

Chapter 2

Linguistic and Cognitive Evidence About Anaphora

Massimo Poesio

Linguistics and psychology provide us with a theoretical analysis of what anaphoric expressions mean, and evidence about how their interpretation is recovered in context—in particular, which information is used. In this chapter we discuss this evidence.

2.1 The Linguistics of Anaphora

2.1.1 *Context Dependence*

The interpretation of many natural language expressions depends on the context of interpretation; in particular, the interpretation of many noun phrases depends on the entities mentioned in the **linguistic context**—the previous utterances and their content. Such dependency on the entities in the linguistic context is particularly obvious in the case of pronouns, whose interpretation entirely depends on them, as illustrated by the following dialogue fragment from the TRAINS dialogues [61]. In this example, the very same expression, personal pronoun *it*, is interpreted in totally different ways in utterances 3.1 (where it refers to engine E2) and 5.4 (where it refers to engine E1). Demonstrative pronouns as well may depend on entities introduced in the linguistic context, as illustrated by demonstrative *that* in 4.3. We will use the term **anaphoric** to indicate expressions that depend on the linguistic context, i.e., on objects explicitly mentioned or objects whose existence can be inferred from what has been said.

Massimo Poesio
University of Essex

- 1.1 M : all right system
 1.2 : we've got a more complicated problem
 1.4 : first thing I'd like you to do
 1.5 : is send engine E2 off with a boxcar to Corning
 to pick up oranges
 1.6 : uh as soon as possible
 2.1 S : okay
 3.1 M : and while it's there it should pick up the tanker
 (2.1) 4.1 S : okay
 4.2 : and that can get
 4.3 : we can get that done by three
 5.1 M : good
 5.3 : can we please send engine E1 over to Dansville
 to pick up a boxcar
 5.4 : and then send it right back to Avon
 6.1 S : okay
 6.2 : it'll get back to Avon at 6

But pronouns are not the only noun phrases whose interpretation depends on the entities in the context. (2.1) also contains the definite NP *the tanker* in 3.1, whose interpretation depends on the **visual context**, which in the TRAINS dialogues is a map of the 'TRAINS world' shared between the participants. *The tanker* has not been mentioned before, but it's on this map, and therefore it is shared and has high salience and can be referred to [22]; this type of context dependence is usually called **(visual) deixis**.¹ Such examples illustrate the fact that the noun phrases in these examples are better viewed as depending on what is usually called the **discourse situation** or **utterance situation** [7], that includes both the linguistic context and the surroundings in which the participants operate.² Following the terminology of Discourse Representation Theory (DRT) [76], we will call the set of entities introduced in the discourse situation U, for 'Universe of Discourse'. There is plenty of work in CL on interpreting references to the visual context [99, 100, 9, 85, 81], but this work falls outside the scope of this book, in which we will focus on anaphora.

More in general, the interpretation of noun phrases depends on the **domain of interpretation**: the particular set of objects under discussion. Indeed, under the most widely accepted theory about their meaning [84], the interpretation of proper names only depends on the domain of interpretation, because proper names are **directly referring**: they are the natural language encoding of logical constants, and therefore the object they are referring to is directly encoded in their semantics (as opposed to being recovered from the discourse situation). Under this view, the process of interpreting proper names would be completely different from that of interpreting pronouns and nominals. The interpretation of the second mention of *Avon* in (2.1), for instance, would not be obtained by finding an antecedent in the discourse situa-

¹ Visual deixis is a type of **exophora**, but this term is not much used in CL.

² There are some constraints on what can be referred to in this way [120]; see below for a discussion of visual focus.

tion, but it would come straight from the lexical semantics of proper name *Avon*—or, more plausibly, through a pragmatic process of identifying the appropriate domain of interpretation, and the object referred to within that domain. (For instance, the proper name *David Mitchell* in (2.2a) refers to a different object than the proper name *David Mitchell* in (2.2b).)

- (2.2) a. David Mitchell (born 12 January 1969) is an English novelist. He has written four novels, two of which were shortlisted for the Booker Prize.
 b. David Mitchell (born 14 July 1974) is a British actor, comedian and writer. He is one half of the comedy duo Mitchell and Webb, alongside Robert Webb, whom he met at Cambridge University.

The conclusion that the two instances of proper name *Avon* are mentions of the same object would be obtained indirectly, through the fact that they both refer to the same object: it would be a genuine 'coreference' task.

There is plenty of work in CL on disambiguating direct references to the domain of interpretation, particularly now that Wikipedia provides unique identifiers for many objects—e.g., the two interpretations of *David Mitchell* above correspond to different Wikipedia pages [13, 27]—but this work, as well, falls outside the scope of this book, and even CL systems concentrating on identifying links between named entities do not identify coreference indirectly through the reference of proper names, whereas systems that attempt to interpret all noun phrases still need to model the context-modifying effect of proper names as they make antecedents available for pronouns and nominals.³

The choice of the domain of interpretation also affects the interpretation of nominals by fixing their **domain of quantification**—the set of objects of the type specified by the nominal complex which are included in the domain of interpretation [97, 24]. For instance, what makes the use of definite NP *the tanker* in (2.1) felicitous is the fact that the domain of quantification of nominal **tanker** consists of a single object (in the TRAINS dialogues the domain of interpretation coincides with the visual context).⁴ The domain of quantification can also be specified by the linguistic context. In the following example, the expression *most employees* is evaluated with respect to the firm mentioned in the first sentence, whereas *the management* is interpreted as the management of the firm.

- (2.3) Kim worked for three years in a large firm. Most employees were friendly, but the management was very distant.

However, we are not aware of much work on identifying the domain of quantification of a nominal apart from [99, 100].

³ We will also note that the direct reference theory of proper names is being challenged again [43].

⁴ Readers may have noticed that the interpretation of expressions like *tanker* is 'context dependent' also in the sense that it depends on the sense of the word *tanker* intended in the circumstances of utterance. We will not be concerned here with this sense of context dependence, but only with expressions that are context dependent in that their interpretation depends on the entities contained in universe of discourse U.

2.1.2 Types of context-dependent expressions

Nominals are not the only expressions whose interpretation is dependent on the linguistic or visual context in the sense above. Other examples include expressions that could be viewed as the analogous for the verbal interpretation domain of pronouns, such as **pro-verbs** like *did* in (2.4a) and **ellipsis** such as **gapping** in (2.4b). But just as pronouns are only the most extreme example of context-dependence among nominals, full verbal expressions have a context-dependent component as well. In (2.4c), for instance, the time of listening to the messages is pragmatically determined by the discourse [96, 30, 76].

- (2.4) a. Kim is making the same mistakes that I did.
 b. Kim brought the wine, and Robin – the cheese.
 c. Kim arrived home. She listened to the messages on her answering machine.

A great deal of interest was paid to ellipsis in the early years of computational linguistics [137, 135, 28] but modern corpus-based work on the interpretation of anaphoric expressions in computational linguistics (and psycholinguistics) tends to focus on the identification of the antecedents of nominal expressions, primarily because of the lack of annotated resources for studying other types of anaphora.⁵ For this reason, we will concentrate on nominal anaphoric expressions in this chapter.

Noun phrases can play four main types of **semantic function**:

Referring Following the terminology used in functional linguistics and natural language generation, we will use the term **referring noun phrases** to indicate noun phrases that introduce new entities in a discourse, or require a link to previously introduced entities. Examples include sentences like *A train arrived soon after*, where *A train* introduces a new discourse entity; or *It left immediately*, where *it* refers to a previously introduced entity. We discuss referring noun phrases and their semantics in greater detail below.⁶

Quantificational Quantificational noun phrases denote relations between the set of objects denoted by the nominal complex and the set of objects denoted by the verbal phrase: e.g., in *Few trains arrived in time*, the quantificational noun phrase *few trains* expresses a relation between the set of trains and the set of objects arriving late—namely, that few of the members of the first set are members of the second set as well:

$$\text{few}(\lambda x.\text{train}(x),\lambda x.\text{arrive-late}(x))$$

Predicative Predicative noun phrases express properties of objects. For instance, in *Kim is a preacher*, the noun phrase *a preacher* expresses a property of Kim (as opposed to referring to a second object).

⁵ A notable exception is the work by Hardt, e.g., [60]. Also, the presence of a VP ellipsis detection and resolution task at SEMEVAL-2010 indicates a renewed interest.

⁶ This sense of 'referring noun phrase' is clearly distinct from the sense in which the term 'referring' is used in the philosophical and semantics literature.

Expletive In languages like English, where verbal arguments always have been filled on syntactic grounds, forms like *it* and *there* can also be used to express semantically vacuous **expletives** as well as pronouns, as in example (2.5).

(2.5) It is half past two.

One should keep in mind that these distinctions are not always easy to make, even for humans [109, 104]. For instance, pronoun *it* in utterance 37.7 in fragment (2.6), also from the TRAINS dialogues, could be interpreted either as an expletive or as a reference to the proposed action of 'going through Dansville'.

- 37.1 M : um
 [5sec]
 37.2 : oh kay
 37.3 : um
 (2.6) 37.4 : ... then I guess we might as well go through Dansville
 37.5 : so
 37.6 : th / dz / cn / dyou /
 37.7 : does it seem like a reasonable alternative to
 37.8 : dealing with the engine that's hanging out in Elmira

An example of difficulty in classifying a noun phrase as predicative or referring is the underlined NP in (2.7), which would seem to be coreferring with Mr. Hoffman, yet appears to be playing more of a predicative role.

- (2.7) Mr. Lieber, the actor who plays Mr. Hoffman, says he was concerned at first that the script would "misrepresent an astute political mind, one that I admired," but that his concerns were allayed.
 The producers, he says, did a good job of depicting someone "who had done so much, but who was also a manic-depressive."

Finally, whether a noun phrase is considered referring or quantificational is often a matter of the particular theory chosen: for instance, in some theories all nominals are considered quantifiers, whereas in DRT and other theories, definites and indefinites are considered of a different type from other nominals.

Predicative noun phrases usually depend less on the universe of discourse *U* than other types of nominals (although they can depend on context in other respects of course). Thus, our interest in predicative NPs in this book will be limited to the fact that as many types of noun phrases can be used referentially in some contexts and predicatively in others, an anaphora resolution system must distinguish between the two types of NPs, hence the distinction must be reflected in anaphoric annotation schemes [102, 111]. Quantificational NPs are often context dependent, but in the sense that their domain of quantification is contextually specified, as discussed above. We will therefore concentrate here on referring expressions, and on the problem of selecting the discourse entity they are associated with, that is generally called **anchor** in the most general case, and **antecedent** in the case the relation between the referring expression and the anchor is one of identity (see below).

There are many varieties of referring noun phrases, which differ primarily according to the rules that govern their anaphoric behavior [113, 19, 57, 39, 37]. Such varieties include:

Reflexives, as in *John bought himself a parrot*;⁷

Pronouns, which in turn can be divided into

- **Definite pronouns** such as *he* or *she*, as in *Ross bought {a radiometer / three kilograms of after-dinner mints} and gave it / them to Nadia for her birthday*. [65]
- **Indefinite pronouns** such as *one* in *Kim bought a t-shirt so Robin decided to buy one as well* [135].
- **Demonstrative pronouns** such as *that* in example (2.1), utterance 4.3.

Nominals, i.e., noun phrases that have a noun as head, such as *a man*, *a woman*, and *the man* in (2.8).

(2.8) A man and a woman came into my shop yesterday. The man wore a baseball hat.

Proper names, such as *Kim* and *Robin* in *Kim and Robin are good friends even though Kim likes sports whereas Robin prefers reading*.

There is a certain degree of cross-lingual variety in the forms that can be used to realize these types of anaphoric expressions. Reflexives and personal pronouns can be realized as **incorporated anaphors** in several Romance languages (e.g., Catalan, Italian, Portuguese, Spanish) and as **zero anaphors** in these languages as well as Japanese. Incorporated anaphors are cases of anaphoric reference in which the anaphoric expression is expressed by an affix of another expression, e.g., a verb, as in the following example from Italian, where clitic suffix *lo* refers back to Giovanni.

(2.9) a. [IT] Giovanni_i e' in ritardo così mi ha chiesto se posso incontrar[lo]_i al cinema.

b. [EN] John_i is late so he_i asked me if I can meet him_i at the movies.

Zero anaphors are cases of anaphoric reference in which one argument is unrealized, as in the following examples from Italian and Japanese.

(2.10)a. [EN] [John]_i went to visit some friends. On the way, [he]_i bought some wine.

b. [IT] [Giovanni]_i andò a far visita a degli amici. Per via, ϕ _i comprò del vino.

c. [JA] [John]_i-wa yujin-o houmon-sita. Tochu-de ϕ _i wain-o ka-tta.

As said above, proper names differ from other referring noun phrases from a semantic point of view, in that they are directly referring rather than referring to an entity introduced in the linguistic context; demonstratives, as well, can be directly referring (both pronouns and nominals) [77]. In this book however we will concentrate on methods for establishing coreference rather than identifying the referent of noun phrases, thus these claims about proper names and demonstratives will be primarily

⁷ Reflexives are also known as 'anaphors' in Binding theory (see below).

of interest in that they suggest that such nominals will often be used to introduce new entities in the linguistic context. But this is true for nominals as well, as shown by (2.11) (from the 1993 TRAINS corpus; reported by J. Gundel) where *the maximum number of boxcars of oranges that I can get to Bath by 7 a.m. tomorrow morning* is not anaphoric—indeed, most studies find that a majority of definite NPs serve this purpose [35, 109]. (We discuss some statistics about the distribution of referring NPs in corpora below.)

- (2.11) S |hello can I help you
 U |yeah I want t- I want to determine
the maximum number of boxcars of oranges that I can get
to Bath by 7 a.m. tomorrow morning
 so hm so I guess all the boxcars will have to go through oran-
 through Corning because that's where the orange juice factory is

Another difference between types of referring expressions intensively discussed in Linguistics is that between reflexives and (personal) pronouns, illustrated by (2.12), in which *herself* must corefer with *Susan*, but *her* cannot, has been investigated in depth in generative syntax, even leading to the development of a whole new Chomskyan paradigm in the '80s (Government and Binding) [113, 19].

- (2.12) Susan considered herself fortunate to meet her.

Several researchers have concerned themselves with the factors influencing the choice among multiple admissible linguistic forms [4, 39, 57, 98, 1, 101]. Gundel *et al.* [57] investigated in depth the difference between personal and demonstrative pronouns (i.e., the difference between *it* and *that*) using corpus data (see also [89, 98]), whereas several papers by Garrod and colleagues (e.g., [39]) discuss behavioral evidence concerning the difference between definites and pronouns and between definites and proper names. We will discuss these differences in the following section.

It is important for the purposes of the following discussion to point out that no form of referring expression is invariably referring or invariably context dependent. Even pronouns can sometimes be non-referring, as shown by the example of expletives.

2.1.3 Relation between anchor and antecedent

The relation between a context-dependent referring expression and its anchor need not be one of identity of reference, as seen so far. Indefinite pronouns *one* and *another* generally stand in an **identity of sense** relation with their anchor: they refer to a different object of the same type, as in (2.13). Definite pronouns may also be used in the same way, as in so-called **paycheck pronouns** from famous example (2.14).

When the anchor is a quantified expression, as in (2.15), a pronoun with that anchor behaves like a variable in a procedure that gets repeatedly called over the

elements specified by the restriction of the quantifier; that the relation between the pronoun and its anchor is not of identity in these cases is seen most clearly when the quantifier is downward entailing, like *no* in this case. We talk in these cases of **bound anaphora**.

Finally, in **associative anaphora**, the context-dependent nominal is related to its anchor by a relation such as part-of, as in (2.16). In these cases, to identify the antecedent a **bridging inference** is generally required [21, 123, 131].

- (2.13) Sally admired Sue's jacket, so she got one for Christmas. [37]
 (2.14) The man who gave his paycheck to his wife is wiser than the man who gave it to his mistress. [79]
 (2.15) No Italian ever believes that the referee treated his team fairly.
 (2.16) We saw a flat yesterday. The kitchen is very spacious but the garden is very small.

Identifying the exact relation between an anaphor and its anchor is not always easy [109, 104, 130]. This difficulty is illustrated by examples like the following, from the WSJ portion of the ARRAU corpus, where the possessive description *its machines* in sentence (2.17e) could refer either to the 'three small personal computers' introduced in sentence 1, or to the entire range of computers sold by Texas Instruments (thus paralleling the reference to the machines sold by Compaq in the previous clause), but it's not clear which. We refer to these cases as cases of **under-specified identity**.⁸

- (2.17)a. Texas Instruments Inc., once a pioneer in portable computer technology, today will make a bid to reassert itself in that business by unveiling three small personal computers.
 b. The announcements are scheduled to be made in Temple, Texas, and include a so-called "notebook" PC that weighs less than seven pounds, has a built-in hard disk drive and is powered by Intel Corp.'s 286 microprocessor.
 c. That introduction comes only two weeks after Compaq Computer Corp., believing it had a lead of three to six months on competitors, introduced the first U.S. notebook computer with such features.
 d. Despite the inevitable comparison with Compaq, however, Texas Instruments' new notebook won't be a direct competitor.
 e. While Compaq sells its machines to businesses through computer retailers, Texas Instruments will be selling most of its machines to the industrial market and to value-added resellers and original-equipment manufacturers.

A range of types of anaphoric references in which the exact semantic relation between anaphor and antecedent is particularly complex to identify, and would require a more sophisticated theory of entities, was discussed by Versley [130] and by Recasens *et al.* [112]. The examples listed by Recasens and colleagues range from cases such as (2.18a), in which different stages of an individual are mentioned [15], to cases such as (2.18b), in which different facets of an individual are considered.

⁸ Recasens [112] used the term **quasi-identity** for these cases.

- (2.18)a. On homecoming night [Postville] feels like Hometown, USA . . . For those who prefer [the old Postville], Mayor John Hyman has a simple answer.
- b. “[Your father]_i was the greatest, but [he]_i was also one of us,” commented an anonymous old lady while she was shaking Alessandro’s hand–[Gassman]’s_i best-known son.
 “I will miss [the actor]_{i1}, but I will be lacking [my father]_{i2} especially,” he said.

2.1.4 Discourse Models

One point that the examples so far should have already made clear is that the universe of discourse *U* used to identify the anchor *Z* of a context-dependent referring expression only includes a subset of the objects of a certain type, among which the entities explicitly mentioned in the previous discourse seem especially prominent: for instance, when interpreting *the man* in (2.8), the only man considered seem to be the one mentioned earlier. (This perception is backed up by psychological research (**author?**) 37). Such considerations are one of the main arguments for the so-called **discourse model** hypothesis [79, 135, 74, 75, 121, 62, 36, 37] and for **dynamic** models of discourse interpretation. The discourse model hypothesis states that context dependent expressions are interpreted with respect to a discourse model which is built up dynamically while processing a discourse, and which includes the objects that have been mentioned (the universe of discourse *U* introduced above). This hypothesis may at first sight seem to be vacuous or even circular, stating that context dependent expressions are interpreted with respect to the context in which they are encountered. But in fact three important claims were made in this literature. First, that the context used to interpret utterances is itself continuously updated, and that this **update potential** needs to be modelled as well. Second, that the objects included in the universe of discourse / discourse model are not limited to those explicitly mentioned. The following examples illustrate the fact that a number of objects that can be ‘constructed’ or ‘inferred’ out of the explicitly mentioned objects can also serve as antecedents for context dependent nominals, including sets of objects like the set of John and Mary in (2.19), or propositions and other abstract objects like the fact that the court does not believe a certain female individual in (2.20). In fact, the implicitly mentioned object may have been introduced in a very indirect way only, as in the case of (2.21), where *the government* clearly refers to the government of Korea, but the country itself has not yet been mentioned either in the text or the title. These implicitly mentioned objects constitute what Grosz [52] called the ‘**implicit focus**’ of a discourse.

- (2.19) John and Mary came to dinner last night. They are a nice couple.
- (2.20) We believe her, the court does not, and that resolves the matter. (NY Times, 5/24/ 00, reported by J. Gundel)

- (2.21) For the Parks and millions of other young Koreans, the long-cherished dream of home ownership has become a cruel illusion. For the government, it has become a highly volatile political issue. [109]

The idea of discourse model, originally formulated by Karttunen [79], was then developed by Sanford and Garrod [121] and Garnham [37] in psycholinguistics, and made more formal, by, among others, Heim [62] and Kamp [75] in theoretical linguistics, and by Webber [135] in computational linguistics.

The theories developed by Heim and Kamp collectively took the name of Discourse Representation Theory (DRT); DRT has become the best known linguistic theory of the semantics of anaphora, and has served as the basis for the most extensive treatment of anaphora proposed in linguistics, [76], as well as many computational models. In DRT, a discourse model is a pair of a set of discourse referents and a set of conditions (statements) about these discourse referents:

$$\langle x_1 \dots x_n, c_1 \dots c_n \rangle$$

represented in the linear notation of Muskens [93] as

$$[x_1 \dots x_n | c_1 \dots c_n].$$

For instance, suppose A addresses utterance (2.22a) to B in an empty discourse model⁹. Then according to DRT update algorithms such as those proposed in [76, 93], when we process this utterance, we update the existing discourse model with information contributed by this utterance: that an entity, engine *e3*, has been mentioned (hence a discourse referent x_1 'representing' that entity gets introduced in the discourse model); and that 'we' (speaker A and addressee B) are supposed to take x_1 . This fact, as well as the fact that x_1 is an engine, are new conditions added to the discourse model. The resulting discourse model is as in (2.22b). Note in particular that interpreting nominal expression *engine E3* has resulted in a new discourse referent being added to the universe of discourse U. (Here and elsewhere we will ignore illocutionary force and simply treat all utterances as statements.)

(2.22)a. We're gonna take engine E3

b. $[x_1 | x_1 = e_3, \mathbf{engine}(x_1), \mathbf{take}(A + B, x_1)]$

This discourse model is the context in which the interpretation of the following utterance takes place. Say that (2.22a) is followed by (2.23a), which contains a pronoun. This pronoun has only one interpretation in the discourse model in (2.22b)—as having discourse entity x_1 as antecedent. Interpreting utterance (2.23a)—i.e., establishing that an instruction to send engine E3 to Corning—leads to a second update of the discourse model; the resulting model is as in (2.23b) and contains, in addition to the discourse entities and the conditions already present in (2.22b), new discourse entities and new conditions on these entities.

(2.23)a. and shove it to Corning

b. $[x_1, x_2, x_3 | x_1 = e_3, x_2 = x_1, x_3 = \mathit{corning}, \mathbf{engine}(x_1), \mathbf{take}(A + B, x_1), \mathbf{send}(A + B, x_2, x_3)]$

⁹ An extreme abstraction!

Two key contributions of dynamic theories of anaphora developed in formal linguistics have been to show that the construction of such discourse models can be characterized in a formal way, and that the resulting interpretations can be assigned a semantics just as in the case of interpretations proposed for other semantic phenomena. The original approach to discourse model construction proposed by Heim [62] and Kamp [75] – and later spelled out in painstaking detail by Kamp and Reyle [76] – was highly idiosyncratic, but later work demonstrated that the methods of syntax-driven meaning composition used in mainstream formal semantics can be used to develop a theory of discourse model construction as well [63, 118, 50, 93].

These formal approaches to discourse model construction center around the idea of **file card**. According to Heim [63], a discourse model can be seen as a collection of file cards, each representing the information about a single discourse entity introduced in the discourse. More precisely, in most recent versions of DRT, mentions of referring expressions are interpreted as follows:

indefinite (a P, some P): a new file card x_i is added to the discourse model and asserted to be of type **p**. This update is formally written $[x_i, |\mathbf{p}(x_i)]$.

proper names: as a result of a reference to object b via a proper name, a new file card x_i is added to the discourse model and asserted to be identical with b . This update is formally written $[x_i, |x_i = b]$. (See for instance proper name *Corning* in (2.23).)

pronouns: a new file card x_i is added to the discourse model and noted as needing resolution via the condition $x_i = ?$. This update is formally written $[x_i, |x_i = ?]$. Resolution leads to this condition being replaced with an equality with the file card of the anchor. (See for instance pronoun *it* in (2.23).)

definite nominals (the P, that P): this is the type of referring expression on which there is the least agreement. Most researchers believe that definite descriptions have a **uniqueness presupposition**: the existence of an object of type P is presupposed instead of asserted, and furthermore this object is meant to be unique [6, 117]. This semantics can be translated as follows: a new file card x_i is added to the discourse model and asserted to be identical with the unique object of type **p** (in the context). This update is formally written $[x_i, |x_i = \iota y.\mathbf{p}(y)]$.

Crucially for what follows, the file card for discourse entity x contains all information that is known in the context about x . Thus for instance after reading the first sentence of example (2.24) our Universe of Discourse will contain an entity x_i whose file card will contain the information that her name is Miss Watson, that she is the sister of the widow, that she is an old maid, etc etc.

(2.24) The widow's sister, Miss Watson, a tolerable slim old maid, with goggles on, had just come to live with her, and took a set at me now with a spelling-book.
She worked me middling hard for about an hour, . . . (from M. Twain, *Huckleberry Finn*).

The notion of file cards, or discourse entities, played a crucial role in work on anaphora resolution of the '80s and early '90s [135, 90, 132, 105] but then took

a back seat to more primitive notions such as single anaphor-antecedent links, although it is now being revived, as we will see in Chapter 9.5.

A crucial feature of these theories is that DRSS are logic representations with their own truth conditions, different although equivalent to traditional first-order logic, and from which inferences can be made. For instance, (2.22b) is equivalent to the pseudo-existential statement that there is an object, this object is identical to e_3 , and that A+B take this object. The existence of a deductive system over these representations is essential because many cases of anaphora resolution require complex inference, as we will see in a moment.

DRT has been used to develop accounts of a range of anaphoric phenomena beyond the simple case of nominal reference to antecedents introduced by nominals, covering reference to events as in (2.25a), to plurals as in (2.25b), or to more abstract objects such as propositions as in (2.25c).

- (2.25)a. John met Mary. That happened at 3 o'clock.
 b. John saw Mary. They had gone to school together.
 c. John met Mary. This fact stroke him as strange

Kamp and Reyle [76] and others provide detailed treatments of anaphora to events and plurals. Their treatment of reference to events is based on the assumption that events are individuals that introduce discourse referents in the common ground, as in (2.26).

- (2.26)a. John met Mary. That happened at 3 o'clock.
 b. $[x_1, x_2, e_1, x_3 | x_1 = john, x_2 = mary, e_1 : \mathbf{meet}(x_1, x_2),$
 $x_3 @ 3pm, x_3 = e_1]$

By contrast, Kamp and Reyle's analysis of plurals, like that of most researchers in the area, is based on the assumption that resolving such references (i.e., finding an anchor for discourse entity x_3 in in (2.27b)) requires bridging inferences on the discourse model as a result of which the model is augmented with new objects. In the case of plurals, these new objects are sets or groups, such as new object x_4 , defined as $x_1 + x_2$ in (2.27c). In the case of propositional references, these new objects are propositions. Resolving the discourse referent x_3 in (2.27e) requires introducing a new propositional variable K_1 , as in (2.27f). As already discussed, one of the key claims of the discourse model hypothesis is that resolving anaphoric references in general requires inferences on the discourse model.

- (2.27)a. John met Mary. They had gone to school together.
 b. $[x_1, x_2, e_1, e_2, x_3 | x_1 = john, x_2 = mary, e_1 : \mathbf{meet}(x_1, x_2),$
 $e_2 : \mathbf{gone-to-school-together}(x_3)]$
 c. $[x_1, x_2, e_1, e_2, x_3, x_4 | x_1 = john, x_2 = mary, e_1 : \mathbf{meet}(x_1, x_2),$
 $e_2 : \mathbf{gone-to-school-together}(x_3), x_4 = x_1 + x_2, x_3 = x_4]$
 d. We believe her, the court does not, and that resolves the matter.
 e. $[x_1, s_1, x_2, s_2 | s_1 : \mathbf{believe}(we, x_1),$
 $\mathbf{court}(x_2), \neg s_2 : \mathbf{believe}(x_2, x_1)]$
 f. $[x_1, s_1, x_2, s_2, x_3, e_1, K_1 | s_1 : \mathbf{believe}(we, x_1),$
 $\mathbf{court}(x_2), K_1 : [\neg s_2 : \mathbf{believe}(x_2, x_1)], \mathbf{matter}(x_4),$
 $e_1 : \mathbf{resolves}(x_3, x_4), x_3 = K_1]$

Little or no work has been done within statistical approaches to anaphora resolution on creating plural objects out of singular mentions as antecedents of plural anaphors. Some research on reference to events has been carried out after the creation of the OntoNotes corpus.

Even richer, if less formalized, models (usually called **mental models** instead of **discourse models** were proposed in psycholinguistics on the basis of work by Bransford *et al.*, Garnham, and Sanford and Garrod, among others [11, 121, 37]. Such models are assumed to encode the results of rich inference and to be more distant from language than the models usually assumed in computational and theoretical linguistics.

2.1.5 Statistics about Anaphora from Corpora

Statistics from anaphorically annotated corpora can give a rough quantitative indication of the relative importance of different types of nominal anaphoric phenomena.

Kabadjov [72] reports several statistics about the relative frequency of different types of nominals in the GNOME corpus and the Vieira-Poesio corpus. The GNOME corpus ([103]; see also Chapter 4) was designed to study local and global salience [108, 107] and in particular, their effect on generation, including text structuring [78], aggregation [18] and determining the form of referring expressions [101]. It consists of texts from three different genres widely studied in NLG: museum labels, pharmaceutical leaflets, and tutorial dialogues.

The subset of the GNOME corpus analyzed by Kabadjov includes 3354 NPs, classified into 28 mutually exclusive types. The five most frequent types are bare-np, the-np and the-pn, pers-pro, pn and a-np, representing 22%, 18%, 10%, 10%, and 8% of the total, respectively.

Concerning the types of relations, the part of the GNOME corpus studied by Kabadjov includes 2075 anaphoric relations; of these, 1161 (56%) are identity relations, whereas the rest are bridging. Among the anaphors, 44% of all anaphors related to their antecedent by an identity relation are pronouns (of which 27% personal pronouns and 17% possessive pronouns), 16% are definite descriptions, and 10% are proper names. Conversely, 97% of possessive pronouns are anaphoric, as are 95% of pers-pro, 38% of proper names, and 30% of definite descriptions.

The anaphoricity (or lack thereof) of pronouns has been studied in a number of papers concerned with detecting expletives. Evans [34] collected statistics from 77 texts from the SUSANNE and BNC corpus chosen to sample a variety of genres, and which contained 3171 examples of *it*. Of these, he classified 67.9% as being nominal anaphoric, 26.8% expletives, 2.2% used in idiomatic / stereotypical constructions, 2% discourse topic mentions, 0.8% clause anaphoric, 0.1% cataphoric. Very similar figures are reported by Boyd *et al.* [10], who studied expletives in text as well (the BNC sampler corpus). Of the 2337 instances of *it* in their corpus, 646 (28%) are expletives. Arguably the most careful analysis of the distribution of pronouns has been carried out by Müller [92], who studied the distribution of third-person pronouns *it*,

this and *that* in multi-party dialogue. Müller asked his coders to classify these pronouns as either 'normal' (i.e. referring to either a nominal or clausal antecedent), 'extrapos-it' and 'prop-it' (two types of expletives), 'vague' (i.e., referring but without a clearly identifiable antecedent), 'discarded' (i.e., included in utterances that were not continued) and 'other'. For *it*, he found that of the around 1,000 cases in his corpus, about 62.5% were classified as referential (of which 57.8% were 'normal' and 4.7% 'vague') and 37.5% as either expletive or discarded (22% as 'discarded', 15.5% as expletive). He also observed however significant disagreements on the classification ($\kappa = .61$).

Type of anaphoric expression	Percentage of total (of anaphors)	Percentage anaphoric	Source
Pronouns	44%		Kabadjov 2007
Personal Pronouns	27%	95%	Kabadjov 2007
<i>it</i>		68–72%	Evans, Boyd <i>et al</i>
Possessive Pronouns	17%	95%	Kabadjov 2007
Definites	16%	30%[Gnome] –40%[WSJ]	Kabadjov Poesio and Vieira
(first mention)		50%[WSJ]	Poesio and Vieira
(bridging)		10%[WSJ]	Poesio and Vieira
Proper names	10%	38%	Kabadjov 2007

Table 2.1 Anaphors and degree of anaphoricity in written text: Summary

The distribution of the antecedents of pronouns— whether they are introduced by NPs or more indirectly— was studied by [98, 32, 14, 58]. Eckert and Strube found that around 22% of the pronouns in their corpus (Switchboard) had a non-NP antecedent, whereas 33% had no antecedent at all. Byron reported that 16% of pronouns in her corpus had non-NP antecedents. Gundel *et al.* analyzed 2000 personal pronouns in the Santa Barbara Corpus of Spoken American English and found that 16% lacked an NP antecedent: around 5% had a non-NP antecedent, 4.5% were expletives, and 4.2% had what Gundel *et al.* call 'inferrable' antecedent, like *she* in the following example, that refers to the mother of the kids just mentioned.

(2.28) [Talking about how the kids across the street threw paint in their yard.]
Those kids are just - And she's pregnant with another one. (2.294)

An extensive study of the uses of definite descriptions was carried out by Poesio and Vieira [109], who were particularly concerned with the percentage of definite descriptions that were first mention, as opposed to anaphoric. Poesio and Vieira carried out two experiments in which definite descriptions were classified according to two slightly different schemes. In both cases, they found that around 50% of definite descriptions were first mention, around 40% were anaphoric, and 10% bridging. However, Poesio and Vieira also raised the issue of agreement on classification, only finding reasonable agreement among their coders on the distinction between first mention and anaphoric ($\kappa = 0.76$) with finer distinctions leading to more disagreements, and the distinction between bridging and first mention in particular being difficult.

2.1.6 *Anaphora vs. Coreference and other Terminological Issues*

We conclude this section on the linguistics of anaphora with some additional discussion of terminological issues, and in particular of the use of the terms 'anaphora' and 'coreference'. As we said above, we use the term **anaphoric** to indicate expressions whose interpretation depends on objects introduced in universe of discourse *U* either by virtue of being explicitly mentioned (like *engine E3* in (2.22)) or by being inferred (as in the cases of plurals and propositional anaphora). As we said, in this book we will primarily be concerned with these expressions and this characterization of the interpretation problem. However, quite a lot of other terms are used in the literature and there is a great degree of confusion about their use, so a few remarks on these issues are in order.

First of all, note that that this use of the term 'anaphoric'— although, we would argue, the most common in linguistics— is not the only use of the term. Many researchers use the term to indicate links at the *textual* level of representation (i.e., between expressions rather than with respect to discourse entities)—indeed, this seems to be the use of the term in the well-known [29]. Other researchers use the term anaphora to indicate the study of pronominal interpretation, reserving the term coreference for the study of anaphoric reference via proper names.

Second, with the first MUC initiative the term **coreference** was introduced for a task which is closely related (although not identical with) the task of anaphoric resolution. As a result, the term 'coreference' has become in CL virtually synonymous with anaphora. Unfortunately, the term coreference has a technical meaning in formal semantics, which has caused all sorts of discussions [29]. To add to the confusion, the term coreference is used in different ways in formal linguistics and in functional linguistics.

As we saw earlier in this Section, in formal semantics the term 'reference' is used to indicate the relation between an expression of the language and an object in the world, if any: proper names are the typical example of expression which is referring in this sense. Two expressions are thus **co-referring** if they refer to the same object. However, not all expressions in the language, and not even all the nominal expressions that we called 'referring' earlier on, are referring in this sense, yet this does not prevent them serving as antecedents of anaphoric expressions. A typical example are expressions occurring in hypothetical or negated contexts, as shown in the examples in (2.29) [95]: neither the hammer mentioned in (2.29a) nor the car mentioned in (2.29a) exist, yet they can happily serve as antecedents of anaphoric expressions.

- (2.29)a. If I had a hammer I would use it to break your head.
 b. I can't buy a car - I wouldn't know where to put it.

Viceversa, there are expressions which are coreferent but are not anaphoric in the sense discussed above —e.g., references to *Barack Obama* in distinct conversations, or in distinct documents, are co-referring (the term used in the case of documents is **cross-document coreference**) but not anaphoric (because distinct universes of discourse are built during each conversation).

This distinction between coreference and anaphora is the reason why computational linguists have generally preferred to avoid the term coreference and introduce other ones [123, 29]. We should however note that in other types of linguistics—particularly in systemic functional grammar and related functional frameworks—the term coreference is used in an entirely different manner [59, 57]. In these frameworks, there is no notion of ‘reference to the world’: all we can do is to refer to objects in our cognitive state—i.e., discourse referents—and therefore the term ‘coreferring’ is synonymous with ‘anaphoric’ in the sense here. And indeed, the use of the term ‘referring expression’ as in this Section comes from this tradition, via NLG. (For further discussion of the notion of ‘reference in the world,’ see [112].)

As the CL use of the term coreference is here to stay, we will note here that the ‘coreference task’ as defined by the MUC guidelines [64] is not the same as coreference either in the sense of formal semantics or in the sense of functional linguistics. Given the focus on applications, most instantiations of the ‘coreference task’ concentrate on entities of a restricted number of semantic classes frequently occurring in newspaper text (persons, organisations, locations, events, vehicles, weapons and facilities in the case of the Automatic Content Extraction (ACE) effort, or include the marking of textual relations that would not necessarily be viewed as ‘coreference’ in linguistics. The most discussed example [29] is that of the relation between *John* and *a fool* in (2.30).

(2.30) John is a fool.

In linguistics, the relation is typically seen as one of predication—being a fool is viewed as a property of John, as discussed earlier in this section. In the MUC / ACE guidelines, the relation is marked as coreference. The problem is that coreference is generally taken to be transitive so these guidelines result in John, ‘Mayor of Buffalo’, and ‘Senator for New York’ being coreferent in (2.31).

(2.31) John was mayor of Buffalo last year and is now Senator for New York.

2.2 The Interpretation of Anaphoric Expressions: Evidence from Corpora and Psycholinguistics

As illustrated by example (2.32), anaphoric expressions can and often are ambiguous in context, and the 'one sense per context' assumption does not apply to this case of ambiguity. Starting with the second sentence, there are two potential antecedents masculine in gender, that become three the next sentence if the system does not recognize that *the skipper of a minesweeper* is an apposition on *his father*. After the fifth sentence, a third potential antecedent appears, the sailor.

- (2.32) Maupin recalls his mother trying to shield him from his father's excesses. "Your father doesn't mean it," she would console him.
 When Maupin was born, his father was in the thick of battle, the skipper of a minesweeper.
 He didn't see his son for two years.
 He learned of his birth from a sailor on another ship, by semaphore.
 "I got very sentimental about six months ago, and asked him to tell me exactly where he was when he found out." (From *The Guardian Weekend*, August 15th, 1998, p. 22.)

Interpreting anaphoric expressions—i.e., resolving this ambiguity—requires a combination of many different types of information, as illustrated by the example above. One of the strongest factors is gender: for instance, *she* in the second sentence is totally unambiguous. Commonsense knowledge can be an equally strong factor: clearly *Maupin* and *his father* cannot corefer if *his* is taken to have *Maupin* as its antecedent. Syntactic constraints also play a role: even if *his son* was replaced with *him* in sentence four (obtaining *he didn't see him*), coreference between subject and object would still be ruled out. Other types of disambiguation depend on factors that appear to behave more like preferences than hard constraints. For instance, the preferred interpretation for pronoun *He* at the beginning of the fourth sentence would seem to be Maupin's father rather than Maupin himself, but that preference appears to be more the result of the preference for pronouns in subject position to refer to antecedents in subject position than a hard constraint or complex reasoning. The same motivation seems to justify the preference for pronoun *He* in the subject position of the following sentence. This difference between **constraints** and **preferences** plays an important role in many computational models of anaphora resolution and is also followed in standard expositions such as [91] so we'll follow it here even though there is not conclusive evidence about the existence of two distinct mechanisms. In this section we will discuss these constraints and preferences and the psychological evidence in their favor; in the following sections we will discuss evidence coming from computational work.

2.2.1 Constraints

Much of the early linguistic work on anaphora focused on the identification of morphological and syntactic **constraints** on the interpretation of anaphoric expressions. Among these constraints the better known are **agreement constraints** (syntactic and semantic) and **binding constraints**. We will now discuss in turn each of these constraints and the evidence from psycholinguistics of their importance in anaphora resolution.

Morphological constraints

Agreement constraints include gender, number and person constraints. We have an example of **gender constraint** in (2.32): *him* in the second sentence can only refer to Maupin or his father, not to his mother. The role of gender matching has been intensively studied in psychology [33, 38, 5]. Such studies demonstrated that gender affects disambiguation very early, and considered also the differences in gender use between languages with semantic gender such as English and languages with syntactic gender such as Italian or Spanish. As we will see in the other Chapters, most modern anaphora resolution systems do incorporate agreement constraints. The problems such systems encounter are that gender is not always used consistently: witness cases like (2.33), an error reported in [128] but due to erroneous use of pronoun *its* to refer to *a customer*:

(2.33) to get a customer's 1100 parcel-a-week load to its doorstep

Even when gender is not used erroneously, systems run into difficulties when pronouns are used to refer to entities referred to using uncommon proper names, as in the examples in (2.34).

(2.34)a. Maja arrived to the airport. [Maja a man] He ...

b. John brought Maja to the airport. [Maja a small dog] It ...

This second problem can be in part addressed by attempting to infer the gender of unknown names [41, 8] but more in general it is clear that people can often infer gender from context (see [25] and other references mentioned by [37], p. 67).

There has been much less psycholinguistic work on the role of number constraints, but several studies have compared the relative difficulty of interpreting plural and singular anaphoric references (e.g., [48]), and Clifton and Ferreira [23] showed that plural pronoun *they* was equally easy to read following a conjoined noun phrase (*Bill and Sue met*) than when the antecedents were syntactically divided (*Bill met Sue*) suggesting that the antecedent for the plural pronoun was found in the discourse model instead of in the syntactic representation. In Computational Linguistics, the main problem with number are nouns which are syntactically singular but semantically plural such as *the Union* in (2.35).

(2.35) The Union said that they would withdraw from negotiations until further notice.

Syntactic constraints

The study of constraints on anaphoric reference played an important role in the development of modern generative linguistics, to the point of giving the name to one of its best-known paradigms, Government and Binding theory [19]. The aim of this work was understanding why pronoun *him* cannot corefer with *John* in (2.36a) (the asterisk indicates that the sentence is ungrammatical under the interpretation specified by the indexing) whereas reflexive *himself* must obligatorily be interpreted as referring to *John* in (2.36b).

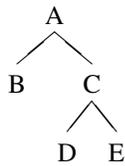
- (2.36)a. *John_i likes him_i.
 b. John_i likes himself_i

Langacker [86] proposed an account based on a relation that he called **command** holding between nodes in a syntactic tree. The definition of the relation was subsequently refined by Lasnik [88] and then by Reinhart [113], who introduced the **c-command** relation, defined as follows:

Definition 2.1. Node A c-commands node B iff

1. $A \neq B$
2. A does not dominate B and B does not dominate A, and
3. every X that dominates A also dominates B.

For instance, in the following tree, A does not c-command anything, B c-commands C, C c-commands B, D c-commands E, and E c-commands D.



The c-command relation is at the heart of the classic definition of what is now called the **binding theory** due to [19], which is articulated around three Principles. Principle A specifies constraints on reflexives and reciprocals (rather misleadingly called ‘anaphors’), and says that they must have a c-commanding antecedent in their **governing category** (the smallest clause or noun phrase in which they are included). Principle B states that pronouns cannot have an antecedent in this governing category. Together, Principles A and B claim that reflexives and pronouns are in complementary distribution. Finally, Principle C states that R-expressions –proper names and nominals–cannot have c-commanding antecedents.

Binding theory subsequently underwent numerous revisions to address empirical limitations of the 1981 version. In [20] the alternative notion of **m-command** was introduced. In HPSG, an alternative definition of **o-command** was introduced based on argument structure instead of phrase structure [110], to account for exceptions to binding theory in so-called picture NPs, as in (2.37).

- (2.37) John was going to get even with Mary. That picture of himself in the paper would really annoy her, as would the other stunts he had planned.

But perhaps the main development after [19] was the proposal by Reinhart and Reuland [115] that some reflexives are **logophors**, i.e., have discourse antecedents—examples being cases like *himself* in (2.38a), which is grammatical, in contrast with the ungrammaticality in (2.38b).

- (2.38)a. Bill_i told us that Elisabeth had invited Charles and himself_i
 b. * Bill_i told us that Elisabeth had invited himself_i

Substantial experimental testing of binding constraints has been carried out over the years. Nicol and Swinney [94] using a priming technique found that only associates of *the doctor* would be primed by *himself* in (2.39a), whereas only associates of *the skier* would be primed by *him* in (2.39b).

- (2.39)a. The boxer told the skier that the doctor for the team would blame himself for the recent injury.
 b. The boxer told the skier that the doctor for the team would blame him for the recent injury.

Gordon and Hendrick [47] found broad support for Principles A and B of binding theory but poor support for Principle C. Runner *et al.* [119] found confirmation that many reflexives in picture NPs behave like logophors.

Semantic constraints

The main semantic constraint on anaphoric reference is the so-called **scope constraint**, that prevents anaphoric reference to antecedents introduced in the scope of downward-entailing operators [79]. Thus, in (2.40a), the reference in the second sentence to the car introduced in the scope of a negation is claimed to be infelicitous. In (2.40b), the car can be referred to within the conditional, but now outside it. (2.40c), illustrates that anaphoric reference to indefinites in the scope of modals is problematic [79, 116].

- (2.40)a. John doesn't have a car. * It is in the garage.
 b. If John has a car, he doesn't use it much. * Let's drive it around the park.
 c. A wolf might have come in. *It ate John first. [116]

Semantic constraints have recently become the object of interest among psycholinguists because ERP experiments¹⁰ is showing that examples of anaphoric reference like those in (2.40) result in so-called 'semantic' violation effects (i.e., N400 effects) - see, e.g., [31] for such effects in cases like (2.40c).

¹⁰ The experimental paradigm of *event-related potentials* look for correlations between text that subjects read and brain activity as measured by EEG.

2.2.2 Preferences

By themselves, linguistic constraints do not eliminate anaphoric ambiguity. None of the constraints discussed above would prevent interpreting *him* in the second sentence of (2.32) as referring to Maupin's father. Neither do these constraints rule out interpreting *He* in the fourth sentence as referring to Maupin instead of his father. Yet these interpretations are clearly **dispreferred**. Much research has been carried out on the factors determining such preferences.

Commonsense knowledge

One such factor is plausibility based on commonsense knowledge. One of the best known illustrations of the effect of plausibility is the minimal pair in (2.41), due to Winograd and also reported in [123]. The only difference between (2.41a) and (2.41b) is the verb in the second clause, but that change is sufficient to change the preference from the council (in (2.41a)) to the women (in (2.41b)).

- (2.41) a. The city council refused the women a permit because they feared violence.
 b. The city council refused the women a permit because they advocated violence.

One type of plausibility effect intensively studied in the literature is the so-called **implicit causality** effect [40, 125]. Garvey and Caramazza [40] observed that subjects, when asked to write a continuation to a sentence like (2.42), would tend to continue in a way consistent with *he* being Bill (i.e., by assuming that the *because* clause explains why Bill is to blame).

- (2.42) John blamed Bill because he ...

Stevenson *et al.* [125] found that these preferences are affected by the thematic structure of the verb (so that agent-patient verbs behave differently from experience-stimulus ones) and by the connective.

In a forced choice experiment, Kehler *et al.* [80] presented subjects with a short discourse and a question uncovering the subjects' interpretation of a pronoun in the second sentence, as in (2.43).

- (2.43) Samuel threatened Justin with a knife, and he blindfolded Erin with a scarf.
Who blindfolded Erin?

Kehler *et al.* found that in discourses with one semantically coherent interpretation, this interpretation was chosen regardless of other salience factors, whereas in sentences where both interpretations were equally plausible, subjects' choice of interpretation more or less reflected general salience.

Another simple form of preference carried by verbs are so-called **selectional restrictions**: restrictions on the type of argument a verb may have. Their effect is shown by minimal pair (2.44), from Mitkov [91]. In (2.44a), the preferred antecedent for *it* is the computer, presumably because *disconnect* prefers an electric

appliance. In (2.44b), however, the preferred antecedent for *it* is the disk, because *copied* prefers an information-carrying device.

- (2.44)a. George removed the disk from the computer and then disconnected it.
 b. George removed the disk from the computer and then copied it.

Because of evidence such as that above the early models of anaphora resolution in CL concentrated on developing theories of commonsense reasoning [17, 136, 70] (this work is surveyed in Chapter 3), but there is clear evidence that other factors are at play as well. In (2.45a), one could argue that it's more plausible for Bill to know the combination of his own safe—yet the interpretation that has John as antecedent of *he* is clearly preferred. And if commonsense reasoning was the only factor determining anaphoric resolution, then (2.45b) should not be funny—the reason it is is that the preferred interpretation for *it* is as referring to the head rather than the bomb.

- (2.45)a. John can open Bill's safe - he knows the combination [68]
 b. If an incendiary bomb drops near you, don't lose your head. Put it in a bucket and cover it with sand [65]

Syntactic Preferences

The next factor obviously playing a role in anaphora resolution is syntactic structure and syntactic preferences. Corpus statistics suggest that in most English corpora, about 60-70% of pronouns occur in subject position, and of these, around 70% have an antecedent also realized in subject position. This preference for pronouns in subject position to refer to antecedents in subject position has been called **subject assignment** and has been extensively studied in psycholinguistics [12, 26].

Researchers also observed a preference for object pronouns to refer to antecedents in object position, suggesting a preference for **parallel** interpretations [122, 73]. Parallelism effects were studied, among others, by Smyth *et al.* [124], who showed that the closer the syntactic function, the stronger the effect; and by Stevenson *et al.* [126], who observed a similar phenomenon, but a much stronger preference for subject pronouns than for object pronouns (80% to 60%).

Researchers including Smyth and Stevenson and colleagues also hypothesized that parallelism might be semantic rather than syntactic in nature; this approach was developed by Hobbs and Kehler [69], among others.

Salience

Another factor that clearly plays a role in anaphora resolution is **salience**, at least in its simplest form of **recency**: generally speaking, more recently introduced entities are more likely antecedents. Hobbs [67] reported that in his corpus, 90% of all pronoun antecedents were in the current sentence, and 98% in the current or the previous sentence, although there was no fixed distance beyond which no antecedent

could be found (one pronominal antecedent was found 9 sentences back). This importance of the antecedents in the current and previous sentence for pronouns has been confirmed by every study of referential distance, if with slightly different figures: e.g., Hitzeman and Poesio [66] found that around 8% of pronoun antecedents in their corpora were not in the current or previous sentence. Distance is less important for other types of anaphoric expressions: e.g., Givon [45] found that 25% of definite antecedents were in the current clause, 60% in the current or previous 20 clauses, but 40% were further apart. Vieira [131] found that a window of 5 was optimal for definites. This is true cross-linguistically [44]

This is not to say, however, that choosing the most recently mentioned antecedent is an effective strategy, as several studies suggest that this strategy would have mediocre results: e.g., Tetreault [128] reports that choosing the most recent antecedent for pronouns that satisfies gender number and binding constraints would result in a 60% accuracy. On the contrary, there is a lot of evidence for a **first mention advantage**—a preference to refer to first mentioned entities in a sentence [42, 46]. Combined, these results provide support for a search strategy like that proposed by Hobbs [67]: going back one sentence at a time, then left-to-right. (See Chapter 3 for discussion.)

A stronger version of the claim that there are differences of salience between entities is the hypothesis that attentional mechanisms of the type found in visual interpretation also affect the interpretation of anaphoric expressions. Authors such as Grosz [52], Linde [89], Sanford and Garrod [121], and others have claimed that linguistic **focusing** mechanisms exist and play an important role in the choice of an antecedent for anaphoric expressions. Gundel *et al.* [57] and others suggested that such mechanisms also affect production, and in particular, the choice of form of referring expression.

The best-known theory of this type is the framework proposed by Grosz and Sidner [54] and articulated in two levels: the **global focus** specifying the articulation of a discourse into segments, and the **local focus** of salience specifying how utterance by utterance the relative salience of entities changes. That discourses are segmented according to 'topics' or the episodic organization of the story is widely accepted and backed up by evidence such as that presented by Anderson *et al.* [3]. Anderson and colleagues presented their subjects with a passage like in Figure 2.1, introducing a main character (in this case, female) and a secondary character (in this case, male) tied to the scenario. This first passage was followed either by a sentence expressing immediate continuation of the episode (*Ten minutes later . . .*) or by one indicating that the story had moved on (*Ten hours later . . .*). Finally, the subjects were presented with either a sentence referring to the main entity, or to one referring to the scenario entity. Anderson *et al.* found an entity x delay effect: after the sentence expressing immediate continuation there was no difference in processing a pronoun referring to the main entity or a pronoun referring to the scenario entity, but when the text indicated a longer delay (and hence, a closure of the previous episode) the pronominal reference to the scenario entity was harder to process.

Grosz and Sidner [54] add the further hypothesis that this segmentation is hierarchical and that it is parasitical upon the intentional structure of the discourse—the

AT THE CINEMA

Jenny found the film rather boring.
 The projectionist had to keep changing reels.
 It was supposed to be a silent classic.
 a. Ten minutes later the film was forgotten
 Ten hours later the film was forgotten
 b. She was fast asleep
 c. He was fast asleep

Fig. 2.1 The materials from [3]

intentions that the participants are trying to achieve. Grosz and Sidner proposed that the global focus is like a stack; by contrast, Walker [133] proposes a cache model. The two models were evaluated by Poesio *et al.* [107] in terms of the way they limit accessibility. Knott *et al.* [83] argued that the intentional structure proposed by Grosz and Sidner, while perhaps appropriate for task-oriented dialogue, is not appropriate for many types of text.

The second level of attention is the so-called **local focus**. According to Grosz and Sidner and other researchers including Linde, Garrod and Sanford, and others, at every moment during a conversation or while reading text some entities are more salient than the others and are preferred antecedents for pronominalization and other types of anaphoric reference. Sidner [123] proposed the first detailed theory of the local focus, articulated around two distinct foci: the **discourse focus**, meant to account for the phenomena normally explained in terms of the notion of 'discourse topic' [55, 114, 129] is usually introduced. In (2.46), the meeting with Ira is the discourse focus and serves as privileged antecedent for certain types of anaphoric reference.

- (2.46) a. I want to schedule a meeting with Ira.
 b. It should be at 3p.m.
 c. We can get together in his office

Sidner also introduced an **actor focus**, supposed to capture some of the effects accounted in previous theories through subject assignment, such (2.47).

- (2.47) John gave a lot of work to Bill. He often helps friends this way.

According to Sidner, the local focus changes after every sentence as a result of mention and coreference. Extremely complex algorithms are provided for both foci and for their use for anaphoric reference.

Centering theory [53] was originally proposed as just a simplified version of Sidner's theory of the local focus [51] but eventually it evolved in a theory of its own –in fact, the dominant paradigm for theorizing about salience in computational linguistics and, to some extent, in psycholinguistics and corpus linguistics as well (see, e.g., the papers in Walker *et al.* (**author?**) 134). According to Centering, every **utterance** updates the local focus by introducing new **forward looking centers** (mentions of discourse entities) and updating the focal structure. Forward looking centers are **ranked**: this means that each utterance has a most highly ranked entity,

called **Preferred Center** (CP), which corresponds broadly to Sidner's actor focus. In addition, Centering hypothesizes the existence of an object playing the role of the discourse topic or discourse focus: the **backward looking center**, defined as follows:

Constraint 3 $CB(U_i)$, the **Backward-Looking Center** of utterance U_i , is the highest ranked element of $CF(U_{i-1})$ that is realized in U_i .

Several psychological experiments have been dedicated to testing the claims of Centering, and in particular those concerning pronominalization, known as Rule 1:

Rule 1 If any CF in an utterance is pronominalized, the CB is.

Hudson and Tanenhaus [71] found a clear preference for subjects, which could however also be accounted for in terms of subject assignment. Gordon and colleagues carried out a series of experiments that, they argued, demonstrated certain features of the theory. Gordon *et al.* [46], for instance, revealed a **repeated name penalty**—a preference for avoiding repeating full names when an entity is mentioned in subject or first mention position, and using pronouns instead. Thus for instance Gordon *et al.* found an increase in reading time when processing sentences b–c of (2.48), with respect to reading sentences b–c of ex:RNP:2 in which the proper name in subject position *Bruno* has been replaced by pronoun *He*.

- (2.48)a. Bruno was the bully of the neighborhood.
 b. Bruno chased Tommy all the way home from school one day.
 c. Bruno watched Tommy hide behind a big tree and start to cry.
 d. Bruno yelled at Tommy so loudly that the neighbors came outside.
- (2.49)a. Bruno was the bully of the neighborhood.
 b. He chased Tommy all the way home from school one day.
 c. He watched Tommy hide behind a big tree and start to cry.
 d. He yelled at Tommy so loudly that the neighbors came outside.

Poesio *et al.* [108] carried out a systematic corpus-based investigation of the claims of Centering, that revealed among other things that entity coherence between utterances is much less strong than expected, so that the majority of utterances do not have a CB. Gundel *et al.* [57] proposed an account of the factors affecting the choice of NP based on a theory of salience with some similarities to Centering but also some important differences. Gundel *et al.* argued that the choice of NP form is the result of a process that, among other factors, takes into account the **cognitive status** of the entities being referred. Gundel *et al.*'s theory distinguishes several levels of 'givenness', including **in focus**, **activated**, **familiar** and several levels of lexical acquaintance. 'Activation' corresponds to Grosz and Sidner's implicit focus, and 'in focus' is related to the notion of CB and CP, except that more than one entity may be in focus and there may also be no entity in focus (for the relation between Gundel *et al.*'s theory and Centering see [56, 106]).

In addition to these **discrete** models of salience, **activation-based** models have also been proposed in which there is no fixed number of foci, but in which all entities have a level of activation [82, 2, 87, 127, 128].

Models that integrate salience and commonsense knowledge have also been proposed, such as Carter's [16]. Carter combined Sidner's theory of focus with Wilks' causal reasoning. Among psychologists, the interaction of Centering with commonsense preferences has been studied by Gordon and Searce [49], who found evidence that pronouns are interpreted according to Centering first and only later is commonsense knowledge used.

2.3 Summary of the Chapter

In this Chapter we introduced, first of all, some terminology and linguistic facts about anaphora that will play a key role throughout the book, providing in particular definitions of linguistic context, anaphora and coreference, and antecedent. We saw that even though virtually all methods discussed in this book will be concerned with identifying the nominal-introduced antecedents of nominal anaphoric expressions, not all anaphoric expressions are nominals, and not all of their antecedents are introduced by nominals. We also saw that not all nominals are anaphoric, or even referring—NPs can be used with a referring, quantificational, predicative, and expletive function. Finally, we introduced the notion of discourse model—the space in which antecedents of anaphoric expressions are searched—and of file card—the properties attributed to a discourse entity in a discourse model.

We next discussed some of the factors affecting the interpretation of anaphoric expressions, distinguishing between constraints—hard factors ruling our certain interpretations—and preferences that simply rank interpretations. Among the constraints we mentioned agreement constraints, syntactic constraints including those from binding theory, and semantic constraints. Among the preferences we discussed lexical and commonsense knowledge, syntactic preferences such as parallelism, and salience preferences, such as recency, first mention advantage, and focusing, distinguishing between local and global focusing effects, and introducing the Grosz / Sidner framework that is among the best known formulations of focusing effects.

Acknowledgments

This work was supported in part by the SENSEI project (FP7-ICT 610916).

References

- [1] Almor, A.: Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review* **106**, 748–765 (1999)

- [2] Alshawi, H.: *Memory and Context for Language Interpretation*. Cambridge University Press, Cambridge (1987)
- [3] Anderson, A., Garrod, S., Sanford, A.: The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology* **35**, 427–440 (1983)
- [4] Ariel, M.: *Accessing Noun-Phrase Antecedents*. Croom Helm Linguistics Series. Routledge (1990)
- [5] Arnold, J.E., Eisenband, J.G., Brown-Schmidt, S., Trueswell, J.C.: The immediate use of gender information: eyetracking evidence of the time-course of pronoun resolution. *Cognition* **76**, B13–B26 (2000)
- [6] Barker, C.: *Possessive descriptions*. Ph.D. thesis, University of California at Santa Cruz, Santa Cruz, CA (1991)
- [7] Barwise, J., Perry, J.: *Situations and Attitudes*. The MIT Press (1983)
- [8] Bergsma, S.: Automatic acquisition of gender information for anaphora resolution. In: *Proc. of 18th Conf. of the Canadian Society for Computational Studies of Intelligence*, pp. 342–353. Victoria, B.C., Canada (2005)
- [9] Beun, R., Cremers, A.: Object reference in a shared domain of conversation. *Pragmatics and Cognition* **6**(1/2), 121–152 (1998)
- [10] Boyd, A., Gegg-Harrison, W., Byron, D.: Identifying non-referential it: a machine learning approach incorporating linguistically motivated patterns. In: *Proceedings of the ACL Workshop on Feature Selection for Machine Learning in NLP*, pp. 40–47. Ann Arbor (2005)
- [11] Bransford, J., Barclay, J.R., Franks, J.J.: Sentence memory: a constructive vs. interpretive approach. *Cognitive Psychology* **3**, 193–209 (1972)
- [12] Broadbent, D.E.: *In Defence of Empirical Psychology*. Methuen (1973)
- [13] Bunescu, R., Pasca, M.: Using encyclopedic knowledge for named entity disambiguation. In: *Proc. EACL* (2006)
- [14] Byron, D.: Resolving pronominal references to abstract entities. In: *Proc. of the ACL*, pp. 80–87 (2002)
- [15] Carlson, G.N.: An unified analysis of the English bare plural. *Linguistics and Philosophy* **1**, 413–457 (1977)
- [16] Carter, D.M.: *Interpreting Anaphors in Natural Language Texts*. Ellis Horwood, Chichester, UK (1987)
- [17] Charniak, E.: *Towards a model of children’s story comprehension*. Ph.D. thesis, MIT (1972). Available as MIT AI Lab TR-266
- [18] Cheng, H.: *Modelling aggregation motivated interactions in descriptive text generation*. Ph.D. thesis, Division of Informatics, the University of Edinburgh, Edinburgh (2001)
- [19] Chomsky, N.: *Lectures on Government and Binding*. Foris, Dordrecht (1981)
- [20] Chomsky, N.: *Barriers*. MIT Press, Cambridge, MA (1986)
- [21] Clark, H.H.: Bridging. In: P.N. Johnson-Laird, P. Wason (eds.) *Thinking: Readings in Cognitive Science*, pp. 411–420. Cambridge University Press, London and New York (1977)

- [22] Clark, H.H., Marshall, C.R.: Definite reference and mutual knowledge. In: A. Joshi, B. Webber, I. Sag (eds.) *Elements of Discourse Understanding*. Cambridge University Press, New York (1981)
- [23] Clifton, C.J., Ferreira, F.: Discourse structure and anaphora: some experimental results. In: M. Coltheart (ed.) *Attention and Performance XII: the psychology of reading*, pp. 635–654. Lawrence Erlbaum, Hove, UK (1987)
- [24] Cooper, R.: The role of situations in generalized quantifiers. In: S. Lapin (ed.) *Handbook of Contemporary Semantic Theory*, chap. 3, pp. 65–86. Blackwell (1996)
- [25] Cornish, F.: Anaphoric pronouns: under linguistic control or signalling particular discourse representations? *Journal of Semantics* **5**(3), 233–260 (1986)
- [26] Crawley, R.J., Stevenson, R.A., Kleinman, D.: The use of heuristic strategies in the comprehension of pronouns. *Journal of Psycholinguistic Research* **19**, 245–264 (1990)
- [27] Csomai, A., Mihalcea, R.: Linking documents to encyclopedic knowledge. *IEEE Intelligent Systems* (2008). Special issue on Natural Language Processing for the Web
- [28] Dalrymple, M., Shieber, S.M., Pereira, F.C.N.: Ellipsis and higher-order unification. *Linguistics and Philosophy* **14**(4), 399–452 (1991)
- [29] van Deemter, K., Kibble, R.: On coreferring: Coreference in MUC and related annotation schemes. *Computational Linguistics* **26**(4), 629–637 (2000). Squib
- [30] Dowty, D.R.: The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy* **9**(1) (1986)
- [31] Dwivedi, V.D., Phillips, N.A., Laguë-Beauvais, M., Baum, S.R.: An electrophysiological study of mood, modal context, and anaphora. *Brain Research* **1117**, 135–153 (2006)
- [32] Eckert, M., Strube, M.: Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics* (2001)
- [33] Ehrlich, K., Rayner, K.: Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior* **22**, 75–87 (1983)
- [34] Evans, R.: Applying machine learning toward an automatic classification of it. *Literary and Linguistic Computing* **16**(1), 45–57 (2001)
- [35] Fraurud, K.: Definiteness and the processing of NPs in natural discourse. *Journal of Semantics* **7**, 395–433 (1990)
- [36] Garnham, A.: *On-line construction of representations of the content of texts*. Reproduced by Indiana University Linguistics Club (1982)
- [37] Garnham, A.: *Mental models and the interpretation of anaphora*. Psychology Press (2001)
- [38] Garnham, A., Oakhill, J.V., Ehrlich, M.F., Carreiras, M.: Representation and process in the interpretation of pronouns. *Journal of Memory and Language* **34**, 41–62 (1995)

- [39] Garrod, S.C.: Resolving pronouns and other anaphoric devices: The case for diversity in discourse processing. In: C. Clifton, L. Frazier, K. Rayner (eds.) *Perspectives in Sentence Processing*. Lawrence Erlbaum (1994)
- [40] Garvey, C., Caramazza, A.: Implicit causality in verbs. *Linguistic Inquiry* **5**, 459–464 (1974)
- [41] Ge, N., Hale, J., Charniak, E.: A statistical approach to anaphora resolution. In: *Proc. WVLC/EMNLP 1998* (1998)
- [42] Gernsbacher, M.A., Hargreaves, D.: Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language* **27**, 699–717 (1988)
- [43] Geurts, B.: Good news about the description theory of names. *Journal of Semantics* **14**(4), 319–348 (1997)
- [44] Givon, T. (ed.): *Topic continuity in discourse : a quantitative cross-language study*. J. Benjamins, Amsterdam and Philadelphia (1983)
- [45] Givon, T.: The grammar of referential coherence as mental processing instructions. *Linguistics* **30** (1992)
- [46] Gordon, P.C., Grosz, B.J., Gillion, L.A.: Pronouns, names, and the centering of attention in discourse. *Cognitive Science* **17**, 311–348 (1993)
- [47] Gordon, P.C., Hendrick, R.: Intuitive knowledge of linguistic coreference. *Cognition* **62**, 325–370 (1997)
- [48] Gordon, P.C., Hendrick, R., Ledoux, K., Yang, C.L.: Processing of reference and the structure of language: an analysis of complex noun phrases. *Language and Cognitive Processes* **14**(4), 353–379 (1999)
- [49] Gordon, P.C., Scarce, K.A.: Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition* **23**, 313–323 (1995)
- [50] Groenendijk, J., Stokhof, M.: Dynamic Predicate Logic. *Linguistics and Philosophy* **14**, 39–100 (1991)
- [51] Grosz, B., Joshi, A., Weinstein, S.: Providing a unified account of definite noun phrases in discourse. In: *Proc. ACL-83*, pp. 44–50. Cambridge, MA (1983)
- [52] Grosz, B.J.: *The representation and use of focus in dialogue understanding*. Ph.D. thesis, Stanford University (1977)
- [53] Grosz, B.J., Joshi, A.K., Weinstein, S.: Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* **21**(2), 202–225 (1995). (The paper originally appeared as an unpublished manuscript in 1986.)
- [54] Grosz, B.J., Sidner, C.L.: Attention, intention, and the structure of discourse. *Computational Linguistics* **12**(3), 175–204 (1986)
- [55] Gundel, J.K.: *The role of topic and comment in linguistic theory*. Ph.D. thesis, University of Texas at Austin (1974). Reprinted by Garland Publishing, New York and London, 1988
- [56] Gundel, J.K.: Centering theory and the givenness hierarchy: Towards a synthesis. In: M.A. Walker, A.K. Joshi, E.F. Prince (eds.) *Centering Theory in Discourse*, chap. 10, pp. 183–198. Oxford University Press (1998)

- [57] Gundel, J.K., Hedberg, N., Zacharski, R.: Cognitive status and the form of referring expressions in discourse. *Language* **69**(2), 274–307 (1993)
- [58] Gundel, J.K., Hedberg, N., Zacharski, R.: Pronouns without explicit antecedents: How do we know when a pronoun is referential? In: *Proc. of DAARC* (2002)
- [59] Halliday, M.A.K., Hasan, R.: *Cohesion in English*. Longman, London (1976)
- [60] Hardt, D.: An empirical approach to VP ellipsis. *Computational Linguistics* **23**(4), 525–541 (1997)
- [61] Heeman, P.A., Allen, J.F.: *The TRAINS-93 dialogues*. TRAINS Technical Note TN 94-2, University of Rochester, Dept. of Computer Science, Rochester, NY (1995)
- [62] Heim, I.: *The semantics of definite and indefinite noun phrases*. Ph.D. thesis, University of Massachusetts at Amherst (1982)
- [63] Heim, I.: File change semantics and the familiarity theory of definiteness. In: R. Bauerle, C. Schwarze, A. von Stechow (eds.) *Meaning, Use and Interpretation of Language*. de Gruyter, Berlin (1983)
- [64] Hirschman, L.: MUC-7 coreference task definition, version 3.0. In: N. Chinchor (ed.) *In Proc. of the 7th Message Understanding Conference* (1998). Available at http://www.muc.saic.com/proceedings/muc_7_toc.html
- [65] Hirst, G.: *Anaphora in Natural Language Understanding: A Survey*. Lecture Notes in Computer Science 119. Springer-Verlag, Berlin (1981)
- [66] Hitzeman, J., Poesio, M.: Long-distance pronominalisation and global focus. In: *Proc. of ACL/COLING*, vol. 1, pp. 550–556. Montreal (1998)
- [67] Hobbs, J.R.: Resolving pronoun references. *Lingua* **44**, 311–338 (1978)
- [68] Hobbs, J.R.: Coherence and coreference. *Cognitive Science* **3**, 67–90 (1979)
- [69] Hobbs, J.R., Kehler, A.: A theory of parallelism and the case of vp ellipsis. In: *Proc. 8th EACL*, pp. 394–401. Madrid (1997)
- [70] Hobbs, J.R., Stickel, M., Appelt, D., Martin, P.: Interpretation as abduction. *Artificial Intelligence Journal* **63**, 69–142 (1993)
- [71] Hudson-D’Zmura, S., Tanenhaus, M.K.: Assigning antecedents to ambiguous pronouns: The role of the center of attention as the default assignment. In: M.A. Walker, A.K. Joshi, E.F. Prince (eds.) *Centering in Discourse*, pp. 199–226. Oxford University Press (1998)
- [72] Kabadjov, M.A.: *Task-oriented evaluation of anaphora resolution*. Ph.D. thesis, University of Essex, Dept. of Computing and Electronic Systems, Colchester, UK (2007)
- [73] Kameyama, M.: *Zero anaphora: The case of Japanese*. Ph.D. thesis, Stanford University, Stanford, CA (1985)
- [74] Kamp, H.: Events, instant and temporal reference. In: R. Bauerle, U. Egli, A. von Stechow (eds.) *Semantics from Different Points of View*, pp. 376–417. Springer-Verlag (1979)
- [75] Kamp, H.: A theory of truth and semantic representation. In: J. Groenendijk, T. Janssen, M. Stokhof (eds.) *Formal Methods in the Study of Language*. Mathematical Centre, Amsterdam (1981)

- [76] Kamp, H., Reyle, U.: *From Discourse to Logic*. D. Reidel, Dordrecht (1993)
- [77] Kaplan, D.: *Demonstratives. an essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals* (1977). Unpublished manuscript, University of California, Los Angeles
- [78] Karamanis, N.: *Entity coherence for descriptive text structuring*. Ph.D. thesis, University of Edinburgh, Informatics (2003)
- [79] Karttunen, L.: *Discourse referents*. In: J. McCawley (ed.) *Syntax and Semantics 7 - Notes from the Linguistic Underground*, pp. 363–385. Academic Press, New York (1976)
- [80] Kehler, A., Kertz, L., Rohde, H., Elman, J.: *Coherence and coreference revisited*. *Journal of Semantics* **25**(1), 1–44 (2008)
- [81] Kelleher, J., Costello, F., van Genabith, J.: *Dynamically updating and inter-relating representations of visual and linguistic discourse*. *Artificial Intelligence* **167**, 62–102 (2005)
- [82] Klapholz, D., Lockman, A.: *Contextual reference resolution*. *American Journal of Computational Linguistics* (1975)
- [83] Knott, A., Oberlander, J., O'Donnell, M., Mellish, C.: *Beyond elaboration: The interaction of relations and focus in coherent text*. In: T. Sanders, J. Schilperoord, W. Spooren (eds.) *Text representation: linguistic and psycholinguistic aspects*, pp. 181–196. John Benjamins, Amsterdam and Philadelphia (2001)
- [84] Kripke, S.A.: *Naming and necessity*. In: D. Davidson, G. Harman (eds.) *Semantics of Natural Language*, pp. 253–355. Reidel, Dordrecht (1972)
- [85] Landragin, F., De Angeli, A., Wolff, F., Lopez, P., Romary, L.: *Relevance and perceptual constraints in multimodal referring actions*. In: K. van Deemter, R. Kibble (eds.) *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, pp. 395–413. CSLI (2002)
- [86] Langacker, R.: *Pronominalization and the chain of command*. In: D. Reibel, S. Schane (eds.) *Modern Studies in English*. Prentice-Hall, Englewood Cliffs, NJ (1969)
- [87] Lappin, S., Leass, H.J.: *An algorithm for pronominal anaphora resolution*. *Computational Linguistics* **20**(4), 535–562 (1994)
- [88] Lasnik, H.: *Remarks on coreference*. *Linguistic Inquiry* **2**(1) (1976)
- [89] Linde, C.: *Focus of attention and the choice of pronouns in discourse*. In: T. Givon (ed.) *Syntax and Semantics 12*. Academic Press (1979)
- [90] Luperfoy, S.: *The representation of multimodal user interface dialogues using discourse pegs*. In: *ACL-92*, pp. 22–31. University of Delaware, Newark (1992)
- [91] Mitkov, R.: *Anaphora Resolution*. Longman (2002)
- [92] Müller, M.C.: *Fully automatic resolution of it, this and that in unrestricted multi-party dialog*. Ph.D. thesis, Universität Tübingen (2008)
- [93] Muskens, R.A.: *Combining Montague Semantics and Discourse Representation*. *Linguistics and Philosophy* **19**, 143–186 (1996)

- [94] Nicol, J., Swinney, D.A.: The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* **18**, 5–19 (1989). Special Issue on Sentence Processing
- [95] Partee, B.H.: Opacity, coreference, and pronouns. In: D. Davidson, G. Harman (eds.) *Semantics for Natural Language*, pp. 415–441. D. Reidel, Dordrecht, Holland (1972)
- [96] Partee, B.H.: Some structural analogies between tenses and pronouns in English. *Journal of Philosophy* **70**, 601–609 (1973)
- [97] Partee, B.H.: Quantificational structures and compositionality. In: E. Bach, E. Jelinek, A. Kratzer, B.H. Partee (eds.) *Quantification in Natural Languages*. Kluwer (1995)
- [98] Passonneau, R.J.: Getting and keeping the center of attention. In: M. Bates, R.M. Weischedel (eds.) *Challenges in Natural Language Processing*, chap. 7, pp. 179–227. Cambridge University Press (1993)
- [99] Poesio, M.: A situation-theoretic formalization of definite description interpretation in plan elaboration dialogues. In: P. Aczel, D. Israel, Y. Katagiri, S. Peters (eds.) *Situation Theory and its Applications*, vol.3, chap. 12, pp. 339–374. CSLI, Stanford (1993)
- [100] Poesio, M.: Discourse interpretation and the scope of operators. Ph.D. thesis, University of Rochester, Department of Computer Science, Rochester, NY (1994)
- [101] Poesio, M.: Annotating a corpus to develop and evaluate discourse entity realization algorithms: issues and preliminary results. In: *Proc. of the 2nd LREC*, pp. 211–218. Athens (2000)
- [102] Poesio, M.: Discourse annotation and semantic annotation in the GNOME corpus. In: *Proc. of the ACL Workshop on Discourse Annotation*, pp. 72–79. Barcelona (2004)
- [103] Poesio, M.: The MATE/GNOME scheme for anaphoric annotation, revisited. In: *Proc. of SIGDIAL*. Boston (2004)
- [104] Poesio, M., Artstein, R.: The reliability of anaphoric annotation, reconsidered: Taking ambiguity into account. In: A. Meyers (ed.) *Proc. of ACL Workshop on Frontiers in Corpus Annotation*, pp. 76–83 (2005)
- [105] Poesio, M., Kabadjov, M.A.: A general-purpose, off the shelf anaphoric resolver. In: *Proc. of LREC*, pp. 653–656. Lisbon (2004)
- [106] Poesio, M., Modjeska, N.N.: Focus, activation, and this-noun phrases: An empirical study. In: A. Branco, R. McEnery, R. Mitkov (eds.) *Anaphora Processing*, pp. 429–442. John Benjamins (2005)
- [107] Poesio, M., Patel, A., Di Eugenio, B.: Discourse structure and anaphora in tutorial dialogues: an empirical analysis of two theories of the global focus. *Research in Language and Computation* **4**, 229–257 (2006). Special Issue on Generation and Dialogue
- [108] Poesio, M., Stevenson, R., Di Eugenio, B., Hitzeman, J.M.: Centering: A parametric theory and its instantiations. *Computational Linguistics* **30**(3), 309–363 (2004)

- [109] Poesio, M., Vieira, R.: A corpus-based investigation of definite description use. *Computational Linguistics* **24**(2), 183–216 (1998). Also available as Research Paper CCS-RP-71, Centre for Cognitive Science, University of Edinburgh
- [110] Pollard, C., Sag, I.A.: *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago, CA (1994)
- [111] Pradhan, S.S., Ramshaw, L., Weischedel, R., MacBride, J., Micciulla, L.: Unrestricted coreference: Identifying entities and events in ontologies. In: Proc. IEEE International Conference on Semantic Computing (ICSC). Irvine, CA (2007)
- [112] Recasens, M., Hovy, E., A., M.M.: Identity, non-identity, and near-identity: Addressing the complexity of coreference. *Lingua* **121**(6), 1138–1152 (2011)
- [113] Reinhart, T.: *The syntactic domain of anaphora*. Ph.D. thesis, MIT, Cambridge, MA (1976)
- [114] Reinhart, T.: *Pragmatics and linguistics: An analysis of sentence topics*. *Philosophica* **27**(1) (1981). Also distributed by Indiana University Linguistics Club
- [115] Reinhart, T., Reuland, E.: Reflexivity. *Linguistic Inquiry* **24**, 657–720 (1993)
- [116] Roberts, C.: Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* **12**, 683–721 (1989)
- [117] Roberts, C.: Uniqueness presuppositions in English definite noun phrases. *Linguistics and Philosophy* **26**(3), 287–350 (2003)
- [118] Rooth, M.: *Noun Phrase Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics*. In: P. Gärdenfors (ed.) *Generalized Quantifiers*, pp. 237–268. D. Reidel, Dordrecht, The Netherlands (1987)
- [119] Runner, J.T., Sussman, R.S., Tanenhaus, M.K.: Assignment of reference to reflexives and pronouns in picture noun phrases: Evidence from eye movements. *Cognition* **81**, 1–13 (2003)
- [120] Sag, I.A., Hankamer, J.: Toward a theory of anaphoric processing. *Linguistics and Philosophy* **7**, 325–345 (1984)
- [121] Sanford, A.J., Garrod, S.C.: *Understanding Written Language*. Wiley, Chichester (1981)
- [122] Sheldon, A.: The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior* **13**(272–281) (1974)
- [123] Sidner, C.L.: *Towards a computational theory of definite anaphora comprehension in English discourse*. Ph.D. thesis, MIT (1979)
- [124] Smyth, R.: Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* **23**, 197–229 (1994)
- [125] Stevenson, R.J., Crawley, R.A., Kleinman, D.: Thematic roles, focus, and the representation of events. *Language and Cognitive Processes* **9**, 519–548 (1994)
- [126] Stevenson, R.J., Nelson, A.W.R., Stenning, K.: The role of parallelism in strategies of pronoun comprehension. *Language and Cognitive Processes* **38**, 393–418 (1995)

- [127] Strube, M.: Never look back: An alternative to centering. In: Proc. of COLING-ACL, pp. 1251–1257. Montreal (1998)
- [128] Tetreault, J.R.: A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics* **27**(4), 507–520 (2001)
- [129] Vallduvi, E.: Information packaging: a survey. Research Paper RP-44, University of Edinburgh, HCRC (1993)
- [130] Versley, Y.: Vagueness and referential ambiguity in a large-scale annotated corpus. *Research on Language and Computation* **6**(3–4), 333–353 (2008)
- [131] Vieira, R.: Definite description resolution in unrestricted texts. Ph.D. thesis, University of Edinburgh, Centre for Cognitive Science (1998)
- [132] Vieira, R., Poesio, M.: An empirically-based system for processing definite descriptions. *Computational Linguistics* **26**(4), 539–593 (2000)
- [133] Walker, M.A.: Centering, anaphora resolution, and discourse structure. In: M.A. Walker, A.K. Joshi, E.F. Prince (eds.) *Centering in Discourse*, chap. 19, pp. 401–435. Oxford University Press (1998)
- [134] Walker, M.A., Joshi, A.K., Prince, E.F. (eds.): *Centering Theory in Discourse*. Clarendon Press / Oxford (1998)
- [135] Webber, B.L.: *A Formal Approach to Discourse Anaphora*. Garland, New York (1979)
- [136] Wilks, Y.A.: An intelligent analyzer and understander of english. *Communications of the ACM* **18**(5), 264–274 (1975). Reprinted in *Readings in Natural Language Processing*, Morgan Kaufmann
- [137] Woods, W.A., Kaplan, R., Nash-Webber, B.: The lunar sciences natural language information system: Final report. Report 2378, BBN, Cambridge, MA (1972)