A CROSS-LINGUISTIC INSIGHT ON AGENTIVE NOUN FORMATION IN ITALIAN AND FRENCH

Bruno Cartoni
Département de Linguistique, Université de Genève
Fiammetta Namer
ATILF-Université de Lorraine
Stéphanie Lignon
ATILF-Université de Lorraine

Abstract
The objective of this paper is to describe a contrastive dictionary-based study on word-formation and to show the advantages of the contrastive approach to improve the description of lexical morphology cross-linguistically. The present article focuses on the formation of agentive nouns. We first describe the methodology used to extract the contrastive data from a bilingual dictionary, then, through different case studies of specific agentive nouns formation, we show the benefit from such contrastive analysis and the new light that it sheds onto each morphological system taken individually.

1 Contrastive approach to morphology

Contrastive lexical morphology is the study of word-formation in two or more languages in parallel in order to highlight cross-linguistic morphological similarities and differences and to shed new light on the languages considered individually (see Lefer, 2011 for an exhaustive overview of contrastive morphology research). It is anchored in the recent tradition of contrastive linguistics (e.g., see James, 1980; Fisiak, 1983; Ringbom, 1994), which addresses different linguistic phenomena in a contrastive perspective.

Studying languages in contrast has often an “applied linguistic” objectives, such as the setup of second language learning method or the constitution of

---

1 This work was supported in 2012 and 2013 by the Switzerland-France Joint Research Program EGIDE ‘Germaine de Staël’ Grant n°26433YL: ‘MOCOCO - MORphologie, COrpus et analyse CONtrastive’.
bilingual lexicography works. But the systematic comparison of two languages can also bring new theoretical insight. In the present study, we show that contrastive approach can ease a “meaning-based approach” (sometime referred as onomasiological) to the study of word-formation, i.e. an approach that first takes an object of study that is defined by semantic types or features, and then looks at the way different languages realize the object formally (through word-formation processes or not). Contrary to classical studies in morphology that usually focus on one particular word-formation process at a time (through the collection of potential complex lexemes coined with the considered process), the bi-directional contrastive analysis proposed here allows gathering all the complex lexemes that take part of a given semantic class, coined or not by a word-formation process. As a consequence, this approach allows considering the lexicon as a whole, and questions interfaces between lexical morphology and other linguistic means of denotation (syntactic paraphrases, simplex lexemes, borrowed lexemes, etc.). Thus, the boundaries of lexical morphology in the constitution of the lexicon are also assessed.

2 Word-formation processes under study: the agentive nouns

In this paper, we focus on nouns denoting agents, as described in lexical semantic studies such as Anscombe 2003, Busa, 1997, Cruse 1973, Van Valin & Wilkins 1996, or within a morphological perspective as in Booij 1986 and Grossmann 1998. Nouns denoting a human gave rise to different propositions of classification. The hierarchy of human nouns being multi-dimensional (according to various factors, such as gender, relation, profession, property, role), we focus here exclusively on characterizing agents, that are named following their behavior, or the ideal they claim to follow, and on classifying agents, or “actors”, that regularly take part of an activity. (see Lo Duca (2004) and Roché (2011 a)).

Two main aspects have been taken into account in order to design the classification of agent nouns. The former is based upon lexical properties of nouns, the latter upon morphological ones. The identification of lexical properties of human nouns is guided by syntactic tests partly inspired by G. Gross’s so-called object classes (see Gross 2009, 2011), and involving verb operators. For instance, the noun used to identify the agent in a clause headed by the French verb ‘pratiquer’ (to practice) or the Italian verb ‘praticare’ denote the actor of a professional or leisure activity (1), whereas the noun compatible with fr : ‘exercer la fonction de’ / it : ‘esercitare/svolgere la funzione di’ (to act as or to hold an office of) refers to a distinguished member of a professional or social hierarchy (2). In (1) and (2), tests and examples are given for French. In (2a), RelA stands for ‘relation adjective’.
(1) a Qui est X ? – X est quelqu’un qui \textit{pratique}, N1 \rightarrow X est un NH
+ prof/leisure [Who is X? – X is someone who practices N1 \rightarrow X is a NH]
b N_{FR} = course, pêche à la ligne, médecine, journalisme, rugby \rightarrow NH_{FR} = coureur, pêcheur à la ligne, médecin, journaliste, rugbyman

(2) a Who is X? - X is quelqu’un qui \textit{exercer la-une fonction}, (RelA + de NH) \rightarrow X est un NH_{fonction} [Who is X? – X is someone who (acts as / holds an office (of)) (RelA + NH) \rightarrow X is a NH_{fonction}]
b RelA_{FR} : dirigeante, politique, ministérielle, officielle, présidentielle, exécutive, judiciaire \rightarrow NH = un dirigeant, homme politique, ministre, officiel, président, exécutif, magistrat
c NH_{FR} : capitaine, ministre, procureur, chef, secrétaire, directeur, président, officier, juge, adjoint

By adapting and extending the initial set of Gross’s operator verbs, we come to the hierarchy presented in Figures 1 and 2. Following Lo Duca (2004) partition, the first distinction is done between characterizing agents (Figure 1), who can be defined by ‘support, are proponents/supporters/followers of X’ for followers, and by ‘have the habit of / are used to X’ for behaviors, where X is morphologically related to the human noun. Actors (Figure 2) are sub-classified in the same way: for instance, among nouns who can be seen as agent of a ‘salient activity’, test-crossings allow to store those matching (1a) with the class ‘Employee or manager’, ‘Manual activity or selling’, ‘Agriculture, livestock farming’ or ‘Sport and leisure’ according to further characterizations.

![Figure 1: Agent Nouns Classification: characterizing agents](image)
The second device which contributes to the distribution of agent nouns within this classification makes use of word formation knowledge obtained by following a Word Based approach (Aronoff 1994, Fradin 2003); it consists in defining each complex word with respect to its base(s), and consequently, to group words whose definition instantiates the same semantic pattern, and whose base belongs to the same semantic type. Table 1 shows what kind of properties can be inferred from morphological analysis. Approximately 70 semantic relations are needed to account for the set of agent nouns we are dealing with in this study.
### Table 1: Semantic features provided by the morphological analysis

<table>
<thead>
<tr>
<th>Agent noun definition wrt its base</th>
<th>NBase Semantic type</th>
<th>Base</th>
<th>Agent Noun_{lg}</th>
<th>Formation pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>He/she whose work or occupation consists of building and/or selling Nbase</td>
<td>object</td>
<td>fiore_{IT}</td>
<td>fioraiot_{IT}</td>
<td>-aiot_{IT}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fleur_{FR}</td>
<td>fleuriste_{FR}</td>
<td>-iste_{FR}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gioiello_{IT}</td>
<td>gioielliere_{IT}</td>
<td>-iere_{IT}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bijou_{FR}</td>
<td>bijoutier_{FR}</td>
<td>-ier_{FR}</td>
</tr>
<tr>
<td>He/she whose work/occupation place is Nbase</td>
<td>location</td>
<td>posteggiatorem_{IT}</td>
<td>posteggiatore_{IT}</td>
<td>-ore_{IT}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>gardien de parking_{FR}</td>
<td>N p N_{FR}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bar_{FR}</td>
<td>barista_{IT}</td>
<td>-ista_{IT}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>barman_{FR}</td>
<td>-ist_{FR}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>borrowing_{FR}</td>
<td></td>
</tr>
<tr>
<td>Musician who uses Nbase</td>
<td>Music Instrument</td>
<td>arpa_{IT}</td>
<td>arpista_{IT}</td>
<td>-ista_{IT}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>harpe_{FR}</td>
<td>harpiste_{FR}</td>
<td>-iste_{FR}</td>
</tr>
<tr>
<td>He/she whose work or occupation consists of doing Nbase</td>
<td>activity</td>
<td>maratona_{IT}</td>
<td>maratoneta_{IT}</td>
<td>-eta_{IT}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>marathon_{FR}</td>
<td>marathonien_{FR}</td>
<td>-ien_{FR}</td>
</tr>
</tbody>
</table>

The multi-dimensional classification described above, including word-formation process, semantic reference to the base and semantico-referential criteria is then applied to the contrastive data extracted from the bilingual dictionary. The extraction methodology is described in the next section.

### 3. Bilingual dictionary as a source of data

Morphological studies usually rely on the collection of complex lexemes, either gathered from large language repository (such as dictionary) or collected in textual corpora. Contrastive morphology follows the same methodology, but the multilingual aspect with such an approach is particularly challenging.

While many contrastive morphology studies are based on multilingual corpora (either comparable or parallel, see for example Cartoni & Lefer 2001, and Lefer 2011), we decide here to focus on a bilingual dictionary, for several reasons. First of all, bilingual dictionaries (like monolingual ones) are exhaustive to some extend: they represent a stable representation of the state of the lexicon at a certain point, and do not depend on the context of production of texts in corpora. And this can be particularly crucial for some word-formation processes. For example, when looking into the French-Italian
parallel corpus extracted from the Europarl corpus (Cartoni and Meyer 2012), - just as (Cartoni & Lefer 2011) did for English-French-Italian contrastive analysis of negative word-formations, only 69 different occurrences of complex lexemes coined in –ista were found in the Italian side (compared to 954 occurrences in –ista of the bilingual dictionary used in this study). In addition, most of them were of the “characterizing agent type” (communista, “communist”, monopolista “monopolist”, probably because of the nature of such corpus (parliamentary debates).

The second advantage of bilingual dictionary is that they can be manipulated more easily than corpus data, where the extraction can require important manipulations.

However, bilingual dictionary have also important drawbacks that one should keep in mind when analyzing the extracted data. First of all, bilingual dictionaries do not contain neologisms, i.e. lexemes that were coined recently and that could bear witness to the productivity of the word-formation rule that coined it (in this respect, extraction from corpus combined with frequency or productivity factor would be better used). The second disadvantage depends on the way bilingual dictionaries are conceived. Rarely based on parallel corpora, bilingual dictionaries’ quality strongly relies on the lexicographers’ work, the way cross-linguistic equivalence is established. In the same vein, practical factors (such as the size of the dictionary and hence, of each entry) may also intervene with the quality of the dictionary. Nonetheless, as we will present in this paper, bilingual dictionaries represent an inestimable source of data for the study of two languages in contrast.

3.1 Extraction methodology

In this study, we rely on the Italian-French bilingual dictionary (Garzanti, 2006), that contains 65,308 entries in the It⇒Fr direction, and 62,046 entries in the Fr⇒It direction. The extraction methodology that was set in place to acquire bilingual data is a recursive one, where every cycle contains four steps.

First, we extracted French entries denoting human nouns and constructed with one of the most frequent suffixation rule -iste, -eur, -ien and -ier (these suffixes were chosen because they were the most frequent translation of Italian entries suffixed in -ista, Cartoni&Namer 2012). In a second step, we duplicate entries that are polysemic and that give rise to different translations. For instance, accessoiriste is translated by Italian trovarobe or accessorista, whether it denotes an employee working in a theater or in a garage.

In the third step, we categorize the nouns according to the classification described in section 2. In parallel, the translations proposed by the dictionary are also classified according to their formal structure. We distinguish lexemes that are no more analyzable (athlète translation of agonista),
morphologically complex lexemes, distinguishing the word formation process that produced it, such as suffixation rules (e.g. suffixed in –aire : mousquetaire, translation of moschettiere), or compounding rules (guardasigilli, translation of chancelier). We also distinguished translation formed following syntactic pattern (it: datore di lavoro, translation of employeur), and translation provided as a definition (fr: personne qui fait du marché noir, as a translation of borsanerista). Finally, in the fourth step, we reiterate the preceding steps for each Italian suffix that was the most frequently used in the translation of French lexemes in –ien, -iste, -ier et –eur.

At the end of this second cycle, we found the initial French suffixes, and also some new suffixes that are used to coin agentive nouns. If they are frequent enough, step 1 to 4 is reiterated.

Figure 3 sketches the iterative process to acquire the parallel data (only most frequent relations are displayed). From a selection of French suffixes, Italian suffixes are individualized. From these, parallel data in the other direction of translation (It->Fr) are acquired. From double extraction, we also uncover non-morphological processes that are also included in the analysis (labeled as “other” in Figure 3).

Figure 3: Iterative process for gathering data (French -> Italian -> French)
Thanks to this methodology, we obtained an exhaustive set of nouns in French and Italian. In total, 2134 entries were gathered for the Fr$\rightarrow$It direction, and 2429 for the Italian$\rightarrow$French direction. Some entries containing more than one translation, a total of 2608 translations were found into French, and 3008 into Italian. In those cases, we reduplicate the entries to obtain 1:1 translation pairs.

Each pair of noun is provided as a translation equivalent by the dictionary, and so can be considered as cross-linguistically equivalent. Each pair is then classified according to the semantico-referential classification (see Figures 1 and 2). Then, each part of the pair is classified according to their word-formation processes (or other means of denotation) and their semantic relation to the base.

4. Benefit from contrastive morphology.

In this section, we present the results of the extraction and classification procedures explained in Section 3. We first provide an overview of the figures that allows the quantification of noun formation processes in the two languages (Section 4.1). This first overview precedes a more detailed presentation of some interesting cases of divergences (Section 4.2).

4.1 Quantifying divergences between languages

The first quantitative analysis is presented in Tables 2 and 3, where we distinguish morphological and non-morphological translation (i.e. translation that is realized though a syntactic construction, a definition, or a non-constructed word). For the sake of clarity, Tables 2 and 3 display only the most frequent configuration.

<table>
<thead>
<tr>
<th>Fr$\rightarrow$ It</th>
<th>Morphological translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-iste (732)</td>
<td>76.3%</td>
</tr>
<tr>
<td>actors (482)</td>
<td>87.4%</td>
</tr>
<tr>
<td>followers (222)</td>
<td>96.4%</td>
</tr>
<tr>
<td>behaviours (28)</td>
<td></td>
</tr>
<tr>
<td>-eur (1482)</td>
<td>76.4%</td>
</tr>
<tr>
<td>actors (1216)</td>
<td>76.2%</td>
</tr>
<tr>
<td>followers (256)</td>
<td></td>
</tr>
<tr>
<td>-ier (319)</td>
<td>88.4%</td>
</tr>
<tr>
<td>actors (310)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Formal type of translation for suffixed nouns in –iste, -ier, -eur according to their semantic type.
This first distinction highlights discrepancies between the two morphological systems, discrepancies that vary according to the type of agents. For example, actors nouns displays more difference between Italian and French, for suffixed noun in –ista, -eur, -ista and -iere (resp. 23.7%, 23.6%, 29.4 and 22.5% of non-morphological translations), while followers are the most homogeneous in their formation in -iste or –ista (resp. 12.6 et 13.2 % only of non-morphological translations).

Another interesting measure to compute is to quantify the divergences of « coverage » of morphological process between languages. The notion of “mutual correspondence” (inspired by the one proposed by Altenberg 2002) measures the differences between two elements that are supposed to be equivalent. Table 4 shows the mutual correspondence for representative pairs of suffixes that are supposed to be equivalent because they are cognate (they more or less share the same forms). For each pairs, we give the proportion (in %) in which one suffix is translated by the other in the two directions of translation, the mean of this two proportions (the mutual correspondence) and the difference between the two proportions. This mutual correspondence allows evaluating the distance in comparison to a null hypothesis, according to which one suffix is always translated by it cognate (a mutual correspondence of 100%) and the difference shows the degree of discrepancies between the two suffixes of the pair.

<table>
<thead>
<tr>
<th>It → Fr</th>
<th>Morphological translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ista (1096)</td>
<td>70.6%</td>
</tr>
<tr>
<td>actors (727)</td>
<td>86.8%</td>
</tr>
<tr>
<td>followers (310)</td>
<td>69.5%</td>
</tr>
<tr>
<td>-ore (175)</td>
<td>84.9%</td>
</tr>
<tr>
<td>actors (1065)</td>
<td>78.5%</td>
</tr>
<tr>
<td>behaviours (93)</td>
<td></td>
</tr>
<tr>
<td>-iere (210)</td>
<td>87.5%</td>
</tr>
<tr>
<td>actors (209)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Formal type of translation for suffixed nouns in –ista, -iere, -ore according to their semantic type.
These quantitative results help triggering several qualitative analyses, and completing or confirming previous studies such as (Fradin, 2003; Lignon et Roché, 2011; Roché 1997, 2004, 2011a,b) for French and (Bisetto, 1996, Dardano 1978, Lo Duca 2004) for Italian. Data in Table 2 also shows in which semantic area morphological divergences are the more important. While followers appear to be coined in a very homogenous way in the two languages with the suffixes -ista and -iste (diff. of 0.3), the differences are much more important for “actors” coined with pairs -iere/-ier, and -ista/-iste. The low proportion (25.8%) of French nouns in Xier translated in Xiere reflects a large availability of the Italian suffix and much more important constraint on the French rules. Another case in point is the important gap in behaviour nouns with the pairs –eur/-ore, which can be explained by the competition in Italian with the suffixe -one, which is very frequently an equivalent of –eur for this type of nouns (critiqueur, criticone).

4.2 Qualitative analysis
From the quantitative analysis presented above, deeper qualitative analysis can be performed. In this section, we overview some of these analysis, with the objective of showing interesting insight contrastive analysis can uncover.

4.2.1 Semantico-referential Divergences
Onomasiological approach provides global picture on semantico-referential classes and the kind of items that are used to denote such class. In Table 5 below, we provide the figure of the different construction in French (by a derivational process, a borrowing (borr.), a compound (comp), etc. - see classification below) according to the semantic type of the Italian source entries. Over-abundant phenomena are highlighted in bold. In the rest of this

<table>
<thead>
<tr>
<th></th>
<th>IT→FR</th>
<th>FR→IT</th>
<th>Mututal correspondance</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>adepte</td>
<td>82.3%</td>
<td>82%</td>
<td>82.15</td>
<td>0.3</td>
</tr>
<tr>
<td>Actor</td>
<td>39.5%</td>
<td>54.4%</td>
<td>46.95</td>
<td>14.9</td>
</tr>
<tr>
<td>behavior</td>
<td>59.3%</td>
<td>85.7%</td>
<td>72.50</td>
<td>26.4</td>
</tr>
<tr>
<td>ore→eur</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actor</td>
<td>72.5%</td>
<td>56.2%</td>
<td>64.35</td>
<td>16.3</td>
</tr>
<tr>
<td>behavior</td>
<td>65.6%</td>
<td>29.2%</td>
<td>47.40</td>
<td>36.4</td>
</tr>
<tr>
<td>iere→ier</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actor</td>
<td>50.7%</td>
<td>25.8%</td>
<td>38.25</td>
<td>24.9</td>
</tr>
</tbody>
</table>

Table 4: Mutual correspondances for three pairs of suffixes IT-FR
paper, we will focus on two specific categories, highlighted in grey in the table.

<table>
<thead>
<tr>
<th>Italian Salient Activities</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>eur</td>
</tr>
<tr>
<td>Artistic and cultural activities</td>
<td>62</td>
</tr>
<tr>
<td>Hunting and Fishing</td>
<td>5</td>
</tr>
<tr>
<td>Illicit activities</td>
<td>47</td>
</tr>
<tr>
<td>Manual activities selling</td>
<td>280</td>
</tr>
<tr>
<td>Sport and Leisure</td>
<td>85</td>
</tr>
<tr>
<td>Agriculture – Livestock farming</td>
<td>54</td>
</tr>
<tr>
<td>Employee or manager</td>
<td>115</td>
</tr>
<tr>
<td>Specialist</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 5: Distribution of the French denotation means according to the semantic-referential type.

In the table 5, the semantico-referential classes come from the classification performed on the Italian entries, and the other column described the different kind of French translations that are found in each class. In this section, we will focus on the “sport and leisure” category. For this category, several observations can be made. First, the two most frequent suffixes are -eur and -iste, both mainly on nominal bases. Second, an important amount of noun phrases are used to denote a sport activity (50 cases), which echoes the lack of “plasticity” of French –iste compared to its Italian counterpart –ista. Finally, there are many borrowings from other languages, and compare to the other semantico-referential classes; “sports and leisure” are the one where the “borrowing” is the most frequent. With respect to these loan words, it is interesting to stress that source Italian lexemes are morphologically constructed on borrowed bases (crossista, judoista), while French simply seems to borrow the actors noun in the foreign language (crossman, judoka).
But looking more closely at the data, it has to be noticed that these borrowings should be considered as “pseudo borrowing”, because they did not exist per se in the source language (mostly English). Similarly, we also noticed an important amount of –eur nouns (85/205), which are derived from nouns. In many cases, the base noun is a loaned word (e.g. bridgeur, footballeur, hockeieur are respectively suffixed on bridge, football, hockey).

Finally, there is an important amount of NP bases (50/205), as shown in examples (3) a to e.

(3)  
   a. coureur de bobsleigh\textsubscript{FR} (<bobbista\textsubscript{IT})
   b. lanceur de javelot\textsubscript{FR} (<giavellottista\textsubscript{IT})
   c. joueur de poker\textsubscript{FR} (<pokerista\textsubscript{IT})
   d. joueur d’êchec\textsubscript{FR} (<scacchista\textsubscript{IT})
   e. joueur de tennis\textsubscript{FR} (<tennista\textsubscript{IT})

The interesting point here is that there are morpho-phonological constraints that prevent the coinage of suffixed lexemes in French (*tennisiste, *javelotiste), cf. 4.2.2. More interestingly, the construction of the French NP follows a very regular pattern, as shown in example (4).

(4) \quad Xista\textsubscript{IT} \rightarrow N1 \text{ prep } N2\textsubscript{FR}
   where N1=supertyp of Xista and N2 = X

Example (4) shows the regular pattern that coins a French noun phrase to translate the Italian complex lexem in –ista. The first noun (N1) of the NP is always a supertyp of Xista (joueur, courreur, ...) and the second noun (N2) is the translation equivalent of the base of the Italian complex lexeme (X).

4.2.2. Morpho-phonological constraints
Morpho-phonological constraints can affect the distribution of affixes in the semantic categories. The hierarchy of these constraints differs in various languages. Nouns of specialists provide a good illustration of this phenomenon. As can be seen from Table 5, French and Italian differ widely in the distribution of the suffixes that are used to coin lexemes belonging to the semantic category of ‘specialists’. In Italian, the suffix –ista is almost the only affix in use. But French displays several suffixes as -iste, of course, but nouns also may end in –ien, -eur, and –isant. Since semantics cannot be invoked to explain such variations, we turn towards a morpho-phonological explanation.
Actually, the onset distribution before suffixes is very different according to the affix value. Figure 5 shows the percentage of the onset distribution before each affix in French.

![Figure 5: Onset repartition (in percentage) before each suffix](image)

From this table, various observations can be put forward:
- Some onsets seem unfavorable to –iste (/ʃ, k, s, t, z/);
  - In those cases, the compensation is not homogeneous:
    - For the onset /ʃ, k, t/, the suffix –eur takes over;
    - For /s/, it is the suffix –ien;
    - And, at least, for /z/, the suffix –ant.

The existence of a dissimilation constraint, which tends to prevent the consecution of two identical or nearly identical phonemes allows us to understand that –iste avoids sibilants.

A quite noticeable fact is that French uses mainly the suffix -ien after /s/ for specialist nouns, whose formation meets dissimilation constraint by employing the suffix -ien instead of the expected –iste (for more details: Lignon & Plénat, 2009)

In synchrony, the use of -ien to form specialist nouns occurs only when the base noun ends with a sibilant; the only specialist nouns built in –ien from bases whose onset value is not /s/ in earlier stages are prior to 16th century (grammairien, historien, chirurgien, etc.). Indeed, before that time, the suffix –ien was the only one available to build specialist nouns.

The sibilant /z/ does not behave like /s/. On the one hand, –iste suffixed nouns from bases ending in /z/ are in minority, due to the dissimilative

constraint. But on the other hand, the suffix used is not –ien. More precisely, such derived nouns are all –isant ending nouns (see §.4.2.3).

Another onset after which the suffix –iste is in minority is /t/; the –iste suffix is not only refractory to sibilants, it is also sensitive to the presence of a /t/, still for dissimilative reasons. The suffix –ant also contains this consonant, even if it occurs only in the feminine forms (–ant/-ante). However, this is probably not the major reason why –ant is never found after a stem ending with /t/, cf. section below.

As Figure 5 shows, the suffix –ien is also conflicting with /t/, but for other reasons than dissimilation. Indeed, /t/ can be assibilated before –ien, but this assimilation is not regular. The speaker can hesitate on the way to pronounce it (/ʃjɛ̃/ or /tʃɛ̃/? ) and generally tends to avoid this hazardous configuration by choosing the only affix capable of bearing with this onset, i.e. the suffix –eur. Furthermore, when applying to verb bases, –eur combines frequently with verbs’ learned stem (also called ‘stem 13’, according to Bonami, Boyé & Kerleroux 2009). The most common form of this ‘stem 13’ ends with /at/. Hence, -eur endings for –eur suffixed nouns are very frequent sequences in the lexicon, which can have a ripple effect on the choice of this suffix.

4.2.3 Impact from “morphological series”

Most Italian nouns of specialist are suffixed with –ista, and, to a lesser extent, with –ore. On the other hand, French has at its disposal a more complete set of affixes to form specialist nouns: -iste is the most frequent one, but these nouns can also end with –isant, -eur and –ien. Affix selection is sometimes ruled by morpho-phonological constraints (e.g. with –ien), but in other occasions, morpho-phonology does not interfere: series effects are at play, and their identification has been made possible through the comparison in French and Italian of semantically homogeneous classes. Namely, this is what happens with -isant. In fact, French nouns denoting specialists of a language or a civilization massively end with –isant (5), whereas the corresponding semantic class of nouns in Italian are always formed by –ista suffixation (6).

(5) Hébraïsant, latinisant, arabisant, italisanisant
(6) Ebraïci(sta, latinista, arabista, italianista

In French, two cases occur with this semantic class of nouns, for a given base X: either both Xiste and Xisant exist (basquisant / basquiste; héroïsant / héroïste), or only Xisant is attested (arabisant).

In the latter case, Xisant refers without any doubt to a specialist, and –isant fulfills the role normally plaid by –iste. But in the former case, Xiste and Xisant do not always share the same meaning: sometimes Xiste refers to the very specialist, whereas Xisant may denote a scholar learning the language in
question, or even an (unskilled) amateur (*latinisant, latiniste*). Otherwise, the
–iste noun denotes an adept (a “basquiste” supports the independence of
Basque Country) and therefore the –isant rival noun refers to the specialist.

5. Conclusion

In this paper, we have presented the advantages of the contrastive approach
for the study of lexical morphology. We have shown that, through a sound
empirically based methodology, contrastive data can be extracted, classified
and analyzed. Through this classification, interesting insights are uncovered.
The main innovative aspect entailed by the approach is the onomasiological
perspective of morphological analyses: the contrastive method allows to
study one language through the lens of the other language, and so to uncover
denotation means in a broad sense, including morphology, syntactic/idiomatic
expression and lexicalized words. In particular, the lexicon of the two
languages can be assessed through different perspectives.
With this methodology, we have built an organized set of lexical data both
morphologically and semantically related. The case studies described in this
paper have shown the important plasticity of some suffixes in Italian, and the
stronger tendency in French to use loan words. Hence, the boundaries of
lexical morphology processes in the construction of the lexicon can be drawn
for each language, and the data shown in this study reveal that this boundary
is not located at the same place in Italian and in French.
With the onomasiological approach, we can also measure gaps and overlaps
between word formation processes, and thus highlight their specificities in
each language. The cases that we have studied have shown, for instance, that
–ista and –iste are the suffixes expected to form nouns denoting specialists.
However, the application of this dominant rule is sometimes compromised by
other parameters, such as well-formedness morpho-phonological constraints
or series effects. In case of mismatch between French and Italian supposedly
semantically equivalent processes, contentious cases are brought to light,
which offers new insights to the morphological analysis in each compared
language.
The proposed methodology, when reiterated for other morphological process,
can open avenues for gathering parallel data and setting the stage for large
scaled contrastive analysis of lexical morphology. If applied to many
different language pairs, this can help in setting a map of word-formation in
different languages.
References


**Dictionary**

Garzanti francese-italiano, italiano-francese. 2006. I grandi dizionari Garzanti. Milano, Garzanti Linguistica