

On the cross-linguistic equivalence of *sentir(e)* in Romance languages: a contrastive study in semantics

Abstract

Recent linguistic studies on perception have focused mainly on verbs referring to the dominant visual and auditory modalities, (e.g. English see/look and hear/listen) and have largely ignored the minor verbs. The present paper seeks to fill this gap by comparing the complex semantics of the cognate verbs sentir(e) in three Romance languages, namely Spanish, French and Italian. Because the objective study of semantics is a problematic issue, we pay special attention to methodological problems and opt for a combined corpus approach involving both a translation corpus and comparable data. Evidence from both corpora indicates that, notwithstanding the fact that the rich polysemy of the three verbs partly coincides, each individual verb has undergone semantic specializations differentiating the morphological cognates.

Keywords

Lexical semantics, empirical approach, contrastive studies, perception verbs, *sentir(e)*, Romance languages

1. Introduction

Over the last two decades, the number of studies of perception verbs in various languages has increased exponentially (cf. for instance Viberg 2001, 2005, 2008; Enghels 2007; Hanegreefs 2008; Vesterinen 2010 among many others). This rich bibliography clearly reflects the enormous attraction this verbal category exerts on linguists. This attraction should not, of course, be surprising, not only because the phenomenon of perception itself is a fundamental cognitive process to all living beings but also because, since “people use language primarily to talk about the world they perceive” (Miller and Johnson-Laird 1976: 119), language and perception are closely related. Hence questions such as “what does the grammatical structure of natural language reveal about the nature of perception and cognition” (Jackendoff 1983: 3).

However, on closer inspection it turns out that the degree of attention paid to the different categories of perception verbs is remarkably disproportional. The great majority of analyses address themselves to visual perception verbs (such as English *to see* and *to look*), occasionally in comparison with those of auditory perception (such as English *to hear* and *to listen*). Only rarely does one find studies on verbs of *inferior* perception modalities like olfaction (cf. Ibarretxe 1999; Fernández Jaén 2006a), taste or touch. To put it another way, “les verbes de perception visuelle se profilent comme les prototypes de la classe sémantique de la perception et servent de modèle à la description des autres verbes de perception” [verbs of visual perception emerge as the prototypes of the semantic class of perception and serve as a model for the description of the other verbs of perception] (Enghels 2007: 5). Once again, this tendency can be explained cognitively: the supremacy of studies of visual – and to a lesser extent auditory – perception correlates with the fact that these dominant perception

modalities act in our Western culture as primary sources of objective information (Viberg 1984: 136; Sweetser 1990: 38). Moreover these verbs occur in a wide range of possible semantic and syntactic contexts, as opposed to other verbs whose cognitive poverty results in a simpler – hence somehow less interesting – profile.

Within the spectrum of perception verbs in the Romance languages, *sentir(e)* ('to feel') constitutes a special case. On the one hand, it certainly does not belong to the category of extensively studied verbs – with much effort we were able to find some mention of it in Franckel and Lebaud 1995 and Dupas 1997, as well as two rather superficial analyses (Badynska-Lipowczan 1996; Valentim 2002). But on the other hand, as will be demonstrated throughout this article, it can hardly be defined as a verb with a poor semantic-syntactic profile.¹

Hence the main objective of this study, which is to contribute to the description of the semantics of the minor perception verbs in the Romance languages, by focusing on *sentir(e)* in Spanish, French and Italian.² These verbs share the same Latin etymology, namely *sentīre*, and are still perfect morphological cognates today.³ Consequently, the question arises as to whether this morphological similarity entails a semantic equivalence, and thus, whether the verbs *sentir(e)* in present day Spanish, French and Italian can be defined as cross-linguistic correspondents.

The outline of this paper is as follows: Section 2 shows how a lexicographic study of these verbs can give a first indication of the *tertium comparationis* but is not sufficient to answer the question of semantic equivalence. Section 3 situates this article within the larger domain of empirical approaches in semantics and further justifies the nature of the data that is used. Consequently, in section 4 we turn to the analysis of a translation corpus, thus applying *mutual correspondence analysis*, a method allowing the detection of how a particular meaning is expressed in different languages. Finally, section 5 gives the results of the study of a comparable corpus and leads to a further refinement of the description of patterns of polysemy of *sentir(e)* in the three Romance languages.

2. In search of the *tertium comparationis*

This article makes use of the notions of *equivalence*, *correspondence*, and *tertium comparationis*. These notions are not undisputed. Meanings and linguistic items in different languages have often been compared, “even though the common ground on which to compare the semantic systems, i.e. their *tertium comparationis*, has never been uncontroversially settled.” (Lewandowska-Tomaszczyk 1999: 53). We agree with the views of Altenberg and Granger (2002: 16) and Divjak (2010a) that complete correspondence or complete overlapping polysemy of cross-linguistic equivalents is very rare in natural language. Therefore, it would not be realistic to try to proceed from a *tertium comparationis* based on the idea of exact identity in meaning of cross-linguistic items. Rather, contrastive lexical studies should start from an assumed similarity. *Sentir(e)* has been defined as a general perception verb in all three languages and it is this classification which constitutes the *tertium comparationis* at its most basic level. Thus, as a starting point the notions of *equivalence* and *correspondence* refer mainly to the formal identity between the three verbs as well as to this basic semantic similarity.

Although the present study is not a lexicographic one, a first step in order to identify the degree of equivalence of the verbs *sentir(e)* is the study of how their semantics has been defined by various dictionaries in the three languages concerned.⁴ At first glance it becomes clear that the dictionaries list a large and quite unorganized inventory of definitions for the verb *sentir(e)*, without distinguishing between major and minor meanings or more or less related meanings. The semantic complexity of the entries can be described by considering the wide range of verbs used to paraphrase *sentir(e)*, the numerous possible objects of these verbs and the particular modalities linked to the verb.

First of all, the dictionaries in all three languages put forward a wide range of verbs paraphrasing the meaning of *sentir(e)*. Among these alternatives, it is possible to distinguish verbs referring to acts of perception (e.g. *percibir* ‘to perceive’), verbs of cognition (e.g. *croire* ‘to believe’), adding in some cases the idea of intuition (*sospechar* ‘to suspect’), and verbs of sentiment (e.g. *lamentar* ‘to regret’; *entristecerse* ‘to sadden’). In accordance with this heterogeneous set of verb classes, the semantic nature of the objects observed is also strikingly diverse. They involve sensations (with or without mention of the five senses), abstract stimuli (*l’importanza* ‘the importance’) and a physical, mental or affective state (*sed* ‘thirst’, *pitié* ‘pity’). Finally, some definitions add specific modalities to the description of the verb (e.g. in Spanish ‘normalmente referido al oído o al tacto, más raro al gusto o al olfato y nunca para la vista’, ‘referring normally to the sense of hearing or touch, rarely to the sense of taste and smell and never to vision’). Thus, as becomes clear from this brief lexicological comparison, the verbs *sentir(e)* cover a wide range of meanings that can be linked to different semantic verb classes which appear in all three languages.

However, the comparison also reveals a number of subtle but significant divergences. One of the possible senses of the Spanish verb is auditory, whereas *Le Nouveau Petit Robert* is explicit that the verb “ne s’emploie pas pour les sensations auditives” (‘not used for auditory sensations’). French, in fact, seems to prefer the olfactory meaning of *sentir*. The *Diccionario del español actual* also explicitly refers to meanings of negative sentiment (with the expression *lo siento*, ‘I’m sorry’), which do not figure in the French or Italian dictionaries. Italian *sentire* stands out in that it seems to have developed the sense of hearing to the extent that it can denote not only the passive but also the active meaning of this sense (as in English ‘to listen’). In addition, there are a considerable number of definitions in Italian without equivalent in the other two dictionaries, such as *consultare un medico* (‘consult a doctor’) and the discourse marker *sentì* (‘listen’).

In sum, this lexicographic study reveals that although at a coarse-grained level of analysis the verbs *sentir(e)* in the three Romance languages may have some meanings in common, each of them seems to exhibit several language-specific features and uses. However, a lexicographic analysis has some drawbacks. In particular, the information provided by the dictionaries does not allow us to measure the precise extent to which the verbs differ in these languages; that is to say, whether they relate more closely to each other in some specific domains than they do in others. It is precisely this type of limitation of the introspective lexicographic analysis that indicates the desirability of employing other methodologies. The empirical studies described in this article will allow us to refine the initial assumption of semantic similarity.

3. Methodology: towards an empirical approach in (contrastive) semantics

The main objective of this paper is to contribute to the study of the semantics of perception verbs. Special attention needs to be paid to the methodology by which this can be achieved. In the last few decades, linguistics in general, and contrastive linguistics in particular, has experienced a significant shift from intuition-based approaches towards the use of corpora and empirical methods. Not only have corpora proven useful for quantitative studies of morphosyntactic aspects of languages, but also within the field of semantics, several authors have recently argued the necessity of pursuing more empirical methods (cf. among others Gibbs 2007; González-Márquez et al. 2007; Divjak 2010a, b; Geeraerts 2010a; Glynn and Fischer 2010; Stefanowitsch 2010; Glynn and Robinson in press). Although some believe that certain research questions are better answered by introspective methods (e.g. Talmy 2007), others have argued that these are not sufficient, but merely represents a first step in what Geeraerts (2010a) refers to as *the empirical cycle*. That is to say, although introspections may be useful sources for generating hypotheses, these need to be tested by means of empirical methods.

However, as Glynn (2010b: 90) points out, the application of such empirical methods to the study of semantics is not straightforward. When applied to semantics, the oft-cited main advantage of corpus linguistics – namely the possibility of quantifying results – raises two questions. First, how can meaning – as an intrinsically subjective and non-observable phenomenon – be investigated by means of quantitative methods? And second, why should we do so anyway? Concerning the former question, the use of corpora for semantic research involves a number of methodological difficulties: there are various means of gathering data, different ways of analyzing data, and a considerable gamut of quantitative methods for assessing the results of the analyses (Glynn 2010a).⁵ With respect to the latter question, Glynn (2010b) argues that quantification permits verification and therefore also the testing of hypotheses. In other words, quantitative methods facilitate the empirical cycle of proposing hypotheses and testing them.

The *empirical turn*, as it has been called by Geeraerts (2006), was originally centered on monolingual corpora, but recent bibliographies reveal that the number of contrastive studies based on parallel and comparable corpora⁶ is undoubtedly on the rise. However, while the use of comparable corpora seems to be generally accepted in linguistic research, this has certainly not always been the case for translated corpora.

Aijmer and Altenberg (1996:12) were among the first to make explicit the possibilities of using parallel corpora for linguistic studies. They claim that these data (a) give new insights into the languages compared, (b) can be used for a range of contrastive purposes which augment our knowledge of language-specific or typological differences, and (c) illuminate differences between source texts and translations. Several authors such as Dyvik (1998, 2005), Johansson (1998), Viberg (1999, 2002, 2005), Salkie (2002), Noël (2003), Santos (2008) among many others have been concerned with the question of what the study of translation may teach us about semantics. In fact parallel corpora are seen as a solution to the problem of analyzing non-observable and hardly operationalizable phenomena such as meaning. These authors all support an approach to translation which views it as a kind of annotation of what the source texts mean, since what translators actually do is evaluate the

interpretation of linguistic elements of the source language in a certain context and try to recreate the same interpretational possibilities in a target language (TL).

However, the use of translations as a source for linguistic studies has not always gone undisputed. Noël (2003:779), for instance, notes the risk that a translation only represents one individual's introspection and that "one is testing the performance of the translator rather than comparing languages [...] or that one is contrasting *ordinary* language with *translated* language, which might well have regularities of its own [...]". In fact, the most frequently cited problem of turning towards this type of empirical data is so-called *translationese* (McEnery and Xiao 2008:22-23); that is, "how can we be sure that, while transposing the source-text, the translators were not consciously or unconsciously influenced by the source-language?" (Van Hoecke and Goyens 1990:124). Empirical evidence has been supplied (for instance by Johansson 1998:13-15) to show that the effect of the original language on the target language is real and that translated language is often characterized by simplification and normalization. This is why more recently, linguists (e.g. Viberg 1999, 2002, 2005; Altenberg and Granger 2002; Gilquin 2008; McEnery and Xiao 2008; Mortier and Degand 2009; Vanderschueren 2010) have argued in favor of a combination of parallel and comparable corpora, where these two are considered as complementary sources of cross-linguistic data.⁷ It is precisely this method which is applied below.

In what follows we first turn to a self-compiled multilingual translation corpus as a heuristic for discovering the potential meaning the verbs *sentir(e)* in each individual language (case study 1). One of the many possible advantages that translation corpora appear to offer is that they can be of great value regarding the problem of the *tertium comparationis*. Translated texts are supposed to share a common meaning, which assures the linguist that he is comparing similar semantic contexts (Divjak 2010b). Second, the study of a comparable corpus of 500 instances of the verb in each language allows us to investigate whether the conclusions reached on the basis of the translation corpus can be invalidated or reinforced (case study 2). This leads naturally to a further refinement of the description of the semantic equivalences and differences between the three morphologically cognate verbs.

4. Case study 1: *Sentir(e)*_{SP/IT/FR}, mutual correspondence in a translated corpus?

4.1 Methodology

The data used to investigate the degree of equivalence between *sentir(e)* in Spanish, French and Italian come from a sizeable, manually aligned and annotated parallel corpus consisting of six fiction texts and their translations in the three languages. The way this corpus has been compiled and handled deserves some additional comment. In the first place, the primary concern has been to maximally limit the effects of the above-mentioned *translationese*, especially since this phenomenon has been reported to be particularly frequent between sister languages (Vanderschueren 2010: 95), as in the case of Spanish, French and Italian. Moreover, *sentir*_{SP}, *sentire*_{IT} and *sentir*_{FR} are also perfect morphological cognates, which considerably increases the risks of simply transferring a lexeme from one language to another. Therefore, we selected source texts written in non-Romance languages such as English and Swedish and instead of comparing original texts and their translations across languages, translations of the same original texts are compared. However, since this kind of multilingual

corpus for the Romance languages does not yet exist, it had to be compiled using source texts that have been translated into all three languages. Table 1 lists the translated texts as well as the number of words they contain.⁸

Table 1. Overview of translation corpus sources

source texts		Spanish translation words	French translation words	Italian translation Words
<i>Män som hatar kvinnor</i>	HNAM	189 800	187 900	174 600
<i>Harry Potter and the Sorcerer's Stone</i>	HPPF	79 000	85 800	83 000
<i>Harry Potter and the Goblet of Fire</i>	HPCF	201 600	217 300	184 000
<i>Harry Potter and the Chamber of Secrets</i>	HPCS	92 300	92 600	88 800
<i>Harry Potter and the Prisoner of Azkaban</i>	HPPR	111 200	119 700	103 400
<i>The Da Vinci Code</i>	DV	154 500	137 600	139 000
TOTAL words		828 400	840 900	772 800
		2 442 100		

To put it another way, we examine how, on an independent basis, Spanish, Italian and French translators react when confronted with the same semantic data in the source text.

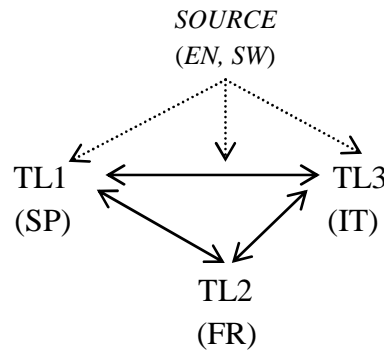


Figure 1. Corpus composition

More precisely, the occurrences of *sentir(e)* had to be extracted from this database and manually matched with the corresponding contexts in the other two languages. This method had to be repeated three times, taking as a starting point *sentir_{SP}*, *sentire_{IT}* or *sentir_{FR}* in each case. In order to explore how the same semantic context possibly conveyed by *sentir(e)* is expressed in the three languages, we apply *Mutual Translation Correspondence Analysis* (MTCA),⁹ which can be summarized as follows:

When TL₁ translates the semantic context of the source text by means of the verb *sentir(e)*, what are the correspondents in TL₂? And conversely: When TL₂ translates the semantic context of the source text by means of the verb *sentir(e)*, what are the different ways the same translation unit is expressed in TL₁?

This strategy of examining mutual correspondences between each pair of languages has to be conducted three times: (1) SP/FR and FR/SP, (2) SP/IT and IT/SP, (3) FR/IT and IT/FR.

To some extent, this procedure could be considered a variation on the *semantic mirrors method* as developed by Dyvik (2005), which consists mainly of collecting the translations in language₁ of the translations in language₂ of a particular word in language₁ in order to obtain corresponding semantic fields in two languages. There are also similarities between MTCA and the method of *translation and back translation*, as applied for instance by Gilquin (2008) and Mortier and Degand (2009), to determine how a particular word or construction tends to

be translated and, through back translation, what it corresponds to in the source language. However, there are three main differences between these methodologies and the one used in this study:

- MTCA does not establish mirrors between a source text and its translations but between two or more different translations of the same source texts, as explained above;
- since the MTCA method is not applied to establish semantic fields but to compare cross-linguistic items, it only creates the semantic mirror of one word, in this case of *sentir*_{TL1}, *sentir*_{TL2} and *sentire*_{TL3};
- MTCA is suitable not only for bilingual but also for multilingual research, as will be shown further on.

4.2 Results and discussion

The extraction of all the occurrences of *sentir(e)* in the three languages left us with 479 examples in Spanish (= target language 1), 371 in French (= target language 2) and 1041 in Italian (= target language 3).¹⁰ According to these numbers, it is clear that the verb is most frequently used in Italian, a fact which is confirmed (indeed, emphasized) when we count the occurrences per 100000 words:

- SP: 479/828400 ~ 57.8/100000
- FR: 371/840900 ~ 44.1/100000
- IT: 1041/772800 ~ 134.7/100000

To put it another way, it seems that a larger number of semantic contexts more frequently yield *sentire* in Italian than they do in Spanish or French. Moreover, only 58 cases of perfect matches between TL₁, TL₂ and TL₃ were detected, almost all situated within the domain of general, non-specified perception:

- (1) a. *Harry sintió como si se le helaran las entrañas.* (HPPF-SP: 211)
- b. *Harry sentit son sang se glacer.* (HPPF-FR: 141)
- c. *Harry sentì le budella congelarglisi dentro la pancia.* (HPPF-IT: 124)
- d. *Harry felt as though his insides had turned to ice.* (HPPF-EN: 170)

These preliminary observations confirm the results of previously conducted studies on cognate or functionally similar verb pairs whose correspondence in translation corpora tend to be surprisingly low (see for instance Viberg (1999) or Altenberg and Granger [2002]). In the remainder of this article we examine the extent to which this phenomenon could be explained by a possible diverging polysemy and different meaning extensions that the verbs tend to develop in different languages.

The translation corpus does provide us with some initial quantitative data as to the degree of correspondence between the three verbs. A closer look at how the occurrences of *sentir(e)* in a particular TL are rendered in the other two TLs reveals very quickly that the equivalents fall into five categories, namely (1) a perfect match with *sentir(e)*, and for French its derived verb *ressentir*, (2) perception verbs other than *sentir(e)*, (3) cognitive verbs, (4) verbs referring to emotion, and finally (5) what could be called cases of *zero correspondence*, which go beyond simple lexical correspondence involving some restructuring of the clause.¹¹ In

what follows the MTCA method systematically maps out the equivalents of *sentir(e)* in a particular TL, in search of the different meanings it can convey and their frequency in each language. The results are presented in three groups, taking the verb in each language as a starting point for comparison with its cognates in the other two languages.

To begin with, Table 2 provides a quantitative overview of how the Italian and French translators react to the semantic contexts where the Spanish translators opt for the verb *sentir*:

Table 2. French and Italian equivalents of Spanish *sentir*
($\chi^2 = 40.6$; $df = 5$, $p < 0.001$)¹²

Equivalents	FRENCH		ITALIAN	
	#	%	#	%
<i>sentir/sentire</i>	126	26.3%	129	26,9%
<i>ressentir</i>	35	7.3%	-	-
cognitive verb	40	8.4%	53	11.1%
perception verb	57	11.9%	55	11.5%
emotion verb	90	18.8%	117	24.4%
zero correspondence	131	27.3%	125	26.1%
Total	479	100%	479	100%

It immediately becomes clear that the proportion of matched cases with French (26.3%) and Italian (26.9%) is essentially equal, except when the translations by the derived verb *ressentir* (7.3%) are also taken into account. In fact, it emerges from the corpus that in French *sentir* competes with *ressentir*, especially when the stimulus refers to an abstract concept such as *la bonheur* ‘happiness’, *la crainte* ‘fear’ etc. (see example 2). Secondly, both languages can resort to other perception verbs such as *éprouver* or *provare* (‘to experience’) when Spanish uses *sentir* (3-4):

- (2) a. *Había comenzado a sentir una punzada de miedo cada vez que mencionaban a Quien-tú-sabes.* (HPPF-SP: 108)
b. *Il commençait à ressentir un frisson de crainte chaque fois qu'on lui parlait de Vous-Savez-Qui.* (HPPF-FR: 73)
c. *He was starting to get a prickle of fear every time You-Know-Who was mentioned.* (HPPF-EN: 85)
- (3) a. *Harry sentía tanta ira que el artículo de El Profeta le temblaba en las manos [...].* (HPCF-SP: 215)
b. *Harry éprouvait une telle fureur que la page de La Gazette du sorcier tremblait entre ses mains.* (HPCF-FR: 304)
c. *Harry was so angry that the Daily Prophet article shook in his hands [...].* (HPCF-EN: 440)
- (4) a. *Mikael..., me faltan palabras para expresar lo que siento leyendo este artículo.* (HNAM-SP: 271)
b. *Mikael... io non ho parole per esprimere ciò che ho provato nel leggere questo articolo.* (HNAM-IT: 308)
‘Mikael ... I don’t have words to express what I felt when I read this article.’¹³

More interesting, though, is the comparison of the equivalent translations using a cognitive verb. It seems that the (rare) cognitive meaning possibly expressed by *sentir_{SP}* is conveyed with more difficulty by *sentire_{IT}*, as is shown by the large list of equivalents (*pensare* ‘to

think’, *avvertire* ‘to comprehend’, *non ignorare* ‘not to ignore’, *sapere* ‘to know’, *credere* ‘to believe’,...) and their somewhat higher occurrence (11.1% v. 8.4%) than their French counterparts (*penser* ‘to think’, *comprendre* ‘to understand’, *croire* ‘to believe’,...). This could be a first indication of the varying frequency of the cognitive meaning that *sentir(e)* can convey in the three languages, which is examined in more detail in the comparable corpus (see Section 5). Finally, in a considerable number of cases *sentir*_{SP} transmits some meaning of regret, never matched with *sentir*_{FR} or *sentire*_{IT} but translated by *désolé*, *excuse(z)-moi* and *scusatemi*, *scusa*, *mi dispiace* ‘(I’m) sorry’ (6):

- (5) a. *A pesar de ser mi prisionero, sientes que controlas la situación; piensas que lo único que haré, si no te mato, es soltarte.* (HNAM-SP: 185)
 b. *Pensi di avere il controllo, nonostante tu sia mio prigioniero, perché sei convinto che l'unica cosa che posso fare se non ti uccido è lasciarti andare.* (HNAM-IT: 211)
 c. *Tu crois que t'as le contrôle bien que tu sois mon prisonnier, parce que tu t'imagines que la seule chose que je puisse faire si je ne te tue pas, c'est te relâcher.* (HNAM-FR: 268)
 ‘You think you have control even though you’re my prisoner, because you think the only thing I can do if I don’t kill you is to let you go.’
 (6) a. *Y era muy probable que algún académico religioso le hubiera seguido hasta el hotel para entablar una discusión con él. Lo siento – dijo Langdon –, pero estoy muy cansado.* (DV-SP: 6)
 b. *Probablement un intégriste survolté que le contenu de la conférence de la veille les symboles païens cachés de la cathédrale de Chartres avait rendu furieux, et qui venait lui chercher noise. Je suis navré, marmonna-t-il, mais je suis fatigué, et...* (DV-FR: 7)
 c. *Probabilmente uno studioso di religioni l'aveva seguito fino all'albergo per insultarlo. Mi dispiace disse Langdon ma sono stanco e...* (DV-IT: 5)
 d. *Most likely, some religious scholar had trailed him home to pick a fight. I'm sorry, Langdon said, but I'm very tired and...* (DV-EN: 6)

Table 3 displays the overall level of correspondence between Spanish and Italian counterparts of *sentir*_{FR}:

Table 3. Spanish and Italian equivalents of French *sentir*
 ($\chi^2 = 17.6$; $df = 3$, $p < 0.001$)

equivalents	SPANISH		ITALIAN	
	#	%	#	%
<i>sentir/sentire</i>	107	28.8%	160	43.1%
cognitive verb	74	19.9%	51	13.7%
perception verb	38	10.2%	36	9.7%
emotion verb	-	-	-	-
zero correspondence	152	41%	124	33.4%
total	371	100%	371	100%

The significant difference in the extent of perfect matches (43.1% v. 28.8%) indicates that the semantics of *sentir*_{FR} is closer to its Italian cognate than to its Spanish one. The data also reveal that Spanish often resorts to other perception verbs, especially *oler* ‘to smell’ (24/38) and to a lesser extent *ver* ‘to see’ (1/38) and *percibir* ‘to perceive’ (10/38), which could

indicate that *sentir*_{SP} is not the most suitable verb to express olfactory perception (7). Half of the Italian perception counterparts are also situated within this domain (*annusare* ‘smell’, *profumare* ‘to scent’, *puzzare* ‘to stink’) (7c), next to *vedere* ‘to see’ (3/36) and *percepire* ‘to perceive’ (8/36) (8):

- (7) a. *Toute la maison sentait le chou et Mrs Figg passait son temps à lui montrer les photos de tous les chats qu'elle avait eus.* (HPPF-FR: 14)
 b. *Toda la casa olía a repollo y la señora Figg le hacía mirar las fotos de todos los gatos que había tenido.* (HPPF-SP: 24)
 c. *Puzzava di cavolo e Mrs Figg lo costringeva a guardare le fotografie di tutti i gatti che aveva posseduto in vita sua.* (HPPF-IT: 14)
 d. *The whole house smelled of cabbage and Mrs.Figg made him look at photographs of all the cats she'd ever owned.* (HPPF-EN: 16)
- (8) a. *Avant le procès, ces dernières semaines, Mikael Blomkvist avait été d'humeur morose, mais elle ne l'avait pas senti aussi sombre et résigné qu'il semblait l'être maintenant à l'heure de la défaite.* (HNAM-FR: 66)
 b. *Durante las semanas anteriores al juicio, Mikael Blomkvist dio la impresión de estar metido en una nube gris, pero nunca lo había visto tan cabizbajo y resignado como ahora, en el momento de la derrota.* (HNAM-SP: 43)
 c. *Nelle ultime settimane, in vista del processo, Mikael Blomkvist si era aggirato quasi avvolto in una nuvola scura, ma non l'aveva mai visto così cupo e rassegnato come sembrava adesso nell'ora della sconfitta.* (HNAM-IT: 48)
 ‘In the weeks before the trial started, Mikael Blomkvist had been walking around under a black cloud. But she had never *seen* him as gloomy and dejected as he seemed to be now in the hour of his defeat.’

The final part of the mutual translation correspondence analysis of *sentir* is summarized in table 4, which gives a survey of the Spanish and French correspondents of *sentire*_{IT}:

Table 4. Spanish and French equivalents of Italian *sentire*
 ($\chi^2 = 20.6$; $df = 3$, $p < 0.001$)

Equivalents	SPANISH		FRENCH	
	#	%	#	%
<i>sentir/sentir</i>	119	11.4%	165	15.9%
cognitive verb	61	5.9%	48	4.6%
perception verb	667	64.1%	583	56%
emotion verb	-	-	-	-
zero correspondence	194	18.6%	245	23.5%
Total	1041	100%	1041	100%

It is clear that the image we get when taking the Italian verb as a starting point is of a quite different nature, in that the number of perfect matches in Spanish and French is much lower but comparable in both languages. However, once again the distance between Italian *sentire* and its French counterpart (15.9% of correspondence) seems to be somehow smaller than with the Spanish one (only 11.4%). These matches are all contexts referring to a general perception process, thus unspecified as to modality (cf. example 1) or tactile perception (9):

- (9) a. *Mikael voltò pagina e si sentì rizzare i peli sulla nuca. Era come se un soffio di vento gelido fosse passato attraverso la stanza.* (HNAM-IT: 346)
 b. *Mikael pasó la hoja y sintió cómo se le ponía el vello de punta, como si un soplo de aire frío hubiese pasado por la habitación.* (HNAM-SP: 304)
 c. *Mikael tourna les pages et sentit les cheveux se dresser dans sa nuque. C'était comme si un courant d'air glacial était entré balayer la pièce.* (HNAM-FR: 430)
 'Mikael turned the page and *felt* the hairs on the back of his neck stand on end. It was as if a cold gust of wind passed through the room.'

Secondly, the number of perception verbs used as equivalents for *sentire_{IT}* is strikingly high in both languages. On closer inspection it emerges that auditory verbs such as *oír/escuchar* 'to hear/to listen' (544/1041) and *entendre/écouter* 'to hear/to listen' (496/1041) are extremely frequent (cf. example 10). This leads us towards the interim conclusion that the wider use of *sentire_{IT}* is mainly due to the fact that this verb, unlike *sentir_{SP}* and *sentir_{FR}*, has adopted the meaning of auditory perception, thus relating it to the domain of the dominant perception modalities:

- (10) a. *Sophie aveva sentito che il cuore accelerava i battiti. La mia famiglia?* (DV-IT: 68)
 b. *De pronto Sophie se oía los latidos de su corazón. ¿Mi familia?* (DV-SP: 63)
 c. *Elle entendit soudain battre son propre cœur. Ma famille?* (DV-FR : 68)
 d. *Sophie suddenly could hear her own heart. My family?* (DV-EN: 64)

4.3 Conclusions

Mutual Translation Correspondence Analysis proved to be especially suitable for the study of the degree of equivalence of cognate verbs in two or more closely related languages. As pointed out before, given that the compared translated texts share a common meaning, we were sure to compare similar semantic contexts. By noting that the equivalents of *sentir(e)* in all three languages fall into the same categories, the possible meanings have been defined more precisely. Moreover, the comparison between the three languages revealed a number of interesting conclusions as to the behavior of the verbs *sentir(e)* vis-à-vis these contexts.

In the first place, since the possible meaning categories have been observed in different proportions, it remains clear that the notion of perfect semantic equivalence cannot be supported, as had already been suggested by the lexicographic analysis. The use of *sentir* in one language rarely corresponds to its congruent verb in another language, as the low degree of perfect mutual translation shows. Although some uses tend to coincide, the verbs in each language seem to display a particular behavior. To sum up:

- *sentire_{IT}* seems to display a rather different profile; *sentir_{SP}* and *sentir_{FR}* seem to match up more frequently with each other;
- *sentir_{SP}* conveys the unique meaning of regret not once observed for its counterparts in the other two languages;
- *sentir_{FR}* seems to be least reluctant to contexts describing cognitive processes;
- the meaning of auditory perception clearly dominates the semantics of *sentire_{IT}* and seems to be much less frequent for *sentir_{SP}* and *sentir_{FR}*.

In short, these observations further weaken the idea of perfect equivalence and provide us with valuable information about possible semantic differentiation of the individual lexemes. However, a translation corpus does not seem to be the best foundation on which to build a hypothesis about the impact and spread of these preferences. Therefore we now turn to the analysis of a wide comparable corpus.

5. Case study 2: degree of equivalence of *sentir(e)*_{SP/IT/FR} in a comparable corpus

5.1 Methodology

In order to assess the viability and usefulness of parallel corpora for the substantiation of semantic claims and, more precisely, for the operationalization of (introspective) hypotheses concerning the semantics of cross-linguistic equivalents, a monolingual corpus needed to be compiled for the three languages. To this end, 1500 occurrences of the verb *sentir(e)* were retrieved – 500 per language – half of which were drawn from literature and the other half from the press.¹⁴ These comparable data were then subjected to a fine-grained qualitative analysis. That is, all the occurrences were manually analyzed and annotated for the semantic categories lurking in the dictionaries and hinted at with more precision by the analysis of the parallel corpus (cf. Sections 2-4 above).

For the sake of clarity, the semantic categories for which the data were annotated are repeated below:

1. general physical perception
2. specific modality of perception (hearing, taste, vision, touch, smell)
3. emotional perception
4. cognitive perception

Following Glynn (2010c) and Zeschel (2010), secondary annotation was employed in order to enhance the reliability of the analysis. As one of the main advantages of empirical research is precisely the possibility of repeating the analysis on the same data set, all the occurrences were coded independently by two linguists. Subsequently, inter-coder agreement was calculated to evaluate the accuracy of the judgments. To determine the degree of this inter-coder agreement, Cohen's Kappa was employed (Cohen 1960). As this takes into account the expected chance agreement, it is generally considered to be a more powerful measure than simple percentage agreement calculation. As a rule of thumb, if the coders are in total agreement then $\kappa = 1$. If there is no agreement other than what would be expected by chance, $\kappa = 0$. In general, a kappa of 0.7 or higher indicates a promising inter-coder agreement.¹⁵ The inter-coder agreement on sense classification for the verb *sentir(e)* in the three languages was approaching 1 ($\kappa = 0.80784437$), which is very high. In the following section, we focus mainly on the results of this classification.¹⁶ The presentation of the findings is organized according to the main typological similarities and contrasts between the languages from both a quantitative and a qualitative perspective.

5.2 Results and discussion

Taking into account the two independent ratings, table 5 shows the frequencies of the different verb senses for each language:

Table 5. Semantic clusters of sentir(e) in a comparable corpus
($\chi^2 = 990$; $df = 8$, $p < 0.001$)

	SP		FR		IT	
	#	%	#	%	#	%
general physical perception	55	11%	88	17.6%	59	11.8%
emotional perception	313	62.6%	25	5%	47	9.4%
cognitive perception	58	11.6%	199	39.8%	30	6%
specific modality of perception	40	8%	58	11.6%	311	62.2%
coder disagreement	34	6.8%	130	26%	53	10.6%
Total	500	100%	500	100%	500	100%

It is striking that although all the basic semantic clusters are present in all three languages, their distribution diverges considerably from language to language.¹⁷ This observation leads to a basic question: how can we interpret these clear frequency differences between the three languages? An additional question is provoked by the category *coder disagreement*, which includes not only cases of actual coder disagreement (see Section 5.1 above) but also instances of the verb which defy assignment to one of the four categories: what are these cases in the coder disagreement category? In what follows, these questions are addressed both intra- and inter-linguistically.

5.2.1 Semantic specializations

Although to a certain extent the basic meanings are present in the three languages, some clear differences arise concerning their respective frequencies.

To begin with, it can be noticed that in Spanish, French and Italian the verb *sentir(e)* often refers to a general act of perception or bodily sensation, without specification of a particular modality:

- (11) a. *¿Cuándo debemos tomar agua? Si nuestro mecanismo de la sed funciona adecuadamente, cuando sentimos sed, en ayunas, con las comidas o entre horas.* (CREA: prensa, 27/11/2003)
 ‘When should we drink water? If our thirst mechanism works properly, when we *feel* thirsty, when fasting, with meals or between meals.’
 b. *Une vraie broche qui lui entrainait dans la chair, et elle sentait ses mains se refroidir.* (Le Monde: 20/03/1998)
 ‘A real pin pierced her flesh, and she *felt* her hands grow cold.’
 c. *L'anziano, però, va in qualche modo costretto a bere, perché sente sempre in ritardo lo stimolo della sete e, più di altri rischia la disidratazione.* (CdS: 18/07/2010)
 ‘The old man, however, in a sense is forced to drink, because he always *feels* too late the stimulus of thirst, and more than others he risks dehydration.’

Nevertheless, judging by the relatively small proportion of occurrences in each of the languages (resp. 11%, 17.6% and 11.8%), this category does not seem to be especially salient in any of them.

Emotional perception, by contrast, clearly predominates in the Spanish corpus (resp. 62.6% v. 5% in the French and 9.4% in the Italian). This group includes the recurrent collocations of

the verb with an abstract noun referring to a particular state of mind or emotion, such as *miedo* ‘fear’, *amor* ‘love’, *ternura* ‘affection’:

- (12) *Darse cuenta de esto le sirvió a Indalecio para sentir cierta ternura, inédita hasta entonces, por aquella chica rica [...].* (CREA: Pombo, A., 2004)
 ‘Realizing this matter helped Indalecio to *feel* a certain affection, hitherto unnoticed, for that rich girl [...].’

As with its use for the expression of general bodily sensations (example 11), sentences such as ‘to feel fear’ and ‘to feel love’, describe an emotion as experienced from within by the subject and therefore, they can be referred to as *internal perception* as opposed to *external* or *sensory perception* triggered by objects and events outside the body (cf. Viberg 2005:129).¹⁸ Related to this category of internal perception are the cases of negative sentiment in which *sentir* adopts the specific meaning of *lamentar* ‘to deplore’, both in its plain verbal use (13a) and in the regular expression *lo siento* ‘I’m sorry’ (13b):

- (13) a. *El camionero que salvó a los niños rechaza hablar de proeza y siente no haber rescatado a sus padres.* (CREA: prensa, 19/12/2004)
 ‘The truck driver who saved the children refuses to speak of heroism and *regrets* not having rescued their parents.’
 b. *Nuestra selección, amigo, y lo siento de verdad, no estará allí.* (CREA: prensa, 14/06/2004)
 ‘Our selection, friend, and *I’m really sorry*, will not be there.’

In addition to its apologetic meaning (13a-b), the semantics of this expression *lo siento* even seems to extend towards the so-called *social emotions* (cf. Damasio 2003: 43-45), when it functions as a kind of adversative marker announcing a negative or contrary opinion to that of the interlocutor (14). Parallel with discourse markers such as *in fact*, this particular use of *lo siento* can thus be described in terms of (*counter*)*expectation* and *adversativity* based on the idea that “speakers, for various reasons, explicitly position themselves and their utterances vis-à-vis expectations raised by preceding discourse or by background knowledge” (Mortier and Degand 2009: 303).

- (14) *El día antes de su llegada el comandante mayor me llamó: - Mañana hay un desfile en honor del general Burguete. Lo siento, pero no tengo a nadie más que a ti para ser cabo de gastadores.* (CREA: Ficción, 1951)
 ‘The day before his arrival the senior commander called me - Tomorrow there is a parade in honor of General Burguete. Sorry, but I have no one else but you to be head of the sappers.’

A preliminary diachronic corpus study of the evolution of the emotional meaning of the Spanish verb *sentir* has shown that there is clear morphosyntactic evidence correlating with this semantic-pragmatic change that might account for an analysis in terms of grammaticalization and (inter)subjectification, as formulated in Company Company (2004), namely loss of syntactic capacities and fixation of the form (tendency towards the use of the first person singular in the simple present), widening of the predication scope and autonomy

of predication (movement to the leftmost or rightmost position of the sentence functioning as an autonomous proposition that has scope over the whole speech act).¹⁹

The cognitive pole of the verb seems to be elaborated most thoroughly in French (resp. 39.8% v. 11.6% in the Spanish and 6% in the Italian). As example (15a) shows, the meaning of *sentir* approximates to that of *se rendre compte* ‘to realise’, *savoir* ‘to know’ or *penser* ‘to think’. Very often within this category, a subtle subjective aspect is added to the verb, pointing rather to the sense of intuition than to pure cognition (15b).

- (15) a. *Les questions de forme, dans le sport comme dans la langue, et cela dans les plus infinis détails de protocole, intéressent en effet et même passionnent les Français. Ils sentent bien que ces questions tiennent à leur raison d’être à la fois commune et singulière.* (Le Monde: 4/08/1998)
‘Questions of form, in sports as in language, and in the most infinite details of protocol, definitely interest and even fascinate the French. They *realize* that these issues are intimately connected with their *raison d’être* both common and unique.’
b. *Tu peux venir avec lui. - Non je préfère être seul pour venir à Avignon. Je t’appellerai si je sens que c’est possible.* (FRANT: Angot, C., 2006)
‘You can come with him. - No, I would rather be alone than come to Avignon. I’ll call you if *I feel/think* that it is possible.’

A second particularity of the French verb turns out to be its rich development of the olfactory meaning. In addition to its use as a passive perception verb (16), *sentir* often appears in a copulative construction, adding a certain appreciation to the olfactory process (17):

- (16) *Avant, nous avions la vue sur les cheminées. Dès qu’il y avait un peu de vent, on sentait l’odeur des corps brûlés.* (Le Monde: 13/03/1998)
‘Before, we had a view of the chimneys. Once there was a little wind, you could *smell* the burnt bodies.’
(17) *Il n’aimait pas son odeur, ça sent le poisson pourri, il ne pouvait pas le faire.* (FRANT: Angot, C., 2006)
‘He did not like her smell; it *smells* like rotten fish, he could not do it.’

Contrary to its Spanish and Italians counterparts, these copulative uses of the French verb are very frequent in both evaluative (18) and metaphorical contexts (19):

- (18) *On était en haillons, on sentait mauvais, ça provoquait toutes les histoires qu’on peut imaginer!* (FRANT: Boltanski C., 2007)
‘We were ragged, we smelled bad, which caused all the stories we can imagine!’
(19) [...] *Qu’ils se jetaient à la face des accusations refoulées depuis des lustres, mais que les électeurs, ces innocents, n’avaient jamais eu le droit d’entendre. Cela sentait les dernières cartouches.* (Le Monde: 22/03/1998)
‘That they hurled reproaches at one another repressed for many years, but that the voters, these innocents, never had the right to hear. It seemed as if they had shot their bolt.’

Moreover, various examples of the French verb suggest that this copulative use even extends towards other domains such as the cognitive (20) and the auditory meanings (21):

- (20) *Et le père Jean, qui vit ici depuis plus longtemps que moi, et connaît bien tous les signes atmosphériques, nous dit au réfectoire: «ça sent l'orage».* (FRANT: Roubaud J., 2006)
 ‘And Father John, who has lived here longer than me, and knows all the atmospheric signs, tells us in the refectory: there’s a storm brewing.’
- (21) *Cependant le ‘français’ de Florent était assez livresque et la langue de Sébastien sent son Bugey natal.* (FRANT: Roubaud J., 2006)
 ‘However the French of Florent was quite bookish and the language of Sebastian sounds like his native Bugey.’

Notwithstanding this particularity of the French verb, it is in Italian that the meaning of *specific perception modality* is dominant (resp. 62.2% v. 8% in Spanish and 11.6% in French). Closer inspection of the data shows that this characteristic can be attributed to the overwhelming frequency of the auditory modality in Italian, where *sentire* can adopt both passive (‘to hear’, 22a) and active meanings (‘to listen’, 22b):

- (22) a. *Quando Umberto Cennamo, custode di uno stabile in via Santa Maria di Costantinopoli alle Mosche, pochi metri dall' incidente, ha sentito il forte rumore di lamiera, ha pensato che fosse caduta la gru che lavora alla metropolitana.* (CdS: 7/08/2010)
 ‘When Umberto Cennamo, caretaker of a building in Via Santa Maria of Constantinople alle Mosche, a few meters from the accident, *heard* the loud noise of metal, he thought that the crane working on the subway had collapsed.’
- b. *E il collega: «Penso sia possibile una mediazione tra chi vuole riposare e chi vuole sentire musica».* (CdS: 6/08/2010)
 ‘And his colleague: I think it is possible to mediate between those who want to rest and those who want to *listen* to music.’

This predominant auditory sense of the Italian verb involves some semantic extensions to the realm of communication. In this respect, *sentire* can adopt not only the general sense of ‘to hear news of somebody’ (23a) but also the more specific meaning of ‘to question/hear witnesses’ (23b) or ‘to consult a doctor’ (23c):

- (23) a. *Voi che informazioni avete? Bisogna chiedere a Giuliano, non lo sento da due giorni.* (CdS: 5/08/10)
 ‘What information do you have? We must ask Julian, I haven’t *heard* of him for two days.’
- b. *Oggi è il giorno di Denis Verdini. Il coordinatore del Pdl verrà sentito dal procuratore aggiunto Giancarlo Capaldo e dal pm Rodolfo Sabelli nell' ambito dell' inchiesta sulla lobby segreta.* (CdS: 26/07/10)
 ‘Today is the day of Denis Verdini. The coordinator of the PDL will be *heard* by the assistant public prosecutor Giancarlo Capaldo and by the public prosecutor Rodolfo Sabelli in the field of the inquiry into the secret lobby.’
- c. *Il medico sportivo che ha eseguito gli esami è ora in ferie e potrà essere sentito solo nei prossimi giorni.* (CdS: 12/08/10)
 ‘The sports doctor who performed the tests is now on vacation and can be *consulted* only in the next few days.’

The elaborate development of the auditory sense of *sentire* even seems to trigger off the grammaticalized use as a discourse marker ('listen'), which turns out to be a particularity of the Italian verb only:

- (24) "*Questo non glielo so dire, dottore*". "*Senti, tu lo conosci un tale che si chiama Angelo Pardo, ha quarantadue anni e fa l'informatore?*". (CA: 7)
 'I cannot say this to him, doctor. *Listen*, do you know a man named Angel Pardo, 42 years old and an adviser?'

This particular use of the Italian verb *sentire* perfectly fits the well-known tendency of (dominant) perception verbs to develop a wide range of discursive uses (cf. *listen*, *look* in English or *oye/oiga* and *mira/mire* in Spanish).²⁰ Indeed, as illustrated in example (25) the Spanish equivalent of this occurrence in the translation corpus corroborates this cross-linguistic grammaticalization pattern:

- (25) a. *Oye, Harry; ¿te importa que me dé una vuelta por el Caldero Chorreante?* (HPPF-SP: 69)
 b. *Senti, Harry, ti piacerebbe se facessi un salto al Paiolo magico a bere un cordiale?* (HPPF-IT: 43)

As becomes clear from these examples, the comparable corpus supplies corroborating evidence for the basic meanings of the verb uncovered through the analysis of the parallel corpus. Moreover, the divergent frequencies found in the former accentuate the cross-linguistic variety, illuminating and highlighting the semantic specializations which the verbs have experienced in each language separately: emotional perception in Spanish, cognitive perception in French and auditory perception in Italian. Nevertheless, the assignment of one of these categories to the verb is not always straightforward, which explains the presence of the category *coder disagreement* in Table 5 above. The next section focuses on these non-classifiable occurrences of the verb.

5.2.2 Towards a fine-grained semantic profile

As shown in Table 5, a striking difference between the data extracted from the parallel corpus and those derived from the comparable corpus is the addition in the latter of a category *coder disagreement*, comprising all instances of the verb resistant to agreed classification. The following examples extracted from the Spanish corpus illustrate this difficulty:

- (26) a. *Le besó la cara rozándole la piel y con una determinación que la sorprendió, la estrechó contra sí y posó sus labios en los de ella, primero suavemente, pero en seguida con fuerza, incluso hasta sentir que le hacía daño.* (CREA: Cásares C., 1996)
 'He kissed her face, brushing the skin with a determination that surprised her, pressed her against him and placed his lips on hers, gently at first, but then with force, even till he *felt / realized* that he was hurting her.'
 b. *Debes de estar impaciente -dije, sintiendo el sabor a mala leche en mi propia voz, una voz insolente que no sabía de dónde venía.* (CREA: Ruiz Zafón C., 2003)
 'You must be impatient, I said, the taste of bad blood (lit. milk) in my own voice, an insolent voice, I didn't know where it came from.'

In example (26a) it is not obvious whether the verb conveys the meaning of a physical perception of touch or a more indirect perception implying a cognitive deduction equivalent to ‘he realized that he was hurting her’. The same type of creativity – introduced consciously or not by the speaker – appears in (26b), where the verb can be interpreted at first glance as a metaphorical use of taste by the presence of the noun phrase *el sabor a mala leche*. Nonetheless, at the same time the expression *estar de mala leche* refers to a particular state of mind, viz. the bad mood of the speaker, whereby a shade of subjectivity is added to the sentence. Moreover, by adding the prepositional phrase *en mi propia voz*, the speaker also hints at the auditory sense possibly conveyed by *sentir*.

These transitions between the semantic categories can also be easily detected in the French and Italian corpora:

- (27) *J'étais d'accord cette fois, car avant mon départ j'avais pris contact avec des gens du FIS et leur avais donné le numéro de notre téléphone satellite sur lequel ils pouvaient nous appeler en toute sécurité. Ils ne l'ont pas fait. Cette histoire de lettre sentait la manipulation.* (Le Monde, 19/02/1998)

‘I agreed this time, because before I left I had contacted people from the FIS and had given them the number of our satellite phone on which they could call us safely. They did not. This story (of a letter) smelled of manipulation.’

- (28) *Da quella vitalità siamo nati e continuiamo a vivere anche noi che non ci riconosciamo nell'emergenza perenne dell'immigrazione, negli scontri etnici, nello scambio truffaldino di favori, nella semplificazione culturale e nella dilagante maleducazione. Vogliamo sentirlo quel cuore, perché vogliamo capire qual è la malattia e insieme curarla, perché la città, a differenza dell'uomo, non ha bisogno di bravi medici ma solo di buoni cittadini per guarire.* (CdS, 27/07/10)

‘From that vitality we were born and we continue to live, not recognizing ourselves in the perennial emergence of immigration, in the ethnic clashes, in the fraudulent exchange of favors, in the cultural simplification and in the prevailing rudeness. We want to *hear* (lit.) that heart, because we want to understand what the disease is and heal it, because the city, unlike Man, does not need good doctors, but only good citizens to heal.’

Without elaborating on these examples, it is clear that in (27) the specific copulative use of the French verb possibly favors metaphorical extensions of the olfactory meaning in the cognitive direction. Similarly, it is not clear whether the Italian verb in (28) refers to a specific auditory perception (‘to hear the heartbeats’), to a general corporal perception, or, if a metaphorical interpretation of the sentence where the heart denotes the city center, induces the more cognitive meaning of ‘to sense/comprehend/know that city center’.

As the examples above suggest, rather than clearly delimited categories, the comparable corpus points to the necessity of an analysis in terms of continuity with clear occurrences of a meaning A gradually turning into clear occurrences of a meaning B and a range of occurrences situated somewhere between these prototypes. It is precisely these hazy boundaries between the categories that account for the coder disagreement.

5.3 Conclusions

By way of conclusion, with respect to the first question of how we can interpret the frequency differences between the three languages, it became clear that the comparable corpus provides corroborating evidence for the semantic clusters of the verb discovered through the analysis of the parallel corpus. Furthermore, the divergent frequencies highlight the cross-linguistic variety, portraying the semantic specialization the verb has experienced in each language separately. In other words, the comparable corpus not only plays a confirmatory role but also, and mainly, a specification role, refining the outcome of the previous analyses.

Secondly, a closer look at the cases of coder disagreement revealed that this category embraces a whole gamut of ambiguous, creative uses of the verb situated somewhere in between the principal meanings, for which neither the dictionary nor the parallel corpus can account. As a consequence, the existence of such a category calls into question the idea of clearly delimited, discreet senses as suggested by the lexicographic approach and the parallel corpus. The principal merit of the comparable corpus thus consists in the visualization of those fuzzy boundaries and transitions between the categories, which turn out to be very fertile areas for the generation of creative, metaphorical uses of the verb.

6. General conclusions

To sum up, this study on the degree of semantic equivalence of the cognate verbs *sentir(e)* in Spanish, French and Italian allows us to claim a number of achievements situated at different levels. In the first place, it provides further insight into the semantics of perception verbs in the Romance languages. It has been shown that *sentir(e)*, formerly treated as an outsider of this verbal category and thus rarely the subject of a thorough analysis, also displays a rich semantic profile, commensurable with the dominant verbs of visual and auditory perception. At the most basic level, the article further illustrates the importance of a well thought-out methodology to facilitate semantic description. Using mutual correspondence analysis on translation data has allowed us to better define the possible meanings conveyed by the verbs in Spanish, French and Italian. Next, we have been able to draw a clearer picture of the principal equivalences and differences by complementing the translation corpus with a comparable corpus. This has allowed us to quantify the presence of the meanings in each language and the continuum between them. The contexts observed in our corpus extend from meanings of direct physical perception (including various modalities of perception), past cognitive perception, to more subjective and emotional values, albeit in different degrees in the three languages.

The twofold case study also contributes to the theory of cross-linguistic equivalence, which has rarely focused on one very complex element in three related languages. The analysis confirms the generally accepted belief that perfect synonymy does not exist (cf. among others Lyons 1968; Lewandowska-Tomaszczyk 1999; Altenberg and Granger 2002), but in addition points towards the *clusters* of meanings that clearly dominate the semantics of the verb in each language. As a member of the category of perception verbs, *sentir(e)* confirms its status of displaying complex but interesting patterns of polysemy that show both cross-linguistic regularities and language specific traits. French *sentir* most dominantly – but certainly not exclusively – covers the field of cognitive (but often intuitive) perception, meaning ‘to think’

or sometimes even ‘to know’. Spanish, on the other hand, has strongly developed the subjective pole of the verb, being the sole language where *sentir* has undergone a grammaticalization pattern in the sense of ‘being sorry’.

Finally, Italian seems to be the language where *sentire* most clearly belongs to the category of perception verbs, referring in the vast majority of the cases to auditory perception. As such, Italian *sentire* can hardly be defined as a perception verb referring to an inferior modality. Rather, being semantically closer to Spanish *oír/escuchar* and French *entendre/écouter* (‘to hear’/‘to listen’), it enters the domain of the dominant perception modalities, as opposed to its French and Spanish cognates. The present study thus gives clear support to the sense modality hierarchy: *sight* > *hearing* > *other modalities* (Viberg 1984, 2001; Evans and Wilkins 2000). The auditory meaning has also been identified as being at the basis of the grammaticalization of the verb in the interjection *senti* (‘listen’). To conclude, both grammaticalization patterns in Spanish and in Italian confirm Viberg’s (1999) conclusion that grammaticalization can drive cognates apart semantically.

Notes

¹This study shows that the verbs *sentir(e)* share the same meaning of *general perception*, hence the translation ‘to feel’. However, it also shows that the verbs can refer to several other meanings. Moreover, for the sake of practical organization, this study does not take into account the constructional profile of the verbs. The syntax-semantics interface will be the subject of future studies.

²In accordance with this objective, the present paper is situated within the domain of contrastive lexical studies (cf. Viberg 2002, 2008) from the perspective of cognitive semantics. Indeed as numerous studies show, one of the central areas of research within these domains has been precisely the investigation of polysemy and (near)synonymy (cf. among many others the recent publications of Viberg 2002, 2008; Divjak and Gries 2009; Glynn and Fischer 2010). Therefore, it should be clear from the outset that the notion of *semantics* as used in the present article refers to the tradition of cognitive semantics as opposed to other theoretical currents in lexical semantics, viz. historical-philological, structuralism, generativist or neostructuralist semantics (cf. Geeraerts 2010b for a comprehensive overview of the main theoretical trends in lexical semantics).

³This is in contrast to other verbs which share the same etymology but nevertheless have a rather different morphology in their actual forms, such as *habēre* v. *avoir* (French), *haber* (Spanish) and *avere* (Italian) or *clamāre* v. *clamer* (French), *llamar* (Spanish) and *chiamare* (Italian). This perfect formal equivalence is what distinguishes our study from other contrastive lexical studies, such as that of Lewandowska-Tomaszczyk (1999), who compares English *persuade* with its Polish equivalents, or Viberg (2002, 2005, 2008), who compares the primary, but morphologically different, equivalents of causative verbs and mental verbs in genetically closely related languages such as Swedish, English and German. Viberg’s (1999) study on Swedish *gå* and English *go* presents a similar starting point of formal equivalence.

⁴The Spanish dictionary consulted was *Diccionario del Español Actual*; for French, we refer to *Le Nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française* and for Italian to the *Grande dizionario Italiano dell’uso*. Since an exhaustive analysis of all the differences and similarities between these dictionaries goes beyond the scope of the present study, this section aspires to sketch a rather general picture of the lexicographic treatment of the verbs.

⁵The recent publications of different volumes with the explicit purpose of establishing the foundations for empirical investigation testify to the rising importance dedicated to these questions (e.g. Gries 2006; González-Márquez et al. 2007; Glynn and Fischer 2010; Glynn and Robinson in press).

⁶The notions of *comparable*, *parallel* and *translation corpus* have received different interpretations (cf. Granger 2003; Noël 2003: 781; McEnery and Xiao 2008: 19-20). In this paper *parallel corpus* is used as a synonym of *translated* or *translation corpus* and refers to a number of source texts and their translations in one (bilingual) or more (multilingual) languages, whereas a *comparable corpus* contains a number of texts that are collected using the same criteria of representativeness (such as genre, period, etc.) but that are not translations of each other.

⁷For a more detailed discussion of the advantages and disadvantages each corpus type offers, see for instance Altenberg and Granger (2002) and Mortier and Degand (2009). This combined corpus approach is not new; a

major characteristic of the *English-Norwegian Parallel Corpus* (ENPC) and *English-Swedish Parallel Corpus* (ESPC) is that they contain originals in both languages that could additionally be compared to translations in both directions.

⁸See also the bibliography for more detailed information. During the compilation of the corpus, it soon became clear that what has been translated varies with the particular language. It seems to be most difficult to find translated texts in Italian. Note that, given that the objective of this study is not primarily to contribute to the study of translations, we did not a priori take into account the quality of the individual translations. It should also be mentioned that the requisite of working with contemporary texts translated into several languages often results in the use of translations of bestsellers, whose quality is hard to define. One could also argue that six different texts do not necessarily guarantee a sufficient representativeness of the translation data, in contrast to, for instance, the ENPC and ESPC, which combine extracts from a higher number of novels with their translations. However, we would like to stress that the translation corpus serves as a heuristic to determine the semantic extensions that the verbs can undergo in the three languages, but that the detailed description of the differences and equivalences is based on an independent highly representative corpus of 1500 tokens. Finally, we consider the method of comparing translations to be especially suitable for morphological cognates, without claiming that the comparison with translated texts without explicit comparison of the source text should be recommended as a general method. The selection of the kind of corpus data should always take into account the research question and the pursued objective.

⁹Note that *Correspondence Analysis* does not refer to the multivariate statistical technique often performed in linguistic studies but to the analysis of the degree of correspondence between two or more languages as to their translations of an original semantic context. In this perspective, it should be mentioned that the notion of *Mutual Correspondence* has already been used by Altenberg (1999) to refer to the degree of equivalence in a bidirectional translation corpus, as becomes clear from the following description: “If an item *x* in a language A is always translated by *y* in language B and, conversely, item *y* in language B is always translated by *x* in language A, they will have a mutual correspondence of 100%. If they are never translated by each other their mutual correspondence will be 0%. In other words, the higher the mutual correspondence value is, the greater the equivalence between the compared items is likely to be” (Altenberg and Granger 2002: 18).

¹⁰Note that the selected examples of *sentir(e)* in the target languages were all instances of a process implying two different participants. Reflexive cases of *sentirse* / *se sentir* / *sentirsi* were not included since they need a more detailed analysis and a proper treatment.

¹¹In the description of our results we focus mainly on syntactically congruent matches (that is, through a verbal expression) and will not go into detail about the cases of *zero correspondence* (Johansson 2002) or *inventive translations* (Salkie 2002) as illustrated by the examples below, which could be the object of a more thorough analysis:

- a. Durante horas estuve convencido de que iba a morir. *Sentía* la angustia de la muerte y no podía hacer absolutamente nada. (HNAM-SP: 328)
 - b. Je suis resté totalement sans défense pendant plusieurs heures. J'étais persuadé que j'allais mourir. *L'angoisse de mourir me bouffait les tripes* et j'étais totalement impuissant. (HNAM-FR: 464)
 - c. Per molte ore ho creduto che sarei morto. *Ero angosciato* e non potevo fare assolutamente nulla. (HNAM-IT: 373)
- ‘For hours I was convinced that I was going to die. I felt the fear of death and I could do absolutely nothing.’

The rather high amount of cases of zero correspondence should not be surprising, given the literary nature of the compiled corpus and the consequently potentially high number of inventive translations, as indicated by Johansson (1998, 2002).

¹²It is not within the scope of this article to resort to advanced statistical techniques. The only measure provided is the result of the well-known Chi-square test, which in this case is significant, indicating that the differences between the languages cannot be attributed to chance. For an application of highly developed statistical techniques to the results of a parallel corpus analysis, see for instance Divjak (2010b).

¹³For the sake of clarity, we give here the English translation of the sentences derived from the Swedish original HNAM.

¹⁴Note that, as was the case for the availability of translations, the availability of representative and varied corpora differs considerably from language to language. While the Spanish data bank *CREA* contains both fiction and journalistic data, the French literary database *FRANTEXT* needed to be complemented by data from the newspaper *Le Monde*. For Italian, data sparseness turned out to be an even thornier problem, which explains our decision to supplement the journalistic database (*Il Corriere della Sera*, *CdS*) with data drawn from two novels (cf. bibliography).

¹⁵More specifically, the strength of agreement is interpreted as follows: < 0.2 = poor; $> 0.2 \leq 0.4$ = fair; $> 0.4 \leq 0.6$ = moderate; $> 0.6 \leq 0.8$ = substantial; $> 0.8 \leq 1$ = very good.

¹⁶It needs to be stressed, however, that the methodology adopted here does not aspire to compete with other recent valuable approaches to semantics, such as collocation or collocation analyses, behavioral profile analysis, semantic maps, and many others. On the contrary, sharing the same aim of investigating semantics on an empirical and objective basis, the present study offers a methodological approach to operationalizing this aim.

¹⁷Indeed, the calculation of the Pearson residuals allowing the identification of the cells most responsible for the significant result corroborates these language-specific preferences. In general, if a Pearson residual in a cell is positive, then the observed frequency in that cell is greater than the expected frequency in that cell, but if the Pearson residual in a cell is negative, then the observed frequency is less than the expected frequency. The more the Pearson residual deviates from 0, the stronger that effect (cf. Gries in press). In this case, the strongest effect is the preference for the emotional meaning in Spanish (residual 16,301), followed by the prominence of the specific perception modality in Italian (residual 14,959) and the preference for cognitive perception in French (residual 10,565). The effect size for table 5 is 0,574.

¹⁸This distinction between *internal* vs. *external perception* has been referred to as *endogenous* vs. *exogenous perception* by Fernández Jaén (2006b: 393).

¹⁹In the wake of Hopper (1998) and following Company Company (2004: 65) we consider ‘grammaticalization’ as an umbrella term referring to “un macrocambio dinámico, un cambio de cambios, que engloba distintas subclases y procesos [...] la rutinización o cristalización del uso, sea cual sea la direccionalidad del cambio” [a dynamic macro change, a change of changes, which includes several subclasses and processes [...] the routinization or crystallization of the use, regardless of the directionality of the change]. Due to space limitations, we will not go into details here concerning the degree of grammaticalization and (inter)subjectification of the expression *lo siento*.

²⁰Cf. Tanghe, S. and Jansegers, M. (in press) for a detailed contrastive analysis of the discourse markers *oye/oiga, mira/mire* in Spanish and *sentí/senta, guarda/guardi* in Italian.

Bibliography

Cited references

- Aijmer, Karin & Bengt Altenberg. 1996. Introduction. In Karin Aijmer, Bengt Altenberg & Mats Johansson (eds.), *Languages in contrast*, 11-16. Lund: Lund University Press.
- Aijmer, Karin & Simon-Vandenberg, Anne-Marie. 2004. A model and a methodology for the study of pragmatic markers: The semantic field of expectation. *Journal of Pragmatics*, 36 (10). 1781-1805.
- Altenberg, Bengt. 1999. Adverbial connectors in English and Swedish: Semantic and lexical correspondences. In Hilde Hasselgård & Signe Oksefjell (eds.) *Out of corpora. Studies in honour of Stig Johansson*, 249-268. Amsterdam: Rodopi.
- Altenberg, Bengt & Granger, Sylviane. 2002. Recent trends in cross-linguistic lexical studies. In Bengt Altenberg & Sylviane Granger (eds.), *Lexis in Contrast. Corpus-based approaches*, 3-48. Amsterdam & Philadelphia: John Benjamins.
- Badynska-Lipowczan, Beata. 1996. Quelques remarques sur *sentir* et *ressentir* en français et en italien. *Neophilologica* 12. 7-13.
- Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20. 37-46.
- Company Company, Concepción. 2004. ¿Gramaticalización o desgramaticalización? El reanálisis y subjetivización de verbos como marcadores discursivos en la historia del español. *Revista de Filología Española* 84. 29-66.
- Damasio, Antonio. 2003. *Looking for Spinoza. Joy, Sorrow, and the Feeling Brain*. Orlando: Harcourt.
- De Mauro, Tullio. 1999. *Grande dizionario Italiano dell'uso*. Torino: Utet.
- Divjak, Dagmar & Gries, Stefan Th. 2009. Corpus-based cognitive semantics: a contrastive study of phasal verbs in English and Russian. In Katarzyna Dziwirek & Barbara Lewandowska-

- Tomaszczyk (eds.), *Studies in Cognitive Corpus Linguistics*, 273-296. Frankfurt am Main: Peter Lang.
- Divjak, Dagmar. 2010a. *Structuring the lexicon. A clustered model for near-synonymy*. Berlin & New York: Mouton de Gruyter.
- Divjak, Dagmar. 2010b. Corpus-based evidence for an idiosyncratic aspect-modality relation in Russian. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: corpus-driven approaches*, 305-330. Berlin & New York: Mouton de Gruyter.
- Dupas, Chantal. 1997. *Perception et langage. Étude linguistique du fonctionnement des verbes de perception auditive et visuelle en anglais et en français*. Louvain & Paris: Peeters.
- Dyvik, Helge. 1998. A translational basis for semantics. In Stig Johansson & Signe Oksefjell (eds.), *Corpora and Crosslinguistic Research: Theory, Method, and Case Studies*, 51-86. Amsterdam: Rodopi.
- Dyvik, Helge. 2005. Translations as a semantic knowledge source. *Proceedings of the Second Baltic Conference on Human Language Technologies*, 27-38. Tallinn: Institute of Cybernetics at Tallinn University of Technology, Institute of the Estonian Language.
- Enghels, Renata. 2007. *Les modalités de perception visuelle et auditive. Différences conceptuelles et répercussions sémantico-syntaxiques en espagnol et en français*. Tübingen: Niemeyer.
- Evans, Nicholas & Wilkins, David. 2000. In the Mind's Ear: The Semantic Extension of Perception Verbs in Australian Languages. *Language* 76. 546-592.
- Fernández Jaén, Jorge. 2006a. Análisis cognitivo del verbo *oler*. *Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística*. 542-561.
- Fernández Jaén, Jorge. 2006b. Verbos de percepción sensorial en español: una clasificación cognitiva. *Interlingüística* 16. 391-405
- Franckel, Jean-Jacques & Daniel Lebaud. 1995. Les échappées du verbe *sentir*. In Janine Bouscaren & Antoine Culioli (eds.), *Langues et langage: problèmes et raisonnement en linguistique: mélanges offerts à Antoine Culioli*, 261-277. Paris: Presses Universitaires de France.
- Geeraerts, Dirk. 2006. Methodology in Cognitive Linguistics. In Gitte Kristiansen, Michel Achard, René Dirven & Francisco Ruiz de Mendoza Ibañez (eds.), *Cognitive Linguistics: Current Applications and Future Perspectives*, 21-49. Berlin & New York: Mouton de Gruyter.
- Geeraerts, Dirk. 2010a. The doctor and the semantician. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: corpus-driven approaches*, 63-78. Berlin & New York: Mouton de Gruyter.
- Geeraerts, Dirk. 2010b. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Gibbs, Raymond. 2007. Why cognitive linguists should care more about empirical methods. In Monica González-Márquez, Irene Mittelberg, Seana Coulson & Michael J. Spivey (eds.), *Methods in Cognitive Linguistics*, 2-18. Amsterdam & Philadelphia: John Benjamins.
- Gilquin, Gaëtanelle. 2008. Causative Make and Faire: A Case of Mismatch. In María de los Ángeles Gómez González, J. Lachlan Mackenzie & Elsa M. González Álvarez (eds.), *Current Trends in Contrastive Linguistics: Functional and Cognitive Perspectives*, 177-201. Amsterdam: John Benjamins.
- Glynn, Dylan. 2010a. Corpus-driven Cognitive Semantics. Introduction to the field. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: corpus-driven approaches*, 1-41. Berlin & New York: Mouton de Gruyter.
- Glynn, Dylan. 2010b. Synonyme, lexical fields, and grammatical constructions. A study in usage-based cognitive semantics. In Hans-Jörg Schmid & Susanne Handle (eds.), *Cognitive Foundations of Linguistic Usage Patterns*, 89-118. Berlin & New York: Mouton de Gruyter.

- Glynn, Dylan. 2010c. Testing the hypothesis: Objectivity and verification in usage-based Cognitive Semantics. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: corpus-driven approaches*, 239-269. Berlin & New York: Mouton de Gruyter.
- Glynn, Dylan & Kerstin Fischer (eds.). 2010. *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. Berlin & New York: Mouton de Gruyter.
- Glynn Dylan & Robinson Justyna (eds.). (in press). *Polysemy and synonymy: corpus methods and applications in Cognitive Linguistics*. Amsterdam & Philadelphia: John Benjamins.
- González-Márquez, Monica, Irene Mittelberg, Seana Coulson & Michael J. Spivey (eds.). 2007. *Methods in Cognitive Linguistics*. Amsterdam & Philadelphia: John Benjamins.
- Granger, Sylviane. 2003. The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies? In Sylviane Granger, Jacques Lerot & Stephanie Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, 17-29. Amsterdam: Rodopi.
- Gries, Stefan. Th. 2006. Corpus-based methods and cognitive semantics: The many senses of 'to run'. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*, 57-99. Berlin & New York: Mouton de Gruyter.
- Gries, Stefan Th. (in press). Frequency tables, effect sizes, and explorations. In Dylan Glynn & Justyna Robinson (eds.), *Polysemy and synonymy: corpus methods and applications in Cognitive Linguistics*. Amsterdam & Philadelphia: John Benjamins.
- Hanegreefs, Hilde. 2008. *Los verbos de percepción visual. Un análisis de corpus en un marco cognitivo*. PhD dissertation: KULeuven.
- Hopper, Paul J. 1998. Emergent grammar. In Michael Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches To Language Structure*, 155-175. Mahwah, NJ: Lawrence Erlbaum.
- Ibarretxe-Antuñano, Iraide B. 1999. Metaphorical mappings in the sense of smell. *Metaphor in Cognitive Linguistics, Selected Papers from the 5th International Cognitive Linguistics Conference*, Amsterdam, July 1997, 29-45. Amsterdam & Philadelphia: John Benjamins.
- Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge: MIT Press.
- Johansson, Stig. 1998. On the role of corpora in cross-linguistic research. In Stig Johansson & Signe Oksefjell (eds.), *Corpora and cross-linguistic research: Theory, method, and case studies*, 3-24. Amsterdam: Rodopi.
- Johansson, Stig. 2002. Towards a multilingual corpus for contrastive analysis and translation studies. In Lars Borin (ed.), *Parallel corpora, parallel words*, 47-59. Amsterdam: Rodopi.
- Lewandowska-Tomaszczyk, Barbara. 1999. *Cognitive Perspectives on Language*. Frankfurt am Main: Peter Lang.
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- McEnery, Tony & Richard Xiao. 2008. Parallel and comparable corpora: what is happening? In Gunilla Anderman & Margaret Rogers (eds.), *Incorporating corpora. The Linguist and the Translator*, 18-31. Clevedon: Multilingual Matters Ltd.
- Miller, George A. & Philip N. Johnson-Laird. 1976. *Language and perception*. Cambridge: Cambridge University Press.
- Mortier Liesbeth & Degand Liesbeth. 2009. Adversative discourse markers in contrast. The need for a combined corpus approach. *International Journal of Corpus Linguistics* 14 (3). 301-329.
- Noël, Dirk. 2003. Translations as evidence for semantics: an illustration. *Linguistics* 41 (4). 757-785.
- Robert, Paul, Alain Rey & Josette Rey-Debove. 2008. *Le Nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française. Nouvelle édition du Petit Robert de Paul Robert*. Paris: Dictionnaires Le Robert.

- Salkie, Raphael. 2002. How can linguists profit from parallel corpora? In Lars Borin (ed.), *Parallel corpora, parallel worlds*, 93-109. Amsterdam: Rodopi.
- Santos, Diana. 2008. Perfect mismatches: 'Result' in English and Portuguese. In Gunilla Anderman & Margaret Rogers (eds.), *Incorporating corpora. The Linguist and the Translator*, 217-242. Clevedon: Multilingual Matters Ltd.
- Seco, Manuel, Andrés Olimpia & Ramos Gabino. 1999. *Diccionario del español actual*. Madrid: Aguilar.
- Stefanowitsch, Anatol. 2010. Empirical Cognitive Semantics: Some thoughts. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: corpus-driven approaches*, 355-380. Berlin & New York: Mouton de Gruyter.
- Sweetser, Eve. 1990. *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Talmy, Leonard. 2007. Foreword. In Monica González-Márquez, Irene Mittelberg, Seana Coulson & Michael J. Spivey (eds.), *Methods in Cognitive Linguistics*. Amsterdam & Philadelphia: John Benjamins.
- Tanghe Sanne & Jansegers Marlies (in press). Los marcadores del discurso derivados de los verbos de percepción: un estudio contrastivo español-italiano. *Actas del II Coloquio Internacional. Marcadores discursivos en las lenguas románicas: un enfoque contrastivo*.
- Valentim, Helena. 2002. Contributo para o estudo do verbo *sentir* - perspectiva enunciativa. *Actas do XVII Encontro Nacional da Associação Portuguesa de Linguística*. 509-516.
- Vanderschueren, Clara. 2010. The use of translations in linguistic argumentation. A case study on Spanish and Portuguese subordinate clauses introduced by *para*. *Languages in Contrast* 10. 76-101.
- Van Hoecke, Willy & Michèle Goyens. 1990. Translation as a witness to semantic change. *Belgian Journal of Linguistics* 5. 109-131.
- Vesterinen, Rainer. 2010. The relation between iconicity and subjectification in Portuguese complementation: Complements of perception and causation verbs. *Cognitive Linguistics* 21. 573-600.
- Viberg, Åke. 1984. The verbs of perception: a typological study. *Linguistics* 21 (1). 123-162.
- Viberg, Åke. 1999. The polysemous cognates Swedish *gå* and English *go*. Universal and language-specific characteristics. *Languages in Contrast* (2). 89-115.
- Viberg, Åke. 2001. The verbs of perception. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals. An International Handbook*, 1294-1309. Berlin & New York: Mouton de Gruyter.
- Viberg, Åke. 2002. Polysemy and disambiguation cues across languages. The case of Swedish *få* and English *get*. In Bengt Altenberg and Sylviane Granger (eds.), *Lexis in contrast*, 119-150. Amsterdam & Philadelphia: John Benjamins.
- Viberg, Åke. 2005. The lexical typological profile of Swedish mental verbs. *Languages in Contrast* 5. 121-157.
- Viberg, Åke. 2008. Swedish verbs of perception from a typological and contrastive perspective. In María de los Ángeles Gómez González, Lachlan Mackenzie and Elsa González Álvarez (eds.), *Languages and Cultures in Contrast: New directions in Contrastive Linguistics*, 123-172. Amsterdam & Philadelphia: John Benjamins.
- Zeschel, Arne. 2010. Exemplars and analogy: Semantic extension in constructional networks. In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: corpus-driven approaches*, 201-219. Berlin & New York: Mouton de Gruyter.

Corpus

Comparable corpus

ATILF: *Base Textuelle Frantext*. www.frantext.fr. [FRANT]

Camilleri, Andrea. 2005. *La luna di carta*. Palermo: Sellerio editore. [CA]

Il Corriere della Sera: archiviostorico.corriere.it/searchresultsArchivio.jsp. [CdS]

Le Monde: www.lemonde.fr.

Lo Gatto, Antonio. 2004. *L'intreccio di universi paralleli*. www.latelanera.com/files/ebook054.pdf

Real Academia Española: *Corpus de Referencia del Español Actual*. www.rae.es [CREA].

Parallel corpus

Brown, Dan. 2003. *The Da Vinci Code*. New York: Doubleday.

[DV-SP] Estrella, Juanjo. 2003. *El código Da Vinci*. Barcelona: Umbriel.

[DV-FR] Roche, Daniel. 2004. *Da Vinci code*. Paris: J.-C. Lattès.

[DV-IT] Valla, Riccardo. 2004. *Il codice da Vinci*. Milano: Mondadori.

Larsson, Stieg. 2005. *Män som hatar kvinnor*. Stockholm: Norstedts Förlag.

[HNAM-SP] Lexell, Martin & Juan José Ortega Román. 2008. *Los hombres que no amaban a las mujeres*. Barcelona: Ediciones Destino.

[HNAM-FR] Giorgetti Cima, Carmen. 2007. *Uomini che odiano le donne*. Venezia: Marsilio.

[HNAM-IT] Grumbach, Lena & Marc De Gouvenain. 2006. *Les hommes qui n'aimaient pas les femmes*. Arles: Actes Sud.

Rowling, Joanne Kate. 1998. *Harry Potter and the sorcerer's stone*. New York: Scholastic Inc.

[HPPF-SP] Dellepiane, Alicia. 1999. *Harry Potter y la piedra filosofal*. Barcelona: Emecé Editores S.A.

[HPPF-FR] Ménard, Jean-François. 1998. *Harry Potter à l'école des sorciers*. Paris: Gallimard Jeunesse.

[HPPF-IT] Astrologo, Marina. 1998. *Harry Potter e la Pietra Filosofale*. Milano: Adriano Salani Editore.

Rowling, Joanne Kate. 2000. *Harry Potter and the goblet of fire*. New York: Scholastic Inc.

[HPCS-SP] Muñoz García, Adolfo & Nieves Martín Azofra. 2001. *Harry Potter y el cáliz del fuego*. Barcelona: Salamandra.

[HPCS-FR] Ménard, Jean-François. 2007. *Harry Potter et la coupe de feu*. Paris: Gallimard Jeunesse.

[HPCS-IT] Masini, Beatrice. 2001. *Harry Potter e il calice di fuoco*. Milano: Adriano Salani Editore.

Rowling, Joanne Kate. 1999. *Harry Potter and the chamber of the secrets*. New York: Scholastic Inc.

[HPCS-SP] Muñoz García, Adolfo & Nieves Martín Azofra. 1999. *Harry Potter y la cámara secreta*. Barcelona: Salamandra.

[HPCS-FR] Ménard, Jean-François. 1999. *Harry Potter et la chambre des secrets*. Paris: Gallimard Jeunesse.

[HPCS-IT] Astrologo, Marina. 1999. *Harry Potter e la camera dei segreti*. Milano: Adriano Salani Editore.

Rowling, Joanne Kate. 1999. *Harry Potter and the prisoner of Azkaban*. New York: Scholastic Inc.

[HPPR-SP] Muñoz García, Adolfo & Martín Azofra, Nieves. 2000. *Harry Potter y el prisionero de Azkaban*. Barcelona: Salamandra.

[HPPR-FR] Ménard, Jean-François. 1999. *Harry Potter et le prisonnier d'Azkaban*. Paris: Gallimard Jeunesse.

[HPPR-IT] Masini, Beatrice. 2000. *Harry Potter e il prigioniero di Azkaban*. Milano: Adriano Salani Editore.