, sinon dans leur typologie, du moins dans leur étendue.

Abstract

Lexicology is not a frequent word in recent Italian lexical studies. It is not any longer a well known domain term. Researchers dealing with ethymology, onomastics or neology will easily admit that they are doing lexicological studies but they generally prefer to mention their specific branch.

The most recent and interesting lexicological studies in Italy are devoted to multiwords and to restricted collocations. Corpus linguistics deeply changed Romance languages lexicological studies by showing how widespread are frequently co-occurrent words.

Mots-clés

Key words multiwords ; restricted collocations ; dictionary of collocations ; corpus linguistics

1. Lexicologie : un terme désuet ?

Le mot lexicologie n'apparaît pas souvent dans les récentes études des chercheurs italiens concernant le lexique.

---

* Traduit de l'italien par Sylvie Pipari. Tous nos remerciements à Manuel Barbera, Salvatore C. Sgroi et M. Prandi pour leurs conseils en phase de rédaction.

La preuve (indirecte) en est les entrées de l'*Enciclopedia dell'italiano* dirigée par Raffaele Simone et al. (2011) : aux côtés des articles consacrés à la phonétique, la syntaxe, la morphologie, figurent les entrées Lexique, Lexicographie, lexicalisation, mais pas Lexicologie. Même le quatrième, et dernier, recueil bibliographique de la Società di Linguistica italiana (voir Iannàccaro (2013)) montre que, à l'aube du troisième millénaire, il y a des chapitres intitulés Lexique, Sémantique, mais pas de Lexicologie.

Par ailleurs, si l'on feuillette les index de ces recueils bibliographiques que la SLI a entamés en 1977 avec le volume dirigé par Gambarara et Ramat, on assiste justement aux vicissitudes du terme *lexicologie*. Dans le premier volume, concernant la période 1965-1975, le chapitre Lexicologie était confié à Aldo Duro ; pour la période 1976-86 nous trouvons le chapitre Lexicologie et lexicographie attribué à M. Pfister (cf. Mioni Cortelazzo (1992)) ; le troisième volume (cf. Lavinio 2002), portant sur la période 1987-1997, ne situe plus *lexicographie* parmi les niveaux d'analyse. On trouve Sémantique, aux côtés de Phonétique et Phonologie, Morphologie, Syntaxe, Linguistique du texte.

Dans un dictionnaire de linguistique comme le Cotticelli Kurras (2007), adaptation du lexique de linguistique en allemand de Bussmann et publié, à l'origine, en 1983, il existe une brève entrée *lexicologie* qui commence de la sorte « secteur de la linguistique, notamment de la →sémantique, qui s'occupe de la recherche et de la description du vocabulaire d'une langue et étudie les expressions linguistiques en particulier par rapport à leur sémantique interne et aux liens entre les différents mots, c'est-à-dire les termes du lexique » (p. 427). Cette définition semble ratifier définitivement un virage qui avait été amorcé dans l'entrée lexicologie du *Dizionario di linguistica* de Beccaria (1989) : « dans la dernière partie du XXe siècle le terme **l[**exicologie**]** a souvent été négligé au profit d'autres mots, tels que morphologie, morphosyntaxe, sémantique, sémantique lexicale. Le phénomène reflète une tendance positive au fond, c'est-à-dire la conviction que le lexique est beaucoup plus structuré et beaucoup plus lié à la syntaxe par rapport à ce que l'on voulait bien admettre par le passé » (p. 428). L'article continue en faisant remarquer que lexicologie et morphologie ont en commun un intérêt pour le signifiant et, en cela, se différencient de la sémantique qui s'attache surtout aux signifiés. Mais que cette différence soit davantage une question de palis terminologiques que le fruit d'un écart effectif des études était déjà clair alors, grâce aux études de sémantique lexicale ; *Lexical semantics* de Cruse avait été publié en 1986 et, dans l'entrée lexicologie de Beccaria, on signalait : « Si elle ne se désintéressait pas du signifiant, la sémantique lexicale pourrait être considérée comme coïncidant avec la **l[**exicologie] ».

Le terme refait surface parfois dans les titres de congrès, par exemple le XLIV Congresso Internazionale di Studi della SLI de 2010, intitulé « Linguistica educativa / Lessico e Lessicologia » en compagnie de branches, justement, fortement intersectorielles.

Il a peut-être perdu concrètement une valeur d'hyperonyme fort : celui qui fait des études d'étymologie, d'onomastique, de néologie admet faire des études de lexicologie, mais il préfère les classer sous ces termes ; il ne ressent pas le besoin, en cette époque du triomphe des interfaces, de les attribuer à un niveau d'analyse qui lui semble réducteur.

Même le terme *lexicographie* qui, plus souvent que d'autres traditionnellement, accompagnait *lexicologie*, a commencé, dans les vingt dernières années, à être plutôt accolé, par ceux qui s'occupent de lexicographie, à *métalexicographie*, revendiquant ainsi une branche de réflexion théorico-critique sur la nature et la préparation des dictionnaires bien plus vaste, quant aux intérêts, et en même temps plus spécialisée, que la lexicologie, car englobant des aspects de modalité de consultation, des moyens de support, des caractéristiques de structure de l'entrée, etc. qui vont bien au-delà de l'étude du lexique d'une langue.

Enfin, l'approche lexicaliste de la description linguistique et la linguistique des corpus ont également contribué à éroder le champ d'usage du terme. Il pourrait sembler paradoxal que deux modes de voir la langue à travers des analyses qui privilégient l'unité lexicale comme porte d'entrée de la description linguistique aient produit l'effet de faire "oublier" la lexicologie, mais, à y regarder de plus près, cet effet est également une conséquence de la crise du structuralisme qui a bouleversé aussi le domaine apparemment plus solide de la sémantique. La crise du structuralisme n'a pas « seulement accentué la nature intersectorielle de la sémantique en attirant l'attention sur les facteurs externes au langage », mais a également concentré « l'intérêt sur les interfaces entre les niveaux d'analyse de la structure interne en érodant la compartimentation autonomiste traditionnelle » (Squartini 2013, p. 521).

Quant à la linguistique des corpus, même si sa déclinaison "à l'italienne" est somme toute en continuité avec la tradition de la linguistique philologique des XIXe-XXe siècles et se rattache au filon *corpus based*, elle a néanmoins changé la façon de faire de la lexicologie. Le linguiste ne rassemble plus les données (du moins il ne les rassemble plus seulement) en interrogeant sa propre compétence et celle d'autres informateurs, comme cela avait lieu dans la recherche dialectologique par exemple, en partant d'une sélection de phénomènes ou de traits décidés par ses soins sur la base de la recherche précédente : « c'est l'usage (comme le montrent les corpus) […] qui fonde la *langue* (en français dans le texte), même si les corpus, étant par définition finis, n'en représentent qu'un sous-ensemble, d'autant plus significatif que le corpus aura été construit de façon avisée (jeu qui ne peut exclure la fameuse introspection)» (Barbera 2013, p. 582). Nombreux sont ceux qui, autrefois, auraient accompli des recherches de lexicologie en recueillant des données "à la main", et qui, maintenant – grâce à la relative facilité avec laquelle des logiciels évolués, comme par exemple Sketch Engine[1], permettent de créer des corpus – réduisent le temps consacré à la construction minutieuse du corpus et analysent ensuite les résultats provenant de ce corpus comme s'il s'agissait de la *langue* et non pas de l'un de ses sous-ensembles. Cependant, malgré ce danger, les études sur le lexique effectuées moyennant l'interrogation d'un corpus présentent l'avantage de montrer au linguiste lui-même ce qu'il n'aurait pas remarqué, car influencé par les lunettes de la tradition grammaticale : sans tomber dans l'extrémisme de l'étude *corpus driven*, il faut souligner que, dans le domaine du lexique, la recherche lexicale à partir de corpus bien construits a mis en lumière dans les langues romanes aussi des rapports significativement fréquents sur l'axe des combinaisons qui, auparavant, étaient passés inaperçus, sinon dans leur typologie, du moins dans leur étendue.

[1] Voir Kilgarriff (2012) et http://www.sketchengine.co.uk

Dans le quatrième volume de *Dictionaries. An International Encyclopedia of Lexicography* (2013) consacré notamment aux récents développements de la lexicologie, les contributions italiennes signées d'une part Marco Baroni et Silvia Bernardini et, de l'autre, Pietro Beltrami se rattachent à la prédilection italienne pour les recherches lexicologico-lexicographiques *corpus based*, tradition qui remonte aux débuts de la lexicographie européenne moderne, entamée par les Académiciens de la Crusca et qui perdure grâce à des ouvrages tel que le *Tesoro della lingua italiana delle origini*, pour s'interroger sur les avantages et les inconvénients du réseau comme corpus.

Les contributions sur la structuration interne du lexique se sont raréfiées, comme le constatait déjà Berruto (2010) ; sont restées des études sur des portions de lexique dans une optique diachronique ou explicitement cognitive. Les frontières de dénotation des différents composants des familles lexicales et des champs lexicaux sont étudiées plutôt du point de vue de la comparaison entre les langues, dans de nombreuses contributions sur les faux amis ou à travers l'analyse des concordances d'un mot et de ses possibles traduisants dans des corpus de différentes langues ou comme chez Marello (2013) entre un corpus de textes écrits par des apprenants d'italien L2 et un corpus de textes rédigés par des italophones, les uns et les autres sollicités par le même support iconique[2].

Observons cependant que cette approche à travers les contextes d'usage conduit rarement à une systématisation en traits ; elle aboutit plutôt à une analyse de cooccurrences typiques ou fréquentes et à une reconnaissance qu'il s'agit non pas de l'acception d'**un** mot polysémique, comme on le disait et on le dit encore dans les chapitres des manuels de linguistique de base, mais du sens que prennent **deux** ou plusieurs mots lorsqu'ils sont associés avec un ou plusieurs autres mots bien définis. Le dictionnaire monolingue devrait organiser la description du signifié dans la glose précisément selon un ordre des signifiés en partant du signifié de l'entrée dans les combinaisons les plus fréquentes, qui souvent – mais pas toujours – sont les combinaisons d'origine, pour passer au fur et à mesure aux combinaisons plus spécialisées, jusqu'à l'enregistrement des expressions idiomatiques[3].

Nous estimons que les contributions italiennes les plus récentes et les plus intéressantes en matière de lexicologie – aussi bien celles qui se déclarent telles que celles qui se donnent une étiquette autre – sont les études consacrées aux unités multilexicales, c'est-à-dire les unités formées de plusieurs mots graphiques, et aux

---

[2] Dans ce cas, l'analyse parallèle de *sembrare* a conduit à centrer chez les apprenants une préférence pour le signifié 'somigliare a' qui agit entre les deux référents, par rapport au signifié 'apparire' avec une complétive. L'analyse a par ailleurs montré chez les apprenants l'usage de *mi sembra*, presque lexicalisé, dans la périphérie de la phrase, avec la fonction de *forse, secondo me*.

[3] Ceci n'avait et n'a pas toujours lieu car domine un modèle abstrait d'organisation du signifié dans la microstructure qui, par exemple, regroupe un signifié et ses emplois figurés. Wiegand et Smit (2013, p. 185) font remarquer que cette distribution des acceptions peut être améliorée grâce à un usage avisé de microstructures spécifiant "which items stand within the textual scope of which item giving the meaning", c'est-à-dire, nous semble-t-il de comprendre à travers la comparaison de microstructures du dictionnaire monolingue allemand Wahrig, avec un usage attentif de discriminateurs de signifié, tant sur le plan des substitutions synonymiques que sur le plan des combinaisons, notamment si l'élément rentre dans un exemple (cf. Marello 1989, p.55-60). Ce filon d'études mené par l'école allemande n'a pas trouvé en Italie un écho suffisant en dehors du domaine de la lexicographie bilingue allemand-italien.

collocations restreintes, c'est-à-dire les expressions linguistiques syntagmatiques qui constituent indubitablement un domaine privilégié de la lexicologie, dans la détermination de leurs frontières dans l'abstrait et en contexte, même si elles ont souvent été abordées par des scientifiques plus concernés par d'autres fronts et/ou ayant des objectifs pratiques, des anglicistes italiens, des traducteurs, des morphologues, des terminologues, des historiens de l'italien, par exemple

Toutefois, il est désolant de constater que, lorsque la recherche part de linguistes travaillant dans le domaine de l'anglais, du français ou de l'allemand, les études théoriques respectivement anglaises, françaises, allemandes sont reines. L'adoption d'outils théorico-méthodologiques écrits dans une autre langue est inusuelle : à part les références anglaises qui, notamment en ce qui concerne l'étude des collocations, ne manquent jamais, les francistes ne lisent pas les germanistes et vice-versa, les anglicistes, quant à eux, ne citent que des textes en anglais. Chacun se cantonne dans les écoles où il s'est formé. Le compte rendu, dans l'*International Journal of Lexicography,* du volume de Simonetta Vietri *Idiomatic Constructions in Italian. A Lexicon-Grammar Approach* a été réalisé par une angliciste italienne qui a lu avec soin l'ouvrage et en a bien saisi l'enjeu, à savoir démontrer que les expressions idiomatiques peuvent subir des variations complexes tout autant que les expressions non idiomatiques. Cependant, au final, il semble que l'on reproche à Vietri de faire du lexique-grammaire[4]. Si, d'un côté, on peut partager ce jugement pour mettre en garde le lecteur-type de la revue, le lexicographe non linguiste attiré par la première partie du titre du livre de Vietri, de l'autre, on constate que le lexique-grammaire est relativement apprécié (ne l'est-il pas encore ou ne l'est-il plus?).

2. Frontières lexicales dans le Groupe Nominal et entre Groupes

Les unités multilexicales appelées également *polyrhématiques* et, plus récemment, expressions *multi-mots*, calque de l'anglais *multiwords,* ou encore *mots syntagmatiques, unités lexicales supérieures* ou *mots complexes*[5] sont plus problématiques à identifier dans les langues romanes qu'en anglais ou en allemand, en raison d'une série de caractéristiques morphologiques évidentes pour le lexicologue-linguiste : alors qu'en anglais ou en allemand le mot qui modifie, délimite, restreint le signifié de la tête est en position adjacente avant la tête, dans les langues romanes l'expression qui modifie vient après la tête et est souvent reliée à celle-ci par des prépositions et des articles qui la rendent très semblable à un syntagme nominal + syntagme prépositionnel. Pour une discussion à notre avis, plus minutieuse lexicologiquement, voir Grossmann Rainer (2004), où ces unités sont considérées sous plusieurs points de vue et dans différentes perspectives, dont les rapports avec les processus de la composition, de l'emprunt et de la néoformation.

---

[4] "In sum this is a very technical book aimed for the specialised reader who is familiar with syntactic parsing and interested in the morpho-syntactic layout of idioms. Consequently, the observations are wholly grammatical and very little attention is placed upon the contextual cues that are nonetheless the triggers for syntactic variation and flexibility " (Pinnavaia 2015, p. 258).

[5] Pour la paternité des différentes manières de baptiser et identifier ces unités lexicales formées par plus d'un mot graphique, voir Masini (2011) e Sgroi (2007).

Comme Barbera (2013), nous estimons qu'il est utile d'accorder une existence métalinguistique à la catégorie unité multilexicale, même s' « il s'agit d'un grand nombre de phénomènes différents entre eux, certains linguistiques (lexicaux, morphologiques, syntaxiques, textuels, etc.), d'autre purement statistiques, d'autres encore linguistiques et statistiques à la fois, quoi qu'il en soit non réductibles l'un à l'autre dans un seul cadre théorique sémantique ; l'apparente unité du phénomène (ou mieux, *des phénomènes*) est une conséquence de l'existence de certains effets pragmatiques unitaires, qui sont par ailleurs capturés a posteriori par la statistique (qui, opérant sur les corpus, travaille ipso facto sur la *parole*. » (Barbera, 2014, p. 63).

L'idiomaticité ou non-compositionnalité du signifié de l'unités multilexicale ou polyrhématique est difficilement défendable pour tous les exemples de polyrhématique fournis et, de fait, comme le fait remarquer Sgroi (2007), la vaste acception de *polyrhématique* (idiomatique et non idiomatique, bi- et pluri-membre) semble l'emporter dans les manuels[6].

2.1. Combien et quels mots font une polyrhématique ?

Distinguer les polyrhématiques formées de deux éléments seulement des composés non univerbés constitue l'un des problèmes abordés à cheval entre lexicologie et morphologie. En effet, quand il y a trois mots (et normalement celui du milieu est une préposition comme *di, a, da*, ex. : *coda di rospo, giacca a vento, ferro da stiro*), il n'est guère difficile de considérer la séquence comme une polyrhématique, mais s'il y a deux mots « les formations qui affichent entre l'unité complexe et sa tête une relation d'hyponymie / hyperonymie et dont le signifié est compositionnel peuvent être définies des composés, alors que les formations qui ne présentent pas ces propriétés pourront être classées parmi les poly-rhématiques » affirme Bisetto (2004) qui cite comme exemple de poly-rhématique *conferenza stampa* et *anima gemella*, qui ne sont pas, respectivement, une conférence et une âme, tandis que *nave cisterna* est bien un navire et serait donc un composé.

Pour paraphraser Voghera, la scientifique qui est le plus souvent citée à ce propos, les polyrhématiques sont des combinaisons de mots que le parlant ressent comme un élément lexical unitaire et qui, sur une échelle qui va du mot monomorphémique à la phrase, se trouvent dans une position intermédiaire entre les composés et les syntagmes libres. Il ne s'agit pas d'un ensemble homogène, mais d'« un ensemble de formations différentes en raison de leur composition interne et du degré de liberté de mouvement manifesté par leurs éléments constitutifs » (Voghera, 2004, p. 59).

Les lexicologues n'ont pas compilé de véritables listes avec des composés d'un côté et des polyrhématiques de l'autre et on est en droit de se demander si cela serait bien utile. Quoi qu'il en soit, les lexicographes, eux, ont dû prendre des décisions, plusieurs même, si, comme nous le rappelle Voghera, le GRADIT enregistre en tant que lemme 63.000 polyrhématiques[7], dont la plupart (84%) nominales. Les lexicographes ont pris des décisions de lexicographes : vu que les unités multilexicales ne sont pas des lemmes autonomes[8] dans les dictionnaires

---

[6] Aprile (2005 p. 139-41), Ježek (2005 p. 35), Thornton (2005 p. 138).

[7] Voghera rappelle que dans le GRADIT près de 67.000 autres polyrhématiques figurent au sein des entrées.

[8] La version digitale des dictionnaires, dépourvue des contraintes d'espace, a permis de rendre plus visibles les polyrhématiques dans les dictionnaires qui les ont organisées comme sous-lemmes ; pour

italiens, les dictionnaires ne s'expriment pas sur leur éventuelle nature de composés, comme ils le font, en revanche, quand le composé est univerbé.

bassorilievo /bassori'ljɛvo/ (bas·so·ri·lie·vo) s.m. TS arte
[av. 1519; comp. di *basso* e *rilievo*]
scultura in rilievo in cui gli elementi emergono di poco dal piano di fondo | la tecnica per realizzare tale scultura GRADIT

Cette information est donnée dans la partie consacrée à l'étymologie, mais seuls les lemmes ont une étymologie, par conséquent les unités multilexicales non lemmes, mais des sous-lemmes dans la meilleure des hypothèses, se voient attribuées uniquement une partie du discours, à savoir *locution* suivie de *adjectivale, nominale, verbale,* etc. et, s'il s'agit d'une locution nominale, d'informations sur le genre et la morphologie du pluriel.

Rares sont les cas où la polyrhématique reçoit une étymologie : comme on le voit avec les entrées *fine settimana, guerra lampo, sangue blu,* ce sont les cas où les polyrhématiques constituent le calque d'une unité d'une autre langue, l'étymologie est diachronique[9] et elle ne met pas en avant le statut de composé.

~ fine settimana loc.s.m.inv. CO
[1932; cfr. ingl. *week end*, 1879]
i giorni di sabato e domenica, in quanto non lavorativi e disponibili per lo svago: *andare al mare per il f. settimana, ci sentiamo nel f. settimana* GRADIT
~ guerra lampo loc.s.f. CO
[1941; cfr. ted. *Blitzkrieg*, 1939]
quella condotta con la massima mobilità e con azioni di sorpresa, che mira a neutralizzare le forze nemiche in brevissimo tempo
~ sangue blu loc.s.m. CO
[1819; cfr. sp. *sangre azul*]
nobiltà di lignaggio

Une autre façon d'enregistrer les polyrhématiques dans les dictionnaires évite leur lemmatisation et consiste à étiqueter le second membre comme adjectif invariable postposé. Confrontons la polyrhématique *verde bottiglia* dans le GRADIT avec l'entrée *bottiglia* du dictionnaire Zingarelli 2015.

~ verde bottiglia loc.agg.inv., loc.s.m.
loc.agg.inv. di colore verde piuttosto scuro, tipico di alcune bottiglie di vetro | loc.s.m. tale colore GRADIT

bottiglia / botˈtiʎʎa/

---

l'instant, De Mauro (2000) et GRADIT sont les plus riches en locutions lemmatisées ; Zingarelli (2015 version en ligne) présente 1129 lemmes et sous-lemmes avec plus d'un mot graphique : la plupart sont des gallicismes, des anglicismes, des latinismes et des locutions conjonctives ou prépositionnelles. Voir à ce propos Marello (2010).

[9] Sgroi (2006) aborde la question de l'étymologie dans les dictionnaires du point de vue de la formation des mots ; Sgroi (2010) traite de l'étymologie synchronique négligée en faveur de l'étymologie lointaine.

[ant. fr. boteille, dal lat. tardo but(t)ĭcula, dim. di bŭttis ‘botte’ ☼ 1365]
A s. f.
1 recipiente spec. cilindrico, generalmente di vetro, […]
B in funzione di agg. inv.
● (posposto al sost.) nella loc. *verde bottiglia*, verde scuro Zingarelli 2015

lampo / ˈlampo/
[da †lampare ☼ 1319]
A s. m.
1 fenomeno luminoso […]
B in funzione di agg. inv.
● (posposto a un sost.) che dura pochissimo; che accade, si svolge, funziona e sim. in brevissimo tempo: *matrimonio lampo; cerimonia lampo; guerra lampo* | ***chiusura lampo, cerniera lampo***, dispositivo per chiudere rapidamente indumenti, […]| ***notizia lampo***, riguardante un evento recentissimo ed espressa con poche parole SIN. flash nel sign. 3 | ***partita, torneo lampo***, nel gioco degli scacchi, quelli in cui a ciascun giocatore è concesso un tempo massimo […]
C s. f. inv.
● (ellitt.) cerniera lampo SIN. Zip Zingarelli 2015

La solution adjectif invariable postposé figure également dans le GRADIT en plus du sous-lemme ; voir l’entrée *lampo* qui a une huitième section “agg. inv. sempre posposto” et une neuvième comme substantif et est suivie aussi des sous-lemmes polyrhématiques *cerniera lampo, chiusura lampo, guerra lampo.* Il semble cependant que cette surabondance dans le GRADIT se borne aux cas où le mot en tant que lemme apparaît postposé et adjacent à un certain nombre de mots ; le dictionnaire Zingarelli, par contre, étiquette *mela* aussi comme adjectif invariable toujours postposé, qui toutefois n’apparaît que dans la locution *verde mela*[10].

lampo […] 1 s.m. fenomeno atmosferico prodotto da scariche elettriche che hanno luogo nell'atmosfera, che si manifesta con un'improvvisa e brevissima emissione di luce: *l. accecante, notturno*; *per tutta la notte si sono susseguiti lampi e tuoni*; *rapido, veloce come un l.*, velocissimo
2a s.m. AU estens
[…]
8 agg.inv. AU sempre posposto, che dura pochissimo, che avviene, si svolge o si compie con la massima rapidità: *matrimonio l., visita l., sciopero l., missione l., notizia l.*: quella immediatamente diffusa al pubblico da una radio o da una una televisione
9 s.f.inv. CO cerniera lampo: *tirare su, chiudere la l. del vestito, dei calzoni*

~ cerniera lampo loc.s.f. CO

[10] Le GRADIT ne cite pas *verde mela* parmi les polyrhématiques de *mela*, comme on pouvait s’y attendre, mais mentionne ce terme dans l’acception 8 s.m. avec d’autres teintes ; parmi les polyrhématiques de *verde*, il donne une place à *verde bottiglia*, qui figure également parmi les polyrhématiques de *bottiglia*, à *verde mare* et à *verde penicillina.*

chiusura per abiti, borse e sim. consistente in una doppia fila di denti di metallo o plastica, che si incastrano per mezzo di un cursore GRADIT

mela B in funzione di agg. inv.
● (posposto al sost.) nella loc. *verde mela*, detto di una tonalità chiara e delicata di verde Zingarelli 2015

3. Binômes vrais et présumés, trinômes : au-delà de l'unité multilexicale ou polyrhématique ?

L'existence d'effets pragmatiques unitaires n'est pas toujours suffisante pour faire d'une expression une unité dont nous sommes convaincus de l'existence métalinguistique comme unité. Récapitulons avec De Mauro (2005, p. 88-91) qui identifie les caractéristiques de la polyrhématique comme suit :

i - être constituée de plus d'un mot

ii - avoir un signifié unitaire

iii - avoir un signifié non compositionnel

iv - avoir une cristallisation lexicale et syntaxique plus ou moins forte

v - avoir une présence significative dans des langages technico-spécialisés

Nous avons déjà remarqué que les manuels ont intégré en particulier les traits (i) et (iv) et qu'ils ont survolé ou beaucoup discuté de (ii) et (iii) ; différentes études *corpus driven* ont également confirmé le trait (v). Néanmoins, si nous faisons abstraction de la profusion d'études réalisées par les linguistes computationnels sur des bases statistiques, dans le cadre linguistique italien les études qui ont tenté d'aborder largement la question de la cristallisation sont rares et appartiennent surtout au domaine lexique-grammaire et à la linguistique des corpus. Naturellement on a beaucoup fait aussi dans le domaine demaurien : à part les études de Voghera et de De Mauro lui-même, voir Russo (2010)[11] et Masini (2012). Masini a mené une recherche digne d'intérêt sur les binômes N + e + N du type *sale e pepe*, *armi e bagagli*. Du point de vue des contenus, les binômes sont souvent liés à un ordre pour ainsi dire naturel : *lampi e tuoni, usa e getta, gratta e vinci, due o tre*. Certains d'entre eux sont plus culturels, comme l'ordre fixe de *uomini e donne, fratello e sorella, gioie e dolori, vincitori e vinti, senza capo né coda*. D'autres encore puisent leur origine dans la littérature : *il gatto e la volpe, la volpe e l'uva, senza macchia e senza paura*.

[11] Le *lessico delle collocazioni* de Russo comprend en réalité dans ce qu'il appelle des collocations de type B et C et D des polyrhématiques bi- et pluri-membres avec une force cohésive croissante entre les éléments, cf. Russo (2010, p. iii-iv).

En réalité, l'inventaire de Masini, propose également, outre les binômes, des trinômes, comme *aglio, olio e peperoncino* ; *Dio, patria e famiglia* ; *pioggia, vento e tempesta* ; *sesso, droga e Rock 'n Roll* ; *vassalli, valvassini e valvassori* ; *vita, morte e miracoli*. Il faut cependant observer que la loupe de la recherche N + e + N (ou SP + e + SP) tend à masquer des expressions qui sont en réalité formées de plusieurs membres, mais pas du même type. Prenons par exemple le pseudo-couple *culo e camicia* : les deux noms apparaissent toujours avec *essere* dans le sens d' « être très unis », normalement dans des entreprises et avec des intentions peu louables de l'avis de celui qui dit de deux individus qu'ils « sont cul et chemise ». De même *d'amore e d'accordo* apparaissent avec *andare*, alors que *di cotte e di crude* est une expression que nous trouvons uniquement avec les verbes *farne, combinarne*.

À notre connaissance, Barbera (2009) incarne la tentative la plus complète d'aborder la question en partant des données statistiques du *Corpus Taurinense*, pour arriver à des réflexions métalinguistiques sur l'attribution des annotations par partie du discours aux unités multilexicales dans un corpus annoté[12].

3.1. Polyrhématiques verbales et collocations

Il est évident que si l'on entend par *collocation* la propriété de deux (ou plusieurs) mots de s'associer fréquemment, les unités multilexicales ou poly-rhématiques sont des collocations. Toutefois, une fois qu'on a opté pour ces termes pris au moins dans leurs caractéristiques (ii = signifié unitaire) et (iv = forte cristallisation), il ne convient pas d'en superposer un troisième qui traditionnellement *n'*engloble *pas* nécessairement ces deux traits, c'est d'autant plus vrai que pour avoir du moins partiellement le trait (iv) il doit s'agir de *collocations restreintes*.

La question est épineuse, notamment pour les polyrhématiques verbales et les collocations restreintes V+SN. Si l'on examine les exemples fournis par Voghera (2004, p.67-68) pour les types de polyrhématiques verbales les plus nombreuses V+(DET)+N, V+SP, l'on voit qu'il devient dirimant, pour les distinguer des collocations restreintes V+(DET)+N, V+SP, d'évoquer la propriété (iii), c'est-à-dire l'idiomaticité, la non-compositionnalité, qui cimente la polyrhématique, en fait une unité[13]. En effet, les composants d'une collocation restreinte, bien qu'ils ne soient pas librement remplaçables sur l'axe paradigmatique en raison de la présence d'une restriction (*fare/ avanzare / *portare avanti un'ipotesi*), sont, dans la plupart des cas, syntaxiquement autonomes (ils admettent des SN singuliers et pluriels avec différents déterminants et l'insertion d'éventuels adjectifs, la passivation, la dislocation à gauche : *avanzare un'ipotesi assurda, avanzare tutte le possibili ipotesi, l'ipotesi è stata avanzata dall'opposizione, l'ipotesi l'ha avanzata l'opposizione*). Certes, il existe des collocations restreintes avec une liberté moindre

[12] Réflexions entamées bien avant par Barbera, Marello (2000). Dans la *Grammatica dell'italiano antico* (cf. Renzi, Salvi, 2010), les locutions sont décrites à différents endroits, mais la question de leur variation interne et celle de l'attribution de partie du discours dans le corpus ne sont pas abordées directement, car elles ne rentrent pas dans les objectifs de l'ouvrage.

[13] Des exemples cités par Voghera (2004, p.67-68) *forzare i tempi, forzare la mano* n'ont pas, selon nous, une opacité suffisante.

et celles-ci ressemblent davantage à des expressions polyrhématiques : alors que *dare una mano di bianco* est une collocation restreinte (*dare / passare/ stendere / *dipingere/ *fare/ *mettere una mano di bianco*), *dare una mano* dans le sens de 'aider' et *dare la mano* comme geste de salutation sont plus proches d'une polyrhématique, même si elles permettent quelques variations (*una mano l'hanno data tutti*)[14].

4. Dictionnaires de collocations : accueillants mais condamnés à l'inachèvement

En cinq ans, près de cinq dictionnaires ont été publiés qui se présentent comme des dictionnaires de collocations italiennes[15], même si parfois le mot ne figure pas dans le titre. Ces ouvrages rassemblent *tout* l'éventail des expressions syntagmatiques où apparait le mot choisi comme lemme ; une attitude qui ne surprend pas, car pour aider dans la fonction onomasiologique (d'encodage), il faut « un dictionnaire spécialisé qu'on appellera dictionnaire de collocations et dont la macrostructure sera faite de bases de collocation sur lesquelles il adressera en microstructure la totalité des collocatifs » (Hausmann - Blumenthal 2006, 5).

Comme d'aucuns l'ont déjà fait remarqué, les dictionnaires généraux sensibles à la combinatoire des mots comme le GRADIT ou les dictionnaires de collocations rendent un précieux service, mais ne peuvent attirer l'attention sur ce qu'on *ne* peut *pas* associer ; à moins que l'on veuille conseiller à l'apprenant/utilisateur de considérer l'absence comme un signal indirect et certain de l'impossibilité de la combinaison.

Cependant, le fait que les verbes, quand ils sont des lemmes dans un dictionnaire de collocations, sont présentés avec les adverbes dont ils sont la base, pose des problèmes de non enregistrement de collocations restreintes V + SN o SP[16].

Ježek (2014) cerne bien le problème lexicologique et sémantique quand il observe que, en déplaçant la perspective la plus commune qui voit N comme terme déterminant du *meaning 'by collocation'* pris par le collocatif – perspective qui, notons-le, est la perspective traditionnelle des dictionnaires de collocations – se

---

[14] Le rôle des verbes support dans les polyrhématiques et les collocations a été largement étudié par Cantarini (2004) dans une comparaison de l'italien et de l'allemand et a montré que les verbes support en soi n'influent guère sur la cristallisation, sauf si on veut soutenir que puisqu'ils sont plus communs et donc plus fréquents, les collocations où ils apparaissent ont une plus forte probabilité de se lexicaliser pour devenir des polyrhématiques verbales. En réalité l'absence ou la fixité de l'article est un signal bien plus efficace de l'impossibilité de passiver ou de faire d'autres variations.

[15] Nous les listons suivant l'ordre de parution : Urzì (2009) ; Russo (2010), Tiberii (2012), Lo Cascio (2012) et Lo Cascio (2013a).

[16] Lo Cascio (2013a) montre bien qu'un bon dictionnaire combinatoire est le fruit d'un énorme travail et qu'une optique de comparaison avec d'autres langues est fort utile justement pour essayer de fournir les combinaisons qui éloignent l'utilisateur de la tentation de transposer littéralement une collocation de la L1, ignorant qu'il s'agit d'une collocation restreinte. Des dictionnaires en commerce, c'est sûrement celui qui possède le plus grand nombre de verbes comme lemme avec un trousseau de collocations V+ SN ou SP. Urzì (2009), le premier qui est sorti, conserve bien la couverture du langage formel dont était parti le fichage de l'auteur et poursuit également un classement interne qui va de la collocation avec verbe support (ou commun) à celle formée avec un lexique appartenant registre élevé, par ex. : *fare/ compiere/ perpetrare un massacro.*

dessine un processus syntagmatique d'ajustement sémantique entre la base et le collocatif, induit par le contact entre la sémantique d'un N inattendu dans le *frame* de référence de V ou A, par rapport auquel N impose une restructuration sémantique. Il ajoute également que cette perspective devrait être comparée « à la perspective opposée, selon laquelle les développements sémantiques du collocatif sont analysables comme le résultat de l'extension métaphorique ou métonymique de la classe de N à de nouveaux domaines référentiels par rapport à ceux attendus dans le frame de référence de V, lesquels, à leur tour, agissent rétroactivement sur V étendant le signifié et les potentialités syntagmatiques » (Ježek 2014, 417)[17].

Voilà donc que les collocations restreintes V + SN (ou SP), notamment si le nom ne figure pas parmi les noms pris en considération par le dictionnaire des collocations en tant que lemmes à côté desquels amasser la totalité des collocatifs, ont de fortes probabilités d'apparaître dans un dictionnaire général. Sinon dans la section polyrhématiques, dans la phraséologie ou même dans les exemples.

En conclusion, nous voudrions mentionner que la recherche des collocations restreintes, un tout petit peu moins peut-être que celle des polyrhématiques, est fortement conditionnée par leur caractéristique (v), c'est-à-dire avoir une présence significative dans des langages technico-spécialisés. Cette caractéristique devrait être opportunément élargie à la diachronie et à la diastratie. La « restructuration sémantique » opérée par la collocation restreinte est un processus qui est reconnu une fois qu'il est entré dans une norme, mais l'obsolescence peut être très rapide : une expression pour laquelle, il y a quelques décennies encore, on n'avait aucun mal à repérer l'origine métaphorique, ou même factuelle,[18] pourrait être totalement opaque pour les jeunes d'aujourd'hui.

Bibliographie
Dictionnaires et lexiques encyclopédiques

BECCARIA Gian Luigi (dir.) (1994, 2004 2ème édition). ***Dizionario di linguistica e di filologia, metrica, retorica***, Torino, Einaudi.

COTTICELLI KURRAS Paola (2007). ***Lessico di linguistica***, Alessandria, Edizioni dell'Orso.

DEMAURO Tullio (2000) ***Il dizionario della lingua italiana***, con CDROM, Torino, Paravia.

[17] Ježek cite Siller-Runggaldier (2006), mais cette position pour le signifié des verbes en général – indépendamment du fait qu'ils se trouvent dans des collocations restreintes V+ SN – est également exposée in Prandi, Cantarini (2013, p. 28-30) où l'on parle de *coercion* dans le sillage de Pustejovsky (1991, 425).

[18] Pour les jeunes d'aujourd'hui, par exemple, l'expression *raccrocher le téléphon* , dans le sens de 'conclure une conversation téléphonique', s'avère beaucoup plus idiomatique que pour les sexagénaires.

GRADIT= DE MAURO Tullio (1999), ***Grande dizionario italiano dell'uso***, ideato e diretto da Tullio de Mauro, con la collaborazione di Giulio C. Lepschy e Edoardo Sanguineti, Torino, U.T.E.T., voll. 6.

LO CASCIO Vincenzo (2012). **Dizionario Combinatorio Compatto Italiano**. Amsterdam: John Benjamins Publishing Company.

LO CASCIO Vincenzo (2013a). ***Dizionario Combinatorio Italiano***, ItalNed Foundation / University of Amsterdam: Amsterdam: John Benjamins.

RUSSO, Domenico (2010). ***MdD Modi di Dire Lessico Italiano delle collocazioni***. Roma: Aracne editrice.

SIMONE Raffaele et al. (diretta da) (2011), ***Enciclopedia dell'Italiano(EncIt),*** Roma, Istituto della Enciclopedia italiana.

TIBERII, Paola (2012)**.** ***Dizionario delle collocazioni***. Bologna: Zanichelli.

URZÌ, Francesco (2009). ***Dizionario delle Combinazioni Lessicali***. Lussemburgo: Convivium.

ZINGARELLI, Nicola (2015). ***Lo Zingarelli. Vocabolario della lingua italiana***. Bologna: Zanichelli. Edizione on line

Ouvrages

APRILE Marcello (2005). ***Dalle parole ai dizionari***, Bologna, Il Mulino.

BARBERA Manuel (2009), ***Schema e storia del "Corpus Taurinense". Linguistica dei corpora dell'italiano antico***, Alessandria, Edizioni dell'Orso.

BARBERA Manuel (2014) : « Le "multiword": consistenza di una categoria» in F: Cugno, L. Mantovani, M. Rivoira, M. S: Specchia (a cura di**),** ***Studi linguistici in onore di Lorenzo Massobrio***, Torino, Istituto dell'Atlante linguistico italiano, p.61-67.

BARBERA Manuel, MARELLO Carla . (2000) : « Les lexies complexes et leur annotation morphosyntactique dans le Corpus Taurinense » , ***Revue française de linguistique appliquée*** 2, 2 "Dossier. Diversité du traitement automatique des langues" p. 57-70.

BARONI Marco, BERNARDINI Silvia (2013): «Corpus Query Tools for lexicography» in R.H. Gouws, U. Heid et al. (eds.) (2013), p.1395-1405.

BELTRAMI Pietro G. (2013): « Theory of dictionary management» in R.H. Gouws, U. Heid et al. (eds.) (2013). p. 524-530**.**

BERRUTO Gaetano (2010): «Semantica strutturale e variazione : una rivisitazione», ***Incontri linguistici***, 33, pp.11-31.

BISETTO Antonietta (2004): «Composti e polirematiche» in M. Grossmann, F. Rainer (a cura di) (2004) p. 38 - 39.

CANTARINI Sibilla (2004). ***Costrutti con verbo supporto. Italiano e tedesco a confronto***, Bologna, Patron.

CANTARINI Sibilla (Hrsg.) (2013). ***Wortschatz, Wortschätze im Vergleich und Wörterbücher: Methoden, Instrumente und neue Perspektiven***, Frankfurt am Main, Peter Lang.

CRUSE D.Alan (1986) . ***Lexical Semantics***, Cambridge, Cambridge University Press.

DE MAURO Tullio (2005). ***La fabbrica delle parole. Il lessico e problemi di lessicologia***, Torino, UTET libreria.

EVERT Stefan (2013): «Tools for the acquisition of lexical combinatorics» in R.H. Gouws, U. Heid et al. (eds.) (2013), p. 1415-1432.

GAMBARARA Daniele e RAMAT Paolo a cura di (1977). ***Dieci anni di linguistica italiana (1965-1975),*** Roma, Bulzoni.

GROSSMANN Maria e RAINER Franz (a cura di) (2004), ***La formazione delle parole in italiano,*** Tübingen, Niemeyer.

GOUWS Rufus H., HEID Ulrich, SCHWEICKARD Wolfgang, WIEGAND Herbert Ernst (eds.) (2013) ***Dictionaries. An International Encyclopedia of Lexicography,*** Berlin/Boston, De Gruyter.

HAUSMANN, Franz Joseph, BLUMENTHAL, Peter (2006): «Présentation : collocations, corpus, dictionnaires. », ***Langue française***, n° 150 (2/2006), p. 3-13.

IANNÀCCARO Gabriele (a cura di) (2013) , ***La linguistica italiana all'alba del terzo millennio (1997-2010),*** Roma, Bulzoni.

JEŽEK Elisabetta (2005): ***Lessico. Classi di parole, strutture, combinazioni.*** Bologna: Il Mulino.

JEŽEK Elisabetta (2014) : «Esistono le collocazioni? Denotazione vs. significato collocazionale in P. Danler e Chr. Konecny (a cura di), ***Dall'architettura della lingua italiana all'architettura linguistica dell'Italia. Saggi in omaggio a Heidi Siller-Runggaldier***, Frankfurt am Main, Peter Lang , p. 409-419.

KILGARRIFF Adam, POMIKÁLEK Jan, JAKUBÍČEK Miloš, WHITELOCK Pete (2012): «Setting Up for Corpus Lexicography», *in* R. Vatvedt Fjeld and J.M. Torjusen (eds.) (2012). ***Proceedings of the 15th EURALEX International Congress.*** 7-11 August 2012. Oslo, Department of Linguistics and Scandinavian Studies, University of Oslo, p. 778-785.

LAVINIO Cristina (a cura di) (2002). ***La linguistica italiana alle soglie del 2000 (1987 – 1997 e oltre)***, Roma, Bulzoni.

LO CASCIO Vincenzo (2013b) «Multiple Beziehungen im Netzwerk: Wortverbindungen und ein-, zwei sparchige und bilingualisierte elektronische Wörterbücher », in S. Cantarini (Hrsg.) 2013, p. 237-266.

MARELLO Carla (1989) ***Dizionari bilingui. Con schede sui dizionari italiani per francese, inglese, spagnolo, tedesco***, Bologna. Zanichelli.

MARELLO C. (2010): « Multilexical Units and Headword Status. A Problematic Issue in Recent Italian Lexicography», in A. Dykstra, T. Schoonheim (eds) (2010), ***Proceedings of the XIV Euralex International Congress (Leeuwarden, 6-10 July 2010),*** Ljouwert, Fryske Akademy.pp.1347-1354; downloadable at http://www.euralex.org/publications/

MARELLO Carla (2013a) «*Sembra che* e subordinate soggettive. Primi sondaggi in italiano L2 scritto » in F. Geymonat (a cura di), ***Linguistica applicata con stile. In traccia di Bice Mortara Garavelli***, Alessandria, Edizioni dell'Orso, 2013, p. 79-94.

MARELLO Carla (2013b) : «Lessico e lessicografia» in G. Iannàccaro (a cura di) (2013), p. 557-580.

MARELLO Carla (2014): « *Dizionari di collocazioni italiane e collocazioni da insegnare nell'uso scritto*» in A. Molino, S. Zanotti, (eds.) (2014), p.183-199.

MASINI Francesca (2011): « polirematiche, parole, » in R. Simone et al. ***EncIt*** (2011), p.1109-1112.

MASINI Francesca (2012). ***Parole sintagmatiche in Italiano***, Cesena-Roma, Caissa.

MIONI Alberto M., CORTELAZZO Michele A. (a cura di) (1992), ***La linguistica italiana degli anni 1976-1986***, Roma, Bulzoni.

MOLINO Alessandra, ZANOTTI Serenella (eds.)(2014) ***Observing Norms, Observing Usage: Lexis in Dictionaries and the Media,*** Bern, Peter Lang.

PINNAVAIA Laura (2015): «Review:Simonetta Vietri Idiomatic Constructions in Italian. A Lexicon-Grammar Approach», ***International Journal of Lexicography,*** 28/2, p. 254-259.

PRANDI Michele, CANTARINI Sibilla (2013) « Metapher und Metonymie als Strategien der Bedeutungserweiterung am Beispiel des Deutschen und Französischen» in S. Cantarini(Hrsg.), p 17- 32.

PUSTEJOVSKY James (1991) «The generative lexicon», ***Computational Linguistics*** 17, 4, p. 409-441.

RENZI Lorenzo, SALVI Gian Paolo (a cura di) (2010) ***Grammatica dell'italiano antico***, Bologna, Il Mulino.

SGROI Salvatore Claudio (2006) «Dizionari a confronto. A proposito della Wortbildung nella lessicografia italiana» in E. Corino, C. Marello, C. Onesti (a cura di) ***Atti del XII Congresso Internazionale di lessicografia***, *Torino, 6-9 settembre 2006* ***Proceedings XII Euralex International Congress***, Alessandria, Edizioni dell'Orso, p. 1181-1192.

SGROI Salvatore Claudio (2010) «Etimologia sincronica *vs* etimologia diacronica (retrospettiva *vs* prospettica, remota *vs* prossima): problemi storico-teorici di terminologia linguistica», ***Studi Italiani di Linguistica Teorica e Applicata***, anno XXXIX, 2010, numero 1, p. 121-145.

SGROI Salvatore Claudio (2010): «Termini-chiave della Wortbildung polisemici e contraddittori e nuove proposte di definizione. A proposito di prefisso, confisso, parasintetico, polirematica», in ***Lessicologia e metalinguaggio*** Atti del convegno del Dipartimento di ricerca linguistica, letteraria e filologica, Università degli studi di Macerata 19-21 dicembre 2005, Roma, Editrice "Il Calamo", p. 557- 677.

SILLER-RUNGGALDIER, Heidi (2006): « "Le collocazioni lessicali: strutture sintagmatiche idiosincratiche? » In E. Cresti (ed.), ***Prospettive nello studio del lessico italiano.*** *Atti del IX Congresso della Società Internazionale di Linguistica e Filologia Italiana (SILFI). Firenze, 14-17 giugno 2006.* Vol. II. Firenze: Firenze University Press, 591-598.

SQUARTINI Mario (2013): «Semantica», in G. Iannàccaro (a cura di) (2013), p. 521-555.

THORNTON Anna Maria (2005) ***Morfologia*** , Roma, Carocci.

VOGHERA, Miriam (2004), «Polirematiche», in M. Grossmann & F. Rainer (2004) p. 56-69.

WIEGAND Herbert Ernst, SMIT Maria (2013) : «Microstructures in printed dictionaries» in R.H. Gouws, U. Heid et al. (eds.) (2013). p. 149-214**.**

Carla Marello
Dipartimento di lingue e letterature straniere e culture moderne
Università di Torino Italia
carla.marello@unito.it