

# SEMANTIC PROSODIES IN ENGLISH AND PORTUGUESE: A CONTRASTIVE STUDY

Tony Berber Sardinha  
Departamento de Linguística  
Pontifícia Universidade Católica de São Paulo

## ABSTRACT

*The present study is aimed at describing the semantic prosody of equivalent items in English and Portuguese. Semantic prosody is the connotation conveyed by the regular co-occurrence of lexical items, as revealed by the exploration of a computer-readable corpus. Although there are several studies dealing with semantic prosody in English, only one has looked at this issue contrasting English and Portuguese. In general, the findings indicate important similarities and differences between the two languages, and point toward inadequacies in contemporary dictionaries. The study was based on the exploration of a Portuguese computer corpus of over 140 million words, one of the largest for Portuguese. The BNC, with 100 million words, was also used in the course of the investigation. The general conclusion of the study presented here is that information on connotation, especially that based on the exploration of corpora, should be part of glossaries and dictionaries, particularly production and bilingual dictionaries, which are tools that translators rely upon regularly.*

**KEY WORDS:** *semantic prosody, connotation, glossaries, dictionaries*

## RESUMEN

*El objetivo del presente estudio es describir la prosodia semántica entre términos equivalentes en inglés y portugués. Lo prosodia semántica se define como las connotaciones que ciertas coocurrencias habituales que acompañan a una palabra, transmiten (tal y como ha quedado atestiguado en varios estudios basados en corpus). Encontramos varios estudios sobre prosodia semántica para el inglés, pero solamente un estudio contrastivo inglés-portugués. En general, los resultados señalan importantes similitudes y divergencias entre ambas lenguas, encontrándose inconsistencias en algunos diccionarios actuales. Este estudio se ha basado en dos corpus lingüísticos: un corpus del portugués con más de 140 millones de palabras (uno de los más exhaustivos en esta lengua) y el BNC, con 100 millones de palabras. Lo principal conclusión de que se ha llegado en este estudio es que la información que conllevan las connotaciones, especialmente la que puede extraerse de corpus lingüísticos, debería incluirse en glosarios y diccionarios, sobre todo a lo hora de compilar diccionarios y en diccionarios bilingües, al ser éstos, utensilios en los que habitualmente confía un(a) traductor(a).*

**PALABRAS CLAVE:** *prosodia semántica, connotación, glosarios, diccionarios*

## I. INTRODUCTION

Lexical patterning has been a common theme in the Corpus Linguistics literature (Francis & Hunston, 1996, 1998; Hunston & Francis, 2000; Partington, 1998; Sinclair, 1991; Stubbs, 1996). Patterning refers to the regularities observed in language due to the consistent repetition of co-occurring units of various kinds (lexical, grammatical, pragmatic, etc.). As Hunston and Francis (Hunston & Francis, 2000) put it:

*'The patterns of a word can be defined as all the words and structures which are regularly associated with the word and which contribute to its meaning. A pattern can be identified if a combination of words occurs relatively frequently, if it is dependent on a particular word choice, and if there is a clear meaning associated with it.'* (p.37)

Observing patterns has become possible due to the exploration of corpora, or large quantities of text in computer readable format. The main motivation for the present study is the methodology for identification and interpretation of lexical patterns in computer corpora available in the Corpus Linguistics literature, notably that in the Birmingham tradition founded by John Sinclair (Raker, Francis, & Tognini-Bonelli, 1993; Hunston & Francis, 2000; Partington, 1998; Sinclair, 1991). The main theoretical underpinning is the notion of *idiom principle* proposed by Sinclair (1991), according to which the language user has available to him or her a large number of pre-constructed phrases which are not readily observable by introspection.

The aim of this study is to look at the issue of translating lexical patterns, with particular reference to issues related to the translatability of semantic prosody across English and Portuguese. The basic methodology is that advocated and implemented by several researchers for the study of corpus-based lexical patterning, metaphor and semantic prosody (Channell, 2000; Deignan, 1999a, b; Louw, 1993; Partington, 1998; Stubbs, 1996). In what follows, a number of items will be analyzed contrastively based on data from a large corpus of Portuguese and previous research. Following this, the results will be discussed briefly in relation to previous research, and the paper ends with general conclusions, limitations, and issues for further research.

## II. SEMANTIC PROSODY

One kind of pattern that is crucial in translations is semantic prosody, or the association of lexical items to a semantic field, interpretable in terms of a connotation. The definition used in the present study is that of Hunston and Francis (Hunston & Francis, 2000, p.137):

'... a word may be said to have a particular semantic prosody if it can be shown to co-occur typically with other words that belong to a particular semantic set.'

The term was coined by John Sinclair (in Sinclair, 1987). Partington (1998, p.68) interprets it as 'the spreading of connotational colouring beyond single word boundaries'. For example, the adjective *impressive* in English collocates with items such as *achievement, talent* and *dignity*. In this manner, it is considered to have a positive semantic prosody (Partington, 1998, p.77). By contrast, *rife* tends to co-occur with words such as *crime, misery* and *diseases*, hence its interpretation as having a negative semantic prosody (Partington, 1998, p.67).

There are several studies in the literature which look at the semantic prosodies of a number of lexical items (Channell, 2000; Deignan, 1999b; Fox, 1998; Hunston, 1995; Louw, 1993; Sinclair, 1991; Stubbs, 1995). The vast majority of these studies focuses on English words. Two studies which investigate the semantic prosody of other languages are Partington (1998), which describes items in Italian, and Rerber Sardinha (1999b), which compares English and Portuguese prosodies (this will be described in more detail below).

In English, Sinclair (1991) analyzes the semantic prosody of *set in* and concludes that this term has a negative connotation (it collocates with *rot, decadence, decay*, for example). Louw (1993) discusses *build up* and shows that as a transitive verb the semantic prosody is favorable (e.g. *build up confidence*), but as an intransitive verb the prosody tends to be negative (*resistance builds up*). Channell (2000) presents a study on several items, including the expression *roam the streets*, which is negative because of its systematic association with items such as *prostitutes, looters* and *vagrant children*. This is in contrast with the prosody for the single word *roam*, which bears a positive or neutral connotation (e.g. *beautiful countryside*). Another expression whose semantic prosody is described by Channell in the same study is *par for the course*, which appears in the dictionary as 'what one would expect to happen or expect sb to do' (Oxford Advance Learner's Dictionary 1995). However, its semantic prosody points toward its use as an indirect complaint, as in *the third was out of order, which is par for the course*. Fox (1998) examines the expression *plain sailing*, which is normally explained as meaning something that is easy to do, but which is typically employed in negative contexts to indicate that which cannot be easily accomplished (e.g. *it wasn't plain sailing, not everything is plain sailing*).

Collocation and semantic prosody are areas of interest for translation studies. The comparison of collocations and semantic prosodies across two (or more) languages may throw light on several questions related to the appropriacy of (so-called) equivalent items. As Partington (1998, p.77) argues:

'Semantic prosody is an important area of research for translation studies. ... it seems to be the case that cognate, or 'look-alike' words in two related languages can have very different semantic prosodies'.

*This is the case, for example, of impressionante, which in Italian has a negative semantic prosody, collocating with items referring to price rises, kidnap attempts, etc. On the other hand, in English impressive shows a positive semantic prosody.*

*In Portuguese, there are very few studies looking at lexical patterning in corpora (e.g. Marques, 1996; Nascimiento & Carvalho, 1996; Neto, 1996; Tagnin, 1999). Of those, only one has focused on semantic prosody, namely Berber Sardinha (1999b). In that study, the author describes the semantic prosody of the expression *tocar para frente* (loosely translatable as *carry on*). According to the author, the Portuguese expression has an unfavorable semantic prosody, indicating something which is done without planning under circumstances where one has no other choice. The author concludes this connotation is inadequate given that the expression was used as a part of a slogan in a government campaign that was run to boost popular support for the government.*

*Semantic prosody creates an expectancy relation between reader and writer, or listener and speaker. Violating semantic prosody patterns text may add 'extra' meaning. This may be intentional, as Louw (1993, p. 164) argues when he discusses how *bent on*, which originally has a negative semantic prosody (byco-occurring with *destroying, harrying, mayhem*, for instance), indicates irony on the part of the author when it is accompanied by favorable items such as *self-improvement*. But the shift in semantic prosody quality may be unintentional. When this happens in translation, it gives rise to a different connotation than that intended by the writer or speaker of the source language text. This should be avoided in translation because it may create misunderstandings with respect to tone or content of the original message.*

### III. PREVIOUS STUDIES: CAUSAR/CAUSE AND ACONTECER/HAPPEN

*As mentioned above, one study which looked at the contrast between semantic prosodies in English and Portuguese is Berber Sardinha (1999b). The study employed a corpus of over 32 million words, of both contemporary written and spoken Brazilian Portuguese. In order to establish whether the co-occurrence of items was significant, statistical measures of association were computed, namely T-Score and Mutual Information. Only those items which passed the statistical tests are reported below.*

*One of the pairs of verbs described was 'cause' and its Portuguese equivalent 'causar'. In English, 'cause' has a negative semantic prosody (Stubbs, 1995, p.43). According to Stubbs (1995, p 43), the most typical collocates of 'cause' include:*

anxiety, concern, crisis, damage, distress, embarrassment, explosion, harm, loss, problem, problems, trouble, AIDS, cancer, death, diseases, illness, etc.

In the Portuguese corpus there were 1723 instances of 'causar' (and derived forms, such as '-causando' and 'causam'). The most frequent collocates in a 4:4 window (4 words on either side of the node) were 'Problemas', 'Danos', 'Morte(s)', 'Prejuízo(s)', 'Impacto', 'Polêmica', 'Mal', 'Prejuizo', e 'Doenças' (respectively, 'problems', 'damage', 'death(s)', 'loss(es)', 'impact', 'polemic', 'evil', 'disease'). It is clearly noticeable that the vast majority of collocations refer to negative events or circumstances. Thus, the author concludes that 'cause' and 'causar' share a negative semantic prosody.

The other verb analyzed by Berber Sardinha (1999b) was 'acontecer', which was compared to 'happen'. The latter has a predominantly negative semantic prosody (Stubbs, 1995, p.47). Typically, the collocates of 'happen' refer to unpleasant situations (accidents, tragedies, crisis); others, however, are unexpected events (e.g. 'miracles') (Stubbs, 1995, p.47).

The results of the collocational analysis indicated that the most frequent collocates in a 4:4 window for 'acontecer' (and derived forms) were 'crime', 'nada', 'coisa(s)', 'acidente', 'evento', 'final', 'jogo', 'mundial', e 'festival' (respectively crime, nothing, thing(s), accident, event, final, match/game, world). The collocates by and large carry either a positive or neutral connotation. Some of the neutral patterns are formed with 'vague' nouns such as 'coisa' ('thing'), 'algo' ('something') and 'caso' ('case'), while others are created by collocates referring to planned events such as 'evento', 'jogo', 'festival'. Only a couple of collocates indicate unfavorable conditions, such as 'crime', and 'acidente' ('accident'), but these are less frequent than the positive or neutral ones. Therefore, 'acontecer' and 'happen' have different semantic prosodies.

#### IV. CORPUS

The corpus used for the present study is the following:

Register	Tokens
Press	137.471.276
Speech*	3.178.882
Literary fiction	1.488.195
Academic prose	502.438
Written business documents	720.257
Total	142.861.043

\*Face-to-face conversation, telephone conversation, interviews, classes, meetings

Two statistical measures were used to calculate the strength of the association between the node and its collocates, namely *T-score (T)* e *Mutual Information (I)*. All results reported below are for collocates which had a score higher than the minimum levels for each test, which are  $T \geq 3$  and  $I \geq 3$  (Stubbs, 1995).

## V. COMMIT/COMETER

The first lexical items selected for study was the pair *commit/cometer*. According to Partington (1998, p.67), *commit* has a negative semantic prosody, collocating with words such as *foul offences* and *crime*. In order to check whether *cometer* has the same kind of semantic profile, the corpus was searched for *comer\**, excluding *comet/cometa*, which resulted in 10335 occurrences. The 15 most frequent collocates, in a 5:5 window are (the figures in brackets are the frequencies for each collocate): imperfection

crime/s (3328), erro/s (errors, 1791), faltas (fouls/faults, 430), irregularidades (wrongdoings/laws, 350), suicidio (suicide, 301), infrações (infractions, 179), atos (acts, 176), infração (infraction, 158), assassinatos (murder, 156), abusos (abuse, 155), equívoco (oversight, 133), atentado (attack, 120), violencia (violence, 119), caso (case, 115), excessos (excess, 105)

As can be noticed, the collocates express essentially negative situations. Hence, the semantic prosody of *commit* and *cometer* are similar.

## VI. SET IN

According to Sinclair (1987, p.155-6), the phrasal verb *set in* has a negative connotation:

The most striking feature of this phrasal verb is the nature of the subjects. In general, they refer to unpleasant states of affairs. Only three refer to the weather: a few are neutral, such as *reaction* and *trend*. The main vocabulary is *rot (3)*, *decay*, *ill-will*, *decadence*, *impoverishment*, *infection*, *prejudice*, *vicious (circle)*, *rigor mortis*, *numbness*, *bitterness* [...] Not one of these is desirable or attractive.

The Cobuild dictionary (Sinclair et al., 1995, CED henceforth) explains its meaning as follows:

*If something unpleasant sets in, it begins and seems likely to continue or develop.*  
(p.1518)

For instance: *Then disappointment sets in as they see the magic is no longer there...*

*Winter is setting in and the population is facing food and fuel shortages*  
(CED, p. 1518)

The negative connotation applies to the phrasal verb only, since *set* may also be followed by the preposition *in*, with the meaning of situated (for example 'set in the countryside').

In the *Dicionário Inglês-Português* by Houaiss (Houaiss, 1982, DIP henceforth), acceptable translations of *set in* are:

'manifestar-se, estabelecer-se; entrar (invemo [winter], frio [cold]); cair (noite [night], nevoeiro [fog])' (p.707).

The first option, *manifestar-se*, has 3901 occurrences in the corpus (the search term was *manifest\**, with *se-* as collocate in L1/R1, excluding *manifestaç\*/manifestante\**). The nouns with acceptable T and I values in a 5:5 window are:

doença (disease, 33), sociedade (society, 7) e povo (people, 5).

A few concordance lines are shown below.

do acabam em depressão. "A doença se manifesta nos quatro primeiros me  
do do que ocorrerá quando a doença se manifestar. O medo da morte e os  
seria suficiente para que a doença se manifestasse nas pessoas. Nos do  
O perfil das vítimas A doença se manifesta principalmente no iníci  
xuais e aleitamento. Mas a doença se manifesta em apenas 3% dos portad  
as de classe social alta. A doença se manifesta em mulher branca, na fa  
Na maioria dos homens, a doença se manifesta entre dois a cinco dias  
tetoras não impediram que a doença se manifestasse no próprio cientista  
ta. "Chegou a hora de a sociedade se manifestar contra esses planos de  
plebiscito: "Se toda a sociedade se manifesta, há risco de o interess  
arência. É preciso que a sociedade se manifeste sobre ele, avalie com m  
cias públicas para que a sociedade se manifeste." A ideia da CET, seg  
fazem parte da vida do nosso povo se manifestam da forma mais bela no  
fci construído sem que o povo pudesse manifestar-se. A população discut  
a consulta do 1963, quando o povo se manifestou sobre a permanência do  
lismo. Ou seja, uma vez que o povo se manifestou pelo presidencialismo

---

<sup>1</sup> This option has the disadvantage of picking up the conjunction 'se', which is different from the particle 'se' which accompanies the verb 'manifestar', but since the corpus was not tagged for part-of-speech, it was not possible to distinguish between the two kinds of 'se' automatically.

There are two main senses of *manifestar-se* in the concordance. When the verb collocates with words such as 'doença', the sense is that of 'affect'. In this case, *manifestar-se* has a negative semantic prosody. But in other cases, the meaning is 'to express oneself', and it does not carry a negative semantic prosody. This is a clear case of the possibility of disambiguating the main senses of a word by looking at its lexical environment (Hunston & Francis, 2000; Sinclair, 1991). The meaning of *manifestar-se* as 'to express oneself' does not appear to be neutral, though. The concordance lines for 'sociedade/povo' + 'manifest\*-se' tend to indicate instances where it was not easy for the public or society to express their views, thus suggesting the particular meaning of expressing oneself in unfavorable circumstances.

The other option recommended by DIP is *estabelecer-se*. The search term was *estabelec\*.se L1/R1,excluding estabelecim\**. The 15 substantivos most frequent collocates in the corpus are (with acceptable levels of T and I, in a 5:5 window):

relação (relation, 38), país (country, 28), mercado (market, 30), governo (government, 16), obra (construction, 15), prazo (deadline, 15), estabelecimentos (establishment, 14), região (region, 14), empresas (companies, 13), necessidades (needs, 13), tipo (type, 11), lei (law, 10), trabalho (work/job, 10), competência (competence, 9)

The collocates do not indicate an unfavorable semantic prosody. The concordance below illustrates some of these cases:

sua materialidade. A relação se estabelece diretamente entre a po  
a USP, diz que a relação que se estabelece entre pai e filho é mu  
reços externos, ou seja, eles se estabelecem no mercado mundial qu  
s dos Estados Unidos procuram se estabelecer no mercado brasileiro  
, bem como os interessados em se estabelecer no país comprando emp  
osa classe média negra tinha se estabelecido no país em consequên

In sum, there is evidence in the corpus to suggest that the negative semantic prosody of *set in* is not captured adequately by the two main options suggested by DIP (*estabelecer-se* e *manifestar-se*). The other options, which are more specific, are: 'cair (noite [night], nevoeiro [fog]); entrar (inverno [winter], frio [cold])' (p.707). Although Sinclair (1987) claims that the association of *set in* with weather and physical terms is rare, it is worth checking in the Portuguese corpus whether the verbs 'cair' and 'entrar' do collocate with the options suggested in dictionary.

For *cair*, the search term was *cair/ cai/ caiu/ caia/ caíra/ cairá/ cairiai caisse/ caindo/ caído*. It should be noted that *night* (the equivalent of *noite*) is not cited by Sinclair (1987) as a collocate of *set in*, but it does seem to have a negative semantic prosody, as the example below indicates (the only one in the RNC):

*Because when the night sets in, strange things happen 0111 in the wilds.*



*Of the terms suggested in the dictionary, noite was the only one that occurred in the neighborhood of cair. In a 5:5 window, noite appeared 69 times in the corpus. Nevoeiro did not occur at all, but nevoa (mist) came out twice in the collocates list. Of these, noite only reached acceptable levels of T and I.*

*The high frequency of the search items for cair exceeded the capacity of the concordancer in WordSmith Tools, whose limit is 16 million lines. As a result, the frequencies of the collocates had to be normalized (Biber, Conrad, & Reppen, 1998) to 16,000. For example, noite occurs 6763 in the corpus, but its normed frequency is 5 (that is, it appears 5 times per 16 thousand words). The frequency of items which appeared less than 16,000 times also had to be normed: for instance, nevoa occurs 189 times, but its normed frequency is 0.03 (that means one occurrence per 800 thousand words)*

*Below is a concordance for cair + noite:*

TV nem golpes de Estado. Ao cair da noite, jantariamos todos juntos, uma chefe regional da GCM. 'Ao cair da noite, há um afluxo de viciados em dr a para acabar. Se, antes de cair na noite, o esquiador optar por um relax no de 976. Quando começa a cair a noite, seu brilho se confunde com o dos guerras na África. quando cai a noite, a revolução e os capitães já es emer sobe. E desce. E sobe. Caía a noite. Era hora de ajustar as bússolas em em hotel fica. Dorme onde cai a noite. Até aí, não ha arce alguma. A a ior espetáculo da Terra. Ao cair da noite, o pai jogou fora as rodelas de usual confusão oriental. Ao cair da noite, estão povoados de jovens desoc  
Pag. 2-6 Com gás; Ac cair da noite: Pelo computador: De olho no vi sair de casa no fim da tarde/cair da noite . passe repelente no corpo . sao sair rom a namorada ou cair na noite cam os companheiros de time. s, Comgás e Petrobrás. Ao cair da noite A idéia da Cesp e acionar a fu "Seresta ao Pé da Serra", ao cair da noite de sexta. Endereços - Bar e da e inquietilharias. Quando cai a noite, não é difícil encontrar um bom p e saberemos, quando a noite cair e a bola começar a rolar. " Antes, ve em atuou em 'Quando a Noite Cai' ", de Patricia Rozema). Com o diretor Fr esistível. Assim que a noite cai, as barzinhos da praia de Iracema e os sh que acontece quando a noite cai sobre a periferia. O festival homenageia b um céu que, quando a noite cai, tem mais estrelas do que os peixes de u e esta/ mas quando a noite/ cai de súbito/ \_assim eu a prefiro/ sem pont

*Two basic patterns can be identified in the concordance: 'ao cair da noite' and 'quando a noite cai'. None has a negative semantic prosody. Although there are a few cases where there is mention of unfavorable circumstances (e.g. afluxo de viciados [influx of addicts], passar repelente contra insetos [apply repellent]), most of the cases refer to pleasant activities (jantar [dinner/dine], relaxar [relax], festival). Therefore, none of these patterns is a suitable translation for set in + night.*

*The last choice offered by the dictionary is set in as entrar. The search term was entrar/ entra entrou/ entrava/ entrara/ entrara entrarla/ entrasse entrando/ entrado. As was the case with cair, the frequency of the search words exceeded the limit of the concordancer, and so the frequencies had to be normalized as well.*

*The options indicated by the dictionary were noite and frio. The two terms feature among collocates of entrar. In a 5:5 window, noite + entrar appears 61 times in the corpus.*

while *frio + entrar* occurs six times. Roth terms reached the minimum values of *T* and *I*. Nevertheless, *none* is a legitimate collocation of *noite* or *frio* with *entrar*, as the concordance below indicates.

In summary, after the examination of all the choices for translations of *set in* given by DIP, the following conclusions were drawn:

- *Manifestar-se* maintains the negative semantic prosody of *set in* only when it forms collocations with *doença* [disease/sickness].
- *Estabelecer-se* collocates with other items which do not reflect the negative semantic prosody of *set in*.
- *Cair* forms regular patterns as 'ao cair da noite' and 'quando a noite cai'. In both cases, though, the semantic prosody is positive, unlike English.
- *Entrar* does not collocate with *noite* or *frio*.
- The sense of these expressions does not express the meaning of continuity and development which characterizes *set in* ('it begins and seems likely to continue or develop', Cobuild dictionary).

Hence, there is no suitable equivalent for *set in* among the choices presented in the dictionary, with respect to its semantic prosody. The question still remains as to what would be the most appropriate verb to preserve the semantic prosody of *set in* in Portuguese. One way to look for an answer is to search the corpus for the translations of the typical collocates of *set in* rather than *for* the verbs.

One of the most typical collocates of the *set in*, according to Sinclair (1987), is *rot*, which can be translated as *podridão*. There are 129 occurrences of this Word in the corpus. None of the verbs in a 5:5 window reached the minimum acceptable levels of *T* e *I*, and therefore there are no collocations for *podridão + verb* in the corpus<sup>1</sup>.

The other major collocate of *set in* is *decadence* ('*decadencia*'). In the corpus, there are 1660 occurrences of '*decadencia*'. The verbs appearing as collocates in a 5:5 window, with acceptable levels of *T* and *I* are:

*está/estão/estava* (*islareiw*s, 102), *entrou* (*entered*, 51), *começou* (*started*, 16), *trata* (*treats*, 12), *retrata* (*pictures*, 10), *vivem* (*live*, 9), *começa* (*start/s*, 8), *evitar* (*avoid*, 7), *veio* (*came*, 7), *caiu* (*fell*, 6), *enfrenta* (*face/s*, 6), *entraram* (*entered*, 6), *reflete* (*reilects*, 6)

The most common verb is *estar*, which is extremely frequent in the corpus: its three forms in the collocate list reached 405,691 occurrences. Despite the high frequency, the collocation with *decadencia* passed both association tests (*T*=4.4, *I*=9.6). This illustrates the

<sup>1</sup> The verb form 'exala' passed the *I* test (with a 15.3 score) but failed the *T* test by a very narrow margin (reaching 1.99, while the acceptable score was 2 or higher).

value of association tests. since without them one might have discarded *estar* as a collocate because of its high frequency in the language. The second most frequent verb form is *entrou*. A few concordance lines for these two collocates are presented below.

período de descanso no inverno. Conforme entra na hibernação, o coração, que  
 oar seus órgãos. (ARF) Moda inverno já entra em pré-liquidação  
 m-]- eu nem- nossa! \*Eu nem quero entrar em detalhes, (inint) frio.  
 ca vez. E pra entrevistar o Parreira! Já entra com o pé frio? Em vez da Gina

The concordance indicates that neither *estar* nor *entrar* + *decadência* capture the exact meaning of *decadence* + *set in*. In Portuguese, these collocations give the idea of the onset (*entrar*) or permanence (*estar*) but neither verb combines the two senses as *set in* does.

In sum, the results indicate that there are no direct equivalents for *set in* in Portuguese. The options offered in the dictionary do not maintain the original semantic prosody of the English verb. Furthermore, the Portuguese collocates which match the semantic prosody of *set in* do not form collocations with verbs which express a sense that is equivalent to *set in*.

então a gente sente que está em decadência, né?  
 rocesso político. Os "ismos estão em decadência, essa é a realidade de hoje  
 vaier menos do que os que estão em decadência cu em extinção''. E-mai  
 as antigas mercearias). Hoje estão em decadência. Nas cadeias de superme  
 co como aquele mundo vai entrando numa decadência incontrollável, numa  
 o mundo ocidental começou a entrar em decadência. E justamente com a limita  
 cesso alquímicos clássicos entraram em decadência, substituídos pela prática

## VII. ENTER INTO A DIALOGUE/MANTER UM DIÁLOGO

The cases illustrated above are for translations of English to Portuguese. However, semantic prosody is a factor for translations into English as well. To illustrate this, see the excerpt below:

'Foi-se o tempo em que era importante saber falar inglês. Hoje isso é imprescindível. Mesmo assim, muita gente ainda vacila na hora de *manter um diálogo* razoável nessa língua. O problema é adquirir fluência em pouco tempo.' ('Exame' magazine, 5 April 3000)

Loose translation:

Gone are the days when it was important to speak English. Nowadays this is indispensable. Even so, lots of people fail when they need to *enter into a reasonable dialogue* in that language. The problem is how to develop fluency in a short period of time.

The expression *manter um diálogo* apparently would not cause problems to a translator, given that there are close equivalents in English in all major dictionaries. For instance, in BBI Dictionary of English Word Combinations, one may look up *dialogue* and find the following phrases: *enter into a dialogue* and *have a dialogue with*. BBI unfortunately does not include semantic prosody information. Since there are no previous studies into the semantic prosody of this expression in the literature, it was necessary to look it up in a corpus. Below is a partial concordance for *enter into (a) dialogue* from the BNC:

ent political need to enter into dialogue with liberal bourgeois organisations  
 odia at the UN and to enter into dialogue with Vietnam came as an immense shock  
 ed his willingness to enter into a dialogue with the dissidents.  
 Moshoeshoe 's wish to enter into dialogue with South African liberation movements

These examples indicate that *enter into (a) dialogue* seems to indicate the meaning of negotiation, or discussion, generally between opposing sides. This interpretation is partly corroborated in Cobuild Dictionary:

*Dialogue is communication or discussion between people or groups of people such as governments or political parties.*

Clearly, this is not most adequate meaning for the expression. The text refers to a conversation, and not to negotiation. The question is whether the problem lies in the choice of the Portuguese expression. Again, there are no previous corpus-based studies into the semantic prosody of this particular expression, and therefore the following exploration was needed.

In the Portuguese corpus there were many instances of *manter + diálogo*, some of which are reproduced in the concordance below.

The chiefdenotation seems to be that of negotiation, and the main connotation can be understood in terms of 'communication under unfavorable circumstances' and perhaps antagonism. This means that both the Portuguese and the English expressions share a semantic prosody. Therefore, there is evidence to suggest that *manter um diálogo* was used improperly in the text cited above.

## Semantic prosodies in English and Portuguese: A contrastive study

adversários. Ele declarou que manterá um 'diálogo respeitoso' mesmo c  
mou ontem a Igreja Católica a manter "um diálogo construtivo" com a  
o assumir o cargo e manter pessoalmente um diálogo intenso com a sorie  
ação do 21. O canal quer também manter um diálogo com seu telespectad  
ça entre uma tuba e um trombone. Manter um diálogo de 'esfriamento sin  
ara uma reunião e vamos procurar manter um diálogo de alto nível, mesm  
e outros bichos. A criança pode manter um diálogo com os fantasmas e  
xt, também norte-americana, para manter um diálogo com a Mega, a agenr  
! Não. É por isso que não da pra manter um diálogo com você, Claudionc  
hor maneira de lidar com o Ira 4 manter um 'diálogo crítico'. Segundo  
e conversas com os clientes para manter um diálogo continuo com eles.  
ma. "É interesse da universidade manter um diálogo entre cinema, espir  
aíses, que tem interesse real em manter um diálogo construtivo entre s  
stoso com os jogadores e procuro manter um diálogo aberto com todos",  
aíses, que têm interesse real em manter um diálogo construtivo entre s  
to. O governo só tem condição de manter um diálogo no momento em que f  
oa confia em si mesma sabe como manter seu diálogo interno nos moment  
que o governo se recusa a manter qualquer dialogo com as tropas da RF  
am ultrapassados". Ele diz que manterá c diálogo com os praças da PM  
os da Comissão Justiça e Paz para manter o diálogo e evitar que o conf  
artheid. Ele se disse favorável a manter o diálogo com a Aliança pela  
estaram ontem intenção de manter aberto o diálogo e evitar a paralisa  
idade surgisse, procuraria manter sempre o diálogo e tentar fazer com  
ai à Opera.) A fulana insiste em manter o diálogo, apesar de os minha  
Segundo Oliveira, a orientação e manter o diálogo nas áreas de desoc  
nomeno do rap, com quem pretendem manter o diálogo político. "Sem nen  
país. A disposição de Mobutu em manter o diálogo e mais uma tentativ  
mpresas. Os camelôs ainda querem manter o diálogo com a prefeitura. O  
Segundo Oliveira, a orientação é manter o diálogo nas áreas de desoc  
Chee-hwa, declarou que pretende manter o diálogo com Taiwan, que é c  
Paulo Renato disse que pretende manter o diálogo com os parlamentare  
país. A disposição de Mobutu em manter o diálogo é mais uma tentativ  
ados. O govern de Minas manterá aberto o dialogo, mas os mentores de  
balhar com elencos locais e manter o mesmo diálogo, apenas com pequena  
e seus filhos andam. Tente manter o melhor diálogo possível com eles.  
com ele; de outro e atacado por não manter diálogo formal com o govern  
al de que o governo está disposto a manter diálogo com os republicanos  
, disse a emissora que Londres iria manter diálogo com membros do Sinn  
tida por esta comissão. "Temos que manter diálogo Entre os tres minis  
um Internacional Israel decide manter diálogo com palestinos 2  
provisada Papa conclama igreja a manter diálogo com 'instituições p  
ustentação parlamentar, enfrentar e manter diálogo com empresários e t  
nt Kabila tinham se comprometido a manter diálogo com o presidente Mo  
ontra o terrorismo Israel decide mancer diálogo com palestinos das  
OTO Papa conclama igreja a manter diálogo com 'instituições p  
açúcar. Nosso interesse é manter o maior diálogo possível. Folha -  
S\$ 5 bilhões. Incapaz de manter canais de diálogo e negociação com os  
e Washington e Londres em manter canais de diálogo com Pequim, princi  
A rainha sabe que precisa manter um bom diálogo com Diana para torn

This in turn would cause problems to someone translating it into English and relying on information available in current dictionaries. A more suitable term in the original text would be *conversar*, which in English might be *keep u conversation*, which would convey the meaning of classroom conversation and avoid the negative connotation associated with *manter um diálogo*.

It must be said, though, that by using *manter um diálogo* the author has captured the meaning of 'troubled communication' which seems to be one of the shades of meaning conveyed by the expression. In a sense, then, the translator was successful in evoking a

particular meaning that was associated with the sen-iantic prosody of the expression, but arguably this may have been overridden by the more pervasive prosody of negotiation under unfavorable conditions, which is inadequate in that context.

## VIII. DISCUSSION

The results indicate that semantic prosodies may vary across Portuguese and English. The present study therefore corroborates other contrastive studies which have identified discrepancies between seemingly equivalent items in different languages (e.g. Partington, 1998). In order to avoid inadequacies, the translator should have access to information on semantic prosodies in the target language, so that s/he could check whether the choices available to him are connotationally appropriate.

The present study also presents evidence that semantic prosody is not directly accessible by intuition. This is the case of 'manter um diálogo', discussed above. Other studies for the English language have indicated that semantic prosodies and semantic profiles found in corpora were unknown (Louw, 1993; Stubbs, 1996). Therefore, it is important that translators have access to corpus-based sen-iantic prosody information in their mother tongue as well.

Information on semantic prosody, valuable as it is, is not available in most reference materials (Berber Sardinha, 1999a). In addition to Cobuild and BBI, mentioned here, Berber Sardinha, 1999a has shown that neither Oxford Learner's Dictionary nor LTP Dictionary of Selected Collocations inform their readers of connotational restrictions. Similar findings are reported by Partington, 1998. The present study also shows that bilingual dictionaries do not seem to do any better in that respect either.

## IX. FINAL COMMENTS

This study has described the semantic prosody associated with a number of lexical items in Portuguese and English, focusing on issues pertaining to the translatability of the semantic prosody of those items. In general, the findings indicate important similarities and differences between the two languages, and point toward inadequacies in contemporary dictionaries. The study was based on the exploration of a Portuguese computer corpus of over 140 million words, one of the largest for Portuguese. The BNC, with 100 million words, was also used in the course of the investigation.

This study has limitations, such as the subjective choice of items selected for study, the lack of a tagged corpus (which would have made it more reliable to find verb collocates, for example), and the limit of 16 thousand occurrences imposed by the concordancer on a few very frequent items.

Further research could take up issues raised by the current study. One of these is the question of how a system for machine translation might incorporate semantic prosody information. Another issue is how to use machine learning techniques for automatically acquiring semantic prosodies from a corpus, thus enabling the researcher to describe a larger number of items than is possible by the methods used in this study. Finally, an issue in lexicography and dictionary design is to what extent it is possible to include semantic prosody information for all items in a dictionary, and if not what options are available – for instance, would a connotation dictionary displaying corpus-based semantic prosody only be a viable alternative on the dictionary market? If not, what would be a principled method for selection of which semantic prosodies to include?

The general conclusion of the study presented here is that information on connotation, especially that based on the exploration of corpora, should be part of glossaries and dictionaries, particularly production and bilingual dictionaries, which are tools that translators rely upon regularly.

## REFERENCES

- Baker, M., Francis, G., & Tognini-Bonelli, E. (Eds.). (1993). *Text and technology - In honour of John Sinclair*. Philadelphia/Amsterdam: John Benjamins.
- Berber Sardinha, A. P. (1999a). Computadores e corpora na produção de dicionários: O que isso tem a ver com o ensino de línguas? Paper presented at the Symposium on 'The English Language Learner and Dictionaries'. ENPULI (National University English Teachers Meeting), 21 July 1999. USP - University of São Paulo, Brazil. Available online at [www.tonvberber.f2s.com](http://www.tonvberber.f2s.com)
- Berber Sardinha, A. P. (1999b). Padrões lexicais e colocações do português. Paper presented at the Computational Processing of Portuguese Symposium, 9th InPLA, PUCSP - Catholic University of São Paulo, Brazil, May 1999. Available online at [www.corpus.f2s.com/tony.pdf](http://www.corpus.f2s.com/tony.pdf)
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics - Investigating language structure and use*. Cambridge: Cambridge University Press.
- Channell, J. (2000). Corpus-based analysis of evaluative lexis. In S. Hunston & G. Thompson (Eds.), *Evaluation in Text - Authorial Stance and the Construction of Discourse* (pp. 39-55). Oxford: Oxford University Press.
- Deignan, A. (1999a). Corpus-based research into metaphor. In L. Cameron & G. Low (Eds.), *Researching and Applying Metaphor* (pp. 203-220). Cambridge: Cambridge University Press.
- Deignan, A. (1999b). Linguistic metaphor and collocation in nonliterary corpus data. *Metaphor and Symbol*, 14(1), 19-36.
- Fox, G. (1998). Using corpus data in the classroom. In B. Tomlinson (Ed.), *Materials development in language teaching* @p. 25-43). Cambridge: Cambridge University Press.
- Francis, G., & Hunston, S. (1996). *Grammar Patterns 1: Verbs*. London: HarperCollins, COBUILD.
- Francis, G., & Hunston, S. (1998). *Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins, COBUILD.
- Houaiss, A. (1982). *Dicionário Inglês-Português (9a edição atualizada)*. Rio de Janeiro: Record.
- Hunston, S. (1995). A corpus study of some English verbs of attribution. *Functions of Language*, 2(2), 133-158.
- Hunston, S., & Francis, G. (2000). *Pattern Grammar - A corpus-driven approach to the lexical grammar of English*. Amsterdam/ Philadelphia: John Benjamins.



- Louw, B. (1993). Irony in the text or insincerity in the writer: the diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Ronelli (Eds.), *Text and technology - Essays in honor of John McH Sinclair*. Amsterdam/Atlanta, GA: John Benjamins.
- Marques, M. L. G. (1996). Os bordões da língua falada e variação. In M. F. B. do Nascimento, M. C. Rodrigues, & J. B. Gonçalves (Eds.), *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística* (Vol. I - Corpora, pp. 309-322). Lisboa: APL.
- Nascimento, M. F. B. do, & Carvalho, A. (1996). Preto e branco ou branco e preto? (Como se combinam os nomes das cores). In M. F. B. do Nascimento, M. C. Rodrigues, & J. B. Gonçalves (Eds.), *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística* (Vol. I - Corpora, pp. 367-380). Lisboa: APL.
- Neto, P. M. (1996). Combinações lexicais num corpus linguístico especializado. In M. F. B. do Nascimento, M. C. Rodrigues, & J. B. Gonçalves (Eds.), *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística* (Vol. I - Corpora, pp. 359-366). Lisboa: APL.
- Partington, A. (1998). *Studies in Corpus Linguistics 2. Patterns and Meanings - Using Corpora for English Language Research and Teaching*. Amsterdam/Philadelphia: John Benjamins.
- Sinclair, J. McH. (Ed.). (1987). *Looking up: An account of the COBUILD Project in lexical computing and the development of the Collins COBUILD English Language Dictionary*. London: Collins.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. McH. et al. (1995). *COBUILD English Dictionary* (New edition ed.). London: Collins COBUILD.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of trouble with quantitative studies. *Functions of Language*, 2(2), 23-56.
- Stubbs, M. (1996). *Text and Corpus Analysis -- Computer-Assisted Studies of Language and Culture*. Oxford: Blackwell.
- Tagg, S. E. O. (1999). Convencionalidade e lexicografia: Um dicionário bilingüe de coligações verbais. Seminar, PUCSP - Catholic University of São Paulo, Brazil, 14 June 1999.