THE DEFINITE ARTICLE IN PORTUGUESE
A DISCOURSE-ORIENTED APPROACH

Célia Lacerda Jaguaribe*

1. Purpose

The purpose of this paper is to offer an analysis of the definite article in texts. It is my view that texts usually contain explicit elements that enable the hearer or reader to construct appropriate contexts for them. One such element is the article. In this sense, the article is a tool used by the speaker or writer to signal an assessment of the hearer’s or reader’s knowledge and probable expectations.

2. Methodology

It is a well-known fact that in any communicative situation—be it oral or written—individuals necessarily make familiarity inferences about what they hear or read. That is, they expect to see connections between what they hear,
or read, and things they have experienced before, and which are, in this sense, known to them. These inferences are shared, or at least, they are assumed by the speaker or writer to be shared with the hearer or reader.

It is this kind of familiarity inferences that has often been discussed under the notion *shared knowledge*-a notion claimed to be relevant to definiteness, as well as, to comprehension in general. Prince (1979) proposed the term *Assumed Familiarity*, and showed, by analyzing a naturally occurring oral narrative, that it represents a range of discrete types of information, which includes what she labels as *unused*, *indeferrable* and *evoked* entities, or referent-
es.

An *unused* entity is a referent which is introduced into a piece of discourse for the first time, but the speaker assumes that it is familiar to the hearer on the basis of the situational environment which the speaker assumes they share. An *indeferrable* entity is a referent that is mentioned in a text for the first time, but which is not entirely new to the hearer in the sense that it can be inferred from other items in the text. An *evoked* entity is a referent which, having first been introduced into a text, recurs at different places, and as such is said to be evoked by the hearer.

The analysis I here present is based on Prince’s framework, and as such it takes as its starting-point the understanding of the hearer or reader. In other words, the elements, or items of a given piece of discourse are analyzed not in terms of their forms, but rather in terms of what is somewhat felt to be familiar to the hearer or reader. It so happens that all such elements here identified can be formally realized by means of the definite article.

The analysis concentrates on written texts, more specifically, vignettes, i.e. short literary sketches, by contemporary Brazilian writers. The reasons for choosing such texts are the following: first, this kind of literary genre is very common in Brazilian literature; second, these texts are good examples of linguistic communicative interactions, with on the one side the writer, who wants to impart some information to a receiver on the other side, i.e. his/her reader(s); third, these vignettes usually portray everyday life situations in which fictitious characters are advanced as real, and engage in conversations, expressed in the form of dialogues or quotes.

The following points need be mentioned in connection with the analysis: First, in analyzing these texts it is important to take into consideration two different kinds of communicative interaction - one between the writer and his/her reader(s), the other between the characters themselves as they talk to one another. Given this distinction, the quote should be also seen as a text,
embedded into a larger one, but, nevertheless, autonomous in itself. Thus,
the analysis should be made under both perspectives, on the one hand, that
of the reader(s), and on the other hand, that of the hearer(s) in the quotes.
Second, it is common in these texts for the writer to introduce fictitious
referents, and act as though he expects the reader to know the particular
entities he is referring to. In creating this atmosphere of pseudo-familiarity,
the writer presumes on the reader’s familiarity with these specific fictitious
referents. This is very common in the opening sentences of narratives, as,
indeed, it happens in the text here presented.

Third, these texts usually contain material that in the last analysis is a
figment of the writer’s imagination - material about which one cannot share
knowledge in the completely conventional sense. Furthermore, readers may
simply not know certain kinds of facts, but still be able to understand the
text at some level. For example, the reader does not, in fact, have to know
the place o Rio de Janeiro [2], introduced at the beginning of the text, in
order to understand it. He only needs to acknowledge that the writer is
acting as if he knows about it.

Finally, the analysis concentrates on single NPs, i.e. those consisting of
the article plus a noun. Both pre-and postmodified NPs, i.e. those in which
other elements precede or follow the noun, are not considered on the grounds
that the occurrence of the definite article in such NPs are further determined
by factors other than those here discussed, such as, semantic and/or syntactic
properties of the lexical units attached to the noun. Also excluded from the
analysis are idiomatic expressions in which the definite article occurs.

3. The Analysis

To illustrate my analysis, I have chosen the text “A Cabra e Francisco”
(The Goat and Francisco) - one of the vignettes in a collection by the
contemporary Brazilian writer Carlos Drummond de Andrade.

In this text, the writer imaginatively creates a story in which a little she-
goat unexpectedly showed up in a hospital in Rio in the middle of the night,
being then operated on by the doorman, after he had found out that she had
a bullet lodged in her neck. Throughout the incident, the doorman talked to
the goat, as if she were human, until, to his amazement, he noticed that the
goat herself could speak, having, indeed, made conversation with him. In
brief, then, the story centers upon these two characters, the goat, and the
doorman. For Case of reference, I shall consider first those referents which
can be identified on the basis of linguistic facts, i.e. their recognition by the
hearer or reader is dependent on circumstances within the discourse itself.
Then, I shall analyze those referents which the hearer or reader can supposedly
identify on the basis of extralinguistic facts that surround the discourse. Within
the former group, are both those referents that I have here described as
evoked and inferrable. Within the latter, are those described as unused
referents.

3.1. Evoked referents

Let us first consider evoked referents. The familiarity shared by speaker/
hearer or writer/reader on what concerns these items, comes from the fact
that either the referent is potentially visible in the situation of utterance, in
which case, it is situationally evoked, or the referent has already been
mentioned in the text. In this case, it is textually evoked.

Situationally evoked referents in the text are: o hospital [12] and o Miguel
Conto [28]. They are situationally evoked in the sense that within this parti-
cular text, i.e. the conversation between the doorman and the goat, they are
potentially visible for both speaker/hearer. In other words, their identification
by the hearer is dependent on what he can immediately perceive rather than
on what has been previously mentioned. Both referents are expressed by
NPs with the definite article.

Textually evoked referents are any subsequent mentions of referents
previously introduced into the discourse. Thus, a cabra [4], [9], [20], [26],
[37], o animal [5], [23], and o bichinho [21] are all subsequent mentions of
the goat previously introduced into the discourse by means of the NP uma
cabrinha malhada (a little spotted she-goat) in the first paragraph.

In the same way, o homem [8], and o porteiro [24] are both subsequent
mentions of o porteiro [3] (the doorman), a referent introduced in the opening
paragraph. Also A Lapa [31], and o Miguel Conto [32] are second mentions
of the previously introduced referents a Lapa [30] and o Miguel Conto [28].
And o hospital [18] is a subsequent mention of o hospital [1], introduced at
the beginning of the text. All these referents stand in a coreferential relation-
ship with an antecedent, and their identification by the hearer or reader is based
on the previous linguistic context. Formally, they are all expressed by NPs
with the definite article.
One can see that in many of these cases the subsequent mention of the referent is lexically different from the first mention, although it is capable of the same reference. Thus, *uma cabrinha malhada* is referred to again not only by means of the NPs *a cabra* [4], [9], [20], [26], [37], but also by means of the NPs *o animal* [5], [23], and *o bichinho* [21]. The latter are identified by the reader on the basis of his/her knowledge of the language, more specifically, knowledge of a class inclusion relationship between *uma cabrinha malhada* and *o animal* [5], [23], as well as, between *uma cabrinha malhada* and *o bichinho* [21]. There is in each case a relationship of hyponymy between these terms, that is, one in which *o animal* [5], [21], and *o bichinho* are superordinate terms. Similarly, there is a class inclusion relationship between *o homem* [8] and its antecedent *o porteiro* [3].

All these instances illustrate the fact that coreferentiality is indeed a semantic phenomenon. That is, as pointed out by Halliday and Hasan (1976:62), anaphoric reference items refer to meaning not to the forms that have gone before.

3.2. Inferrable referents

There are other items in the text whose identification by the hearer or reader is also determined by the linguistic environment, i.e. by what has been previously mentioned. However, unlike those referents analyzed as evoked, they do not stand in a coreferential relationship with an antecedent, rather they are inferred from items previously mentioned. Within this group, are those referents here described as inferrables. This is the case, for example, of *o porteiro* [3], first mentioned in the text, and which is related to the referent *o hospital* [1], introduced at the beginning of the text. One can say that such a referent has brought up a particular frame for the reader, a “hospital” frame as it were, which has implicitly introduced the subsequent items *o porteiro* [3], as well as *os doutores* [15], *a sala de cirurgia* [16], and *a farmácia* [17], later mentioned in the text. The identification of these referents on the part of the reader is determined by some sort of associative relation between these items and the previously mentioned referent *o hospital* [1]. Furthermore, the item *o porteiro* [3] leads the reader to infer yet another frame, namely one consisting of items usually associated with people, say, a “human being” frame in which items like *o braço* [6], *a mão* [34], *o lado* [36] are seen as natural and expected parts. In fact, the degree of associative relation between on the one hand, *o braço* [6], *a mão* [34], *o lado* [36], and on the
other, *o porteiro* [3] is quite high, given that the mention of a person immediately brings into consideration all his/her body parts.²

It is also likely that the first mention *o porteiro* [3] brings up for the reader a different sort of frame, namely, one concerning the physical aspects of the activity undertaken by people such as doormen³. Thus, the subsequent mention *a cabeça* [11] can be inferred from the first mention *o porteiro* [3] plus the knowledge that it is common for doormen to have a place where to sit on (at least, in some cultures). Other inerrerae referents in the text are: *o cheiro* [7], *o pescoço* [14], *os olhos* [35], *as barbas* [138], all of them associatively related to the initially introduced referent *uma cabrinha molhada*. Again, as in the case of evoked referents, all these referents are formally realized by NPs with the definite article.

### 3.3. Unused referents

Let us now consider those referents which the hearer or reader can supposedly identify on the basis of his knowledge of extralinguistic facts that surround the discourse, i.e. those referents here described as unused. Again, they are also expressed by NPs with the definite article.

In the text, the referents expressed by the NPs *o Rio de Janeiro* [2], *o Icô* [25], *a Lapa* [30], *o Hospital Veterinário* [29], as well as *o dia* [13], [19], *a noite* [22], *o cão* [41], and also *as cabras* [33], and *a cabra* [40] (i.e. goats in general) are not present in the situation of utterance, i.e. they are not potentially visible referents, nor have they been mentioned in the text before. Nevertheless, they can presumably be identified by the hearer/reader on the basis of the extralinguistic context.

Thus, in the case of *o Rio de Janeiro* [2], *o Icô* [25], *a Lapa* [30], *o Hospital Veterinário* [29], it is knowledge of the existence of places such as cities, suburbs, and hospitals, respectively, that underlie the identification of these referents on the part of the hearer/reader. This does not necessarily mean that the hearer/reader has specific knowledge of these referents, but rather that he can acknowledge that the writer is acting as if he knows about them. This acknowledgment is based on some general knowledge that things like cities, suburbs, and hospitals exist. They are, therefore, part of the universe of discourse shared by both speaker or writer and hearer or reader.

In the case of *o dia* [13], [19], and *a noite* [22], these are referents identified by the hearer/reader on the basis of some general knowledge concerning time division, as conventionally established in western cultures.
We can mention, in passing, that these referents are the so-called unique referents, whose identification is always made on extralinguistic grounds no matter what the situation. The referent o céu [41] can also be viewed as a unique referent, whose identification by the hearer/reader is made on knowledge of certain religious beliefs or creeds - knowledge presumably shared by both parts, i.e. the speaker or writer, and the hearer or reader. An interesting aspect to be noted here is that the abstract concept heaven is an extension of the concrete referent sky, being both expressed by the same form, i.e. o céu. As discussed in Chafe (1972:52), these unique referents can be regarded as known sets which have but one member. Thus, if the hearer/reader knows, say, the concept céu (sky), he cannot help but know which member of the set is being talked about, since there is only one.

Similar to these unique are the referents expressed by the NPs as cabras [33], and a cabra [40], mentioned in the text. Since they refer in a unique way to a set or class, it is knowledge of the existence of such sets or classes that underlies the identification of these generic referents on the part of the hearer/reader. This is directly related to the fact that generic concepts, as discussed in Du Bois (1980:226), are directly available in the speech situation, being, therefore, not processed via the mediation of some prior mention, but rather, directly. In this sense, then, the speaker or writer may presuppose that the hearer or reader is familiar with these sets, or is informed about them.

3.4. Further distinctions

Let us examine now some other referents in the text which are also expressed by NPs with the definite article. At the beginning of the text, the NP o hospital [1] expresses a referent which is introduced to the reader for the first time. There is nothing in the linguistic, nor the extralinguistic context that may lead the writer to assume that the reader is familiar with the introduced referent. Nevertheless, he uses the definite article. In so doing, the writer pretends that there is a context in which just one particular hospital exists, and he acts as though he expects the reader to know the particular entity he is referring to. In this sense, such a referent is not altogether different from those discussed under the label unused. It differs only in that the reader’s presupposed familiarity is based on an invented rather than a genuine context. In isolation, a sentence like: ‘o hospital como o Rio de Janeiro dorme’ would not make much sense, and it would very likely provoke a wh-question on the
part of the hearer, such as: 'que hospital?' (what hospital?). In the context of
the narrative, however, it makes perfect sense. It is evident that this sort of
first mention reference is an example of a well-known type of literary device,
whose effect is to place the reader immediately within a context which the
writer creates.

Examples such as this point to the fact that the concepts of existence
and uniqueness that underlie the use of the definite article should not be
considered in absolute terms. Existence in such cases can only be defined in
terms of various ‘possible worlds’, and uniqueness is always relative to a
given communicative situation.

At the end of the text, there is a referent expressed by an NP consisting
of the definite article plus a proper name, i.e. *o Ariano Suassuna* [39]. This is
another instance of an unused referent. The speaker/writer presupposes the
hearer-reader’s familiarity with such a referent, as evidenced by the fact that
there is no further explanation by the speaker/writer concerning this item.
Given the uniqueness feature inherent in proper names, i.e. a proper name
usually designates a single, determined entity, one would not expect the
definite article to occur with a proper name. It occurs, however, in the above
mentioned NP, viz. *o Ariano Suassuna* [39]. In pragmatic terms, this
phenomenon may be seen as a strategy used by the speaker/writer to signal
closer familiarity with the referent. That is, in addition to the uniqueness
boundary already present in the proper name itself, the speaker/writer imposes
yet another boundary, a close to self boundary, as it were (Acton 1977).
Thus, in mentioning *o Ariano Suassuna*, the speaker/writer invites the hearer/
reader to share a closer degree of familiarity with the referent.

There are two other items in the text that deserve additional comments,
namely: *a razão* [10], and *o diálogo* [27]. They are both first mentions, but
the reader is able to identify them on the basis of what has been previously
mentioned. Thus, in the case of *a razão* [10], it is introduced into the text
after the doorman has explained to the goat why she is not allowed into the
building. The entity expressed by the NP *a razão* [10] refers to this
explanation, being also inferred from the previously mentioned event of the
goat being forced to leave the building. Similarly, the NP *o diálogo* [27]
refers to the conversation held by the doorman and the goat, and in this
sense it can be said to be an evoked entity. Lyons (1977:672) discusses a
similar example to show that a potential referent is salient in the universe of
discourse, even though it is not present in the situation of utterance and has
not been mentioned previously by either the speaker or the addressee.” Such
cases illustrate the fact that anaphoric reference items need not refer to meanings introduced through a direct reference, but may refer to meanings introduced indirectly, as in the case of *a razão* [10] and *o diálogo* [27].

4. Implications

The analysis here presented has several bearings on the study of the linguistic phenomenon in general. First, it permits a generalization of anaphora as a process which involves not only linguistic, but also extralinguistic facts. In this sense, an item is anaphoric not only when it stands in a coreferential relationship with some item previously mentioned, as in the case of evoked referents, but also when it refers to something whose knowledge the speaker or writer assumes the hearer or reader shares with him/her on the basis of their experiential background. In other words, it refers to something which is, as pointed out by Lyons (1977:672), “in the intersubjective experience of common memory of speaker and addressee” - something which has, as it were, been previously established in the universe of discourse common to both parts. It is this broader view of anaphora that underlies the introduction of unused referents into discourse.

The second aspect of the analysis is the generalization of notions like existence, referentiality, uniqueness, specificity, and genericness, which underlie the use of the definite article, as belonging to a larger class of distinctions, namely, that of familiarity inferences on the part of the hearer/reader. By taking into consideration these familiarity inferences, one is able to account for both the specific and nonspecific uses of the definite article in a unified way, and to bring under the right perspective both the uniqueness and inclusiveness features of the definite article, which most analyses solely based on semantic distinctions claim to be essential characteristics. Thus, it is clear in the text that in the case of, for example, *a cadeira* [11], *a sala de cirurgia* [16], *a farmácia* [17], and *o braço* [6], *a mão* [34], the reference is neither specific, nor unique. The hospital may certainly have more than one chair, operating room, or pharmacy; and the doorman should be physically normal with both his hands and arms, since it has not been mentioned otherwise. In such cases, the items introduced by the definite article are not linked to an actual referent, nor to a verbal description, but to a conceptual representation. In other words, the writer’s reference is rather to a type than to an individual thing, and in this sense, the choice of, say, a particular chair, operating room,
pharmacy, etc. is immaterial. What is at issue here is the establishment of a link to the concepts chair, operating room, pharmacy, as part of the relevant frame, namely, the "hospital" frame.

One sees, then, the relevance of the notion frame in the analysis of the definite article, and this brings us to the third aspect of the analysis, namely the need in linguistics for a more complex notion of frame in order to account for the several kinds of inference underlying the process of decoding on the part of the hearer/reader. In this connection, a distinction is made between, on the one hand, inference determined by facts that are basically linguistic, i.e. having to do with the inherent properties of the meaning of the word, and on the other hand, inference determined by extralinguistic facts in which culture certainly plays a decisive role. In the first case, are those inferrables whose degree of associative relationship with the entity from which they are inferred is very high. A case in point is possession relations, in particular, those defined as inalienables. For example, the referents expressed by the NPs o braço [6], and a mão [34] are inferred from the referent o porteiro, plus the fact that humans are known, among other things, to have both hands and arms. Putting it still another way, one can say that the mention o porteiro [3] has, as it were, foregrounded a frame for the reader in which items like o braço and a mão are taken for granted.

In the second case are those inferrables for which this associative relationship is rather loose. Thus, the entity expressed by the NP a cadeira [11] can also be inferred from o porteiro [3] through the foregrounding of another frame, one concerning the activities usually associated with doormen. However, the degree of inference here is much weaker than in the case of o braço and a mão. And it is less determined by some semantic properties of the trigger, i.e. the referent that triggers these associations, than by other facts in which cultural expectations come into play. Thus, whereas it is most likely for human beings to have hands and arms, having a chair where to sit on is a contingent rather than an inherent property of being a doorman. One can see, then, that the notion frame and its underlying associative relations are very much dependent on probabilities. This does not just mean the possibility that a particular thing exists in a frame, but, indeed, as pointed out by Rhodes (1981), the probability of finding a frame with such a characteristic - a situation which in the last analysis is very much tied up with cultural facts.

Finally, the analysis hints on the old and much debated issue of how language and culture are intertwined, the former being in many instances a
reflection of the latter. In this connection, one may want to inquire into the facts that determine the use of the definite article with abstract nouns, a phenomenon typical of the Portuguese language, and, in fact, of Romance languages in general, as opposed to other languages, such as English, in which these nouns are characterized by the absence of the article, as, for example, o amor (love), a alegria (joy), a riqueza (wealth), a paz (peace), a felicidade (happiness), etc. The Greco-Roman tradition of personifying such abstract concepts in the form of gods made them concrete entities, which became part of the public record of these cultures. According to Rhodes (1981), the public record includes all those things in a society on which people's interactions are based, i.e. things that make up their common universe of discourse. As part of this common universe, these entities constituted information shared by the members of these communities. It is no surprise, therefore, that the languages that have flourished under the Greco-Roman tradition still reflect such distinctions.

To conclude, I would like to acknowledge the fact that the analysis here presented is to a certain extent subjective. It is my view, however, that psycholinguistic experimentation can both provide this kind of analysis with greater objectivity, and shed more light on the processes that involve the hearer/reader's understanding and comprehension.

Notas

1. These referents are fully discussed in connection with the definite article in Lacerda (1983).

2. Downing (1980:93) points out that "these parts are seen as affiliated to their wholes, and this affiliation is expressed by the use of possessives, and by use of definite articles which can be appropriately used only because the
whole has already been mentioned setting up the expectation that the part also exists.

3. Jones (1980:46) discusses the fact that the frames of a person's knowledge are related to one another in different and complex ways. Furthermore, "the possibility that frames may be members of more than one class of frames indicates one aspect of the incredible complexity of knowledge structures."

References


DU BOIS, John W. "Beyond Definiteness: the Trace of Identity in Discourse." In: *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative*


ANEXO

A CABRA E FRANCISCO
by
Carlos Drummond de Andrade


- Está bem, filhinha. Agora pode ir passear. Depois você volta sim?

Ela não se mexe, séria.

- Vai, cabrinha, vai. Seja camarada. Preciso sonhar outras coisas. É a única hora em que sou dono de tudo, entende?


Essa coisa pegajosa... O que: sangue?! Por que não me disse logo cabrinha de Deus? Por que ficou me olhando assim feito boba? Tem razão: eu é que não entendi, devia ter morado logo. E como vai ser? [15] Os doutores daqui são um estouro, mas cabra é diferente, não sei se eles topam. Sabe de uma coisa? Eu mesmo vou te operar!


- Marcolina. Dou-lhe este nome em lembrança de uma cabra que tive quando garoto, [25] no Icó. Está satisfeita, Marcolina?

- Muito, Francisco.

Sem reparar que [26] a cabra aceitaria [27] o diálogo, e sabia o seu nome, Francisco continuou:


- E daí?

- Daí, preferi ficar por aqui mesmo e me entregar a seus cuidados.

- Você me conhecia?

- Não posso explicar mais do que isso, Francisco. [33] As cabras não sabem muito sobre essas coisas. Sei que estou bem a seu lado, que você me salvou. Obrigada, Francisco.


Aí Francisco levou um susto, saltou para [36] o lado.

- Que negócio é esse: cabra faltando?! Nunca vi coisa igual na minha vida. E logo comigo, meu pai do céu!

[37] A cabra descerrou um olho sonolento, e por cima [38] das barras parecia esboçar um sorriso:

- Pois você não se chama Francisco, não tem o nome do santo que mais gostava de animais neste mundo? Que tem isso, trocar umas palavrinhas com você? Olhe, amanhã vou pedir [39] ao Ariano Suassuna que escreva um auto [40] da cabra, em que você vai para [41] o céu, ouviu?