CATEGORIAL MORPHOLOGY: A SEMANTIC PERSPECTIVE ON THE MORPHOLOGICAL LEVEL*

LIVY REAL

Abstract

This work points out a semantic approach to the morphological level, strongly based on Bayer (1997) and Hoeksema (1985). Based on Categorial Grammar, a syntactic-semantic tool commonly used by semanticists and computational linguists, I look at the morphological level through the same syntactic rules adopted by the model. In this way, it is possible to gain a theoretical improvement since I look at morphological, semantic and syntactic levels through the same formal structure. I will briefly introduce Categorial Grammar showing its application in Brazilian Portuguese. I have chosen a nominalizer suffix, -ura, present in such words as abertura ‘opening’ and assadura ‘baking/rash’.

1. Introduction

This paper proposes a categorial view of the morphological level. Since I consider Categorial Grammar (CG) a very interesting and insightful, but underused, tool, I check its application on the morphological level to see what kind of problems/insights (if any) can arise from this methodology. I expect to obtain a good description of the morphological phenomena using the same mechanisms that are applied to the syntactic level in CG. I think that we will have a great theoretical improvement, since we will be able to consider two different linguistic levels through the same rules.

As some scholars from CG argue (e.g. Wood 1993), CG seems to give some insights about linguistic phenomena that are quite different from the usual, for example, regarding CG approaches, one could discard such traditional linguistic notions as subject, predicate and verb. So, also, I am looking forward to gaining similar insights on the morphological problem that I have chosen to describe here.

I have chosen a very specific morphological phenomenon of Brazilian Portuguese, the possible meanings of the suffix -ura. Since Portuguese is a language with weak morphology, I work with derivation, unlike Hoeksema (1985), who made the first attempt to use CG on morphology using Dutch and its composition phenomena.

In the next section, there is a brief explanation about CG mechanisms. I concentrate efforts on making explicit some concepts that are relevant to this work (such as type theory) and I leave out some other interesting points (comparisons between CG and phrase-structure grammar, syntactic phenomena, etc.).

In section 3, I show how it is possible to use the rules from CG on the phrasal level. I explain my proposal in more depth in section 4 and, next, I introduce the suffix -ura, its meanings and uses. In Section 5, I show an ontology proposal (Moens & Steedman 1988) for linguistic expressions that seems to be necessary to explain some facts that come from the semantic analysis of -ura.

* I am very thankful for the help of Álvaro Fujihara, José Borges Neto, Luís Arthur Pagani and Maximiliano Guimarães from Federal University of Paraná (UFPR) and Ricardo Andrade from University of São Paulo (USP).
Finally, in section 6, I give a brief review of this work and I argue that it is possible to use CG on the morphological level without problems and with some insights about morphological structure, probably because of the semantics-morphology parallelism that is obtained in CG.

2. Categorial Grammar: a brief explanation

Normally used by semanticists and syntacticians, Categorial Grammar (CG) is a tool that easily shows the parallelism between syntactic and semantic relations among expressions. It is based on a highly informative lexicon and on just six pairs of rules that structure all the possible lexical interactions, such as the commutation process.

In CG, it is not necessary to postulate arbitrary categories of lexical items, since there is the possibility, from some basic categories, of discovering all the others inductively. This mechanism to ‘discover’ categories comes from the fact that CG considers the linguistic expression in terms of functions and its arguments. For instance, a sentence like *Peter runs* is understood as an application of a function (*runs*) to an argument (*Peter*), thus, semantically, this sentence is considered as a function run(Peter). It means that CG does not use the idea of constituency in a linguistic expression’s constituent structure as others theories do, such as Transformational-Generative Grammar (TGG) or Head-driven Phrase Structure Grammar (HPSG). It is possible to formalize it as:

(1) If X and Y are basic categories,
then a function from X to Y is also a category,
there is no need for X and Y to be atomic.

The last line above gives the system's recursivity, i.e., if X and Y are categories, X[Z will be too, and also (X[Z])(X[Z]), and so on. To combine expressions, CG uses a ‘function’ that is very much like the cancellation of fractions, with the value as numerator and argument as denominator in the functor category, and its application to its argument working like multiplication (Wood 1993: 11).

For instance, if X and X[Y are categories and they must be combined, it will be:

\[
\frac{X}{X} \cdot \frac{Y}{X} = Y
\]

To construct these categories it is also necessary to have a word-order operator that gives the most relevant difference between the CG semantics and its syntax, showing us when the two categories are complex, whether each of them is either an argument or a function. The notation used to make explicit this word-order operator is two different slashes that will replace the vertical slash above. When an expression takes its argument on the left, like intransitive verbs in languages like Portuguese and English, the vertical slash may be replaced by a backslash: X\Y. A forward slash is used when the argument is taken on the right side: X/Y. In both cases, the function must be an unsaturated expression, i.e., the function must be ‘incomplete’.

As I have said, there is a very close parallel between syntax and semantics in CG. This parallel is maintained through the adopted lexical categories. These categories are semantically motivated: the syntactic categories for CG have a strict relation to the semantic types of each expression. As CG adopts a Montagovian semantics, which assumes only two semantic types *e* and *t*, that express respectively entities and truth
values, it is expected that all the lexical syntactic categories have a parallelism with these two types. It is possible, for example, to assume we have just two basic categories N and S. Following the way of defining categories proposed above, we have:

(2) Since N and S are basic categories,
A function from S to N is also a category,
A function from N to S is also a category,
A function from N\S to N is also a category,
A function from (N\S)\N to S\S is also a category...

In a syntactic model, as described above, it is possible to define the syntactic-semantic relation as:

(3) \[ \text{TYPE}(N) = e \]
\[ \text{TYPE}(S) = t \]
\[ \text{TYPE}(A \setminus B) = \langle \text{TYPE}(A), \text{TYPE}(B) \rangle \]
\[ \text{TYPE}(B/A) = \langle \text{TYPE}(A), \text{TYPE}(B) \rangle \]

Note that above I have introduced parentheses. They are used to indicate what part of the complex expression is the main connective. In CG, there are some logical rules that can change the order of the items in parentheses, with, of course, a change in the structural semantics. This rule is called rule 4 (Moortgat 1988, Wood 1993, Borges 1999). Considering this rule and that here I am concerned with morphology, in which the complexity of categories is less than in syntax, I will use parentheses just when they are necessary to understand morphological processes.

Let us take the sentence *Peter smokes*, belonging to the category S, as a real example. Assuming that *Peter* is a noun (N), it is possible to arrive at the category of *smokes* without needing to understand *smokes* as a verb of some specific kind (such as intransitive verb, monoargumental, VP, etc.). Through this strategy, *smokes* is an expression that from an N at its left side produces an S, and therefore it is a term of category N\S.

So, summarizing, categories are inductively constructed from some basic categories (here I will adopt just S and N), using / and \ to indicate the formation order: N\S is an expression that takes an N on the left to form an S and an N/S is the opposite, an expression that selecting an S on the right side forms an N. Also it is worth remembering the relation between the syntactic categories and semantic types (<N,e> and <S, t>).

An advantage brought by CG is the absence of a hierarchy of rule application, since its formal mechanism gives the possibility of always arriving at the same final results independently of the order of rule application. Considering a sentence like *Peter loves Mary*, the idea of subject and predicate (or NP and VP) disappears, since *loves* would be an NS/N, no matter whether we choose to begin this analysis with *Peter* or *Mary*. Also there is no distinction, sometimes problematic, between deep structure and

---

1 The amount and the kind of basic categories in CG is a task for linguistic description (Hoeksema 1985). Logically, the most economical way to describe linguistic expressions is assuming just two categories as we do here, but some linguists argue that to describe natural languages we may need more. Hoeksema (1988), describing Finnish, Chamorro and Atayal, contends that we need a lexical feature to express case, and Borges (1999) argues that for syntactic description we need at least three basic categories (N - noun, NC – common noun, S – sentence). Here, as my goal is to give an introduction to CG and to apply it to derivational morphology, I adopt just two basic categories, N and S.

2 To see an explanation of this mechanism see Wood (1993).
superficial structure, since this distinction does not exist in the model. Let us see an example similar to *Peter smokes* in the formal model of CG:

**Figure 1: Pedro corre ‘Peter runs’**

```
[ exp Pedro corre. 
  cat S 
  reg R1 

[ con [ exp Pedro ] 
  cat N 
  reg lex ] 
  exp corre ]
```

In the analysis above I use the notation adopted by Pagani (2003), Feature Structure (FS), although the most current notation is the Pravitz diagram. In FS, ‘features [on the right side] are represented by atomic symbols, but values [left side] can be represented by an atomic symbol or another FS’ (Pagani 2003: 394). In this representation, *exp* is the expression, *cat* the syntactic category, *reg* the rule of expression formation, *con* the lexical entry constitution and *den* the semantic denotation. This last feature will be used in the next sections to express the semantic differences presented in morphology, but for now my goal is just to explain the CG mechanism.

Analyzing *Pedro corre*, *exp* is the expression itself; *reg* shows that this expression is made through R1, the commutation rule that simply combines two expressions; in *cat*, *S* indicates the sentence category (or the *truth value* category, in terms of semantic types); and *con* brings together the expressions that form this sentence: *Pedro* and *corre*. Two new expressions appear in *con*, making explicit the whole process of sentence formation. *Pedro* is a *N*(oun) and is lexicalized (lex) and *corre* an expression that selects a *N* on the left (*Pedro*) and forms an *S* (*Pedro corre*). These three structures (*Pedro*, *corre* and *Pedro corre*) also exhibit the *den* feature. Formally, *corre* expresses an event with a single participant, *Pedro* is an individual (*p*) and *Pedro corre* denotes an event in which *Pedro* is the participant.

In this way, with a few primitive categories, it is possible to find the category of the expression *corre* without postulating a new category to indicate the relation created by this expression, usually called ‘intransitive verbs’. The extension of this analysis to other word class expressions is easily recognizable. Let us see the representation of a sentence with a bi-argumental verb:

---

3 All of my examples are in Portuguese, since my following morphological analysis uses Brazilian Portuguese as an example and to arrive at some of my insights on the morphological level, we need to analyze the same expressions on the syntactic level.
Here the category of assa is N\S/N, i.e., an expression that selects an N on each side to form a sentence. It is very interesting to see the relations between these categories – N\S is the category of both intransitive verbs and transitive verbs that have already selected their complement. This is to say that these expressions have the same behavior, even if they are composed of different items. Both corre ‘runs’ and assa bolos ‘bakes cakes’ work in the same way in a sentence – selecting a noun on the left, they form a sentence – and because of that, they have the same categorial status. Thinking in traditional terms, the predicate (N\S) can be both corre and assa bolos.

3. Categorization on other levels

Applying the same rules to smaller expressions, these analyses are possible also on the morpho-lexical level, as I proposed in a previous work (Real 2006) and the first
attempt at categorial morphology (Hoeksema 1985). I understand that one of the improvements brought by CG to morphological studies is related to the possibility of seeing the lexicon compositionally, which is very interesting to a study that searches for lexical regularities.

In the present study, I work with the suffix –ura, present in nominalizations from participles in Brazilian Portuguese. I argue that, through the CG tool, I can reach a good transparency between the morphological level and the semantic one that other theories cannot capture. Some studies, linked in some way to Principles and Parameters (Chomsky 1981) and Distributed Morphology (Halle & Marantz 1993, Marantz 1997), see the same suffix, -ura (in fritura, ‘frying/fried food’) as a suffix of nominal categories (e.g. França & Lemle 2006). Since to determinate the morpheme’s category they need to consider just the final category of the word formed, the morphemes will not express in themselves all the word-formation process. This analysis is consistent with that model, since Generative Grammar starts from complex structures to arrive at the smallest elements. Within such a model of analysis, the difference is not necessarily explicit, e.g., between a suffix that nominalizes verbs and a suffix that nominalizes adjectives. Both of them are suffixes with nominal syntactic category; at best it can be made explicit what kind of category these suffixes select, but this is not relevant to the model.

In CG, the compositional view of complex structures, be they phrases, sentences or words, is always relevant and clear, because it is from the minimum items that it is possible to arrive at complex ones, first starting from lexical items. I will show in the next sections how it is possible to do it at different linguistic levels.

3.1. Phrasal Level

Let us see the application of CG on the phrasal level. In an expression like bolo assado ‘baked cake’, I may analyze assado ‘baked’, bolo ‘cake’ and the relation between these expressions. Bolo is a N(oun), information given by the lexicon (lex), assado is a word that was formed from the suffix –do ‘-ed’, that forms participles and adjectives, and from the verb assar ‘to bake’, whose category is also known from the lexicon. To find the category of assado, and consequently of the suffix -do, I may consider its function on the sentence denotation.

The normal use of assado ‘baked’ appears in sentences like Eu comi um bolo assado ‘I ate a baked cake’ or Bolo assado é delicioso ‘Baked cake is delicious’. In these contexts, assado works as an adjective, i.e., it forms a complex noun selecting another noun. We can see that assado does not change the syntactic category of its argument. In that case, the category NN is needed for assado, since from an N it gives another N.
Adverbs have similar categories, taking an $X$ to form a complex $X$. For example, in a sentence such as *Obviously I ate a baked cake,* obviously has a category $S\backslash S$ since it forms from another $S$ (‘I ate a baked cake’) another $S$. *Obviously* is the kind of expression normally called a sentence adverb, since it seems to modify a sentence. For example, obviously is different from fast, which in a sentence like *Peter runs fast* could receive a category $(N\backslash N)/N$ since here it modifies the verb/verbal phrase, and not the sentence. Since this discussion is not my main focus, I will leave it aside.

### 4. Categorial Morphology

Let us see the application of CG on the morphological level. If we have the category of *assado* ‘baked’ (NN) and the category of *assar* ‘to bake’, CG’s formal mechanism gives the possibility of finding the category of the suffix -do ‘-ed’ through a simple mathematical calculation, similar to the calculation of fractions. In this way, we can apply the same rules of the syntactic and phrasal levels to the morphological level.

*Assado*, $NN$, is the result of the multiplication of the categories of *assar* ($NS/N$) and -do, whose category we do not know yet. Then -do is defined from the categories of *assar* and *assado*: -do is an item that selects an expression $NS/N$ on the left side and forms an NN, so it is an expression with the category $(NS/N)(NN)$. Despite the size of the category, what it says is very simple: from a transitive verb, the suffix -do forms an expression that functions as an adjective.

All these notions are very intuitive and can be understood without a formal representation. However, a formalized application of the theory is possible and desirable not only to ensure its computability, but also to make it possible to see some insights that intuitively we would not, such as the possibility of seeing every category from lexical entries with just two primitives. A formalization can also work as a metalinguage to make sure that it is possible to speak about the human language using another tool, since it is a well-known fact that natural languages present some problems when used as a theoretical tool, such as ambiguities, which can disappear if some formalization is used to describe them. However, we should bear in mind that the formalization presented here is simple and quite intuitive.

The feature structure of *assado* would be:

---

5 Here, I have introduced the *den* feature that expresses the semantic denotation of expressions. I will discuss what semantics we adopt below and introduce it here and not before for didactic reasons.

6 For a discussion about the semantics of adverbials, see Guimarães (2007).

7 It is interesting to note that this kind of application of a syntactic theory at a morphological level has been proposed before (Selkirk 1982, Hoeksema 1985, Halle & Marantz 1993).
Categorial morphology: a semantic perspective on the morphological level

Real

Figure 4: Assado ‘baked’

\[
\begin{aligned}
\text{exp} & \quad \text{assado} \\
\text{cat} & \quad N\backslash N \\
\text{reg} & \quad \text{R1} \\
\text{con} & \quad \begin{cases}
\text{exp} & \quad \text{assar} \\
\text{cat} & \quad N\backslash S/N \\
\text{reg} & \quad \text{lex} \\
\text{sem} & \quad \lambda x \lambda y \lambda e [\text{assar}^*(x)(y)(e)] \\
\end{cases}, \\
\text{exp} & \quad -\text{do} \\
\text{cat} & \quad (N\backslash S/N) \backslash (N\backslash N) \\
\text{reg} & \quad \text{lex} \\
\text{sem} & \quad \exists e \exists y [\text{BECOME}(e)(y)]
\end{cases}
\]

5. Suffix –ura

As shown above, it is possible to find, with the same tool, categories not only for words, sentences and phrases, but for every lexical item we wish, for example, morphemes. In this work, I have chosen the suffix –ura from Brazilian Portuguese to demonstrate the functioning of the model. I will find its category with the same process described above and then I will analyze other words where the same suffix is present to verify the applicability of the new-found category.

So, I started from assadura ‘baking/rash’, not only because we have already found the categories of assar ‘to bake’ and assado ‘baked’, but because this is a special case concerning the semantics of nominalizations. In the following I will see how to deal with it.8

Assadura is formed by R1 (the rule that unifies two expressions: assado ‘baked’ and –ura (the suffix)) and has the category N, since it is a normal noun, assadura brings together two expressions –ura and assado, and the latter brings two more expressions: assar ‘to bake’ and –do ‘–ed’. Assado ‘baked’, as we saw, is an expression that from an atomic N forms a complex N. Then assado has a category N\N, formed by assar and –do. Assar, NS/N, is an expression that needs two Ns, one on each side, to form an S, as in Pedro assa o bolo ‘Peter bakes the cake’.9

-Do, then, is a suffix that from an N\S/N forms an N\N. Here is the most interesting point brought from CG to morphology: it is not necessary to postulate a new category for the suffix considering only one feature. From the selected word and the formed word category, I find the functor’s category (the suffix –do): if the suffix selects an N\S/N on the left forming an N\N, its category is (N\S/N)(N\N).

---

8 Other cases of nominalization involving the same suffix are words like alvura ‘whiteness’ and gordura ‘fatness’, nominalized from adjectives. This kind of nominalization, however, has a simpler solution than the cases I have chosen to deal with in the present work as I showed in Real (2006). The most important point is the different behavior of the suffix –ura when it is nominalizing adjectives and when it is nominalizing participles: when selecting an adjective, semantically it always results on the property of that adjective: the property of a white thing is its whiteness. As we saw, defining the semantics of nouns formed by –ura and participles are more complicated. Syntactically speaking, this model allows a generalization between the two different behaviors assumed by –ura: it nominalizes participles and nouns, but through this tool, we can see it nominalizes just N\N categories, since we can understand these two kinds of expressions as included in the same syntactic category. As nouns formed from adjectives having the same structure as nouns composed by participles and their semantics is simpler and was explored in Real (2006), I have chosen not to expose these nominalizations with –ura from adjectives and just note that –ura has more behaviors than those analyzed here.

9 Here I just do not speak about articles, nouns and their relation; we could just understand articles as N/N expressions, but much literature has been written about the role of determiners in CG. A reference list can be found, for example, in Heim (1988) and Heim & Kratzer (1998).
Now, with both categories (of *assado* and *assadura*), through the same mechanism, it is easy to find the category of –ura, that selects an \(N\backslash N\) (*assado*) and forms an \(N\) (*assadura*) and, thus, its category is \((N\backslash N)\backslash N\).

**Figure 5: -Ura**

<table>
<thead>
<tr>
<th>EXP</th>
<th>-ura</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>((N\backslash N)\backslash N)</td>
</tr>
<tr>
<td>REG</td>
<td>lex</td>
</tr>
</tbody>
</table>

And, consequently, *assadura*:

**Figure 6: Assadura**

<table>
<thead>
<tr>
<th>EXP</th>
<th>assadura</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>N</td>
</tr>
<tr>
<td>REG</td>
<td>R1</td>
</tr>
<tr>
<td>CON</td>
<td>{}</td>
</tr>
<tr>
<td>EXP</td>
<td>assar</td>
</tr>
<tr>
<td>CAT</td>
<td>(N\backslash S/N)</td>
</tr>
<tr>
<td>REG</td>
<td>lex</td>
</tr>
<tr>
<td>EXP</td>
<td>-ura</td>
</tr>
<tr>
<td>CAT</td>
<td>((N\backslash N)\backslash N)</td>
</tr>
<tr>
<td>REG</td>
<td>lex</td>
</tr>
</tbody>
</table>

It is worth noting that in the descriptions above I did not introduce the formalization of the expressions, because I have not discussed the different meanings of *assadura*: the process of baking (‘baking’) and its result (‘rash/baked’). In Real (2008), I showed that the same phenomena appear in many expressions formed by the suffix (-ura) as in *pintura* ‘painting/picture’, *abertura* ‘opening’, *fritura* ‘frying/fried food’ and I tried to explain these two possible interpretations looking at the semantics of the verb forming these expressions. However, these two possibilities apparently come neither from the verb’s semantic structure, nor from its thematic configuration (Real 2008). Before assuming that this behavior has pragmatic causes, I have tried to find an explanation for it on the semantic structure of the suffix.

In Real (2008), I analyzed the same process through the Subatomic Semantics proposed by Parsons (1991). However, there is no adequate treatment of nouns that can represent events or individuals in this model, despite Parsons’ famous treatment of nominalizations. Therefore, in the present work, I try to find a possible semantic treatment of this phenomenon based on the proposal of Bayer (1997).

Bayer (1997) suggests that there are more necessary basic semantic types than the two proposed by Montague (1973) (\(<e>\) for entities and \(<t>\) for truth values) to deal with natural language. According to the author, nouns that express an event may have a

---

10 In Portuguese ‘assadura’ can have different meanings: ‘baking’ (the process of baking), and ‘rash’. As ‘to produce a rash’ in Portuguese can be understood as ‘to bake’ here I consider the sense that ‘rash’ is the result of ‘baking’, even if in English the sense is not the same. In the next pages, I will discuss how to understand these two different meanings in Portuguese and a possible relation between them assuming that these two readings are, in fact, produced by an ambiguity present in the suffix ‘–ura’.
different semantic type from $e^{11}$, and therefore he proposes the use of a third semantic type: $\epsilon$. Adopting a new semantic type that characterizes specific events, Bayer, without using a complex formal approach, is able to make explicit the difference between eventive nouns and non-eventive nouns, eventive sentences and non-eventive sentences, and eventive nouns and eventive sentences.

As the relations between expressions and the transparency between them are the base of CG, if I am considering the semantics of the suffix –ura, I must reconsider all the semantics of the expressions formed by it (such as abertura) and see how compatible the semantics of these expressions are in relation to the semantics of the suffix. As speakers have a stronger intuition about words than about morphemes, I will check at the end of my experiment if besides being computable my proposal is also intuitive.

Thus, with Bayer’s semantics, the expressions that I distinguish are assadura (process) and assadura (resultive), as in:

(4) A assadura do bolo levou duas horas. (process)
   ‘The baking of the cake took two hours.’

(5) A assadura do bebê estava muito vermelha. (resultative)
   ‘The baby’s rash was very red.’

I call assadura in (4) assadura-p and in (5) assadura-r.

According to Bayer (1997), assadura-p expresses an event and, because of that, its semantic type is $\epsilon$. On the other hand, assadura-r has the Montagovian semantic type $e$, of individuals/things. Since they have different semantic types, the two expressions will have very different internal semantics, despite having exactly the same phonological and syntactic features$^{12}$. The most important innovation of Bayer (1997) is the fact that expressions of the same category do not have exactly the same semantic type (or equivalent semantic types).

-Ura-p (in assadura-p) is then a suffix that forms an expression typed $\epsilon$ (assadura-p) from an expression typed <e,e> (such as assado, ‘baked’) . On the other hand, ura-r, from the same expression, forms an expression of type $<e>$ (or $<<e,e>e>$, which is logically equivalent).

In Real (2008), I showed that the Subatomic Semantics proposed by Parsons (1991) solves some semantic questions, such as causality and quantification of events, although it does not make explicit the difference between eventive nouns (like abertura-p) and eventive verbs (like abrir ‘to open’). Since I am adopting here Bayer’s proposal, a Davidsonian one (i.e. that does not use thematic roles as a tool), using Parsons’ semantics with no arrangements will be impossible. This is because Parsons bases his semantics on thematic roles, but the formalization proposed by Bayer (1997) is too simple to explain the phenomena I am working with here, and that is why I am going to

---

$^{11}$ This idea is quite similar to what Pagani (2007) defends: that we cannot consider entities and events as the same type. Pagani argues that, logically, events are different from entities and they must be treated as second order expressions. Dowty (1979) and Parsons (1991) treat both events and entities as the same semantic types ($e$). On the other hand Schein (1993) and Pietroski (2005), in the same theoretical framework that Dowty (1979) and Parsons (1991) used – neo-Davidsonian semantics – also pointed out the necessity of second order logic to describe natural language semantics.

$^{12}$ As Bayer (1997) just argues for the new semantic type, but does not really discuss its influence on the syntactic level (and here, I note the strong CG parallelism between these two levels), I take all the expressions with the semantic type $\epsilon$ as a syntactic category $N$, but it seems a very important element to be considered by CG scholars.
work with an adaptation of Subatomic Semantics (Parsons 1991) for Bayer’s Type Theory (Bayer 1997). These two different approaches give a kind of Davidsonian semantics, but with a more complex typification and some operators (such as CAUSE, DO, BECOME) that easily express the causality notion, observed by some scholars as the most relevant relation to some ontological semantics (Moens & Steedman 1988, Pinker 2007).

Before analyzing the two different meanings of *assadura*, it is necessary to make sure we are dealing with two different words and not just the same word assuming different meanings in different contexts. To do so, I use one of the tests proposed by Carpenter (1997: 14-16) to distinguish vagueness from ambiguity. According to this test, if a word is vague, it is just one lexical item, but it has in its lexical structure some degree of underspecification and its meaning must be defined in context. On the other hand, if an expression is ambiguous, it corresponds with two different lexical items and must be characterized differently in the lexicon. Although Carpenter (1997) shows other tests to check ambiguity and vagueness, I reproduce here just one, the ellipsis test. In Real (2009), I reproduce more tests and check more words formed with the suffix –ura.

In this test, the expression to be tested appears in a sentence and it is then immediately followed by an ellipsis related to it. If the gap refers obligatorily to the same sense assumed by the explicit expression, the expression is ambiguous. It is vague if, at the same time, it admits more than one sense, as in:

(6) Mary considered the newspaper interesting, and Paul too colorful.

In the sentence above, *newspaper* refers to the informational content in a newspaper, but its ellipsis refers to the physical object, the pages, the paper. So *newspaper* is a vague expression, since it allows two readings in the same structure. Let us see if the same happens with *assadura*:

(7) Maria achou a assadura vermelha demais, e Paulo demorada demais.\(^{13}\)
Maria thought the rash too red, and Paulo long too.

Therefore, I will consider *assadura* as an ambiguous word, i.e., as a phonological expression that does not necessarily have the same syntactic-semantic structure since the ellipsis of ‘assadura’ and ‘assadura’ itself must have the same reading. This point of view is also put forward by Pustejovsky (1995: 27-28), who does not assume this ambiguity in his *Generative Lexicon*, but argues that in a model that assumes a lexicon as I do (which he calls *Sense Enumerative Lexicon – SEL*, that without ‘transformations’), this would be the only solution to capture the differences between the two readings of *assadura*. As I have shown, *assadura*-p and *assadura*-r have the same syntactic structure, i.e., they are formed by the same expressions and have the same category (a category which was found intuitively considering its normal use: a noun).

In this way, the only feature that differentiates *assadura*-r and *assadura*-p is the semantic feature, and, as pointed out above, I assume a Davidsonian semantics, even if

---

\(^{13}\) Unfortunately, we cannot translate this sentence and capture the same phenomenon, since in English the two readings from *assadura* are expressed by different words. The literal translation of the sentence (7) is: ‘Mary considered the rash too red, and Paul the baking too slow’. Another test that is considered by some semanticists is to compare the tested word to its equivalent in another language. If we consider this test, *assadura* will be obviously ambiguous, as in English the two ideas that *assadura* can generate are expressed by two different words (‘baking’ and ‘rash’).
it has strong bases in Subatomic Semantics (Parsons 1991). It means that in relation to the semantic types (the link between the syntactic category and the logical form in a lexical item), \textit{assadura-p} and \textit{assadura-r} are different: the former has the type \textlangle \varepsilon \rangle and the latter the type \textlangle e \rangle.

To define exactly what happens in these two behaviors, I adopt the events ontology proposed by Moens & Steedman (1988). Now I will briefly explain what the main concepts in this ontology are, and then consider the phenomena involving the suffix \textit{–ura} on this ontology.

5.1. Ontology

I choose this ontology and not any other, because here concepts like causality are well explained and, moreover, because this ontology does not pretend to be an ontology for verbs, i.e., it does not work with verbs as primitive units; it takes sentences, events or states as primitive units. I will also apply it here because this ontology works with sub-units. As we will see, through this mechanism, it will be easier to link the ontology adopted here and the formalization proposed by Parsons (1991), since it also works with sub-units and with the idea of compositionality.

Although Moens & Steedman (1988) do not propose a strong formalization, I think that if I am working both with the idea of compositionality, in a weaker version of Frege’s principle (as argued in Hoeksema (1985)), I will have an easier integration between this ontology and the morpho-lexical level.

In the ontology proposed, the authors classify kinds of events as \textit{eventives} and \textit{statives}. In stative events, it is possible to distinguish four possible kinds: \textit{lexical state}, \textit{progressive state}, \textit{consequent state} and \textit{habitual state}. On the other hand, events are classified by their atomicity ([atomic] or [extended]) and by their consequences ([+ consequent] or [-consequent]). For example, a point event is an event [-consequent] and [atomic], while a process event is a [-consequent] and [extended] event. Let us see all this classification on the table below:

\begin{center}
\begin{tabular}{|c|c|c|}
\hline
\textbf{EVENTS} & \textbf{STATES} \\
\hline
\textbf{atomic} & \textbf{extended} & \\
\hline
\textbf{+conseq} & \textbf{CULMINATION} & \textbf{CULMINATED PROCESS} \\
\hline
recognize, spot, win the race & build a house, eat a sandwich & understand, love, know, resemble \\
\hline
\textbf{-conseq} & \textbf{POINT} & \textbf{PROCESS} \\
\hline
hiccup, tap, wink & run, swim, walk, play the piano & \\
\hline
\end{tabular}
\end{center}

Let us see some examples and how Moens and Steedman categorize them:

(8) Sandra hiccupped.
(9) Sandra was hiccupping.

The sentence in (8) is considered by the authors as a point sentence: it is a [-consequent] (since it does not need a culmination) and [atomic] event. The sentence in (9) is a process (even if it is already finished).

Another interesting issue about this theory is the notion of CONTINGENCY. The authors base their theory on the idea of causality to create relations between eventualities. In this way, they can abandon the notion of temporality as the main feature in the characterization of these kinds of relations to adopt the notion of causal dependence, which is not necessarily temporal. In this way, the relations in (10)-(12) are easily captured, even if they are not temporal relations:

(10) When they built the 39th Street bridge, a local architect drew up the plans.

(11) When they built the 39th Street bridge, they used the best materials.

(12) When they built the 39th Street bridge, they solved most of their traffic problems.

(Moens & Steedman 1988: 15)

Also Moens & Steedman (1988) propose a notion to capture similarities between kinds of eventualities: the NUCLEUS. In this approach, each eventuality is composed by a nucleus, a complex structure composed of three other elements: PREPARATORY PROCESS, CULMINATION and CONSEQUENTIAL STATE.

Figure 8: Focus (Moens & Steedman 1988: 18)

Every eventuality is composed by a nucleus that makes the aspectual and temporal relations explicit, but some part of the nucleus can be focalized and this focus is based on the notion of contingency. Let us see how it appears:

(13) Sandra baked potatoes.

(14) Sandra had baked potatoes.

In both (13) and (14), the preparatory processes are the same (buying potatoes, washing them, preparing the kitchen, etc.); the culmination is the exact moment that the potatoes finished ‘baking’ and become ‘baked’. Even if in (14) the process does not necessarily reach the culmination, it is worth remembering that every eventuality has the entire nucleus, so it does not matter if some part of it is focused, since all three nuclei’s parts are present in every eventuality.

In this ontology, the difference between these two sentences is the focus. In (13), the culmination is the focus, while in (14) it is the preparatory process. Through the notion of nucleus, the authors can characterize both eventualities and their parts. It is
also through this concept that it is possible to unify kinds of events: when a state is considered, just a given aspect of its nucleus is focused, the other aspects necessarily remain there, and because of that, it is possible to refer morphologically or syntactically to the other aspects.

So, considering the notions presented (contingency, nucleus and kinds of eventualities), the authors propose a treatment of all possible relations between eventualities that the English language allows. Below, I present a graphic of these possibilities proposed by the authors:

**Figure 9: Relations between eventualities (Moens & Steedman 1988: 18)**

It is not possible, for example, to arrive at a progressive state from a punctual event straightforwardly. To make this transition it is necessary first to transfer the focus (syntactically or morphologically) from a punctual event to a process one, to then arrive at a progressive state.

Proposing a treatment of relations between eventualities, the authors establish the possible ways in which these changes can occur, without allowing these shifts to occur randomly. For example, through the iteration process from a punctual event, it is possible to arrive at a process event. In the latter, a culminated process is obtained or some suffix is added bringing progression and a progressive state is obtained.

Now I can return to the examples (8) and (9) and see how this approach connects these two sentences:

(8) Sandra hiccupped.

(9) Sandra was hiccupping.
To arrive at (9) (a progressive state) the sentence (8) (a punctual event) must receive some iteration element, in this case, the progressive morpheme. The phenomenon shown above can also be seen in the following table adapted from the table proposed by the authors:

5.2. Applying the ontology to -ura

Now I will try to explain the differences found between the two morphemes – *ura*. Both of them nominalize *assado* ‘baked’, that in Subatomic Semantics (Parsons, 1991) can be formalized as:

(15) \[ \lambda y \lambda x \exists e \left[ assar^* (y)(x)(e) \right] \land \exists y \left[ BECOME (e)(y) \right] \]

i.e., *assado* is an expression that denotes the existence of an event with two participants where one of them obligatorily becomes *assado* ‘baked’.
Bearing in mind the ontology presented above, *assado* represents a process that has already culminated: it is formed from *assar* (a process without culmination) and *–do* (an element that includes the culmination of the process), as can be seen in the next figure:

**Figure 12: Assado (adapted from Moens & Steedman 1988: 18)**

![Ontology Diagram](image)

*Assadura*-p and *assadura*-r are formed from *assado*, but have different meanings. Through this ontology, I can understand *assadura*-p as a process, just like *assar*: it is an expression that denotes an event with two participants, the event of *assar* (*‘bake’*):

(16) **Assadura-p:**

$$\lambda y \lambda x \lambda e [a s s a r' (y)(x)(e)]$$

In this sense, *assadura*-p has the same formal structure as *assar*, but they are different in relation to the syntactic category. Yet it is interesting to see that (17) below is grammatical: what was done was the nominalization of a verb, keeping all its semantic characteristics:

(17) O assar dos bolinhos demorou muito.

‘The baking of the cakes took long.’

As *assar* and *assadura*-p seem to have a similar behavior and are of the same kind of eventuality (a process), I understand the process to form *assadura*-p as:
The word formation process for *assadura* can be understood as:

(18) \((-\text{culmination}(-\text{ura}(+\text{culmination}(\text{-do(process(assar))})))))\)

where \(-\text{ura}\) just introduces the idea that the culmination was not necessarily reached.

On the other hand, considering *assadura-r*, that is also formed by *assado*, but that refers itself to the individual/thing that is formed after the baking process (*assar*), I would have a logical form as

(19) \(\lambda y \lambda x \exists e [\text{assar}^r(y)(x)(e) \land \exists w [\text{BECOME}(e)(w) \land \exists z [\text{CAUSE}(w, z)]]]\)

i.e., *assadura-r* denotes an event with two participants where one, when it has become baked (*assado*), produces a resultative individual/thing.\(^{14}\) In the ontology proposed by Moens & Steedman (1988), *assadura-r* does not have an explanation as obvious as *assadura-p* has. It seems that from a culminated process (*assado*), *ura-p* introduces an individual or a resultative state, although in this ontology there is no connection between the culminated process and the resultative state, as shown in Figure 9. This possibly happens because this framework was proposed in English or had no morphological concerns. To express the relation found between *assado* and *assadura-r*, a figure like the following is expected:

---

\(^{14}\) I would like to make clear that here I use the terms *participant* and *individuals* freely, as I did not discuss the nature of these elements, mainly, about the nature of the resultative thing after the process (the rash).
Here, the link between culminated process and resultative state would be possible. Even if in the framework this relation was not explicit, I think that this ontology helps to understand the relations between these phenomena and that through the semantics I adopt, it is possible to arrive at a good formalization for these two suffixes. I will show in the next sub-section how it is possible to formalize these two expressions.

5.3. A logical form for –ura

Obviously, since I have different denotations for the two different assadura, they will be composed of ura-p and ura-r + assado, the difference being in these morphemes’ denotation. This is to say that, if I understand that the morphological level has a semantic component, some features that would be explained just on the pragmatic level can be already sketched. Thinking about this semantic component on the morphological level, I can find more than syntactic categories of morphemes; it is possible even to define them formally:

\[(20)\]  \(Ura-p:\)
\[\lambda e[\neg C'ul(e)]\]

\[(21)\]  \(Ura-r:\)
\[\exists w \exists e[BECOME(e, w) \land \exists z[CAUSE(w, z)]]\]
Ura-p removes the culmination from the event that was added by the participle morpheme (–do, ‘-ed’), after this morphological process, the event has the same denotation as when it was in the infinitival form. Even if CG does not use transformations as a theoretical possibility since it is a monotonic grammar, it is possible to cancel an operation on the logical level. It is worth noting that I am not canceling-deleting a linguistic item; I am canceling semantic information that was removed from a later operation. In addition, both of them remain in the formalization, i.e., in a word such as assadura-p it is present in both the culmination and its cancellation. 15 Assar was an event without a culmination when expressed by assar ‘to bake’, gained the culmination in assado ‘baked’ after being selected by –do, and lost it when –ura-p nominalized assado, returning to its first denotation, denoting only the event without other elements as resultative individuals or culminations.

Concerning ura-r, I can say that the suffix does not remove the culmination from the event because it expresses exactly what came after this culmination: a resultative individual. Ura-r links the event (that could be assar ‘to bake’, abrir ‘to open’, etc.) to the emerging of a new individual, expressed by the operator CAUSE (Parsons 1991).

6. Conclusion

In the present work, I briefly introduced the notion of Categorial Grammar, explaining how its formal mechanism functions and how this tool can be applied to different linguistic levels, such as the morphological one. To show this possibility, I considered the morpheme –ura of Brazilian Portuguese, mainly its occurrences in nouns formed from participles. Syntactically speaking, I used the classical mechanism of CG, shown, for instance, by Hoeksema (1985). Concerning the semantic level, I used a combination of Subatomic Semantics proposed by Parsons (1991) and the typification proposal of Bayer (1997), that argues for a Davidsonian semantics, i.e., with no thematic holes.

After applying a test proposed by Carpenter (1997) that distinguishes vagueness from ambiguity, it was made clear that the words formed by –ura are ambiguous. As these words are formed by the same word-base and by the morpheme –ura, I considered the morpheme –ura as two different lexical items. Considering this and the fact that both of these meanings have the same syntactic category and phonological representation, I tried to build a formalization that could show the semantic role in the word formation process to make explicit what produces these differences. As a result, I gave the following feature structures:

**Figure 15: Ura-p**

\[
\begin{array}{cc}
\text{exp} & \text{–ura} \\
\text{cat} & (np\backslash np)\backslash np \\
\text{reg} & \text{lex} \\
\text{sem} & \lambda e \neg(\text{cul}(e))
\end{array}
\]

15 Here one can argue that with a monotonic logic, as CG pretends to be, we could not remove any introduced information. However the ontology adopted by Moens & Steedman (1988) gives the possibility of canceling some information (as the culmination) without deleting all of the information. Both the culmination and its cancellation are in the formula.
From the presented characterization, I showed Categorial Grammar can be applied to the morphological level, and also that its application demonstrates syntactic-semantic behaviors that that would probably not be captured without a strategy based on a bottom-up compositional analysis like the one presented here. This strategy shows, for example, that even on the morphological level it is possible to distinguish between different semantic behaviors and that I have certain theoretical gains in doing so, as it does not leave the explanation of some phenomena to pragmatics and uses the same rules at different linguistic levels, e.g. the syntactic and the semantic ones. This application shows that understanding the semantic component just on the speech interface, after all other morpho-syntactic processes, is less explicative than thinking of the semantic component spread on different linguistic levels.

References


http://www.rci.rutgers.edu/~mabaker/papersupdate.html


