

The French verb *affecter* in 3 types of news articles: a geometric representation with semantactic areas

Jacques François

Université de Caen & CRISCO, FRE 205

Abstract

Verb polysemy is linked with verb ‘polytaxy’, i.e. predicate frame variation (or in the terminology of Levin & Rappaport-Hovav 2005, ‘argument alternation’). The geometric representation of polysemy delivered by the DICTIONNAIRE ÉLECTRONIQUE DES SYNONYMES of CRISCO (see www.crisco.unicaen.fr) is a two- or three-dimensional visualization of the multidimensional semantic space of the item under study based on sense coordinates specified by cliques of synonyms, i.e. points of intersection of the semantic domain of a set of synonyms. The synonyms assembled in a particular clique generally share semantactic properties, i.e. they govern at least one common syntactic structure in the same sense. Therefore we can locate semantactic areas, i.e. pairs {context-sensitive sense; sense-giving context}. As to the French verb *affecter*, some of its predicate frames are salient in particular text types. The salient frames for three columns (France, International, Society) of the French newspaper *Le Monde* in 2005 are described.

Keywords: polysemy, verb, French, argument realization, electronic lexicography.

1. Syntactic-semantic classification of 200 occurrences of the verb *affecter* in 3 news columns of the newspaper *Le Monde* in 2005

In the digitalized version of the French daily newspaper *Le Monde*, the political and societal topics are arranged in three columns: FRANCE (politics), INTERNATIONAL and SOCIETY (concerning France in the first place). In ca. 6 millions words the verb *affecter* occurs 298 times (absolute frequency) i.e. with a relative frequency of $0,49 \cdot 10^{-4}$. By comparing the relative frequency of this verb in the three columns, one notices that this frequency is similar in FRANCE and SOCIETY with ca. $0,67 \cdot 10^{-4}$ and reduced by half in INTERNATIONAL with $0,33 \cdot 10^{-4}$ (see Table 1). At this level of investigation, that observation is difficult to interpret. As a matter of fact, some of the predicate frames associated with the 298 occurrences are only partially explicit as to the properties of the arguments, and only 200 fully explicit occurrences have been retained.

Distinguishing predicate frames (or argument structures) is not an easy task because

- i. a complement may be facultative,
- ii. the verb often occurs in the passive or in the medio-passive voice without mention of an agent,
- iii. the common subcategorization *qc* (something) vs. *qn* (someone) vs. *qp* (in some place) is poor, unable to distinguish physical entities and entities referring to various types of states of affairs (state, event, process, activity, action) and
- iv. the underlying ontology for selection restrictions is not set (for instance, two predicate frames *qn affecte qc<PSYCH:feeling>* and *qn affecte qc<nonPSYCH>* must be distin-

guished, illustrated by resp. *Marie affecte la colère* \approx Marie feigns anger, and *Les femmes africaines ne souhaitent pas affecter les liens du mariage* [*Le Monde* 2005, SOCIETY] \approx African women are not willing to question the conjugal relation).

		Size in words	Occurrences of <i>affecter</i>	Frequency *10 ⁻⁴	Occurrences with fully interpretable predicate frame
3 columns of <i>Le Monde</i>	France	ca. 2 175 000	147	0,676	96
	International	ca. 3 211 000	105	0,327	75
	Society	ca. 694 300	46	0,662	29
Total		6 080 300	298	0,490	200

Table 1: Distribution of the occurrences of the verb *affecter* across the 3 columns investigated

Finally, the 200 occurrences retained deliver 7 types of predicate frames, distinguished at four levels:

Level 1: VALENCY (e.g. *qn affecte qc* \rightarrow feign vs. *qn affecte qc (à qc/qn/qqp)* \rightarrow allocate / appoint)

Level 2: CATEGORIZATION (e.g. *qn affecte qc* \rightarrow feign vs. *qn affecte de INF* \rightarrow pretend to)

Level 3: SUBCATEGORIZATION (e.g. *qc affecte qc* \rightarrow influence vs. *qc affecte qn* \rightarrow move vs. *qn affecte qc* \rightarrow feign)

Level 4: restrictions of selection (see above).

The most represented predicate frame is *qc affecte qc* (\rightarrow sth influences sth) with 75 of the 200 occurrences (see Table 2) Three frames are very poorly represented in the threefold corpus: *qn affecte qc*<psych:feeling> \rightarrow to feign (4 occurrences), *qn affecte de INF* \rightarrow to feign (3) and *qn affecte qc*<nonpsych> \rightarrow to question (2), corroborating the data from *Le Monde* 2002 investigated in François (2007: 259), resp. 2, 1 and 0 from 100 occurrences (against resp. 26, 19 and 9 from 100 occurrences taken from contemporaneous French novels, revealing a significant text type contrast).

Predicate frames	frequency	
	absolute	relative*10 ⁻⁴
<i>qc affecte qc</i>	75	1,233
<i>qn affecte qc (à qc/qn/qqp)</i>	55	0,905
<i>qn affecte qn (à qn/qc/qqp)</i>	39	0,641
<i>qc affecte qn</i>	22	0,361
<i>qn affecte qc</i> <PSYCH:feeling>	4	0,066
<i>qn affecte de INF</i>	3	0,049
<i>qn affecte qc</i> <nonPSYCH>	2	0,033

Table 2: Absolute and relative frequency of the 7 types of predicate frames distinguished

In Table 3, the 200 occurrences distributed along 7 types of predicate frames are distributed across the three columns. The Table reveals that:

- i. the 3 last types are non significant in each column;

- ii. the proportion of uses in the passive voice is high in FRANCE and SOCIETY and significant in INTERNATIONAL;
- iii. the distribution among the 4 significant frame types significantly differs in the 3 columns.

	absolute and relative frequency (*10 ⁻⁴)	qc affecte qc	qn affecte qn	qn affecte qn (à qn/qc/qqp)	qn affecte qc (à qc/qn/qqp)	qn affecte qc<nonPSYCH>	qn affecte qc<PSYCH>	qn affecte de INF	TOTAL	PASSIVE VOICE
FRANCE	abs	23	11	14	43	0	2	3	96	51
	relative	0,11	0,05	0,06	0,20	0	0,01	0,01	0,44	
INTERNATIONAL	abs	49	10	6	9	1	0	0	75	30
	relative	0,15	0,03	0,02	0,03	0	0	0	0,23	
SOCIETY	abs	3	1	19	4	1	1	0	29	17
	relative	0,04	0,01	0,27	0,06	0,01	0,01	0	0,42	
3 columns		75	22	39	55	3	2	4	200	

Table 3: Distribution of the predicate frame occurrences of *affecter* across the 3 columns

That last observation is further illustrated in Table 4 on which Figure 1 is based: among the four frame types

- i. *qc affecte qc* → INFLUENCE, is moderately represented in FRANCE and INTERNATIONAL and poorly in SOCIETY
- ii. *qn affecte qn* → MOVE, is poorly represented in the three columns, with a light overrepresentation in FRANCE
- iii. *qn affecte qn (à qn/à qc/qp)* → APPOINT, is highly represented in FRANCE and poorly in the two other columns, and
- iv. *qn affecte qc (à qc/à qn/qp)* → ALLOCATE, is highly represented in FRANCE and poorly elsewhere.

Table 4 and Figure 1 display the distribution of the 4 significant constructions across the 3 columns. The prevalence of *qn affecte qn (à qn/à qc/qp)* in SOCIETY and *qn affecte qc (à qc/à qn/qp)* in FRANCE stands out.

Some illustrations of the four main constructions of *affecter* in *Le Monde* 2005:

- i. *qc affecte qc*, SYN: *influencer*, 75 occurrences, prevalent in FRANCE (0,11* 10⁻⁴) and INTERNATIONAL (0,15* 10⁻⁴)
 - 51 occurrences in the active voice, e.g.

De fait, la zone franche de Manaus, créée en 1967 pour compenser la crise économique qui **AFFECTAIT** la principale activité de l'époque, la récolte du caoutchouc, est un succès

Relative frequency (value * 10 ⁻⁴)		France	International	Society
<i>qc affecte qc</i>	SYN <i>influencer</i>	0,11	0,15	0,04
<i>qc affecte qn</i>	SYN <i>toucher</i>	0,05	0,03	0,01
<i>qn affecte qn (à qn/à qc/qp)</i>	SYN <i>nommer</i>	0,06	0,02	0,27
<i>qn affecte qc (à qc/à qn/qp)</i>	SYN <i>destiner</i>	0,20	0,03	0,06

Table 4: The most represented frame types in each of the 3 columns of *Le Monde*

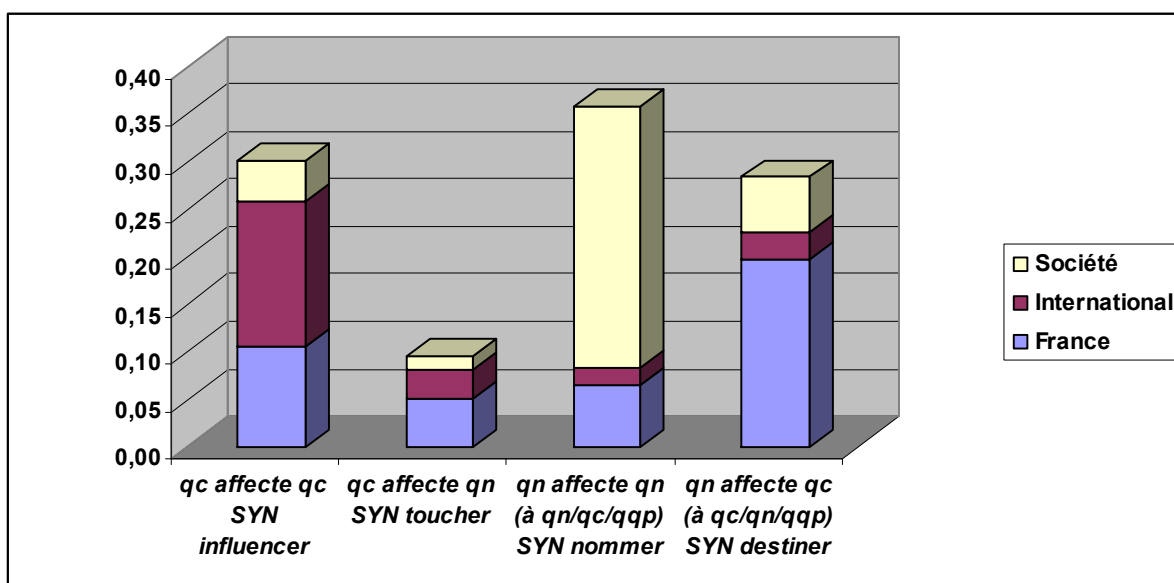


Figure 1: Distribution of the 4 significant constructions across the three columns of *Le Monde* 2005

Un groupe de représentants républicains a proposé des coupes budgétaires drastiques **AFFECTANT** les services publics, la compagnie ferroviaire Amtrak, ...

- 24 occurrences in the passive voice, e.g.

Selon lui, 29 000 tentes et 118 000 couvertures ont été distribuées dans les zones **AFFECTÉES**, alors que les besoins sont estimés à 100 000 tentes et 2 millions de couvertures.

Les entreprises privées qui opèrent dans la région nord-est ont également été **AFFECTÉES**, leur production ayant chuté de 340 000 à 150 000 barils.

- ii. *qc affecte qn*, SYN: *toucher*, 22 occurrences, barely represented in 'France' (0,05* 10⁻⁴)

- 11 occurrences in the active voice, e.g.

Les données statistiques issues de l'Insee confirment que cette pathologie des temps modernes **AFFECTE** davantage les classes sociales défavorisées.

L'ensemble des activités de soins sont concernées par ces événements. Toutefois, ils **AFFECTENT** plus fréquemment des personnes fragiles (sujets âgés, sévérité de l'état de santé ...)

- 11 occurrences in the passive voice, e.g.

Le "Vieux", néanmoins, semble **AFFECTÉ** par ce conflit l'opposant à ses anciens collaborateurs.

un grand appartement hébergeant six malades **AFFECTÉS** par des « handicaps profonds »...

- iii. *qn affecte qn* (à *qn/qc/qqp*), SYN: *nommer*, 39 occurrences, prevalent in 'Society' ($0,27 * 10^{-4}$)
- 9 occurrences in the active voice, e.g.

le ministère a répondu que les rectorats pourront, en fonction des besoins, **AFFECTER** les assistants d'éducation à ces tâches.

La loi du 13 août a également prévu que les inspections d'académie gardent la responsabilité d' **AFFECTER** les élèves et donc d'accorder, ou non, les éventuelles dérogations.
 - 30 occurrences in the passive voice, e.g.

Elle a été **AFFECTÉE**, en 2004, auprès d'un jeune autiste sévère, scolarisé dans un établissement ordinaire.

'examen des faits devrait débiter mardi 5 avril, avec l'audition de trois des enquêteurs **AFFECTÉS** à l'époque à la brigade des mineurs d'Angers.
- iv. *qn affecte qc* (à *qc/à qn/qp*), SYN: *destiner*, 55 occurrences, prominent in FRANCE ($0,20 * 10^{-4}$)
- 23 occurrences in the active voice, e.g.

Deuxième question : le Parlement a débattu, un vote a eu lieu, **AFFECTANT** les dividendes des sociétés d'économie mixte à l'Agence de financement des infrastructures de transport

le gouvernement devra décider s'il **AFFECTE** à celle-ci une nouvelle ressource fiscale
 - 32 occurrences in the passive voice, e.g.

Enfin, les élus peuvent, sans dat -butoir, céder la charge de tout autre bâtiment s'il reste **AFFECTÉ** à "une mission de service public" ou à "une opération d'intérêt général relevant de sa compétence"

Les cendres peuvent être dispersées, soit dans un espace **AFFECTÉ** à cet effet dans le cimetière ou attendant au crématorium, soit en pleine nature.

It turns out that the preference for the passive voice is strong for *qn affecte qn* / *qn est affecté par qn* (à *qc/à qn/qp*) with 77% of the occurrences and moderate for *qn affecte qc* / *qc est affecté par qn* (à *qc/à qn/qp*) with 58%. The active and passive figures are identical for *qc affecte qn* and the active voice is strongly preferred for *qc affecte qc* with 68%. Such preferences seem to be related to the valency of the construction: the passive voice is prevalent in constructions with 3 arguments and the active voice in constructions with only 2 arguments.

2. The geometric representation of the semantic space of the verb *affecter* and the semantactic areas of the 4 significant constructions

The *Dictionnaire Electronique des Synonymes* of CRISCO in Caen (www.crisco.unicaen.fr) is based on synonymy links delivered by 7 dictionaries (2 extensive dictionaries of the French language, 2 synonymy dictionaries of the 19th and 3 synonymy dictionaries of the 20th century). These 450000 links concerning 45000 items, once symmetrized, build a large relational network. For any item under study, the *DES* delivers a multidimensional semantic space consisting in a scatter of points. These points represent the cliques of synonyms, i.e. the smallest points of intersection among the semantic spaces of related synonyms, and the space of the item encompasses as many dimensions as synonyms involved. The *DES* delivers for the item under study a default geometric representation on the surface offering the best distribution of the scatter of points.

For *affecter*, the 4 main synonyms are located on the surface 1*2:

- i. in the middle left region for **altérer** ($\Leftrightarrow qc\ affecte\ qc$) and **toucher** ($\Leftrightarrow qc\ affecte\ qn$)
- ii. in the upper right region for **nommer** ($\Leftrightarrow qn\ affecte\ qn$ (à $qn/à\ qc/qp$))
- iii. in the lower right region for **destiner** ($\Leftrightarrow qn\ affecte\ qc$ (à $qn/à\ qc/qp$))

At the left end of the visualization one finds the clique $\{affecter, affliger, attrister, chagriner, navrer, peiner\}$ which is semantically and syntactically homogenous¹ in its link to the use $qc\ affecte\ qn$. But this clique does not include the most general synonym related to that construction, namely *toucher*. Therefore the clique appears indeed outside of the space of *toucher* but in the same region. Apparently in the same area the extreme clique of the space of synonym *altérer* typically linked with the use $qc\ affecte\ qc$ excludes *toucher*, although it seems to belong to the space of this synonym on the surface $1*2$. It means that we have to do with 2 different layers of representation which are projectionally mixed up on this surface (they would appear at distance on another surface of projection). The extreme clique in the upper right region is $\{affecter, afficher, montrer, offrir, présenter, révéler\}$ linked with the use $qn\ affecte\ qc$ and the extreme clique in the lower right region is $\{affecter, appliquer, attribuer, consacrer, destiner\}$ concerned with the use $qn\ affecte\ qc$ (à $qn/à\ qc/qp$). Inbetween you find the space of synonym *nommer*. Its upper extreme clique contains *révéler*, like the upper right clique and its lower extreme clique contains, as well as *désigner*, like the upper clique of *destiner*. This suggests that the three concerned clique are located on related layers. On the other hand, the lower extreme clique of *nommer* consists of very different synonyms $\{affecter, installer, nommer, titulariser\}$ apparently pertaining to another layer.

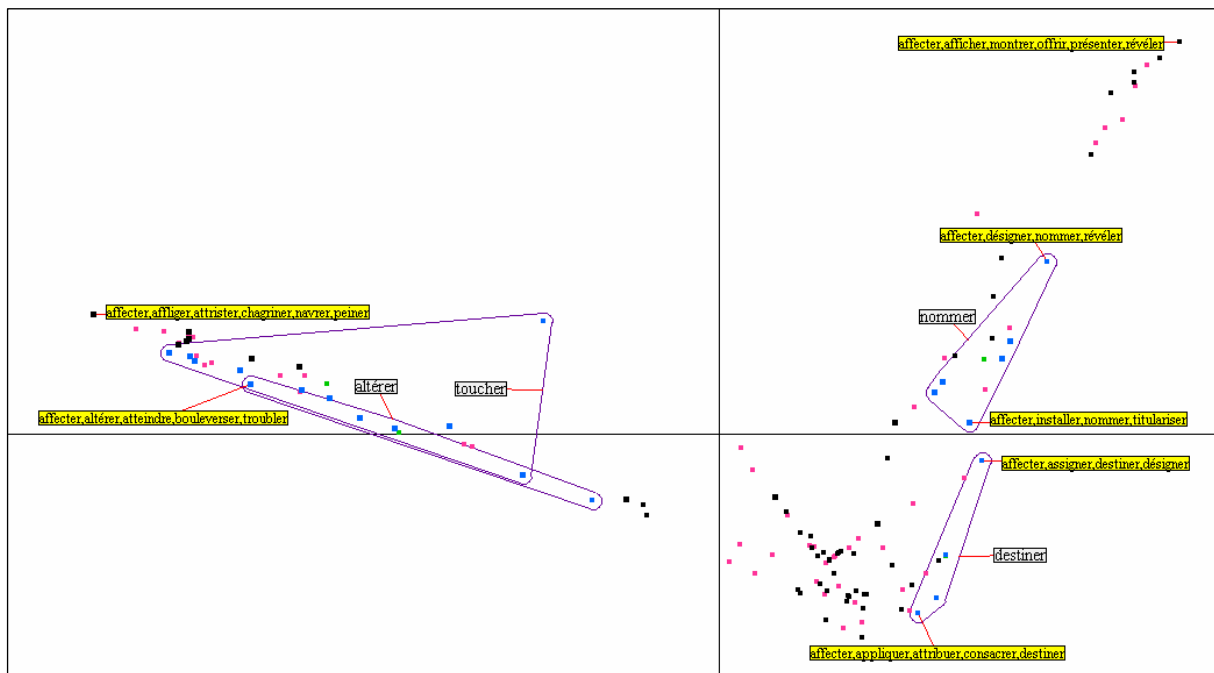


Figure 2: Projection of the semantic space of affecter on the surface $1x2$

- Cliques of synonyms compatible with the 4 significant constructions

¹ Extreme cliques turn out to be always homogenous. Heterogenous cliques, i.e. cliques whose some components share one syntactic and/or semantic property and some other components another incompatible property, are located at the interface between two semantactic areas, preventing the fragmentation of the semantic space of the item under study.

i. *qc affecte qc*, SYN: *altérer*, Engl. *to influence*

One single clique is associated to the frame *qc affecte qc* (sth influences sth), namely

2 : affaiblir, affecter, altérer, atteindre

That clique includes *atteindre* that pertains to the 4 of the 13 cliques associated below with *qc affecte qn* (sth moves so). Therefore there is only one semantactic area for *qc/qn affecte qn* with clique 2 getting the special status of sense coordinate of \boxed{qc} *affecte qc*.

ii. *qc affecte qn*, SYN: *toucher*, Engl. *to move*

The semantactic area of *qc affecte qn* located in the middle left region of the semantic space of *affecter* is marked out by 13 cliques.

10 : affecter, affliger, atteindre, frapper, toucher, troubler, émouvoir

11 : affecter, affliger, attrister, chagriner, navrer, peiner

12 : affecter, affliger, attrister, frapper, émouvoir

13 : affecter, affliger, chagriner, troubler

34 : affecter, attaquer, atteindre, frapper, toucher

38 : affecter, atteindre, blesser, frapper, toucher, émouvoir

39 : affecter, atteindre, bouleverser, frapper, toucher, troubler, émouvoir

44 : affecter, blesser, intéresser, toucher, émouvoir

45 : affecter, blesser, navrer

47 : affecter, bouleverser, frapper, impressionner, toucher, troubler, émouvoir

48 : affecter, chagriner, gêner, navrer

49 : affecter, chagriner, gêner, troubler

67 : affecter, frapper, marquer, toucher

Mentioned in 7 cliques, the synonym *toucher* plays a central role, but its space is not exactly correlated with that of *qc affecte qn*, since on the one hand its extreme clique at the lower right end {*affecter, attaquer, jouer, toucher*} contains the synonym *jouer*, which cannot replace *affecter* in that frame, and on the other hand 3 of the 13 cliques are located outside the space of *toucher* beyond its left end.

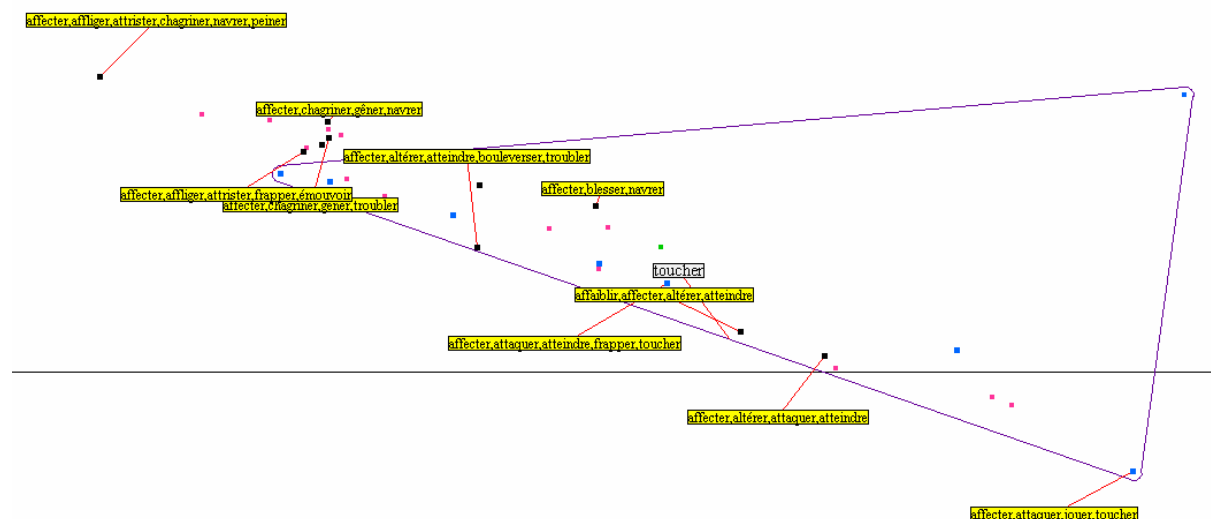


Figure 3: Zoom on the semantactic area of *affecter* ⇔ *toucher*

iii. *qn affecte qn* (à *qn*/à *qc/qp*), SYN: *nommer*, Engl. *appoint*

The next type of use is represented by *qn affecte qn* (à *qn*/à *qc/qp*.) Engl. *to appoint*. 7 synonymy cliques consist of synonyms sharing this frame in the same sense. Here, the most mentioned synonym (in 5 of the 7 cliques) is *nommer*, followed by *désigner* (in 3 cliques).

The semantactic area does not overlap the upper part of the space of *nommer*, because the clique {*affecter*, *désigner*, *nommer*, *révéler*} at the upper end of its space, as well as the clique {*affecter*, *assigner*, *désigner*, *marquer*} contain one excluded synonym, resp. *révéler* and *marquer*.

50 : affecter, choisir, désigner, nommer, prendre
 57 : affecter, déplacer, muter
 58 : affecter, déplacer, nommer
 60 : affecter, désigner, nommer, qualifier

62 : affecter, désigner, nommer, titulariser
 68 : affecter, installer, mettre en place
 70 : affecter, installer, nommer, titulariser

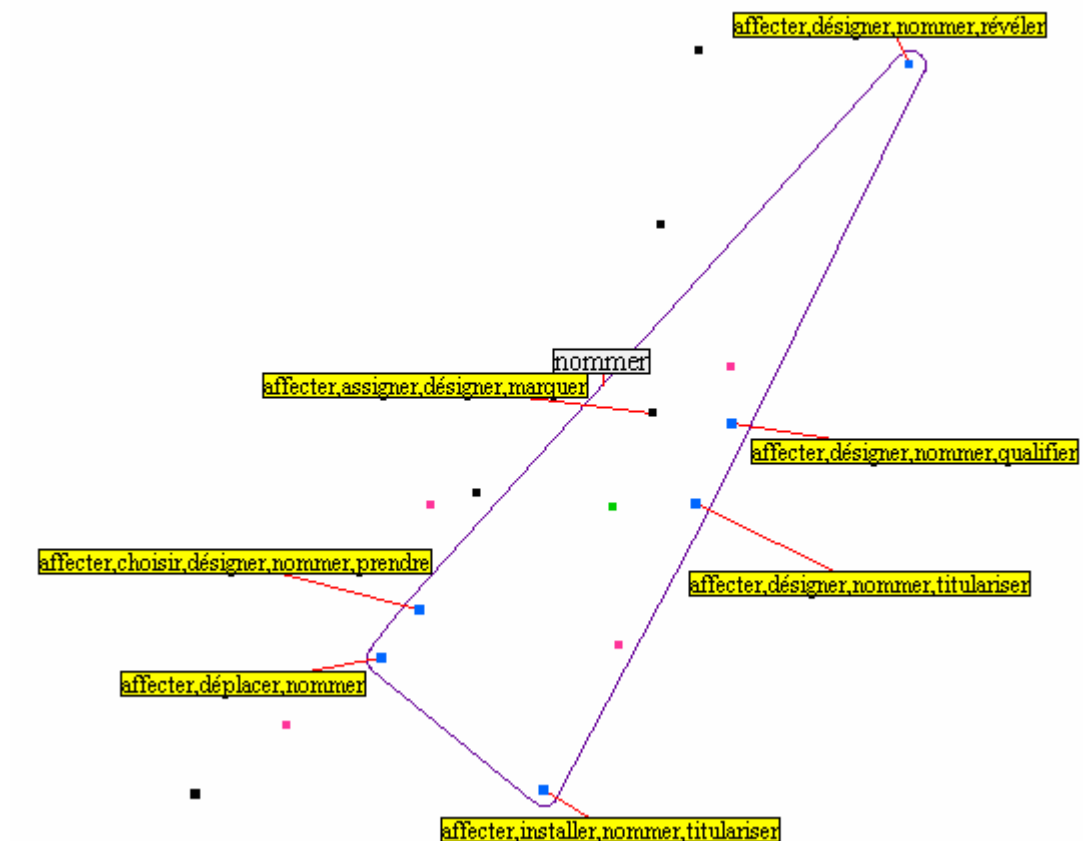


Figure 3: Zoom on the semantactic area of *affecter* ⇔ *nommer*

iv. *qn affecte qc (à qc/à qn/qpp)*, SYN: *attribuer* / *destiner*, Engl. *allocate*

The last type of use is that of *qn affecte qc (à qc/à qn/qpp)*, Engl. *to allocate* represented by 7 synonymy cliques. The most mentioned synonym is *attribuer* followed by *destiner*. *Appliquer* occupies the lower part of the semantactic area, *mettre* the left part, *classer* the upper part and *destiner* the right part.

24 : affecter, appliquer, attribuer, consacrer, destiner
 25 : affecter, appliquer, attribuer, consacrer, mettre
 26 : affecter, appliquer, attribuer, imputer, mettre
 28 : affecter, assigner, attribuer, classer

29 : affecter, assigner, attribuer, destiner
 40 : affecter, attribuer, classer, mettre
 41 : affecter, attribuer, destiner, réserver

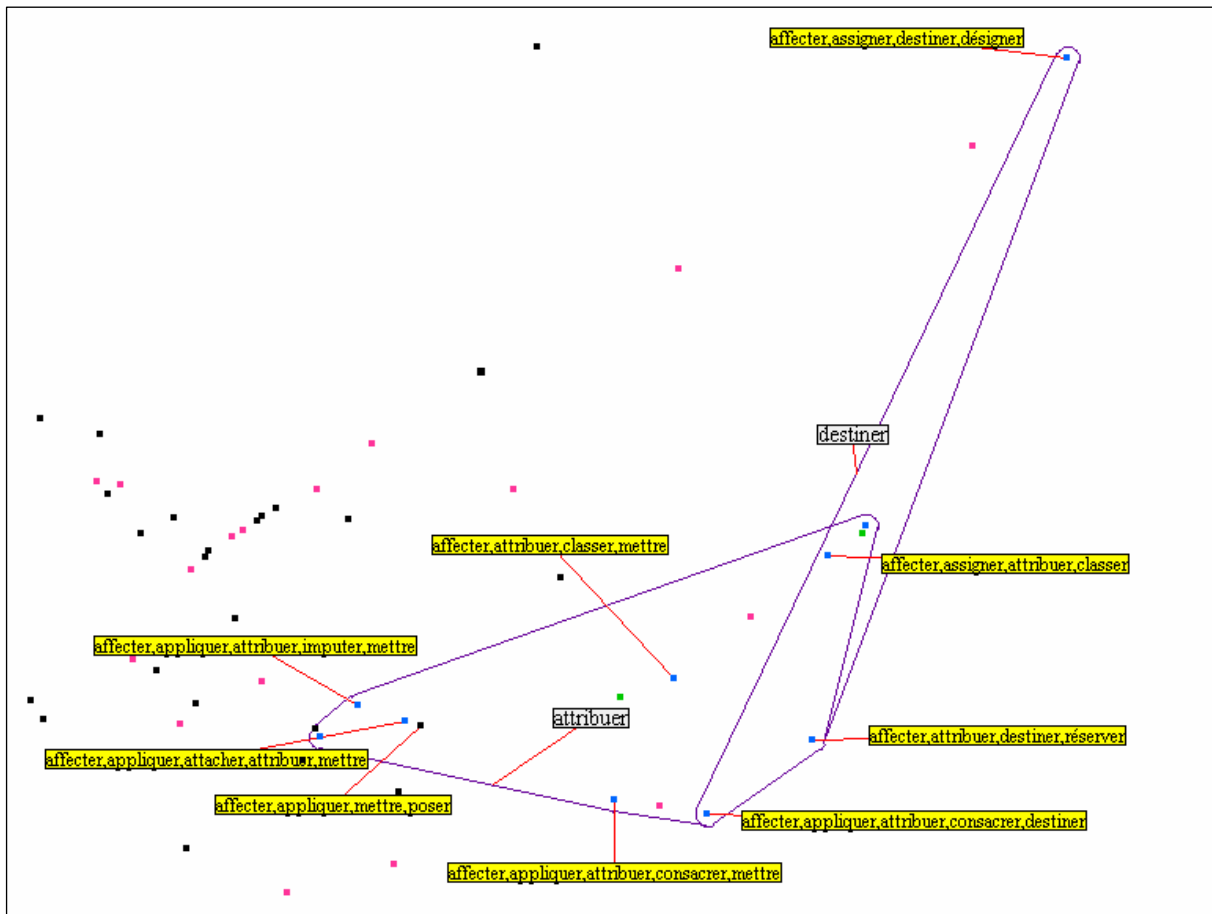


Figure 3: Zoom on the semantactic area of *affecter* ↔ *attribuer-destiner*

3. Concluding remarks

The purpose of this paper was to relate two series of observations concerning verb polysemy and polytaxy (or ‘argument alternation’) in French: on the one hand the distribution of various verb senses related to various predicate frames in 3 different types of news articles, and on the other hand the location of these verb senses and frames in the projectional geometric representation of the semantic space of the verb item under study using the *Dictionnaire Électronique des Synonymes* of CRISCO. Both topic are illustrated by the French verb *affecter* which enjoys a rich polysemy linked with a polytaxy described in terms of valency and subcategorization.

In next future a tridimensional visualization of the semantic space of the item under study shall be available on the site of CRISCO, delivering a more vivid representation of the different layers (preventing for instance the damaging overlap on the surface 1*2 between the two areas of *affecter* ↔ *nommer* and *affecter* ↔ *destiner* in the right region of the semantic space of *affecter*).

References

- François J. (1996). Semantische und syntaktische Gliederungskriterien in der Verbpolysemie: Entwurf eines Kalkülverfahrens am Beispiel von frz. *affecter*. In: E. Weigand & F. Hundsnurscher (eds. 1996), *Lexical structures and language use*, Vol.2: Session Papers. Tübingen, Niemeyer, p.53-64.

- François J. (2005). Polysémie verbale et cadres participatifs: *demander* et ses synonymes. In F. Lambert & H. Nølke (eds.), *La syntaxe au cœur de la grammaire*. Collection Rivages linguistiques. Presses Universitaires de Rennes, p.85-98.
- François J. (2007). *Pour une cartographie de la polysémie verbale*. Louvain, Peeters [Bibliothèque de la Société de Linguistique de Paris].
- François J. (to appear). Combining a semantic map and a dual text corpus analysis for analyzing verb polysemy (illustrated by French *jouer*). In *Proceedings of the colloquium Ambiguity and Disambiguation, Sfax, 5-6 May 2006*.
- Levin B. & Rappaport-Hovav M. (2005). *Argument realization*. Oxford, New-York: Oxford University Press.
- Victorri B. & Fuchs C. (1996). *La polysémie – Construction dynamique du sens*. Paris, Hermès.
- Victorri, B., Jacquet G. and Venant F. (2005). Polysémie lexicale. Chapter 3 of P. Enjalbert (ed. 2005), *Sémantique et traitement automatique du langage*. Paris: Hermès-Lavoisier, p.99-132.
- Victorri B. & Ploux S. (1998). Construction d'espaces sémantiques à l'aide de dictionnaires de synonymie. In *TAL* 39, p.161-182.