Imperfeito: a broad-coverage study[*]

Diana Santos
INESC

1. Introduction

Imperfeito is an important tense in Portuguese, both because of its frequency and because of its undisputable aspectual import. Several researchers have emphasized different properties of Imperfeito. In this study, I set out to investigate actual uses of Imperfeito in real text, to assess the various proposals. It is my hope that the work presented here can be classified as *theoretically informed corpus analysis*, benefitting from theoretical insights and quantitative results alike.

This paper addresses two different problems: (i) the meaning of Imperfeito in itself, i.e., what could be concluded from the individual analyses that made up the study, and (ii) the problems posed by a feature-based classification method for the semantics of real text. After a brief description of the study in Section 2, I devote Section 3 to the questions of classification, and Section 4 to the linguistic results concerning Imperfeito proper. In Section 5, I discuss the general description of a tense.

2. The empirical study

The study had two parts, albeit only the first one will be described here:

(i) annotation of two texts in Portuguese with a set of features describing possible reasons for the use of Imperfeito

(ii) comparison of the corresponding English tenses

The two texts in question were one Portuguese original (Sena, 1984) and the translation of an English text (Steinbeck/Dionísio, 1977, Chapter 1).

All occurrences of Imperfeito were annotated[1]. The annotation described the meaning/use of the clause or sentence in Imperfeito. This was a crucial feature of the annotation, i.e., it was not meant to describe exclusively the import of Imperfeito. In Section 2.1 below, this choice will be motivated at length.

On the other hand, the use of two different texts was necessary in order to evaluate the generality of the conclusions. Furthermore, given that the two texts were idealized (planned, generated) in two different languages, I wanted to see whether this would give rise to significant differences. While obviously more data have to be consulted for a definitive conclusion, this was, I believe, a reasonable starting point.

2.1 Methodological and theoretical assumptions

I highlight here some properties of this work which in my opinion distinguish it from other studies, and explain the label "broad-coverage study".

1. I consider real text to be the only defensible input for language engineering, as opposed to the constructed sentences typical of theoretical linguistics. While much recent serious work in semantics has resorted to real text (see e.g. Sandström (1993), Caenepeel (1989)), it is still a

[1] With the exception of *ser que* constructions, corresponding to English it-clefts, e.g. *era o cabo quem falava*; *era Joana que se levantava em silêncio*. There is a perfect agreement between the tenses of the verb *ser* and the clefted clause, therefore taking these numbers into consideration would distort the quantitative distribution, in addition to not being clear what classification those occurrences should get.

selected, filtered, set of examples which constitutes the evidence for their investigation (as it has always been the case of traditional descriptive grammar, exemplified e.g. by Sten (1973)). In this work, I took a more radical path in not performing any filtering whatsoever.

In other words, instead of the best examples, or of the protypical instances of a category or use, I will consider **every** instance of Imperfeito, which is what any NLP system has to do.

2. It is customary, and has led to interesting insights and advances in linguistic theory as a whole, to divide aspectual import at least among the (verb) lexicon, the arguments and the temporal adverbials, in the computation of the final aspect of the sentence. Since Verkuyl (1972), aspectual calculi have tried to compositionally arrive at the final sentence aspect (according to the intuitions of their authors), departing from the aspectual import of their components. However, I have never seen a principled account of the analysis the other way around, i.e., given first the sentential aspect, "divide" or explain it then through sentence internal features. I think that this way of proceeding has specific advantages, first of all, because the aspect values from which it departs are more consensual. -- It is well known that a speaker has more intuitions about the meaning of a complete sentence than about its constituents. Therefore, this method[2] can more easily be applied to real text.

In addition, this is motivated because one of the most conspicuous properties of natural language in context is that several "components" cooperate to the same end. In other words, redundancy is a key property of an informative message.

2.2 Previous studies

The only studies using parallel[3] corpora for tense and aspect I am aware of are Slobin (1994) and my own, reported in Santos (1994).

In the latter work, whose goal was to look for statistical regularities for tense translation, I annotated all tense transfers (i.e., not only those involving Imperfeitos) in a considerably larger corpus, balanced between original Portuguese and original English text.

Apart from large tendencies regarding the use of tenses in the two languages and the translations among them, it soon became clear that such a broad description would be of little help regarding translation subtleties, though. Rather, Santos (1994)'s main contentions were against the hypothesis that tense alone could be helpful in the task to align translations at the clause level, given the high frequency of clause misalignment. (See the paper for a detailed presentation of this phenomenon). In fact, it became clear that the number of factors at stake required a closer look at the individual translations and uses of tenses, and this was one of the motives to carry out the study reported in the present paper.

If Santos (1994) is therefore too general, Slobin (1994), on the other hand, is too specific, since it only compared sentences describing 'motion events' in a corpus of English and Spanish novels and their corresponding translations, using Talmy's (1985) typology as frame of reference. Even though Slobin does not explicitly mention aspect anywhere in the paper, I believe that his findings are relevant to the more general problem of comparing the tense and aspect systems of the two languages (see Santos (subm.) for a comprehensive demonstration of such view).

Another relevant study in this connection is Leiria (1991), who studies the production of European Portuguese texts by non-native learners of the language. She compares the

---

[2] I am speaking here about human analyses of aspect, not about computational ones.
[3] 'Parallel' in the sense of being composed by texts and their translations. There is another use of the term 'parallel' which refers to original texts in two or more languages sharing e.g. size, genre, subject, etc.

correct (input) text to the ones produced by her students, as far as the use of Imperfeito and Perfeito is concerned. Her work is relevant to my concerns here for two reasons. She uses a quantitative analysis of corpora, and she presents a classification of verbs in context, not only a prototypical one[4]. While the specific details concerning language acquisition do not concern us here, Leiria (1991) poses an interesting question, namely, what is the relative weight of inherent (i.e., lexical) aspect versus context in the choice of the tenses. She concludes that verbs with clear inherent aspectual values lead non-native speakers to choose tenses which agree in aspectual values, while verbs with several aspectual profiles (or, in other words, with less marked lexical aspect) cause more difficulties to learners.

The main shortcoming of Leiria (1991) is the same of Santos (1994): merging several phenomena under the same label makes it difficult to draw non-trivial conclusions.

Finally, Dahl (1985) attempted to characterize the uses of tenses and aspects in a broad range of languages, through the collecting of answers by native informants to a specially designed questionnaire, and his methodology and conclusions were influential in sorting out of the results of this work. However, Dahl does not use real text, and, therefore, is unable to extract frequency or even plausibility of usage of the tenses that he studies. Moreover, he uses mainly data occurring in simple main clauses, and, therefore, many usages of tenses and aspects, that may be only used in narrative contexts, cannot be elicited that way.

2.3 The labels

After an initial systematization of the traditional descriptions of Imperfeito, inspired by descriptive Portuguese grammars (e.g., Cunha & Cintra, 1987) and based on Sten's (1973) extensive material and my own data, the following features were selected for the classification of the occurrences of Imperfeito in real text:

HAB: habituality, a regularity turned property. I have argued for the importance of this (generally acknowlegded[5]) value of Imperfeito elsewhere, see Santos (1993). One example is *Tibério bebia o sangue dos escravos*. It is important to note that I do not grant a separate status to indefinite frequency and habituality, proposed e.g. in Antona & Tsujii (1993).

GRAD: gradual situation, either homogeneously or making reference to a definite end. During the classification process, I realized that I used this feature to signal two different things: (i) inherent or described graduality de per se, e.g., *uma nuvem vermelha que se dissipava*, and (ii) occurrence not yet arrived at its goal[6], e.g., *uma saudade antecipada daquelas mãos que morriam*. All instances of GRAD are also marked EXT.

EXT: extended situation, whose endpoints do not matter. This feature, again, can signal either one single situation, or the successive repetition of extended situations without definite endpoints, cf. respectively *o imperador dormia* and *Quintílio Vero torcia as mãos*.

---

[4] *Se fizermos um levantamento dos verbos utilizados pelos diferentes autores para exemplificarem um dado tipo, verificamos que a escolha recai invariavelmente sobre verbos e situações que não apresentam dúvidas* [...] *quando se trata, como é aqui o caso, de classificar um conjunto de verbos não previamente seleccionados, muitos são os que apresentam propriedades menos bem definidas* (Leiria, 1991:172).

[5] See Cunha & Cintra (1987:450), Leiria (1991:178), Sten (1973:100), among others.

[6] This last case, incidently, is the main definition of Cunha & Cintra (1987:380): *considerá-la* [a acção expressa pelo verbo, DMS] *como concluída, isto é, observada no seu término; ou pode considerá-la como não concluída, ou seja, observada na sua duração, na sua repetição.*

PROP: essential property (in opposition to a contingent one). The importance of the opposition of this value with TS (below) for Portuguese was also stressed in Santos (1993). While it is uncontroversial that states (permanent or temporary) tend to be described by Imperfeito, Cunha & Cintra explicitly mention permanency as a key factor: *factos passados concebidos como contínuos ou permanentes* (1987:451, my emphasis). Examples are *sorriu dos termos que a inversão impunha*, *os deuses amavam-no*.

Additionally, PROP was used to signal the specification of Manner in Talmy's (1985) typology. For example, in *a voz uivava tanto, que deve ser verdade*, an extended action is described and a property of such action is being ascribed, while in *a roseira que o envolvia*, the description of the position is done solely through manner.

TS: temporary state, as opposed to permanent, characterized two different kinds of situations: statements unambiguously temporary, like *Como estava velho!*; and bare specifications of location, such as *tropeçaram nos rolos que estavam no chão*. One should note that the Imperfeito progressive in Portuguese is the prototypical example of a temporary state, and one could suggest it was in privative opposition to plain Imperfeito (but see section 5.1 for an alternative analysis).

IND: explicit marking of perspective. One of the uses of Imperfeito is to signal that things are being described through the view, perception and feelings of one character of the plot (or of the narrator[7]). In this connection, the phenomenon named in Portuguese 'discurso indirecto livre' (henceforth free indirect speech) is extremely relevant, cf. Cunha & Cintra (1987:452): *Relevância particular tem o Imperfeito do Indicativo no chamado Discurso Indirecto Livre, em que autor e personagem se confundem*. Sten also refers the use of Imperfeito to render *the contents of an observation, a dream, a vision* (Sten, 1973:102, my translation). Sandström claims that the concept of IND is pertinent for English as well: some sentences of narrative discourse *can be, and sometimes must be, interpreted as denoting the content of an act of perception, contemplation, or speech, on the part of a protagonist* (Sandström, 1993:178). Examples are: *Ou chamava-o, mais uma vez, para aconselhá-lo* [...] *?*, and [Kino] *afastou a manta do nariz* [...]. *Ao pé das silvas, dois galos cumprimentavam-se* [...]. In Section 4.1, I touch upon the matter of the differences between strict and free indirect speech.

COND: modal, conditional. This use of Imperfeito is related to the following construct "*se* <clause in Imperfeito do Conjuntivo>, <clause in Imperfeito>*" (for example, *se ele me pedisse, eu casava*), but, in many cases, the *se*-clause is only implicit (see also Oliveira (1986)). COND describes an explicit non-"real", hypothetical action. Only one instance of this use was found in the two texts, and, symptomatically, in direct speech: *se as abrisse* [...] *seria alguém* [...] *a quem eu dava a morte*.

PIT: "pitoresco". This is a clearly marked use of Imperfeito only found in planned writing, in connection with a specified definite temporal location. This use of Imperfeito (see Sten (1973:99ff)) is traditionally analysed as expressing vividness, but it may also be a perspectivation marker (see section 4.5). An example is *Momentos depois,* [...] *assomava à porta da biblioteca*.

---

[7] Of course, the question of perspective is a complex one. See Caenepeel (1989) for a discussion of personalised narrators; narrators situated inside or outside the fictional world; internal or external narrative perspective; directly presented or represented perspective, and the like.

PLUR: this marker corresponds to extension in space rather than extension in time. The name originated in the fact that plural events are always distributed in space even though temporally they may overlap. Contrarily to the other features, this one was considered to be required during the analysis of the second text, motivated by examples like *Os da frente passavam a palavra aos de trás* and *chegaram ao sítio onde as cabanas se acabavam e a cidade de pedra e cimento começava.* This may be related to the fact that, in some sentences with plural participants, there is an event plurality distinct from the number of participants (see Krifka (1990)). PLUR is meant to indicate a distributed event (or series of events), with a vague plurality.

Several other uses were not considered, such as **politeness**, because this is a feature related to interaction and, since the texts considered had relatively few direct speech, it would not be a relevant feature for most ocurrences, and **future of the past**, which I believe to be a misconception: the examples invoked to illustrate this use are better placed under a perspective analysis, since they are pure instances of indirect speech, cf. Camara (1964) and Sten (1973:102-3). Also, the use of Imperfeito in *se*-clauses, amply discussed by Sten (1973:104ff), was not considered, since the only occurrence found was already clearly IND and PROP, not requiring a special label: *e que viria ele, na verdade, ali fazer, se o tempo não lhe chegava para tratar as pessoas ricas nas casas de pedra e de cimento da cidade?*

On the other hand, one of the most invoked parameters in the description of Imperfeito is its use for **background** instead of foreground (BACK was therefore originally included in the classification set). However, as soon as the classification work began, I realized that this label was too difficult to assign, possibly indicating that the notion of background is simply a consequence of the kind of prototypical uses of Imperfeito, as noted e.g. in Fleischman (1985). Even though one should study the correlation between syntactic backgrounding and Imperfeito, this does not require human annotation, and therefore this label was excluded from the classification set.

Finally, some functions of Imperfeito are traditionally described in connection with particular discourse structures: for example, Cunha & Cintra refer the use of Imperfeito *para indicar, entre acções simultâneas, a que se estava processando quando sobreveio a outra* (Cunha & Cintra, 1987:450). This is obviously not a property of Imperfeito clauses, but of a particular combination of such clauses with other grammatical devices (tenses and connectives). Incidentally, I believe this particular effect is unproblematically subsumable under EXT, and all *quando*-clauses in Imperfeito were classified thus.

After the whole study, I realized that the most obvious feature had been forgotten: namely, the notion of **past**. Going through all instances once again, I could find, in Text 1, 18 occurrences where pastness seemed relevant, all in direct speech, except for *Como ele gostava de Salerno!*, which is indirect speech, but recognizably past due to the context. In Text 2, no occurrences clearly conveyed past, except maybe the unique occurrence of Imperfeito in direct speech, namely *Aquilo -- dizia ele, é que era civilização*. This gave me an a posteriori justification not to have included a feature PAST in the classification set, namely, the same as for politeness: it was only pertinent to direct speech.

3. The classification

In this section, I describe from the point of view of the classification process the attribution of one or several causes for the use of Imperfeito in the two texts.

Looking at the labels chosen, one easily remarks that they do not belong to the same linguistic levels of description. (The same happens in any characterization of Imperfeito or, for this matter, of French Imparfait.)

In fact, (i) EXT, HAB, GRAD or PLUR concern features of the occurrence which is being described, albeit seen through the eyes of a narrator, (ii) in addition to that, PROP and TS may also be considered to concern the evaluation of a situation by the narrator as essential or temporary, (iii) COND and IND relate to explicit cognitive features of the text producer, respectively, the explicit consideration of non-actual situations, and the explicit introduction of another (i.e., different from the narrator) subject's consciousness, and finally, (iv) PIT marks the explicit use of a formal device regarding text elaboration.

This leads to the important conclusion that one should expect that several distinct motives for using Imperfeito should co-occur in a sentence. Furthermore, if one does not grant an a priori priority to any such level, there is no guarantee that in every instance any one label can always be regarded as primary.

However, even when facing classifications at the same level, where the reader would think that only one label should be assigned, in fact, more often than not more than one label was employed. One of the first conclusions of this annotation task was therefore that a clear understanding of clusters of features was at least as important as the analysis of cases classified by a single label.

For a discussion of the problems involved and specific examples, I direct the reader to Santos (subm.2). Here, I will only present the results obtained, which are of two kinds: those concerning the distribution and frequency of the several cases (immediately below) and those related to the linguistic properties elicited, in Section 4.

Table 1 displays the general quantitative results, concerning 298 clauses in Imperfeito. The numbers in parentheses correspond to exact classification, irrespective of other levels (see footnote 8 for clarification).

The first relevant figure is that only 64 occurrences[8] (21%) received a single classification, demonstrating unambiguously the complexity of the language system. When more than one label was assigned, cases of ambiguity were much less frequent than cases of undecidedness or reinforcement: only 12 cases[9] (4,03%) were considered ambiguous. However, I should note that in theory no ambiguous cases should pop up. But it was already convincingly argued in Sampson (1987) that it is a myth that a human reader can always select only one interpretation (see also Santos (subm.)).

annotation	Text1	Text2	Total
HABitual	24 (4)	41 (21)	65 (25)
HAB+PROP	15 (6)	7 (5)	22 (11)
HABvPROP		5	
(HAB+PROP)+EXT	7	2	9

[8] This figure is not directly obtainable from any of the tables, given that the numbers in parentheses in Table 1 correspond to exact classification in that level, i.e., for example the 4 HAB only instances in Text 1 may co-occur with IND or PIT, from another level. Only PIT, IND and COND only numbers correspond to the only annotation of the occurrence of Imperfeito.

[9] This number was calculated thus: 4 HABvEXT + 5 HABvPROP + 2 HABvEXT+PROP + 1 EXTvPROP.

HAB+EXT		10 (5)	
HABvEXT	13 (4)		
(HABvEXT)+PROP	2		
EXTended action	48 (15)	63 (30)	111 (45)
EXT+PROP	26 (15)	7	33 (22)
EXTvPROP		1	
GRADual action (+EXT)	7	9	16
PLUR	6 (3)	25 (10)	31 (13)
PLUR+HAB	2	4	6
PLUR+EXT	2	11	13
PROPerty	74 (44)	63 (43)	137 (87)
TStemporary state	15 (13)	19	34 (32)
CONDitional	1		
INDirect speech	18 (2)	32 (2)	50 (4)
PIToresco	6 (2)	5 (1)	11 (3)
Imperfeitos	130	168	298
Tensed clauses (approx.)	596	535	1131
Words	4410	3051	7461

Table 1

As far as joint classifications across levels are concerned, I present in Table 2 the distribution regarding the first level, namely, the one concerning the features EXT, GRAD, PLUR and HAB. Taking all levels into account, 200 occurrences[10] (67%) have (any number of) labels of only one level.

	Text1	Text2	Total
Level 1 only	37	63	100
Level 1 plus PROP or TS	32	20	52
Level 1 plus IND	4	13	17
Level 1 plus PIT	4	5	9
Total	66	103	169

Table 2

As far as the same level is concerned, again in Table 3 is displayed the distribution in level 1, which in fact exhausts all possibilities, given that in the other levels no co-occurrences exist. This gives a total of 51 occurrences out of 169 which have more than one label, corresponding to 30%.

	Text1	Text2	Total
Only one label	43	75	118
Two labels	22	22	44
Three labels	1	6	7
Total	66	103	169

Table 3

---

[10] This number was calculated thus: 100 level 1-only + 1 COND + 4 IND only + 3 PIT only + the number of TS only and PROP only instances taking out those marked IND or PIT, respectively 28 and 64.

4. Properties of Imperfeito

I turn now to the most relevant findings of this study in terms of the specification of the uses and of the systematic connections between labels.

4.1 Perspective and Imperfeito

The number of instances marked IND clearly shows that Imperfeito is used to signal that the sentences in question are to be read as "perspectivized", independently of (or in addition to) the properties of the situation described. In order to keep Table 1 relatively simple, no complex cooccurrence labels were considered involving IND or PIT. It is striking, however, that, in Text 1, out of 18 IND, 5 cases are marked TS and 8 PROP, making thus up 72%. (In Text 2, of 32 cases marked IND, 20 are PROP, and one TS, corresponding thus to 66%). There is therefore a strong correlation between IND and stativeness, which is undoubtedly related to the claim made by Caenepeel (1989), in a dissertation devoted to the connections between aspect and perspective, that every stative sentence conveys some perspective.

The three IND cases which are marked GRAD-EXT correspond to strict indirect speech. This remark seems relevant, since it is very probable that strict indirect speech and free indirect speech have different properties. In fact, while free indirect speech is a device to which Imperfeito is tightly connected, as mentioned above, in strict indirect speech, on the other hand, Imperfeito only has the role of rendering present tensed direct speech.

One could thus claim that the properties of Imperfeito in strict indirect speech are simply the reflection of those of Presente, and that they do not concern us here. However, I do not agree with such a view, since there are many similarities between Imperfeito and Presente. Rather than a coincidence, this is a richness of the language, because Imperfeito tends to convey, while reporting Presente, the same features that it would transmit alone. One should however expect an imperfect match, in that there are cases (like the one above) where the use of strict indirect speech enlarges the normal range of application of Imperfeito. Incidentally, this is exactly the case of the "future of the past" use (cf. 2.3. above), in my view an instance of strict indirect speech rendering the futurate use of Presente.

4.2 Properties described as habits

The combination of the features HAB and PROP seems to be significant in Portuguese, signalling cases which could be conveying an habitual action or an attitude depicted by such an action. Examples are: *mas tu rias deles*; *queixava-se, sim, das dores que o não largavam*; *do entusiasmo com que discutias, pela noite dentro*; *que lições não dava ele às crianças*.

In addition, the HAB+PROP classification co-occurs with the label EXT in the following instances: *o grande Tibério, que lembrava com saudade*; *não sabia o que fazia*; *uma noite cerrada em que nada se via*; *até ele, [...] conspirava*; *Cláudio, que tremia de ser imperador*; *queixava-se, sim, das dores que o não largavam*. The first three cases can be considered special in that the verb itself has relevant aspectual properties, and has long been identified as such (see e.g. Santos (1991a,b), where I called such verbs **acquisitions**). For example, the first sentence describes a mental property ('remember') that could go on permanently, or

only 'occur' when people consciously directed their attention
to it.

Given the definition I gave of HAB in the first place, one
may argue that PROP is an integral part of HAB, or, at least,
one connotation of it. However, it was clear that HAB-only
instances were a distint case from HAB+PROP. Examples of cases
marked HAB alone are: *numa das conspirações que descobria
todos os dias*; *crucificava todos os dias o Senhor em mim
mesmo*; *dos míseros reais que raramente dava como esmola*; *vinha
sempre bater à porta humildemente*. Not only the objective
tests for the two cases were found to be different enough (see
section 4.2.7.1 below), it was also striking to see that
HAB+PROP cases in Text 2 never convey the attitude described
by the habit. They either mark mental verbs, like *sabia*,
conhecia, *lembrava*, *sentia* (acquisitions, as noted above); or,
in the remaining three cases, forming a single unit, they
represent the thoughts of a character about another (and are
thus also IND), and a comparison is explicitly stated,
therefore suggesting that we face a situation where PROP and
HAB are independently motivated, and thus do not constitute a
HAB+PROP case.

This absence of "true" HAB+PROP in Text 2 seems to
indicate that English does not have any category relating to
the Portuguese complex category HAB+PROP. This is additionally
corroborated by the fact that, in the translation of Text 1,
out of 13 HAB+PROP cases, 11 were rendered by past simple,
while the two remaining cases had specific unrelated reasons
to be translated with the help of modals. Given that past
simple is the unmarked, most common past tense in English, and
moreover the sentences had no distinguishing properties, no
formal distinction could be found.

4.3 Habituality and perspective

One interesting correlation is that most, if not all, instances marked HAB (or
HAB+PROP) occur either in direct speech or in free indirect speech, in other words, they
presuppose a human evaluation. Incidently, the same is also true of most PROP only (non-
Manner) instances: 31 out of 36 PROPs in Text 1 were found in direct or indirect speech.

This may indicate that HAB does not mainly reflect a temporal pattern, but a belief, an
opinion, a viewpoint, and ought to be grouped with IND and COND (or, alternatively, with
PROP and TS). Since it is known that English does not systematically distinguish between
permanent and temporary states, while it does have formal markers for conditional uses (e.g.,
the modal *would*) and for perspective (for example, the progressive), the comparison of
Portuguese with English indicates that we should group HAB with PROP and TS. Thus one
could distinguish Portuguese, with formal means of expressing a threefold stative character,
from English, without such expressive means on the state realm. (That English and
Portuguese differ in their partitioning of semantic domains and in the relative importance they

give to different sides of a situation is, I believe, uncontroversial. See Santos (subm.) for more differences between the two languages).

4.4 Manner and independence of tense

Whenever manner of actions was specified, the sentence received the classification EXT+PROP. In this context, it was clear that the label PROP did not come from tense. In every instance (underlined), it was either present lexically in the verb or in an adjunct: *No clarão indistinto que difuso vinha*; [<utterance>] -- *e sorria.*; *a voz uivava tanto, que deve ser verdade*; *teve [...] um frio que lhe lambia as pernas depiladas*; *seguindo Átis que corria*.

It is well known (see Talmy (1985), Slobin (1994)) that English gives more importance to manner than languages like Portuguese, which are verb-framed rather than satellite-framed. It is therefore not surprising that tense (a grammatical device) is not related to manner in any way in Portuguese (while, in English, the contrary may be the case).

4.5 Pitoresco: some remarks

The Pitoresco usage of Imperfeito is always specified by a definite (often punctual) temporal specification (see Sten (1973:99f), Kamp & Rohrer (1983:256) about French Imparfait). Semelfactivity (i.e., once-only occurrence) is expressed through that temporal specification, and therefore the Imperfeito clause refers to a unique situation in progress. If the situation in itself is normally seen as extended, and therefore marked EXT, no special effect is conveyed. However, if the situation is normally conceived as punctual, then the use of Imperfeito brings about a special colour. So, even though one could argue that `PIT is just a stylistic device of the basic semantics of the imperfective, there is a striking objective difference between an EXT case and marking something that is not extended as extended.`

I was able to identify two slightly different grounds for PIT in the texts studied, namely, vividness in the description, and the signalling of perspective, exemplified respectively by *Momentos depois, [...] assomava à porta da biblioteca*; *E, de repente, saía-se com a mais inesperada das ideias* and *o imperador, que [...] agora fingia dormir*; *Kino baixou os olhos. Agora a aurora alastrava rapidamente*. In the two last sentences, the use of *agora* unequivocally signals perspective from the characters' side (it is not deictic in that it does not correspond to the moments of reading or writing, but to the 'now' of the narrative)[11].

Even though Pitoresco is generally referred as a (Portuguese-specific?) style device, it appears in the two texts. This, I believe, demonstrates I was right in using it as a label.

4.6 Global comparison of the two texts

There are three global differences between the two texts: different original language, different author[12], different genre (short story versus first chapter of a novel). I assume that the last difference is not particularly relevant here, given that the short story is also divided in three parts and can be looked upon as a small novel. Furthermore, while a larger text is liable to give more attention to detail, more descriptions of the characters, of the background and of their thoughts, and contain more episodes whose aim is to illustrate relevant features

---

[11] It is interesting to see that Mateus Silva, analysing a similar example, namely, *O barco entrava agora definitivamente no mar alto*, proposes that this Imperfeito sentence be interpreted as durative present, *entrava = está a entrar* (Mateus Silva, 1992:497).

[12] This is a consequence of the first difference, since, to my knowledge, there are no famous writers bilingual and with comparable works in the two languages.

(originating parallel episodic structures, according to Sandström (1993)), the English text chosen corresponds to a relatively straightforward narrative, making little use of these ingredients. In fact, Steinbeck's realist style makes the book simple in narrative structure. In turn, Sena's writing style in the short story also follows a relatively simple narrative structure, and the result is that the styles of the two authors are not too dissimilar.

In Table 4, I present a quantitative abstract of the main categories. However, it should be borne in mind that there is no claim to the statistical relevance of the differences[13], given that with two texts only it is not possible to estimate a statistical distribution.

Still, due to the fact that a novel has more narrative density than a short story, it was foreseeable that there are more HAB and IND cases in Text 2 than in Text 1. Also, and since this seems to be a particular richness of Portuguese Imperfeito, HAB+PROP is much rarer in Text 2 than in Text 1. On the other hand, the apparently more frequent use of HAB in Portuguese text coming from English is explained, in my view, by the introduction (by the translation) of a vagueness not present in the English text (or at least not so conspicuously): in fact, since habituality is not marked by tense in English, all EXT-HAB instances are interpretable as simply EXT in Text 2.

annotation	Text 1	Text 2
HABitual	24 (18%)	41 (24%)
EXTended	48 (37%)	63 (37%)
PROPerty	74 (57%)	63 (37%)
TStemporary state	15 (12%)	19 (11%)
PLURal	6 (4,6%)	25 (15%)
INDirect speech	18 (14%)	32 (19%)

Table 4

One conspicuous difference between the two texts is the kind of perspective itself. While in Text 1 it is mainly reported thoughts, and therefore conceivable as internal speech, in Text 2 there are two salient kinds of perspective: (i) either simple perspective of perception, i.e., a stylistic "cinematographic" device of presenting the story through the eyes or senses of one character; or (ii) the presentation of the words of the narrator explaining the character's feelings and opinions. (I chose to signal only the second case as IND.) The difference between IND in Text 1 and Text 2 (ii) has obviously to do with the properties of the main characters of the two texts: an intelectual and analytic mind in his mature and old ages (Marcus Sempronius) versus a non-literate young man (Kino): Therefore, if IND in Text 1 is closely related to free indirect speech, in few cases does the reader of Text 2 have the impression that the things perspectivized would be verbalizable by the character himself, rather, they belong to the omniscient narrator that explains "his" characters[14].

However, the difference (not quantified) between IND in Text 1 and (i) is, in my view, to be ascribed to the language difference, namely, the relative preference of Portuguese to

---

[13] I thank Isabel Hub Faria for raising this point after the oral presentation of the paper.

[14] This is even explicitly acknowledged by the narrator: *and if he had been able to speak it*, referring to Kino.

express opinions rather than facts, as opposed to English (see again Santos (subm.) for a defence of this claim). This can be relevant for a typology of narrative, cf. Slobin's remark *typologies of grammar have consequences for "typologies of rhetoric."* (1994:15).

5. Discussion: the description of a tense

Is it possible to assign Imperfeito a general unified meaning?

The relationship between the various "aspects" of meaning has been much discussed in the literature, and one approach that is often followed is to divide the meaning of one linguistic construct into basic and secondary (or derived) meanings. What counts as basic is in general what remains valid across contexts / secondary meanings. However, if, theoretically, such descriptions can be appealing, in reality it is exactly the difficulty at getting at one operational common meaning that gave rise to such fragmentation (between basic and secondary meanings) in the first place. Let us look at the following plausible candidates to encompass all uses of Imperfeito:

Remoteness, bringing together the notions of modality (COND, PROP, and possibly HAB), distance from present (past), distance between the speakers (politeness), distance through mediation by someone else's consciousness (IND).

Extendedness, encompassing something that remains, that takes time (EXT), that is permanent (PROP), that is in progress (GRAD), that has vague spatial contours (PLUR), that is recurrent through time (HAB).

One possible description of the meaning of Imperfeito would thus be remoteness in the modal axis and extendedness in the temporal axis. Interesting and enlightening this formulation may be, for practical text understanding there would not be much advantage in it: In any case (of a real occurrence of Imperfeito), to understand it, one would still have to answer the following question(s): remote regarding which direction? extended in what way? which brings us back (at least approximately) to the labels used above[15].

Another relevant observation has been produced in the work by Sweetser (1990), namely the identification of a *persistent parallelism between formal markers of aspects of content, aspects of the speaker's reasoning, and aspects of the current speech act* (1989:21). This gives us the key to understand the perspective use of Imperfeito, which clearly belongs to the epistemic, rather than to the content, level. At the epistemic level, the subject matter is precisely the thoughts and reasoning processes of the actors in the narrative. I think that Imperfeito signals precisely epistemic mode.[16]

Finally, and to satisfactorily explain PIT, I have to note that there are two complementary ways to handle natural language semantics. One is to look at the (common) properties of the situations described (reference) and study referential properties; the other is to look at the language system and see in what various relations do the linguistic devices enter. It is customary to divide these relations in syntagmatic and paradigmatic ones, the first

---

[15] Interestingly, Taylor (1989) argues precisely the same way about the English about the English past tense, rejecting a view of remoteness as common core to past reference, counterfactuality and politeness. Rather, he uses the notion of a polysemous prototypical category, described in terms of family resemblances, *a polysemous category whose various meanings are linked, some through metaphor, some through metonymy, to a central sense* (Taylor, 1989:147). He analyses past reference as the central sense of past tense; counterfactuality as metonymy (conventionalization of the implicature that past often implies that something no longer holds); and politeness as metaphor: time as space, involvement as distance/proximity in space.

[16] Of Of course, in the cases where the thoughts of a character describe something which in the real-world would also be described by Imperfeito, those occurrences are pragmatically ambiguous (i.e., about which level they relate to).

regarding relationship with other categories, the second concerning relations to elements in the same category.

Formal semantics traditionally uses referential properties, e.g. habituality (in English) has been defined thus: *habituals simply posit the existence of some set of events of the specified kind* (Ramsay, 1992:229).

On the other hand, most linguistic tests to elicit semantic categories play on co-occurrence with other (linguistic expressions denoting other) semantic categories, and are thus syntagmatic in nature (cf. the ubiquitous co-occurrence tests with specific adverbials to decide on aspectual properties).

Nevertheless, paradigmatic relations are very important as well, even though they rely less on objective observation and more on semantic competence from the part of the analyst. Typical cases are tests looking for sameness/difference of meaning by employing another tense, or order change between clauses, cf. Sandström (1993:5): *A technique I have employed here and there is to manipulate authentic examples in various ways to see in what way interpretation is affected by changes in the surface form.*

```
Even though for the "more referential" labels (those I
dubbed of level 1, e.g. EXT or GRAD) a referential
characterization will work and a paradigmatic one will
possibly be superfluous, for labels such as PIT, which have no
referential import, an intra-linguistic comparison of
alternative ways of expression becomes essential. I will in
fact give an exclusively paradigmatic account of PIT: the
Pitoresco use of Imperfeito contrasts with a normal use of
Perfeito in the description of a telic punctual and
semelfactive event, producing a closeness effect (as if the
reader is brought to the middle of the plot).

For a throrough description of the paradigmatic relations
Imperfeito establishes with other Portuguese tenses, see
Santos (in prep.).
```

6. Objective classification tests

```
A necessary condition for the usefulness of a study like
this is its repeatibility by other researchers, as well as its
applicability to a range of new other texts. Therefore, a
(relatively) objective set of tests has to be provided to make
these two goals attainable.
```

Prior to presenting a set of classificatory devices to decide on each label, though, I should state that these tests are an after-product of the actual classifications, i.e., first, I employed my native competence and my judgement of the semantic properties during the classification process, and then later I tried to come up with a set of tests that would allow me (and anyone else) to arrive at roughly the same labels. It should therefore be emphasized that, the correctness of the tests to be presented is independent of that of the above results.

```
Again, recall that the tests are supposed to apply to
clauses in Imperfeito:
```

If the sentence makes reference to a goal: GRAD+EXT; if it denotes an extended situation: EXT; if it expresses a property (or attitude): PROP. If that property (or attitude) is observable through repeated action: HAB+PROP. If the sentence expresses a set of events, or

one event with more than one participant (and the participants are not synchronized): PLUR. If the tense cooccurs with frequency adverbs or corresponding nominal frequency determiners: HAB. If it can be replaced by Perfeito: PIT. If it can be replaced by the progressive, then it is not HAB, and should be marked EXT. If the context is not felicitous with Perfeito, COND. Finally, if the only change in the replacement by Perfeito is that simultaneity gives way to sequence, it should be marked EXT.

7. Imperfeito and discourse

The last criterion suggested raises the question of the relationship of the use of this tense and discourse progression or structure.

When the situations described are atelic, extended, and semelfactive, the oppositions effected by Imperfeito versus Perfeito cannot be referential in nature (Bache (1982:68f) already identified durative, atelic and non-stative situations as the only ones which allowed freedom of aspect choice). If, on top of that, no modal operation is credible, the distinction can be pragmatically used for a totally different purpose. In fact, I believe it is used to convey temporal information among sentences. That this is pragmatic and not semantic is a matter settled by the possibility of explicitly cancelling the presupposed temporal order, as is the case of English in "they married and had children although in the opposite order", cf. Comrie (1985:27). An example of cancelling the simultaneity value between two actions described in Imperfeito could thus be *Ele cantava e ela corria, mas nunca ao mesmo tempo*.

8. Conclusion

While a formalizaton of the data presented here is under way, I found it important to have a broad description of Imperfeito in real text, as unbiased as possible by a priori theoretical requirements. With the study presented in this paper, I gathered a significant number of hypotheses and intuitions about the meanings of Portuguese sentences, which could not be gotten if I had relied only in my introspection. I formulated as well some ideas on concrete differences between English and Portuguese, that directed Part 2 of the study.

Among the most important conclusions of this work, I include the verification in real text of the conspicuous use of Imperfeito for signalling perspective, the Portuguese tendency to mingle habits and properties through that same tense, and the capacity of the language to convey several distinct values with the same formal mark. Furthermore, several distinctions not usually in the literature were found to be relevant to the classification task, which allowed me to refine the traditional labels and to assess roughly their frequency of occurrence.

References

Antona, Margherita & Jun-ichi Tsujii. "Treatment of tense and Aspect in Translation from Italian to Greek - An Example of Treatment of Implicit Information in Knowledge-based Transfer MT", *Proceedings of tTMI'93* (Kyoto, July 14-16, 1993), 131-53.

Bache, Carl. "Aspect and Aktionsart: towards a semantic distinction", *J. Linguistics* 18 (1982), 57-72.

Caenepeel, Mimo. "Aspect, Temporal Ordering and Perspective in Narrative Fiction", Phd dissertation, University of Edinburgh, 1989.

Camara Jr., Joaquim Mattoso. "Une catégorie verbale: le futur du passé", in Lund, H.G. (ed.), *Proceedings of the Ninth International Congress of Linguists,* Haag, 1964, 547-51.

Comrie, Bernard. *Tense,* Cambridge University Press, 1985.

Cunha, Celso & Lindley Cintra. *Nova Gramática do Português Contemporâneo,* Edições João Sá da Costa, Lisboa, 1987.

Fleischman, Suzanne. "Discourse functions of tense-aspect oppositions in narrative: toward a theory of grounding", *Linguistics* 23 (1985), 851-82.

Kamp, Hans & Christian Rohrer, "Tense in Texts", in Bäuerle, R., Schwarze, C. & von Stechow, A. (eds.), *Meaning, Use and Interpretation of Language,* Berlin: de Gruyter, 1983, 251-269.

Krifka, Manfred. "Four Thousand Ships Passed Through the Lock: Object-Induced Measure Functions on Events", *Linguistics and Philosophy* 13, 1990, 487-520.

Leech, Geoffrey. *Semantics: The Study of Meaning,* Penguin Books, second edition, 1981.

Leiria, Isabel Maria Caetano. "A aquisição por falantes de Português-europeu língua não-materna dos aspectos verbais expressos pelos Pretéritos Perfeito e Imperfeito", Tese de Mestrado, Faculdade de Letras de Lisboa, 1991.

Mateus Silva, Helena. "Retroacção aspecto/tempo na textualidade", *Actas do VIII Encontro da Associação Portuguesa de Linguística* (Lisboa, 1-3 Outubro de 1992), 488-99.

Oliveira, Fátima. "Algumas considerações acerca do P. Imperfeito", *Actas do 2.o Encontro da Associação Portuguesa de Linguística* (Lisboa, 1-3 Outubro 1986), 78-95.

Ramsay, Allan. "Generic NPs and Habitual VPs", *Proceedings of COLING-92* (Nantes, 23-28/7/92), Vol I, 226-31.

Sampson, Geoffrey. "MT: A nonconformist's view of the state of the art", Margaret King (ed.), *Machine Translation Today: The State of the Art (Proceedings of the Third Lugano Tutorial, 2-7 April 1984),* Edinburgh University Press, 1987, 91-108.

Sandström, Görel. "*When*-clauses and the temporal interpretation of narrative discourse", PhD dissertation, Department of General Linguistics, University of Umeå, Report nr. 34, DGL-UUM-R-34, May 1993.

Santos, Diana. "Para uma classificação aspectual dos verbos portugueses", *Actas do VII Encontro da Associação Portuguesa de Linguística* (Lisboa, 7-8 October 1991), 389-401.

Santos, Diana. "Integrating tense, aspect and genericity", *Actas do IX Encontro da Associação Portuguesa de Linguística* (Coimbra, 29 September - 1st October 1993), 391-405.

Santos, Diana. "Bilingual alignment and tense", *Proceedings of the Second Annual Workshop on Very Large Corpora* (Kyoto, August 4th, 1994), extended version as INESC Report AR/10-94.

Santos, Diana. "Lexical mismatches and tense and aspect", August 1994, submitted.

Santos, Diana. "Classification of Imperfeito in real text", October 1994, submitted.

Santos, Diana. "Tense and Aspect in English and Portuguese: a contrastive semantical study", Phd dissertation, in preparation.

Sena, Jorge de. *Antigas e Novas Andanças do Demónio*, Edições 70, 5ª edição, 1984.

Slobin, Dan I. "Two Ways to Travel: Verbs of Motion in English and Spanish", in M.Shibatani & S. A. Thompson (eds.), *Essays in Semantics,* Oxford University Press, 1994.

Steinbeck, John. *The Pearl*, Bantam Books, 1975. Dionísio, Mário (translator). *A pérola*, Publicações Europa-América, 1977.

Sten, Holger. *L'emploi des temps en portugais moderne,* Det Kongelige Danske Videnskabernes Selskab, Historisk-filosofiske Meddelelser 46, 1, Munskgaard: København 1973.

Sweetser, Eve E. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*, Cambridge University Press, 1990.

Talmy, Leonard. "Lexicalization patterns: semantic structure in lexical forms", in T. Shopen (ed.), *Language typology and semantic description, vol.3: Grammatical categories and the lexicon,* Cambridge University Press, 1957, 57-149.

Taylor, John R. *Linguistic categorization: Prototypes in Linguistic Theory*, Clarendon Press, Oxford, 1989.

Verkuyl, Henk. *On the Compositional Nature of the Aspects*, Dordrecht: D. Reidel, 1972.