

# An Integrating Framework for Anaphora Resolution

Dan Cristea and Gabriela Eugenia Dima  
“Al.I.Cuza” University of Iași  
Faculties of Computer Science and Letters  
*{dcristea, g.dima}@infoiasi.ro*

## 1. Introduction

The linguistic phenomenon of anaphora has been monopolising the attention of linguists and computational linguists for some time now, leading to various interpretations. Basically, anaphora is the phenomenon of reiteration of an entity (called “antecedent”) by a reference (called “anaphor”) that points back to that entity. For practical reasons, we will call “referential expressions” (REs) both participants in an anaphoric relation. Actually, during the reading of a text, it is very likely that an anaphor become, in its turn, an antecedent for another co-referential anaphor that follows it. In such case we will use the term “anaphoric chain” to denote the textual relation of the co-referential REs. The process of identifying the antecedent of an anaphor is called “anaphora resolution” (AR). Its automatisation represents one of the main preoccupations of computational linguists, as anaphora resolution is extremely important in most NLP applications, from machine translation to automatic summarisation and information extraction.

In order to develop an anaphora resolution system, it is necessary to deeply understand the nature of the referential process in discourse and the problems behind it. Since 1976, Haliday and Hassan stressed that the anaphoric relation is a semantic, and not a textual one. Although it is unanimously agreed that semantic features are essential for anaphora resolution, the authors of automatic systems devised so far preferred to avoid the extensive use of semantic information [Lappin and Leass, 1994], [Mitkov, 1997], [Kameyama, 1997]. This choice, motivated by the difficulty and complexity of achieving a correct semantic approach, has had as a consequence the fact that an algorithm of anaphora resolution with a very high degree of success has not been found yet.

We therefore believe that considering semantic features when dealing with anaphora resolution is the key point to the improvement of the current automatic systems. Our intention in proposing an anaphora resolution framework is to provide a background for a solution that takes into account the semantic nature of discourse constituents and could function in some of the cases that often constitute a problem for the automatic systems developed so far.

Thus, besides stressing the importance of analysing the word and the discourse beyond their formal appearance, the paper includes a presentation of some of the problems raised by co-referential anaphora resolution and proposes a three-layer framework, which could integrate most existing approaches to anaphora resolution.

In the Sections 2 to 7 we will discuss different discourse phenomena that interfere with the identifications of antecedents. As this analysis evolves, the need of a layered general representation of co-referentiality will be justified. In Section 8 we will formalise the notion of model and propose a general architecture of a system that could accommodate any such model. Our frame applies only to incremental types of text processing therefore rejecting a-priori any bi-directional or looking ahead models.

## 2. Discourse is more than mere words

Words are much more than mere conventions that designate things from the real world. Starting with the school of Ferdinand de Saussure, linguists have acknowledged the conceptual nature of words. Saussure's famous dichotomy defining the linguistic sign (**signified/significant**) introduces the idea of a concept present behind every word in the mind of a speaker<sup>1</sup>. Both Saussure and the linguists that have followed and developed his theories do not consider the entities from the real world as playing any part in the act of speaking; in this regard, Hjelmslev (1961) stresses that the nature of the signified is a purely mental one. A different approach, that of Ogden and Richards (1923), leads to a similar conclusion. Though the semiotic triangle proposed<sup>2</sup> does include, as a third component, a material **referent**, its role is considerably minimised. Moreover, in analysing the three components of the triangle (symbol, reference, referent), Ullmann (1962) states that the study of the referent and its relation with the reference should be the concern of philosophy or other sciences, while linguistics should only deal with the reference-symbol relation.

Various attempts to include the referent as a real material part of the speech act have led to terminology confusions, which have often caused a perspective confusion as well. In this respect, Tanaka, in his Ph.D. thesis (1999, p. 230), is trying to make a compromise between various theories, proposing *to define the notion of «referent» simply and vaguely as «what the addresser refers to», and not to equate it with a verbal expression nor a mental representation*. To justify this approach, he presents the following example, when a person **X** refers to a well-known personality (in this case the former South-African president Nelson Mandela) that he also actually sees with his eyes:

*X: Look! That's Mandela.*

*Y: Where?*

*X: Over there. At the corner. [Tanaka, 1999, p. 229]*

Tanaka sustains that **X** is not pointing to the mental representation of the South African president in **Y**'s memory, but directly to the referent (the president itself). The question that **Y** asks, *where?*, is meant to identify someone existent in the immediate context and has nothing to do with the mental representation **Y** has on Mandela. On the contrary, we would say that Tanaka is mistaking a representation of the person that includes the place where the person is located for the person itself. It is likely that both **X** and **Y** would have mental representations of the referent Mandela. These can be noted as [**Mandela-X**] and [**Mandela-Y**]. **X**'s exclamation is relative to his own representation, to which he has

suddenly added a new attribute (**place**) with a determined value (**there**), updating it to [**Mandela-X: place=there**]. At the same time, the question asked by **Y** is not intended to identify the concept of Mandela, already familiar as [**Mandela-Y**], but to complete it, as it had evolved to [**Mandela-Y: place=unknown**].

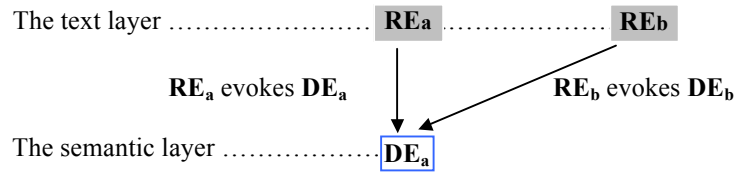
The above situation is not much different from what happens at a dinner table, for instance. When one says: *Would you pass me the salt, please!*, his mind has built a representation of the saltcellar and is operating with this one, rather than with the image of the saltcellar that is lying on the table. Actually, at the moment he makes his request, he might not even know how the saltcellar in that particular restaurant looks like. Nevertheless, when he expresses his request, he unconsciously knows that his table partner has a similar representation in his own mind, and believes this representation will be used by the partner to identify the requested object and to support the actual change of its position towards him. He also has the inner belief that this will cause his own senses to perceive the movement, and, as a result, to update the representation he has of the saltcellar. One problem that arises from the presence of a mental concept behind the rigid form of a word is that any mental concept is closely related to the mind that has conceived it. Nevertheless, when the two people having dinner speak about “salt”, even without knowing the exact representation of the concept in each other’s mind, they manage to refer to the same thing. Their understanding is based on the existence of some steady semantic features that characterise the concept. These provide the minimal necessary frame, which identifies that concept as unique, distinguishing it from the multitude of other concepts.

Sometimes the distinction is not possible at word level, requiring the presence of a context in order for someone to be able to establish which set of characteristics is the one referred to. This is mainly the case of polysemantic words. For instance, in the sentence *The board has decided to increase the capital of the company*, when reading *The board*, several representations could be automatically constructed in the mind of the reader: [**plank**], [**table**], [**committee**], [**display panel**], [**printed circuit**], etc., each with its own set of semantic characteristics. The ambiguity is resolved only later when the verb “has decided” allows the selection of the correct representation on the basis of its own semantic features: **making a decision** → **mental activity** → **animate agent**. As such, the meaning candidates of the word *board* are filtered out to that particular concept that implies an animate agent: [**committee**].

### 3. Representation of anaphoric relations

To complete the set of semantic characteristics attributed to a discourse entity (DE) it is imperious to have in view the co-referentiality of notions in a discourse. In this respect, endophoric references, in the form of anaphora/cataphora, are essential. Halliday and Hassan consider the anaphoric function as crucial in creating cohesive links within sentences. As we share their opinion that *the reference items must match the semantic properties of the item referred to* [Halliday and Hassan, 1976, p. 32], we believe that an

anaphora model should necessarily take into account the semantic representation of the words involved in discourse.



**Figure 1: Two-layer representation of a co-referential anaphoric relation**

**Figure 1** displays the most common situation of two referential expressions, which are co-referential. The antecedent and the anaphor are located on the text layer (**RE<sub>a</sub>**, respectively **RE<sub>b</sub>**). On a deeper semantic layer, their common representation is marked as **DE<sub>a</sub>**<sup>3</sup>. As the reading moves on, a semantic representation is first born when **RE<sub>a</sub>** is encountered. Then, at a later moment, when **RE<sub>b</sub>** is read, it evokes the DE already built by **RE<sub>a</sub>**. Any subsequent co-referring referential expression will in its turn evoke the same DE<sup>4</sup>. One way of representing discourse entities in NLP systems is as feature structures, which are lists of attribute-value pairs. The exact configuration of these attributes, as well as their types (range of accepted values) should be evidenced by the anaphora resolution model.

#### 4. Time and discourse

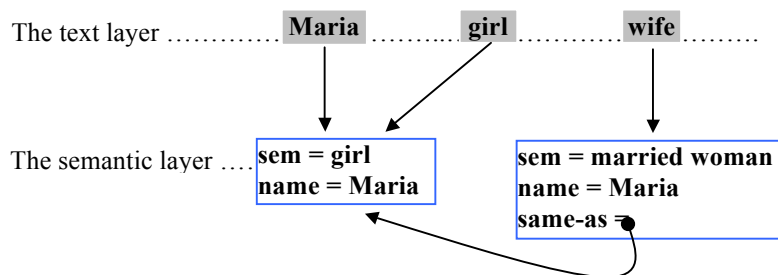
Another issue to be considered is the dynamic nature of discourse and of its entities. The reading of a text can be projected on three distinct time axes: the **real time axis**, including the process of reading and whatever else the reader might do between two pages or two chapters of a book, the **discourse axis**, which takes into account only the text, in its linear processing by the reader's mind and the **story time axis**, which relates only to the events described in the book. Thus, it becomes obvious that a discourse is different than the text it originates in. Libraries are depository of texts, not of discourses. A discourse is a text in the process of reading, a text interpreted by a human mind, during the lecture. When the reading process comes to an end, the discourse is also closed and only a representation of it remains in the reader's memory.

The anaphora resolution process takes place on the discourse axis, being similar to the progress made by the reader when "processing" one page after another. As a consequence, on this axis the entities that participate in a discourse are not constant in time, as new or different characteristics may appear at different moments and they are added to the initial ones. Although this is not always obvious, and sometimes not even relevant, in many cases it becomes drastic to recognise the changes in the entities introduced in discourses in order to establish a correct anaphoric chain. In a discourse such as:

*Maria was dreaming. The girl had just finished high-school and was engaged to be married. She imagined her wedding night in the arms of her beloved and imagined herself as his loving and caring wife. He would have made her a woman, and a mother, and they would have grown old together in peace and happiness.*

Maria is referred to by a series of nouns (girl, wife, woman, mother), which clearly describe a progress from one stage to another.

To mark these changes, a different notation on the semantic layer between the first element of an anaphoric chain and each subsequent anaphor could be used in order to preserve the information relative to their successive realisation in time. **Figure 2** shows a proposal for such a notation. Nevertheless, the **same-as** link should not be included unless a difference in the content of the semantic representations is inevitable.



**Figure 2: Representation of entities variable in time**

## 5. Agreement/disagreement in morphological features

The complexity of semantic analysis may prove to be useless in the case of pronominal anaphora. Pronouns express very general notions, lacking almost completely semantic features. For instance, the following sets of characteristics can be identified for the 3<sup>rd</sup> person personal pronouns in English: *he* [+animate, +male, +singular], *she* [+animate, +female, +singular], *it* [+inanimate, +singular], *they* [+plural]. It is obvious that these characteristics are relatively poor; besides, they are not even restrictive as a *boat* can be referred to as *she*, while a *baby* is often referred to as *it*. Moreover, in languages that have grammatical gender distinction for common nouns (French, Romanian, etc.), the co-reference depends on morphological criteria. There is no semantic reason in French, for example, that *voiture* (En. *car*) should be referred to by a feminine pronoun, while its synonym *automobile* (En. *automobile*) should be referred to by a masculine pronoun (En. *he*). Still, the language does not accept a different type of anaphor than the one mentioned.

Barlow (1999) argues that morphological criteria are not real restrictions and, therefore, may lead to errors in the resolution of anaphora. This could, indeed, be a problem in English, where concord/agreement rules take more often into account semantic properties of words, embracing a multitude of aspects difficult to synthesise. It is very common, for

instance to co-refer a collective noun (therefore with the grammatical number of singular) with a plural pronoun having in view the fact that the semantic representation of the noun includes more than one agents:

*I think that they are a manufacturer?*

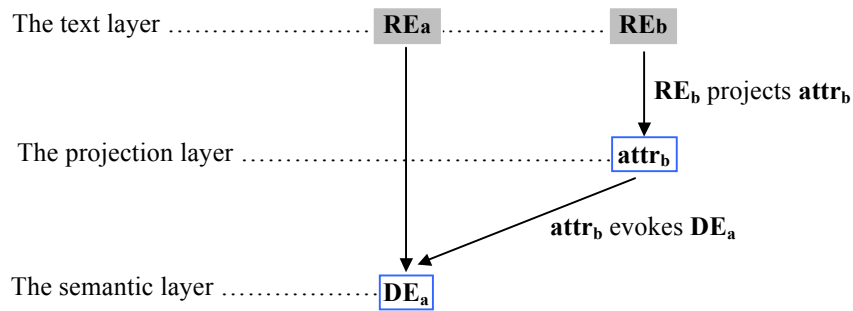
On the contrary, in other languages, such as Romanian or French, morphological restriction may constitute the only criterion that could lead to a correct anaphora resolution:

*Un camion (m.) a heurté une voiture (f.). Celle-ci (f.) a été complètement détruite.  
(A truck hit a car. This was completely destroyed).*

The resolution of the pronoun *celle-ci* in the above example is immediate, as gender match (feminine) indicates the noun *voiture* as its only possible antecedent, while eliminating the other semantically possible candidate, *camion*. The lack of grammatical gender in English does not permit such a quick process, the discourse remaining ambiguous.

There are cases, as Barlow noticed, when concord rules are apparently refuted by instances in which concord seems to be unobserved. The Spanish example quoted by Barlow, *Su Majestad suprema ... , él se mostró muy emocionado* [Barlow, 1998, p. 37], in which an apparently feminine phrase (*su majestad suprema*) takes a masculine reference (*él*), raises a problem beyond regular concord: that of gender ambivalence. Most languages acknowledging gender distinction have a number of nouns or phrases that can be referred to by both masculine and feminine pronouns, according to the natural gender of the person designated. This is, for instance, the case of nouns naming professions (*doctor, teacher, professor*). A solution in this case may be the introduction of the semantic category of natural gender, which could be retrieved from the context, and would allow a correct resolution, or, at least, would not prevent it.

Thus, it is possible to identify a set of morphologic (possibly also syntactic) restrictions for each language, which would be useful in eliminating candidates that are not acceptable in the system of a particular language. In any case, the identification of these characteristics for the purpose of anaphora resolution suggests the necessity to introduce an intermediate layer in the representation: that of the *morpho-syntactic-semantic attributes* as feature structures. Consequently, any characteristic the text layer contributes to **projects down** to this layer, as in **Figure 3**. Therefore, the anaphora interpretation becomes a two-step process<sup>5</sup>. The evoking mechanism seems to be based on this intermediate layer, and not on the text itself, and is more or less a language-dependent set of rules.



**Figure 3: Three-layer representation of co-referring REs**

The introduction of a third, intermediate layer is not meant to modify the previous (classic) two-layer model, but only to refine it when aspects related to processing are to be considered. As shown in the figure in relation with RE<sub>a</sub>, when the resolution process is completed, the intermediate layer should be disposed of.

## 6. Cataphora

Cataphora is usually defined as the referential relation in which the element referred to is anticipated by the referring element, usually a pronoun.

Scepticism about the existence of cataphora has been expressed by a number of scholars such as Bolinger (1977), Stockwell (1995), Cornish (1996), etc. Sometimes this scepticism has even embraced a categorical denial. See for instance Kuno (1972), who believes that all seemingly cataphoric pronouns must have their co-referential expressions somewhere in the preceding text.

This firm position is refuted by Carden (1982), as Tanaka (1999) notes, who provides evidence, within the 800 examples he collected from corpora, of pronouns that represent the first mention of the discourse entity in the text. Examples of the type “when he realises something, X (= he) does something else” are frequent in newspaper articles and television programmes:

*When he became president, George Bush renewed his appeal for a "kinder, gentler nation." (Compton's Interactive Encyclopedia, 1995, title word Bush)*

Gordon and Hendricks (1997), found that pronoun – name co-reference is frequent in situation when a subordinate clause or a prepositional phrase, containing the pronoun, precedes the main clause, containing the co-referent noun. Also in literary texts writers often prefer to introduce a character with a pronoun, which has a sum of characteristics attached, and only later his name is mentioned:

*“From the corner of the divan of Persian saddle/bags on which he was lying, smoking, as was his custom, innumerable cigarettes, Lord Henry Wotton could just catch the gleam of the honey-sweet and honey-coloured blossoms of a laburnum...”* (O. Wilde – *The Picture of Dorian Gray*)

Another sceptical attitude towards cataphora is expressed by researchers who stress the sequential linear processing of discourse: since texts are linearly processed by humans, from their beginning to their end, there is no way one would look towards the end of the text in order to recuperate an antecedent. In such a view, which coincides also with ours, the moment the pronoun is processed, an empty discourse entity is introduced. The subsequent co-references are linked to this entity and eventually add new features.

The idea that we operate with mental representations, nowadays supported by most researchers, is actually much older. In old literature (both drama and poetry), the actual plot used to be introduced by a prologue in which the events were either briefly narrated, or mimed, so that the audience could be familiarised with the protagonists and their deeds. These introductions were judged necessary to avoid possible confusion created by the intricacy of the plot. Practically, it was a simple way of allowing the public (or the reader) to construct general mental representations, which were to be completed later, with characteristics resulting from the facts presented in the play or in the poem<sup>6</sup>.

In modern terms, we could say that such prologues would provide a mental representation for each entity in the form of a general frame, including only minimal characteristics. This representation is later updated and its complete resolution is expected at the end of the play/poem. For instance, when Shakespeare states that his story is about *a pair of star-cross'd lovers*, he practically introduces two semantic entities, a **she** and a **he**, who are characterised by the only relation of being in love. Though there is no precise mention about to whom the writer refers, two distinct entities are already shaped and details given further will always refer back to these entities. The public or the reader do not need to wait until a proper description of him or her is given but they can very well operate with these two general concepts for as long as the writer considers it appropriate.

The preference, or even the necessity of the human mind to create mental concepts is also demonstrated by the analysis of a particular phenomenon of the Romanian language, namely the anticipation of the direct or indirect object by a pronoun. In Romanian, the sentence:

*I taught Gabriel to read.*

is normally translated as

*l-am învățat pe Gabriel să citească. (Him I taught Gabriel to read)*

where l is a personal pronoun, 3<sup>rd</sup> person singular, masculine, accusative, anticipating *Gabriel*. Though this anticipation is not compulsory, grammar books consider the construction without anticipation as archaic, recording the tendency of the language

towards an extensive use of this process<sup>7</sup>. It is a fact that a Romanian speaker would consider an utterance without the anticipatory pronoun (*\*Am învățat pe Gabriel să citească.*) as odd or even incorrect. On the contrary, if the nominal direct object is missing (*L-am învățat să citească.*), the sentence is perfectly acceptable, although the amount of information is considerably reduced. This is due to the fact that the noun is perceived as only adding details or completing the meaning of the pronoun. The noun loses its function of central semantic unit, which is taken by the entity first mentioned in the discourse, in this case by the pronoun.

The main function of such anticipation (which occasionally occurs in other Romance languages as well) is to provide the reader/listener with a notion that allows him to produce in his mind a general representation of an entity, which may (or may not) be disclosed later. The disclosure constitutes a mere addition of data, that go back to complete the set of characteristics of an already existent entity.

In fact, the notion of cataphora is acceptable only when it refers to a morpho-syntactic reality. From a morpho-syntactic viewpoint, the pronoun is a part of speech which replaces/substitutes a noun. Hence, for functional reasons, it is necessary to analyse the syntactic relation between the noun and its substitute and, therefore, the existence of the term “cataphora” (or “backwards anaphora”, as it is also called) in this context is particularly helpful. Nevertheless, such a positional distinction becomes empty of meaning in an approach that tries to emphasise the cognitive processes involved in the interpretation of texts. Since, as already mentioned, it is impossible to detach the discourse from the moment of its interpretation, the association of a unique directionality to the process of discourse interpretation, the one of its linear span (of its reading or of its writing), becomes evident. As a consequence, on this axis, the direction of the anaphoric relation is always towards the beginning of the text.

## **7. Anaphora and/or cataphora: the resolution moment**

In some cases, the anaphora resolution moment may be delayed until other discourse elements intervene to elucidate the anaphoric co-reference. This is the case in the following example [Tanaka, 1999, p. 221]:

*Police officer David Cheshire went to Dillard's home. Putting his ear next to Dillard's head, Cheshire heard the music also.*

The disambiguation moment of the pronoun *his* is the moment the reader processes *Dillard's head*. An inference allows the recuperation of [**Cheshire**] instead of [**Dillard**], since they were the only characters in the story and, by pragmatic knowledge, the system should recognise that a man cannot put part of his head next to the head itself. Therefore, the resolution moment is not that of the pronoun reading, and neither that of the succeeding co-referential proper noun reading, but intermediate. The proper noun reading will, perhaps, strengthen the belief that the antecedent is [**Cheshire**], as inferred.

Also, sometimes the pronoun resolution may be based on an inference without an explicit restating [Tanaka, 1999, p. 252]:

*The government contended Jacobson, 48, former big-time horse trainer turned East Side real estate operator, killed Tupper because Miss Cain, **his** live-in girlfriend of five years, moved from his apartment to **Tupper's** just down the hall.*

It becomes obvious from these examples that pronoun anaphora resolution might not be immediate even in the case of classical anaphora. When a reader cannot be sure about the antecedent at the encounter point (*his*), the final resolution will be deferred until he has enough information to disambiguate it from the later input (in this case, *Tupper's*). It is therefore correct to assume that, in general, there is a distinction between the point when a reader encounters a pronoun and initiates its interpretation (**initiation point**), and the point when the reader completes the interpretation of the pronoun (**completion point**). As Sanford and Garrod (1989) note, the gap between the two points can be nil, when a reader resolves a pronoun at the very moment he encounters it, or it can be extended to the end of the phrase, clause, or sentence in which the pronoun is included. During the gap between the initiation point and the completion point, humans retain the information obtained by processing the co-text of the pronoun in some kind of temporary location of memory until the resolution is stabilised.

When automatic processing is to be considered, the amount of indecision could be significant due to possible lack of resolution capabilities, or absence of enough knowledge in the implemented processors. Consideration of the processing weakness in a practical setting becomes important for the overall behaviour of the system to such an extent that it must be considered from the beginning a designing factor (parameter) and, as such, included into the framework. The delayed resolution could be supported in our representation by the introduction in the feature structure on the projection layer of an attribute called **candidates**, which stores the list of potential antecedents when the feature structure on the projection layer is retained after unsuccessful evoking. This attribute is essential in the evoking process, as it will be shown in Section 9.

## **8. Automatic anaphora resolution: a framework**

After enumerating some of the complex aspects involved in the identification of anaphoric links, this section describes a general framework for anaphora resolution, capable to integrate or accommodate a large majority of the existing approaches. The central notion in the framework we propose is that of **anaphora resolution model**. Such a model has four components, as follows.

**Component 1** consists of a **set of primary attributes** that fill the projection layer and are then reported to the semantic layer. Each specific attribute will contribute to establishing a relation involving the pair antecedent-anaphor.

The attributes, as mentioned in most of the existent approaches, could be classified according to the following categories:

**a. morphological:**

- number;
- lexical gender;
- person.

All the approaches use morphological criteria to filter out antecedents. However, as already mentioned, the elimination of possible referential links based on mismatches in morphological features may lead to errors. Though we do not share Barlow's view (1998) in this respect, that morphology should be ignored, a less categorical approach would be preferable.

**b. syntactical:**

- full syntactic description of REs as constituents of a syntactic tree [Lappin and Leass, 1994];
- marking of the syntactic role for subject position or obliqueness (the subcategorisation function in respect to the verb) of the REs: all CT based approaches [Grosz, Joshi and Weinstein, 1995], [Brennan, Friedman and Pollard, 1987], syntactic domain based approaches [Chomsky, 1981], [Reinhart, 1981], [Gordon and Hendricks, 1998], [Kennedy and Boguraev, 1996];
- quality of being adjunct, embedded or complement of a preposition [Kennedy and Boguraev, 1996];
- inclusion or not in an existential construction [Kennedy and Boguraev, 1996];
- syntactic patterns in which the RE is involved, that can lead to the determination of syntactic parallelism [Kennedy and Boguraev, 1996], [Mitkov, 1997].

**c. semantic:**

- position of the head of the RE in a conceptual hierarchy (hypo/hypernymy): all models using WordNet [Poesio, Vieira and Teufel, 1997]. Features as animacy, sex (or natural gender) and concreteness could be considered simplified semantic tags derived from a conceptual hierarchy;
- inclusion in a synonymy class that is determined by the context;
- semantic roles, out of which selectional restrictions, inferential links, pragmatic limitations, semantic parallelism and object preference can be verified.

**d. positional:**

- offset of the first token of the RE (a NP) in the text [Kennedy and Boguraev, 1996];
- inclusion in an utterance, sentence or clause, considered as a discourse unit [Azzam, Humphreys and Gaizauskas, 1998]. This feature allows the calculation of the proximity between the anaphor and the antecedent. These can be either intra-unit (c-coindexation criteria can apply only for intra-sentence RE pairs) or inter-unit (many models use the number of units between REs, in the order established by the domain of referential accessibility, see Component 4 below).

**e. surface realisation (type):**

- the domain of this feature contains: zero-pronoun (also called zero-anaphora or non-text string), clitic pronoun, full pronoun, reflexive pronoun, possessive pronoun, demonstrative pronoun, reciprocal pronoun, expletive "it", bare noun

(undetermined NP), indefinite determined NP, definite determined NP, proper noun (name).

**f. other:**

- inclusion or not of the RE in a specific lexical field, dominant in the text (this is called “domain concept” in [Mitkov, 1997]);
- frequency of the term in the text [Mitkov 1997];
- occurrence of the term in a heading [Mitkov 1997].

**Component 2** includes a **set of knowledge sources** fetching values for the attributes during text processing. The type of processing that we stick to is an incremental one, considering text in the order it is read.

What we understand by a knowledge source is a virtual processor able to fill in values for one single attribute on the projection layer, for instance number, or gender, or part of speech, or syntactic role. Practically, current processors simultaneously give values for more than one such attribute. Thus, a morpho-syntactic description tagger represents several knowledge sources as it provides more than one attribute for the head word of the RE [Tufis, 2000].

At least two knowledge sources are fundamental to all possible models of anaphora resolution: a part of speech tagger, to associate tags to every token of the text, and a shallow parser, capable to recognize REs by grouping tokens into NPs. These should be structured as head-modifier compounds. As there are NPs which cannot be REs (i.e. *the bucket* within the verbal phrase *to chick the bucket*), a parser should reject such noun phrases, by running a set of regular expressions to discover phrasal units.

Besides the two processors above, currently existing systems use additional knowledge sources. For instance, Kennedy and Boguraev (1996) introduce a marker of syntactic function and a set of patterns which recognises the expletive “it” (near specific sets of verbs or as subject of adjectives with clausal complements). Azzam, Humphreys and Gaizauskas (1998) use a syntactic analyser, a semantic analyser, and an elementary events finder. Gordon and Hendrick (1998) employ a surface realisation identifier and a syntactic parser, while Hobbs (1978) requires, for his semantic approach, a syntactic analyser, a surface realisation identifier and a set of axioms to determine semantic roles and relations of lexical items.

**Component 3** contains a **set of heuristics or rules** intended to co-operate in order to answer one or both of the following two questions, in this order: (1) Does a RE introduce a new discourse entity? (2) If not, which one of the existing DE does it co-refer? The heuristics/rules are intended to perform the evoking phase between an existent structure belonging to the projection layer and discourse entities from the semantic layer.

In accordance with most authors, we propose to accomplish this process by two types of rules:

- demolishing rules (applied first), which rule out a possible candidate. These rules lead to a filtering phase that eliminates from among the candidates those discourse entities that cannot possibly be referred to by the RE under investigation;
- promoting/demoting rules, which increase/decrease a salience factor associated with an attribute. These rules allow a subsequent selection phase, in which the best candidate is chosen from the ones remaining after the demolishing rules have been applied or a new entity is introduced.

The above rules are sometimes based directly on the values of the attributes brought by the knowledge sources. Other time, they are applied through binary relations, which should be calculated on the basis of primary attributes.

For instance, in order to rule out possible candidates, Kennedy and Boguraev (1996) implement conditions that prevent a pronoun to co-refer a constituent (NP) which contains it. Thus, in *the child of his brother*, *his* is neither *child*, nor *brother*, but a different entity. For the remaining candidates, they compute the salience by weighing a set of attribute-values pairs. The weights are linguistically and experimentally justified (cf. [Keenan and Comrie, 1977], [Lappin and Leass, 1994]).

Gordon and Hendricks (1997) show that the antecedent's syntactic prominence (notion related to the height in a syntactic tree) influence the selection of the co-referential candidate. In Gordon and Hendricks (1998), the salience of the relations between names and pronouns is calculated by using a graduation of surface realisation pairs: name-pronoun > name-name > pronoun-name. Therefore, if the surface realisation of the anaphor is a name, then a candidate whose surface realisation is a name will weigh higher than one whose surface realisation is a pronoun. Although not explicitly mentioned, it is probable that, if the surface realisation of the anaphor is a pronoun, then a candidate whose surface realisation is a name will weigh higher than one whose surface realisation is a pronoun.

**Component 4** contains a set of rules that configure the domain of referential accessibility. This component has two tasks: to establish the semantic DEs that are closed to being referred and to order the remaining ones according to prominence criteria.

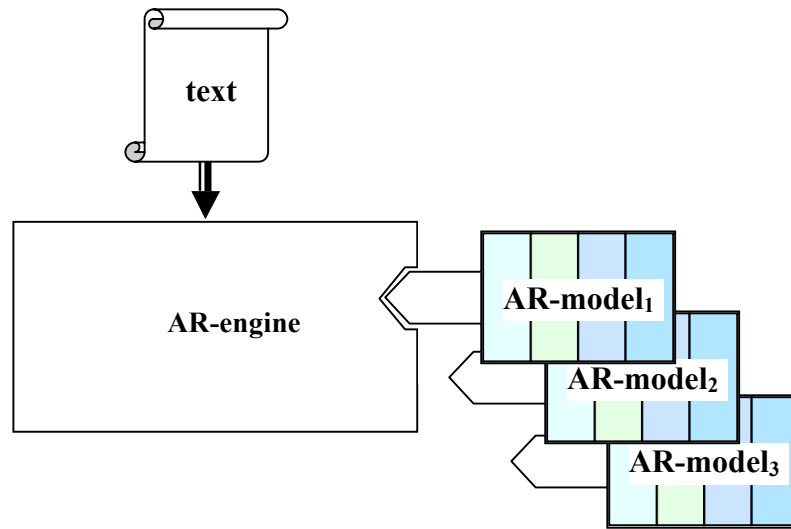
With respect to the first task, one filtering criterion could be given by the c-command relation, according to which it is impossible that a RE be co-referent with another RE that precedes it and c-commands it [Hobbs, 1978], [Gordon and Hendrick, 1998]<sup>8</sup>. This filtering criterion applies only to REs which are in the same sentence.

For the second task, different strategies could be imagined. The naïve one is, of course, the linear order, which places all the discourse entities found so far in the linear order of the text. Following this ordering, among the set of discourse entities with respect to which the current anaphor is computed the same salience score, the one closest to it is selected. In the attentional-based approaches [Grosz and Sidner, 1986] the accessibility in the current discourse unit is given by the top-down order of states in a focus stack. Focus-based approaches [Sidner, 1981], [Azzam, Humphreys and Gaizauskas, 1998] use

registers for current focus, alternate foci list, which are updated after each sentence and define an order in which to look for the antecedent. Gordon and Hendrick (1998) speak about “ordering of entities in the discourse order that determines the accessibility of those entities as referents for subsequent expressions” in their Discourse Prominence Representation approach. Vein theory (VT) [Cristea, Ide and Romary 1998] orders the units of the preceding discourse according to an extended domain of referential accessibility (E-DRA) [Cristea, Ide, Marcu and Tablan 2000]. In VT the DRA of a unit is given by the discourse structure and identifies a sequence of discourse units, called vein, in which, most naturally, the REs contained in that unit could find their antecedents. The E-DRA is the DRA plus the remaining units, as the least prominent.

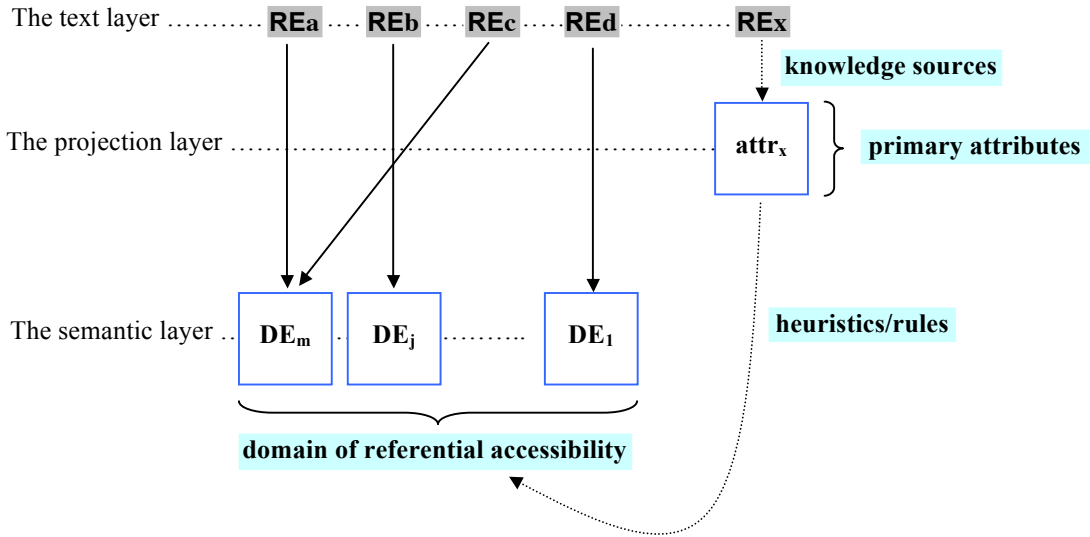
### 9. Behaviour of the AR-engine

In modern approaches, NLP applications which require anaphora resolution perform it by a distinct module which could be plugged-in the system. In our view, the AR module is a specialised engine that, implementing the framework described above, does its job helped by a specific AR-model (as in Figure 4).



**Figure 4. AR-engine accepts plug-in AR-models**

Based on the three-layer representation proposed in Section 5, the four components of an AR-model, can be accommodated into the framework as shown in Figure 5.



**Figure 5. Integration of the four components of an AR-model into the framework**

The text is processed sequentially. On the text layer the input tokens are marked due to the contribution of the two compulsory knowledge sources (POS tagger and shallow parser). At the beginning of the processing, when no text is read yet, there is no feature structure built on the projection layer and the space for discourse entities is empty. At a certain moment during processing (see Figure 5), the module processes a certain RE ( $RE_x$  in the figure) which belongs to a certain discourse unit. The figure shows the case when all preceding REs have been resolved. As the projection layer is only meant to support REs under resolution, no structure is maintained on this layer for the resolved REs. On the contrary, the semantic layer contains structures corresponding to all discourse entities already found ( $DE_1, \dots, DE_m$ ), ordered according to the prominence criteria given by the current conditions pertinent to Component 4. At the same time, these discourse entities define equivalence classes among referential expressions on the text layer, so that all co-referential expressions refer to the same DE. There is one link that is established between each RE and a specific DE, as in [Kennedy and Boguraev, 1996].

The moment  $RE_x$  is read, or shortly afterwards (to enable the existent knowledge sources to acquire enough information), a representation on the projection layer is built ( $attr_x$ ). The processing moment of a RE may be delayed in case some knowledge sources require more text than the span of tokens included in the RE itself or in the immediate preceding context. For instance, some semantic features of REs are identifiable only after the processing of the verb they relate to, while zero pronouns can be identified only after the verb is located.

Suppose the feature structure  $attr_x$  on the projection layer has the list of attribute-value pairs:  $a_i = v_i, i \in \{1, n\}$  and  $DE_1, \dots, DE_m$  is the ordered list of accessible discourse entities at the current moment, with  $DE_1$  – the most prominent. Suppose the attribute-value list of  $DE_j$  is:  $a_i^j = v_i^j, i \in \{1, n\}$ . The **candidates** attribute of the  $attr_x$  feature structure is a  $m$ -long vector of numeric value pairs: the first number is an index of a corresponding DE



with previous discourse entities), and the domain of referential accessibility (able to filter and to order the list of discourse entities with which the projected features of the current RE are confronted).

The framework is, in itself, language independent. The adjustment to one language or another is done by setting the list of attributes, defining the knowledge sources capable to fill them and applying evoking heuristics/rules specific to each language. There is no reason to believe that the set of rules to define the domain of accessibility would be language specific.

The framework allows the implementation of any number of knowledge sources, thus configuring an AR-system towards robustness. Knowledge sources can be combined to work in correlation but the lack of one or more of them shouldn't block the system. It is however obvious that the more knowledge sources are used and the more accurate data they contribute, the more precise the resolution would be.

We believe that the proposed framework could also be upgraded in order to accommodate training for automatic acquisition of parts of Component 3, making possible the automatic calculus of heuristics or weighing factors out of training corpora.

Last but not least, in [Cristea, Ide and Romary, 1998], [Ide and Cristea, 2000] it is shown that anaphora resolution cannot be dissociated from the disclosure of discourse structure. These must rather be considered as pair processes, one benefiting from the other. In this respect, the framework we propose is compatible with this supposition.

## References

- Azzam, S., Humphreys, K., and Gaizauskas, R. 1998. Evaluating a Focus-Based Approach to Anaphora Resolution. *Proceedings of the 17th Coling and the 36th Annual Meeting of the ACL (COLING-ACL'98)*. Montreal, Canada.
- Barlow, M. 1998. Features Mismatches and Anaphora Resolution. *New Approaches to Discourse Anaphora*. S. Botley and T. McEnery (eds). Technical Papers Vol 11.
- Bolinger, D. 1977. *Pronouns and Repeated Nouns*. Bloomington, Indiana, Indiana University Linguistics Club.
- Brennan, S.E., Friedman, M.E., and Pollard, C.J. 1987. A Centering Approach to Pronouns. *Proceedings of the 25<sup>th</sup> Annual Meeting of the ACL*, Stanford
- Carden, G. 1982. Backwards Anaphora in Discourse Context. *Journal of Linguistics* 18.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht, the Netherlands Foris Publishers.
- Cornish, F. 1996. 'Antecedentless' Anaphors: Deixis, Anaphora, or what? Some Evidence from English and French. *Journal of Linguistics*, 32.
- Cristea, D., Ide, N., and Romary, L. 1998. Veins Theory: A Model of Global Discourse

Cohesion and Coherence, *Proceedings of the 17<sup>th</sup> Coling and the 36<sup>th</sup> Annual Meeting of the ACL (COLING-ACL'98)*. Montreal, Canada.

Cristea, D., Ide, N., Marcu, D., and Tablan, V. 2000. Discourse Structure and Co-Reference: An Empirical Study, *Proceedings of the 18<sup>th</sup> COLING 2000*, Luxembourg.

Gordon, P.C. and Hendrick, R. 1997. Intuitive knowledge of linguistic coreference. *Cognition*, 62.

Gordon, P.C. and Hendrick, R. 1998. The Representation and Processing of Coreference in Discourse. *Cognitive Science*, 22.

Grosz, B.J., Joshi, A.K., and Weinstein, S. 1995. Centering: a Framework for Modelling the Local Coherence of Discourse, *Computational Linguistics*, 21 (2).

Grosz, B.J. and Sidner, C.L. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12.

Halliday, M.A.K. and Hassan, Ruqaiya. 1976. *Cohesion in English*, Longman, London and New York.

Hjelmslev, L. 1961. *Prolegomena to a Theory of Language*, Madison, Wisconsin University Press.

Hobbs, J.R. 1978. Resolving pronoun references. *Lingua*, 44. Also in B. Grosz, K. Sparck-Jones and B. Webber, eds., *Readings in Natural Language Processing*, Morgan Kaufmann, Los Altos, 1986.

Ide, N. and Cristea, D. 2000. A Hierarchical Account of Referential Accessibility, *Proceedings of the ACL 2000 conference*, Hong Kong.

Kameyama, M. 1997. Recognizing Referential Links: an Information Extraction Perspective, *Proceedings of a Workshop "Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts"*, Madrid.

Keenan, E. and Comrie, B. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8.

Kennedy, C. and Boguraev, B. 1996. Anaphora for Everyone: Pronominal Anaphora Resolution without a Parser, *16<sup>th</sup> International Conference on Computational Linguistics*, vol.1.

Kuno, S. 1972. Functional Sentence Perspective: A Case Study from Japanese and English. *Linguistic Inquiry* 3.

Lappin, Y., Shalom, and Leass, Herbert J. 1994. An Algorithm for Pronominal Anaphora Resolution in *Computational Linguistics*, vol. 20, n. 4.

Mitkov, R.. 1997. Factors in Anaphora Resolution: They Are not the Only Things that Matter. A Case Study Based on Two Different Approaches in Ruslan Mitkov and Branimir Boguraev (eds.) *Proceedings of the Workshop "Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts"*, Universidad Nacional de Educación a Distancia, Madrid.



---

<sup>3</sup> This type of semantic representation is called “center” in the centering theory [Grosz, Joshi and Weinstein, 1995].

<sup>4</sup> For other types of anaphoric relations (bridge or functional anaphora), the corresponding DEs of the anaphor and antecedent are different.

<sup>5</sup> The composition of the *projects* and *evokes* relations yields the *realises* relation of centering [Grosz, Joshi and Weinstein, 1995].

<sup>6</sup> Going not further back in time than to Shakespeare, such a prologue can be found, for instance, at the beginning of *The Tragedy of Romeo and Juliet*:

*“Two households both alike in dignity  
(In fair Verona, where we lay our scene)  
From ancient grudge break to new mutiny,  
Where civil blood makes civil hands unclean.  
From forth the fatal loins of these two foes  
A pair of star-cross’d lovers take their life,  
Whose misadventur’d piteous overthrows  
Doth with their death bury their parents’ strife.  
The fearful passage of their death-mark’d love  
And the continuance of their parent’ rage,  
Which, but their children’s end, nought could remove,  
Is now the two hours’ traffic of our stage;  
The which, if you with patient ears attend,  
What here shall miss, our toil shall strive to mend.”*

Practically, with very few details, the whole action of the play is summarised. Thus, the very first moment Romeo says on stage that he is desperately in love, the audience already knows that he is one of the two lovers and that he would eventually take his own life. The mental concept of Romeo is constructed since the prologue, and the audience does not need to know how he looks like or how he dresses to be able to recognise him later on.

<sup>7</sup> cf. *Gramatica Limbii Romane*, Ed. Academiei, 1963; Avram, M., *Gramatica pentru toti*, Ed. Academiei, 1986

<sup>8</sup> Actually the c-command relation takes place between the corresponding REs seen as constituents of a syntactic tree.