Diana Santos
INESC

Sobre a classificação aspectual dos verbos portugueses

This is the English version of a paper presented at 7º Encontro da Associação Portuguesa de Linguística, 7th-8th October 1991, Lisbon, that will appear in its Proceedings soon. Everything that was added because of the English rendering will appear in this font.

It should be emphasized that any claims in this paper concern simply Portuguese, and not the English translations presented as a reading aid. Most often than not, the English words that are the best translations of the Portuguese verbs presented here have a different aspectual profile.

1 The problem

It is well known that different verbs present, while subject to the same morpho-syntactical mechanisms, different behavior (and different senses).

The attempt to establish a verbal classification that allows identification of those differences in what concerns the tense and aspect system of a language has already a considerable history.

This task should answer satisfactorily to the three following points:

• To furnish a satisfactory interpretation of the proposed classes (in order to prevent an ad-hoc sistematization). That is, to give a semantic characterization of the aspectual classes.
To present objective criteria that allow the establishment (at least for the great majority of the verbs of one language) of whataspectual class a verb belongs.

To explain and formalize the interrelationship between the proposed classification and the grammatical mechanisms associated with tense and aspect. That means that verbal classes should not be separated from a global theoretical environment.

This paper proposes an aspectual classification of the Portuguese verbs, tackling explicitly the two first points. The third, i.e., an unified theory of tense and aspect, will not be described, for shortage of space.

2 Previous considerations

The first point to clarify concerns the object of the classification, which is the VERB.

Many were the authors who proposed other linguistic entities as the object of an aspectual classification: verb phrases or sentences without tense (tenseless sentences, propositions, radicals) ([Vlach 81, Moens 87, Galton 84]). For us, the verb is entitled to a classification by itself, independently of the fact, known by everyone, that by no means is the verb the only constituent with aspectual import, nor the aspect of an expression containing a verb automatically inherits its aspectual character.

In short, we are interested in a lexical classification. The fact that other words (e.g. nouns or adjectives) can also be classified in terms of aspect (see [Verkuyl 72] or [Krifka 91]) is a consideration which is orthogonal to the subject of this paper, and should not be confused with the above argument.

Regarding terminology, the Vendlerian names ([Vendler 67], see also [Mourelatos 81] for an overview), that have been used by the vast majority of the scholars in this field, have been more a source of confusion than of clarification, since to the same names different concepts have been attributed. (For a detailed justification of this statement, see [Verkuyl 89].) This implies, for instance, that the discriminating criteria vary from author to author, as well as their explanation.

We will be using in this paper an original terminology.

The presence of concepts such as causality and agentivity has been invariably permanent in verbal aspect theories ([Dowty 79, Moens 87, Verkuyl 89], etc.). As clear examples, let
us point out the existence or absence of result to distinguish events from states and activities; and intentionality (or existence of agent) to draw the line between states and activities.

Obviously, we are not denying the strong relevance of result or causation considerations to the analysis of natural language. We only think (as, in fact, other authors — [Pustejowsky/Bergler 87]) that it may be useful to separate the two domains, both from a methodological point of view and because the distinction “Has this action a result?” is extremely artificial: just look at the pairs lembra/esquecer, construir/destruir, escrever/ler, crescer/manter-se, (remember/forget, write/read, grow/maintain), for example.

Thus, our proposal attempts to classify the verbs following temporal-only criteria, such as temporal localization, kind of localization, and, in case it is an interval, whether that has privileged points.

Finally, before presenting the actual work, we present our perspective on natural language processing, that motivated this investigation.

• At first, we consider that any serious study in linguistics must concern language as a whole, and not base itself in a carefully selected choice of examples. This automatically implies, for the matter in consideration, that any verbal subcategorization should make sense to any verb, no matter how little “interesting” this verb may be. On the other hand, we do not content ourselves which give up on the most difficult cases (and, therefore, most illuminating for the proposal of any theory) under the pretext that they are statistically irrelevant.

• Moreover, we think it to be fundamental that a theory can be verified by an average native speaker, and not only by the theory+’s creator. That is, one given hypothesis on the language should be objectively measurable, and not depend on more or less developed intuitions.

• From the start, it should be mentioned that we are interested in linguistic criteria. We are not concerned with the description of the outside world, with actions described by the verbs, but with the task of determining how each language, as a cognitive system, sees and treats those actions.

• Finally, and because we come from the NLP field, we are also interested in what can be automatically detected (without requiring the human critical sense).
The phenomenon that concerns us now, that is, tense and aspect, is actually especially adequate to a broad-coverage treatment, since, contrarily to other phenomena, time is omnipresent in natural language (recall Quine’s [Quine 60] statement on the bias of language concerning time).

3 Aspectual classes

We consider important to draw the following distinctions:

1. Verbs that do not convey any temporal information. These verbs express properties or situations. Consequently, they have no restrictions regarding their localization in time\(^1\): STATES.

   \(pertencer\ a\), \(possuir\), \(amar\), \(parecer-se\ com\), \(ser + A DJ\). (belong, own, love, resemble, be + ADJ)

2. Verbs that correspond to something punctual, that is, something that happens in a particular instant, and not in an interval. Their placement in an interval forces a repetitive interpretation: CHANGES\(^2\).

   \(reconhecer\), \(descobrir\), \(encontrar\), \(estremecer\), \(deitar\). (recognize, discover, find/meet, tremble, lie down)

3. Verbs that correspond to an interval, with no privileged instants. That is, what they describe stretches in time (it can never be only true in an instant), but there is no point or points of that interval that are special, nor any difference between different points: ACTIVITIES. (In fact, it doesn’t even make sense to talk about the verb with respect to those points, the verb only carries a meaning with respect to intervals.)

   \(correr\), \(crescer\), \(ler\), \(amarelecer\), \(procurar\), \(vaguar\), \(afastar\), \(fazer\). (run, grow, read, yellow, look for, perambulate, move away, make)

We encompass in this class verbs that are considered by some as corresponding to the repetition of a simpler event.

\(^1\)It is interesting to note, in this connection, that [Link 87]’s definition of states is that they cannot be localized in space.

\(^2\)Often, these verbs also define a change of state, and thus their name. However, as we said above, we will not be concerned here with causal phenomena.
cintilar, cumbalear, saltitar, martelar, tremer, cacarejar, vibrar, baloiçar, tremeluzir, abanar. (twinkle, stumble, hop repeatedly, hammer, shake, cackle, vibrate, swing, flicker, to shake (something))

4. Verbs that correspond to an interval with a privileged point. They can therefore be used to refer to the interval, to the point, or don’t care. That point corresponds to the moment after which it can no longer be avoided that the situation described by the verb obtains. It may be the last point of the interval (morrer, fechar (die, close)), the first one (cair, abrir (fall, open)), or be in the middle (sair (leave)):

morrer, cair, abrir, sair, fechar, ganhar. (die, fall, open, leave, close, win)

5. Verbs that correspond to a structured interval, that is, that designate an interval decomposable in subevents, and that has an intrinsic beginning and (or) end:

construir, produzir, atravessar, percorrer, pintar, jantar, preencher. (build, produce, cross, roam about, paint, have dinner, fulfill)

Apart from the above classes, there are verbs that present a basic multiplicity of senses, perfectly related, that is, non-attributable to historical or orthographical hazard (“logical polisemy”, according to [Pustejowsky 89]). Contrarily to the cases above, for which morpho-syntactic structure does nothing else that select the viewpoint under which the situation must be considered, the verbs belonging to the following classes have the two senses on equal footing (which selection, by the way, is done by the very same mechanisms / differences of context). The occurrence of one of the senses does not imply — does not force — the occurrence of the other, however.

6. Verbs that correspond to a change and to the property resulting from it: ACQUISITIONS.

conhecer, saber, lembrar-se de, perceber, ver, ficar, vestir (roupa), acreditar em. (meet, know, remember, understand, see, stay, wear, believe)

7. Verbs that correspond to an activity and the property that comes with it: INTERVAL STATES.
viver, morar, trabalhar, gostar de. (live, live, work, like)

8. Verbs that correspond to a change and an activity composed of several changes:

SERIES.

tossir, espirrar, gritar, pigarrear, respirar, acenar, gargarejar, balançar. (cough, sneeze, shout/cry, clear the throat, breathe, wave, gargle, oscillate)

It is important to clarify that the distinction done here (between, for instance, PEAKS and WORKS on one side and ACQUISITIONS in the other) is rarely done about aspecual phenomena. Basically, we distinguish between the existence of a complex temporal pattern associated to a word, and a vague description, that is, a metalinguistic property of words that a priori do not carry a fixed aspecual description.

The first case allows the speaker to choose the perspective under which s/he wants to refer to the event described by the verb, which corresponds to what [Smith 91] calls “viewpoint aspect”.

The second case has constituted a motive of perplexity for several authors, and, in fact, phenomena related to the aspecual classes we present as composed have had several references in the literature, as will be shown in the following.

About acquisitions, some authors had already highlighted the existence of a class of verbs with properties both of achievements and of their resulting states, for instance [Dowty 79] for English and [Heinämäki 84] for Finnish. In Portuguese, verbs with such behavior can be easily disambiguated, in the past tense, by the simple choice of the verbal tense Perfeito or Imperfeito, contrarily to the languages just mentioned, in which the distinction is not realized morphosyntactically.

Ele conheceu-a / ele conhecia-a. (He met her / he knew her)

Ele soube da história / ele sabia da história. (He learned about the matter / he knew the matter)

On the verbs covering properties of both states and activities, such a class had already been proposed in [Carlson 81], and [Dowty 79]’s subchapter 3.8.2 contains a detailed discussion of such verbs in English. For Portuguese, see the equivalence between the truth conditions of the present progressive and the simple present

3And also sit, lie, hang, hide, wear in English.

4For a more detailed contrastive study, see [Santos 91].
Ele está a viver no Porto. ⇔ Ele agora vive no Porto. (He is living in Oporto. ⇔ He lives now in Oporto.)

Ele está a usar lentes de contacto. ⇔ Ele agora usa lentes de contacto. (He is wearing contact lenses. ⇔ He wears now contact lenses.)

which is true neither for states

Ele está-se a parecer com o pai. Ele parece-se com o pai. (He is resembling his father. He resembles his father.)

Ele está a gostar do filme. Ele gosta do filme. He is liking the film. He likes the film.

nor for activities

Ele está a correr. Ele corre (agora). (He is running. He runs (now).)

Ele está a mexer-se. Ele meze-se. (He is moving. He moves.)

Finally, the repetition of the same action, even though classified as a different class by [Freed 79], has not been generally connected with vagueness regarding its denotation. In particular, the iterative interpretation of expressions has been traditionally considered as a property of the morphosyntactic structure and not lexical.

Series can be detected, for instance, through the cooccurrence with the aspectualizer acabar de, contrarily to activities,

Ele acabou de tossir. (He has just coughed)

* Ele acabou de crescer. (* He has just grown)

and the acceptance of the progressive with a clause of the kind Há X tempo que... (It's X temporal units since...), contrarily to changes.

Há duas horas que está a tossir. (It's 2 hours since he is coughing.)

* Há duas horas que ele está a encontrar a Maria. (* It's 2 hours since he is finding Maria.)

At last, it is important to clarify that there are some verbs that do not have associated any aspectual information, such as performative or causal verbs (of 3rd order in [Passoneau 88]'s terminology): afirmar, prometer, causar, resultar (state, promise, cause, result).
4 Selection criteria

To situate our work in a general setting, let us mention that there are three kinds of objective criteria:

1. Those that require human linguistic competence;
2. Those based in text corpora;
3. Those based in existing dictionaries or thesauri.

The order above intends to reflect, on one hand, an increasing difficulty of stipulation of those criteria; on the other hand, an increasing ease to apply them (considering that the task is to classify every verb and not a dozen). In this paper, we will only present criteria of the first type, and will postpone for a future study their evaluation and rewriting taking into consideration the two other types.

We begin by noting that the criteria do not pretend to be exhaustive, but only sufficient. The methodology followed was the search for grammatical expressions that had some relationship with the concepts under study, therefore the aspectualizers (which focus precisely on subparts of an action) and tenses were the most used. On purpose, we did not enter in consideration with criteria that covered simply a verb subclass (as the so fashionable aspectual distinction between unaccusative and inergative verbs, inside the intransitive [Sanfilippo 91]).

For lack of space, the vast majority of the examples was omitted. It is however hoped that the readers can reconstruct them easily.

We have also decided, for clarity’s sake, to present the criteria in a descriptive, as opposed to algorithmic, way, and therefore the order is of no significance. To apply them it would thus be necessary to reformulated them in a form “If x and if y and if z then A, etc.”.

For obvious reasons, the criteria themselves are skipped, since they cannot make any sense to people who do not speak Portuguese.

After enumerating this set of criteria, it is essential to recall that they concern verbs, and thus, they should be applied with care. In particular, examples such as reconhecer o terreno (‘‘to recognize the field’’, meaning to identify the surround-
ings, used for a dog, or a detective) sair às 5 horas (to leave at 5), or construir casas (to build houses) do NOT violate the criteria above, for instance. In fact, these verbal expressions do not correspond to a change, a peak and a work, respectively, but to a work, a series and an activity. This result can, moreover, be objectively expressed by that calculus that, given the verbal classification and the syntactical context, produces a new classification.

We do not claim to have exhausted the possible sources of criteria with this study, not to have contemplated all the pertinent phenomena in the tense and aspect field in Portuguese. In particular, we put aside the bigger or smaller facility with which a verb acquires an habitual connotation, and also the unicity property (impossibility of repetition) that some verbs carry, which are features that allow one to infer others from the tense and aspect realm.

5 Discussion

The first point that may seem strange is the fact that we resorted to so many classes, since in the literature, though the Vendlerian quadrilogy is a fundamental reference, many authors use simply a tri- or even bipartition.

The reason for our approach is that there are objective data that verbs of different classes behave differently with the same syntactical mechanisms, therefore, it is necessary to distinguish them at the lexical level.

We postpone for another paper the discussion of the evaluation of these criteria, simply saying for now that this evaluation will only be possible through a corpora-based study, covering a large number of verbs.

How to treat idioms (by definition, expressions which meaning is more than the “sum” of its parts)? Do they also suffer aspect changes? “Dar o corpo ao manifesto” (‘‘give the body to the manifesto’’, have a lot of effort), “dar voz à alegria” (‘‘give voice to happiness’’, express one’s contentment), “dar largas à dor” (‘‘give wideness to pain’’, show deep sorrow), “dar o braço a torcer” (‘‘give the arm to twist’’, accept that one’s wrong), “dar a mão” (‘‘give the hand’’, give hands), “dar uma gargalhada” (‘‘give a laughter’’), “dar um pontapé” (‘‘give a kick’’).
kick), for instance, are or aren’t related with the aspect of the verb “dar” (kick)? We have previously [Santos 90] called the attention to the fact that the distribution of these expressions versus its lexicalization in a single verb varies from language to language and does not obey to any semantic criterion; therefore there is no reason to treat these expressions as special. We shall thus assume that, in principle, their aspectual characterization is equivalent to that of the lexical verb, with the consequent regular modifications that the respective syntactic structure implies. This subject will however be focussed more detailedly in [Santos in prep.].

Finally, we repeat that our aspectual classification is not but one of the elements of a calculus that allows us to get the aspectual characteristics of a text (sentence, expression or paragraph). It presupposes therefore that we know also how to formalize the influence of the morphosyntactic structure on top of the lexical classification proposed. Even though this paper does not concern this issue, it should be emphasized, on one hand, that there is already a lot of work on this subject, and therefore the previous statement does not stand as an utopic wish. In particular, the classes that were the subject of this paper have risen from a deep study of a calculus which will be the subject of other papers, in which a formal description of these classes will be given. We chose not to present their formalization here, not only because many of the options taken were either fundamented by non-lexical phenomena or by evidence from other languages, but also because we think that this proposal should be valid, at the level of the description of Portuguese, independently of the chosen formalization.

Acknowledgements

I thank the Junta Nacional de Investigação Científica e Tecnológica the PhD scholarship that allowed this work, and IBM Portuguesa for all their support.

This work is due to a great extent to the discussions I had with my supervisor, Lauri Carlson, on tense and aspect in general and on its realization in Portuguese, and many of the ideas vehiculated here are his own, though I am the only responsible for this article.

I would also like to thank all members of the NLP Group at INESC for their support and constructive criticism, and Henriqueta Costa Campos for the extensive bibliography that put at my disposal.
References


