Identity, non-identity, and near-identity:
Addressing the complexity of coreference

Marta Recasens[a,*], Eduard Hovy[b], M. Antònia Martí[a]

[a]*CLiC, Department of Linguistics, University of Barcelona, Gran Via 585, 08007 Barcelona, Spain*
[b]*USC Information Sciences Institute, 4676 Admiralty Way, Marina del Rey, CA 90292, United States*

Abstract

This article examines the mainstream categorical definition of coreference as "identity of reference." It argues that coreference is best handled when identity is treated as a continuum, ranging from full identity to non-identity, with room for near-identity relations to explain currently problematic cases. This middle ground is needed to account for those linguistic expressions in real text that stand in relations that are neither full coreference nor non-coreference, a situation that has led to contradictory treatment of cases in previous coreference annotation efforts. We discuss key issues for coreference such as conceptual categorization, individuation, criteria of identity, and the discourse model construct. We redefine coreference as a scalar relation between two (or more) linguistic expressions that refer to discourse entities considered to be at the same granularity level relevant to the linguistic and pragmatic context. We view coreference relations in terms of mental space theory and discuss a large number of real life examples that show near-identity at different degrees.

Keywords: Coreference, Discourse, Categorization, Near-identity, Refocusing, Neutralization

1. Introduction

Coreference phenomena have been treated by theoretical linguists who study the relation between anaphoric pronouns or definite descriptions and their antecedents, by discourse analysts who research factors contributing to coherence, by psycholinguists interested in the knowledge involved in the interpretation of coreferent expressions, by logicians and language philosophers who analyze propositions in terms of existence and truth conditions, and by computational linguists who attempt to build coreference resolution systems that automatically identify coreferent expressions in a text. Despite the varied interests, common to most of them is the understanding of coreference as "identity of reference," namely a relation holding between linguistic expressions that refer to the same entity. This apparently straightforward definition, however, hides a number of unexamined assumptions about reference and identity that we set out to explore in this article.

---

*Corresponding author. Tel.: +34 934035671; fax: +34 933189822.
E-mail address: mrecasens@ub.edu (M. Recasens).

The shortcomings of the current definition become especially apparent when real corpora are annotated with coreference information, because the traditional either/or definition of coreference cannot accommodate problematic cases of referential ambiguity (Versley, 2008; Poesio and Artstein, 2005; Van Deemter and Kibble, 2000). The coding guidelines of corpora such as ACE (Doddington et al., 2004) and OntoNotes (Pradhan et al., 2007) give instructions on how to deal with the most frequent cases of metonymy, but they are often *ad hoc*, to suit the needs of the task at hand, rather than based on a solid theoretical model. As a result, there appear to be inconsistencies between the annotations of the texts that are included in both corpora. Unlike ACE, which does not distinguish between a nation, its physical location, its government, or its people (they are all subsumable under the umbrella concept of "geo-political entity"), OntoNotes annotates mentions of these different aspects as non-coreferent. For this reason, *the people of France* and *France* are two mentions of the same entity in ACE, but two mentions of two different entities in OntoNotes. The guidelines are less clear for other complex cases such as evolving entities or overlapping sets. Compare the two inconsistent annotations for (1) and (2), where coreferent noun phrases (NPs) are printed in italics, and (a) and (b) are extracted from the ACE and OntoNotes corpora, respectively.

(1)   a.   On homecoming night *Postville* feels like Hometown, USA, but a look around *this town of 2,000* shows *it*'s become a miniature Ellis Island. *This* was an all-white, all-Christian community ... For those who prefer the old Postville, Mayor John Hyman has a simple answer.

      b.   On homecoming night *Postville* feels like Hometown, USA, but a look around *this town of 2,000* shows *it*'s become a miniature Ellis Island. *This* was an all-white, all-Christian community ... For those who prefer *the old Postville*, Mayor John Hyman has a simple answer.

(2)   a.   Last night in Tel Aviv, *Jews* attacked a restaurant that employs Palestinians. "We want war," *the crowd* chanted.

      b.   Last night in Tel Aviv, Jews attacked a restaurant that employs Palestinians. "*We want war,*" *the crowd* chanted.

The complexity exemplified by (1) and (2) arises when two references denote "almost" the same thing, either for a single individual—*Postville* and *the old Postville* (1)—or across two groups—*Jews*, *we*, and *the crowd* (2). To address similar examples, Versley (2008) proposes to use dot objects (Pustejovsky, 1995) in an extension of Poesio et al.'s (2006) Justified Sloppiness Hypothesis that he calls "Generalised Sloppiness Hypothesis." According to Versley (2008), referential ambiguity can be solved in terms of coercion: a dot object is introduced for referring expressions that have two possible interpretations (e.g., physical and informational), and the context predicate coerces the mentions into the appropriate granularity level, resulting in a coreference interpretation. This model makes clear the difficulties imposed by the definition of coreference itself, but it does not explain the difference between the coreference interpretations of (1) and (2).

We further the work of Poesio et al. (2006) and Versley (2008) by distinguishing the two directions along the axis of granularity in which categorization operates: one toward lesser granularity and one toward greater granularity. These directions broadly correspond to the processes of representational contracting and stretching that Fauconnier and Turner (2002) refer to as compression and decompression. We claim that some borderline identity relations are "more" identical than others and so that coreference is best viewed as a continuum ranging from identity to non-identity, with room for near-identity relations to handle currently problematic cases. Near-

identity relations, which do not fall neatly into either coreference or non-coreference, occur when entities share most but not all feature values. We provide several real data examples to illustrate different cases that elicit a near-identity relation and account for them in terms of two cognitive operations of categorization: refocusing and neutralization. The former pushes toward non-identity and creates new referential indices by highlighting the value difference(s) between two discourse entities (DEs), whereas the latter pushes toward identity and reduces the potential number of referential indices by neutralizing the value differences.

The goal of this article is to develop a more encompassing theoretical account of coreference phenomena that explains under what circumstances linguistic expressions are interpreted as coreferent, or quasi-coreferent. To this end, our approach tries to integrate Jackendoff's (1983; 2002) conceptual semantics, Fauconnier's (1985; 1997) mental space theory, Fauconnier and Turner's (2002) conceptual blending, Geach's (1962) relative identity, and Barker's (2010) claim that variability in tolerance for degrees of similarity is a matter of semantic interpretation or pragmatics. The framework presented here reviews, from the perspective of coreference, key issues such as conceptual categorization, individuation, criteria of identity, and the role of context, to help interpret real data. In brief, we redefine coreference as a scalar relation between DEs conceived of as the same at the granularity level relevant to the linguistic and pragmatic context. Such an understanding has consequences for the various branches of linguistics, from theoretical to psycho- and computational linguistics.

2. Background

Since coreference touches on subjects such as reference, categorization, and identity about which an extensive philosophical and linguistic literature exists, we can partly build on previous research. Only partly, however, because, as this section will reveal, there is a gap between real data and much previous theoretical work—which mostly uses prefabricated examples—that makes it unable to account for the problems exhibited by naturally occurring data. Two recent exceptions are the works of Poesio et al. (2006) and Versley (2008). In this section, we discuss the main drawbacks of existing accounts while reviewing the main ideas from previous work that are relevant to our account of coreference, which will be fully presented in the next section. Throughout we make explicit the assumptions and commitments underlying our approach. In order to avoid getting into deeply philosophical discussions, we will limit ourselves to the key ideas that serve as the basis to develop our coreference framework.

We will start by defining the projected world as opposed to what we call "the world," and then explore the elements and processes involved in the construction of the projected world. Entering the domain of language, we will consider the language-specific counterparts to concepts—i.e., discourse entities (DEs)—and to the projected world—i.e., the discourse model. Finally, we will get to our main subject of interest: identity relations and coreference, which play a key role in organizing DEs in the discourse model.

2.1. What reference is about

The realist theory that views reference as about the real world has underlain traditional theories of meaning from the theory of mediated reference (Frege, 1892), where a distinction is drawn between *sense* (intension) and *reference* (extension), to the theory of direct reference (Russell, 1905), where meaning is equated with reference. Common to them is the assumption that the target of linguistic reference is the objective, *real world*, whether directly or mediated by a sense.

It was not until the advent of cognitive semantics in the 1980s that this view began to be questioned in semantics.[1] Jackendoff (1983) argues for a conceptualist theory of reference according to which the information conveyed by language refers to entities in the world as conceptualized by the language user. He calls this world the *projected world*. The projected world is the world constructed by our perceptual systems in response to whatever is "out there."

We need to distinguish between the real world as the source of environmental input and the projected world as the experienced world. In fact, the study of language does not need to take the real world into account but only the projected world, as direct access to the former is barred to us and so our linguistic expressions must necessarily refer to the latter. An immediate corollary is that language is necessarily subjective. That does not however imply unprincipled variability. The fact that the processes by which we construct the projected world are universal makes our projections compatible to a major extent, thus enabling communication. By dissociating our account of coreference from real-world referents, we can abandon the requirements imposed by identity judgments in terms of an objective, unique, world that often result in dead-end contradictions. Instead, the way entities are built in language is closely tied to our cognitive apparatus rather than to intrinsic properties of the entities themselves. The discourse model parallels the projected world.

2.2. *Categorizing the projected world*

Once we have replaced the real world with the projected world, we need to consider what forms and provides structure and regular behavior to the projected world, which brings us immediately to mental information, conceptual structures, categories, and the like. At this point we start treading on thin ice for much remains unknown when it comes to the brain. Concepts and categories are closely intertwined, the former referring to all the knowledge that one has about categories—collections of instances which are treated as if they were the same. By arguing against the classical Aristotelian view that categories are defined by necessary and sufficient conditions—Wittgenstein (1953) being a precedent—Jackendoff (1983) claims that categories in the projected world are determined by complex perceptual and cognitive principles. Entities are not given by the external physical world, but it is the human cognitive apparatus that carves up the projected world into seemingly distinct and distinguishable categories, thus making divisions where there are none in the world.

Jackendoff (1983) argues that for an entity to be projected there must be a corresponding conceptual constituent. We *construct* entities from the environmental input according to the concepts that we have experienced and learned. The situation itself, our previous experience, our intentions or needs, can make certain features more salient than others and lead us toward a particular individuation. A key point in the categorization process is that it is graded rather than categorical. We are born with an "ability to conceptualize the world at different granularities and to switch among these granularities" (Hobbs, 1985). This flexibility lies at the basis of the categorization operations that we present in Section 3.4. We will see that although the notions of *type*, *token*, *role*, *value*, and so forth, are useful, the distinction is not so clear-cut when dealing with real data. Fauconnier's (1985; 1997) mental space theory is helpful in understanding the abstract mental structures that we construct while we think and talk, and in which referential structure can be projected. A preliminary application of mental space theory to complex coreference phenomena occurs in Versley (2008).

---

[1]Before, in the 18th century, the philosopher Kant had distinguished the *noumenal world* (the real world) from the *phenomenal world* (the world we perceive).

2.3. Building DEs

It is by connecting to conceptual structures that language acquires meaning, and there can be no reference without conceptualization: "A language user cannot refer to an entity without having some conceptualization of it" (Jackendoff, 2002). Note, however, that being in the real world is not a necessary condition for reference, and an entity's being in the real world is not sufficient for reference either. The crucial feature for linguistic reference is to have a referential index, which brings about the construction of a *discourse referent* (Karttunen, 1976) or a DE (Webber, 1979). These are the instances we talk about by means of referring expressions, believing that they are objects "out there."

As a discourse evolves, DEs grow in number and populate the *discourse model*, which is a temporary mental "toy" replica of the projected world built by language users specifically for interpreting a particular discourse. Apart from the collection of DEs, the discourse model includes the information that is said about them, i.e., their properties and the relations they participate in, and this information accumulates as the discourse progresses. Thus, categorization and meaning are constructed incrementally in an active process in which the speaker integrates contextual and conceptual information with more abstract information (Poesio et al., 2006). Properties may validly be changed or introduced in the discourse that are clearly untrue of the original "real-world" referents. It is between DEs that coreference relations occur.

2.4. Identity in the discourse model

Identity judgments between DEs become coreference judgments. As already hinted, we view coreference as the relation between expressions that refer to the same DE in the discourse model. Our approach to identity—and "sameness"—lies within the domain of discourse and distances itself from logical or philosophical ones, where applying an *absolute* notion of identity to the ever-changing physical world results in a number of paradoxes (Theseus's Ship, Heraclitus' river, Chrysippus' Paradox, the Statue and the Clay, etc.).

As pointed out by Fauconnier (1997, p. 51), "a natural-language sentence is a completely different kind of thing from a sentence in a logical calculus." Mathematical formulas give structural information explicitly and unambiguously. In contrast, language expressions do not have a meaning in themselves but only a meaning *potential*. The so-called Leibniz's Law[2] fails in opaque contexts as exemplified by (3), where James Bond, the top British spy, has been introduced to Ursula as Earl Grey, the wealthiest tea importer. If the wealthiest tea importer is actually the very ugly Lord Lipton, then (3-a) is true, whereas (3-b) is false. Note that although the two names/descriptions are true of the same referent, one cannot be substituted for the other *salva veritate* due to their being embedded in Ursula's beliefs.

(3)   a.   Ursula thinks the wealthiest tea importer is handsome.
            b.   Ursula thinks Lord Lipton is handsome.

In response to the notion of absolute identity, Geach (1967) argues that there is only *relative* identity.[3] An identity judgment must always be accompanied by some particular standard of

---

[2]Leibniz's Law or the Principle of the Identity of Indiscernibles state, respectively, that,

   For all x and y, if $x = y$, then x and y have the same properties.

   For all x and y, if x and y have the same properties, then $x = y$.

[3]We still believe, however, that absolute identity exists at least as a mental concept relative to which the more useful notion of relative identity is understood.

sameness. That in accordance with which we judge corresponds to Geach's (1962:39) *criterion of identity*, which he identifies as a common noun *A* so that "*x* is the same *A* as *y* but *x* and *y* are different *G*s." Reprising example (1) from Section 1, for which a notion of absolute identity produces two contradictory annotations, we find in Geach's relative identity a satisfactory explanation: the old and the new Postville both refer to the "same city" but to two different temporal instances: the city of Postville at $time_1$ (a white, Christian community) and the city of Postville at $time_2$ (with 2,000 citizens from varied nationalities).

However, unlike Geach (1962) and like Barker (2010), we claim that criteria of identity are not part of the lexical meaning of nouns. Rather, pragmatics is a major factor to take into consideration, as depending on the context, nouns can shift between what Barker (2010) calls an *individual-level* or type reading, and a *stage-level* or token reading. The latter replaces a set of individuals with a set of stages, where an individual is made up of a series of temporal slices or stages. Under a stage-level reading, we talk about entities in a temporal way (e.g., the old Postville versus the new Postville), whereas under an individual reading, we take in all times at once (e.g., the city of Postville). It is also from this perspective that the identity between "coreferent" discourse referents that evolve through discourse is considered by Charolles and Schnedecker (1993).

The different granularity levels at which we categorize—and thus at which DEs can be construed—make it possible for us to conceive of identity relations at different degrees, coarser or finer. Linguistic studies that elaborate on the use of the words *same* and *different* (Nunberg, 1984; Baker, 2003; Barker, 2010; Lasersohn, 2000) coincide in that identity judgments take into consideration only those properties that are relevant to the pragmatic purpose, that is, "when we say that *a* and *b* are the same, we mean simply that they are the same for purposes of argument" (Nunberg, 1984:207). The choice of granularity also depends on the textual context, e.g., the type that the predicate selects for (Versley, 2008). Both the textual and pragmatic context help resolve the choice between the different mental space configurations that are theoretically possible for a discourse (Fauconnier, 1997).

2.5. Summary

We can conclude this section with the following major assumptions,

1. There is no unique physical world to which referring expressions all point, but a host of individual worlds projected by our minds.
2. DEs are constructed based on the concepts and categories responsible for building the projected world, thus with the same potential range of individuation.
3. The discourse model is the mental space dynamically constructed for discourse understanding, and so is the space where coreference takes place.
4. Coreference relations between DEs depend on criteria of identity largely determined by the linguistic and pragmatic context.

3. Coreference along a continuum

The different elements presented in the previous section are integrated here into a single model by utilizing the mental space framework and conceptual blending. We start by redefining coreference as it is generally understood, followed by a description of the mental space framework and formal notation. This will provide the tools to present our continuum model for coreference as well as the operations of refocusing and neutralization that we use to account for coreference in real data.

3.1. Definition

The mainstream definition of coreference can be phrased as

> Coreference is a relation holding between two (or more) linguistic expressions that refer to the same entity in the real world.

This definition presents two major problems: its assumption that "sameness" is a straightforward relation, and its commitment to the "real world" as the domain of entities to which language refers. We propose the following alternative definition that forms the basis of our coreference model:

> Coreference is a scalar relation holding between two (or more) linguistic expressions that refer to DEs considered to be at the same granularity level relevant to the linguistic and pragmatic context.

Note that there are three keywords in this new definition. First, following the lead of Webber (1979) and Versley (2008) among others, we no longer allude to the real world; rather, we place the coreference phenomenon within the discourse model, thus "DEs." Second, these entities are constructs of conceptualization mechanisms and, since there are degrees of individuation, the identity relation only holds at a certain "granularity level." Last, the granularity level is set at the value that is relevant to the textual and pragmatic factors of the particular discourse.

3.2. Fauconnier's mental spaces

The general structure of our framework draws on Fauconnier's (1985, 1997) mental space theory. Its value lies in the tools it provides for making explicit the construction of meaning from the (underspecified) forms of language, as these themselves contain little of what goes into meaning construction. By operating at the conceptual level and unlike truth-conditional approaches, mental spaces allow for a broad range of potential meanings that are narrowed down conveniently as a function of the discourse context. Our main two focuses will be mental space elements—high-order mental entities corresponding to DEs that are named by NPs—and the connections between them. Showing how establishing connections and accessing elements affects coreference judgments constitutes a major contribution of this article.

Following usual notational conventions, we use circles to diagram mental spaces—the cognitive domains between which mappings and links are automatically established as we think and talk. They contain elements (represented by lower case letters) and connectors (represented by lines) that relate elements across spaces based on identity, analogy, representation, etc. The starting point for any mental space configuration is the base space, and subordinate mental spaces are set up in the presence of "space builders," i.e., language forms that point to conceptual domains (perspectives) like time, beliefs, wishes, movies, or pictures. Counterparts of elements created in other spaces are represented by the same letter with a subscript number. In mental space theory, the possibility of using a term from one space to refer to a linked element in another domain is known as the Access Principle.

Example (4), shown in Fig. 1, is borrowed from Fauconnier (1985) and provides a succinct explanation of how mental space configurations are built up.

(4)   In the movie Orson Welles played Hitchcock, who played a man at the bus stop.
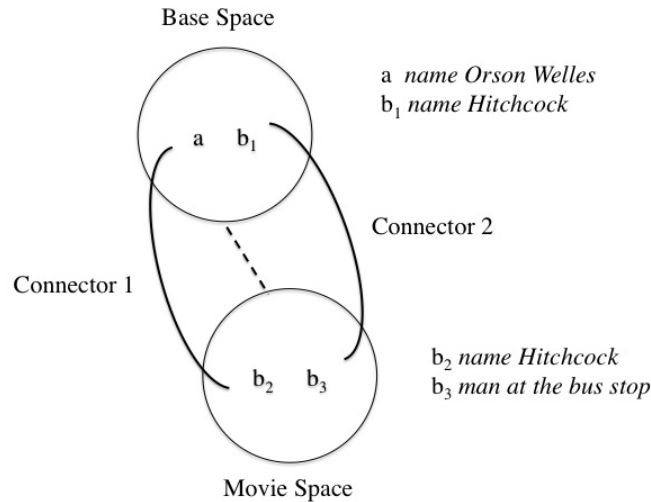
Figure 1: Mental space configuration of (4).

The base is always placed at the top and linked to its child spaces by a subordination relation. In this case, the base represents the reality space with the two DEs introduced by *Orson Welles* and *Hitchcock*. In addition, the two characters played by these two actors appear in the movie space, giving rise to two additional DEs. Then, Orson Welles-the-person is linked with Hitchcock-the-character (Connector 1), and Hitchcock-the-person is linked with the man at the bus stop (Connector 2). The two connectors exemplify actor-character relations.

Note that we could add a third connector linking Hitchcock-the-person (b_1) with Hitchcock-the-character (b_2), as this is a link—of the representation type—that we would make for a coherent discourse. With such a framework then, the different granularity levels at which DEs can be conceived can be easily represented by adding subordinate mental spaces with counterparts to DEs in a previous space (i.e., DEs constructed earlier in the ongoing linguistic exchange). By setting up a movie space, the discourse context in (4) turns the granularity level of person versus representation into a relevant one. In the diagrams we only show the mental spaces that are activated according to the discourse interests. That is to say, the same elements placed in another discourse could give rise to a different mental space configuration.

The theory of conceptual blending presented in Fauconnier and Turner (2002) is a development of mental space theory that aims to model the set of operations involved in structuring a network of mental spaces, and responsible for the development of emergent structure and novel conceptualizations. Given that the problematic cases of coreference that are the focus of this article often involve two or more mental spaces, conceptual blending is of special interest. In fact, a key component of blending, compression (and its opposite: decompression), is at the center of the two categorization operations discussed below. The blended mental space usually compresses selected aspects of structure from each input space through "vital relations" based on time, space, identity, role, part-whole, and representation, as in the case of (4). Not coincidentally, identity is pointed out as the primary vital relation: "Human mental life is unthinkable without continual compression and decompression involving identity" (Fauconnier and Turner, 2002, p. 115).

3.3. Continuum

A mental space, representing a coherent perspective on some portion of the (possibly partly imaginary) world, contains the entities (and events, etc.) present in that portion. Each entity is conceptualized by discourse participants with a set of associated features with specific values characteristic to the particular space. Coreference is evident for the two end points: entities with the same feature values corefer, while entities with different feature values do not corefer. There is, however, a wide range of in-between situations, namely entities that share most but not all feature values, and this is our main concern in this article and the reason for assuming a continuum model of coreference. One arrives to this middle-ground domain of *near-identity* by exclusion: if a relation does not fall into either identity or non-identity, then we are confronted by a near-identity relation, and depending on the number of shared/different feature values, the near-identity relation will be nearer or further away from identity. We claim that, in front of a near-identity relation, one of two cognitive operations of categorization takes place that puts emphasis on either identity or non-identity. They are presented in Section 3.4.

Throughout a discourse, some DEs are mentioned multiple times and new features might be introduced, old features might be omitted or their values changed, etc. The speaker states a series of feature–value pairs that the hearer is able to recognize or know as (supposedly) true of the DE (at that time), enough to pick it out uniquely. The problem of coreference is determining whether a new expression in a discourse refers to the "same" prior DE, or whether it introduces a new (albeit possibly very closely related) one. Rephrased in terms of the mental space framework, the problem of coreference is determining the (in)compatibility between the feature values of the various elements and counterparts in other spaces. As will be clear, feature values can be different but only potentially incompatible, where the decision is a contextual one. In our continuum model, the configuration of mental spaces is guided by two main principles:

1. Linguistic expressions (e.g., temporal phrases) that involve a change in a feature value (e.g., time) function as space builders.
2. The pragmatic needs suggest a preference for feature value compatibility (or not), and hence for identity, near-identity, or non-identity of reference.

As we will show in Section 4 using real data, there are different features whose change in value requires the building of a new mental space, but not all near-identity relations are equally strong. The perceived degree of near-identity mainly depends on (i) the near-identity operation that is involved and, closely related, (ii) the number of shared features.

3.4. Near-identity operations: refocusing and neutralization

The operations that we call "neutralization" and "refocusing" instantiate Fauconnier and Turner's (2002) compression and decompression in the context of coreference. In order to apply them to the particular situation of linking DEs, we define them in terms of features and values. Refocusing and neutralization work in opposite directions and imply a change in one or more feature values. While the former highlights the difference in values thereby shifting toward greater granularity, the latter neutralizes individual differences thereby shifting toward lesser granularity. Refocusing concentrates on the lowest ontological level and can be thought of as an operation that creates new instances of a role (i.e., new referential indices). Neutralization, on the other hand, concentrates on the role shared by all the instances and can be thought of as an operation that blends two or more similar entities into an underspecified representation of the inputs (Poesio et al., 2006), and so it reduces the potential number of referential indices. Finally, we will
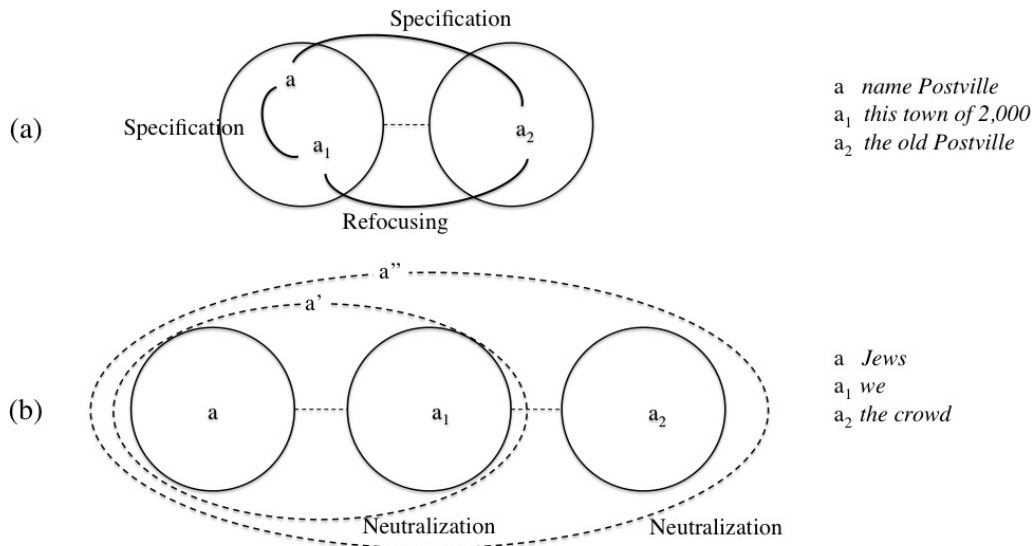
Figure 2: Mental space configurations of (5) and (6).

use the term "specification" to refer to the cases in which one or more feature value(s) are added but do not require refocusing or neutralization to establish identity, because the new values are fully consistent with the original's existing (assumed) values.

The near-identity operations are best illustrated with the Postville and Jews examples from Section 1, repeated in (5) and (6).

(5)   On homecoming night *Postville* feels like Hometown, USA, but a look around *this town of 2,000* shows it's become a miniature Ellis Island. This was an all-white, all-Christian community ... For those who prefer *the old Postville*, Mayor John Hyman has a simple answer.

(6)   Last night in Tel Aviv, *Jews* attacked a restaurant that employs Palestinians. "*We* want war," *the crowd* chanted.

In (5), one entity is *Postville*, whose name feature carries the value "Postville" (Fig. 2(a)). The second mention (*this town of 2,000*) predicates a new property of an existing entity. Since mental spaces are defined as a particular (value-defined) perspective over the constituent entities, etc., it is in the nature of the theory of mental spaces that when one changes a value for a feature, one must, by construction, generate a new space. The citizens number feature specifies detail that is consistent with the existing DE as defined so far. This value augmentation is what we call specification. The past time value of the third mention (*the old Postville*), however, clashes with the implicit time feature of the previous DE, which carries the value "the present." This value replacement occurs with refocusing. Changing the time value from "the present" to "the past" for the Postville entity automatically brings into existence the new-Postville space that contains the updated Postville entity.

In other cases, some features may be underspecified or take multiple values, as occurs with the Jews example (6). The introductory entity *Jews* is a conceptual set and hence has a members

feature with values {$person_1$, $person_2$, ..., $person_n$}. The subsequent mentions *we* and *the crowd* also have a members feature, but *we* is used generically to refer to a larger group of which the group involved in the incident (*Jews*, *the crowd*) is presented as a subgroup that considers itself representative. However, the key issue here is not whether every member of the collection is present in all three values, but rather the set itself. For the purposes of this discourse, it is irrelevant whether those who chanted are a subgroup or all of those who attacked the restaurant. Thus, we say that the three mentions have been neutralized by losing a distinctive value. The neutralized blends are represented by dashed circles in Fig. 2(b).

These two examples serve to illustrate the role of context. When the feature value changes for communicative purposes, like in (5), where the city of Postville is split in temporal slices to draw a distinction between the old and the new city, then we are in front of a refocusing shift between a_1, a_2, etc. (Fig. 2(a)). In contrast, a neutralization shift occurs when the change in value at the instance level is irrelevant in such a way that the feature ceases to be distinctive and the intersection of the features of a, a_1, a_2, etc., successively creates the increasingly underspecified variations a', a'', etc. The last created referential index serves as the only available index, overriding the previous one (Fig. 2(b)).

The compression and decompression nature of neutralization and refocusing is also well illustrated by examples that do not involve coreference but implicitly involve more than one referent. Sweetser (1997) uses (7) to make the point that polylexemic change predicates (7-a) allow both role and individual readings of the subject *your apartment*, while monolexemic change predicates (7-b) only allow an individual reading of the subject. An individual reading involves a single apartment that was smaller at an earlier time and bigger at a later time. A role reading, on the other hand, concerns not any individual holder of the role *your apartment* but rather a set of temporally successive apartments that have filled the same role.

(7)   a.   *Your apartment* keeps getting bigger every time I visit.
         b.   *Your apartment* keeps growing every time I visit.

Sweetser (1997) explains the two kinds of readings using mental spaces. What she calls a "role reading" can be explained in terms of compression. It is possible to compress time across the different apartments as they play the same role. As a result, the variability across more than one entity is conceptualized as change in a single entity, so to speak. In contrast with (5) and Fig. 2(a), the discourse in (7-a) favors the loss of the distinctive values other than size. Tobin (2010) comments on the same example from the perspective of conceptual blending arguing that a new, "blended" mental space is created where the multiple apartments are compressed.

The distinction made by Barker (2010) between stage-level and individual-level readings (see Section 2.4), exemplified in (8-a) and (8-b), respectively, can be seen in terms of near-identity in that (8-a) and (5) are similar by way of what we have called refocusing. Both are grounded in a distinction between the different instantiations of the passenger role (e.g., John A. Smith as a passenger at $time_1$ is different from John A. Smith as a passenger at $time_2$) or, in the latter example, between the different temporal slices of Postville.

(8)   a.   Easyjet served *10 million passengers* last year.
         b.   Easyjet served *10 million people* last year.

In contrast, (8-b) bears resemblance to (6) in that there is a compression across specific instances into one type, thus reducing the number of referential indices (e.g., all the temporal slices of John A. Smith are reduced to one John A. Smith). The temporal dimension is being neutralized. As

Barker (2010) points out, (8-a) can be true at the same time that (8-b) is false, as two passengers might count as one person.

4. Empirical evidence for near-identity

This section illustrates the ability of the near-identity model to account for empirical data. We begin by describing the corpus of real examples used in the bottom-up approach. We apply the operations of refocusing and neutralization outlined above to interpret these examples, and divide them into three large groups according to the features responsible for the granularity shift, which is in turn closely related to the semantic type of the entity. Then, we consider our claims about the centrality of context, and about the different degrees of near-identity, and we finish by addressing the form/function correlation.

4.1. Corpus

We created a working data set by selecting a total of 60 text excerpts from three coreferentially annotated corpora—ACE (Doddington et al., 2004), OntoNotes (Pradhan et al., 2007) and AnCora (Recasens and Martí, 2010)—as well as from the Web, a television show, and real conversation. We manually selected examples that included problematic coreference relations. By *problematic* we mean those cases that involved disagreements during the coreference annotation process of the AnCora corpus (Recasens and Martí, 2010) or cases encountered in the other sources that could be argued either way—coreferent or non-coreferent—according to the authors. Moreover, we undertook a preliminary annotation exercise with six coders that confirmed the ambiguity of these examples (Recasens et al., 2010).

We used this data set as a starting point for a bottom-up approach to coreference. At this point it became evident that using the notion of near-identity (and its corresponding operations) within a continuum of coreference proved helpful in explaining the hardest relations from a uniform perspective. Accounting for coreference through the coexistence of different mental spaces and their complex interplay as discourse unfolds generally overcomes the shortcomings of the traditional categorical definition.

4.2. Case-by-case analysis

The real data presented here constituted the starting point of this paper. They provided us with valuable insight into the intricacies of coreference and, after organizing them into separate groups, two general patterns of near-identity stood out, namely what we called neutralization and refocusing. The groups illustrated are exemplary, not exhaustive, and we limit our discussion to the perspective of refocusing and neutralization, although we do not imply that the operations we present take place singly. We readily agree that additional vital relations as proposed by Fauconnier and Turner (2002) might also be pertinent.

4.2.1. Sequences of individuals in time and space

As already discussed in relation to the Postville example (5), one of the reasons for different degrees of DE granularity is the spatiotemporal dimension. Depending on the discourse, an entity can be seen as a unique individual or as a set of temporal slices. The context in (9) leads to interpreting the second mention of *the Australian Open* as a set of instances, unlike the first mention, which implicitly refers to a particular celebration of the championship at a particular time in a particular place. A refocusing operation is at work in order to understand that

the performance of the two tennis players in the championship is the best in relation to *all* the championships.

(9)   Juan Carlos Ferrero and Francisco Clavet, the two last hopes of Spanish male tennis in *the Australian Open*, were eliminated today in the third round ... It is his best performance in the *Australian Open, where he had never progressed past the second round.*

As a result of the inherent temporal dimension of DEs, they can be split into a set of temporal slices *ad infinitum*, thereby generating additional type levels. Thus, the Australian Open can be split into the 2010 Australian Open, the 2009 Australian Open, the 2008 Australian Open, and so on; the Australian Open 2010 can in turn be split into the Australian Open 2010 first-round, the Australian Open 2010 second-round, and so on. Conversely, Fauconnier and Turner (2002) observe that it is very common to compress analogies between many specific individuals into an identity relation and, ultimately, into a single individual species. Consider (10) and (11) from our corpus.

(10)   Fishermen on this Canadian island province have shared tales of their catch. Lobster in recent years. But not too long ago, another delicacy—salmon. Oh, yeah, we used to get *salmon* in the spring, but we don't see *it* anymore. I think *they* are pretty well wiped out.

(11)   Half of children under the age of 5 get *the flu*. While unvaccinated kids bring *it* home and infect brothers and sisters, a vaccinated child helps reduce the risk by 80%.

Although the mentions refer to instances, as implied by the use of verbs like *get* and *bring* (i.e., one cannot get salmon as a kind), the focus is on the roles of *salmon* and *flu* that are filled by any instance(s) in the context, rather than on the instances. These are similar enough to disregard the details at a more granular level, thus licensing the use of pronouns. Truly, however, it is just *near*-identity, as neutralization must first take place to make the distinctive values ignorable. The change in space is especially evident in (12), where a set of tokens differing in their location values (e.g., the plant in the South of France versus the plant in Catalonia) are neutralized so that coreference can be established at the type level to state that the *same* plant colonized, entered, and spread.[4]

(12)   *The plant* colonized the South of France, from where *it* entered Catalonia in the 80s, spreading quickly ... Also, *it* presents an important population in the high basin of the Segre River.

Another interesting example from the corpus is (13). It illustrates a locational split between the first mention whose place feature carries the value "South" and the second mention whose value is "North." Unlike (12), however, the context in (13) favors a refocusing operation rather than neutralization due to the contrast drawn by the context between the celebratory mood in one place and the neutral mood in the other place. Role/value relations involving numerical parameters (14) are frequent examples of refocusing operations due to time and/or space shifts. (14-a) explicitly states that the CPI values are different because the first one corresponds to the entire Spain whereas the second corresponds to only Catalonia. In (14-b), all the GDP values correspond to Catalonia but at different moments.

---

[4]This is a case of Talmy's (1978) fictive motion: *The telephone poles* get taller as you go down the road.

(13) I once crossed that border into Ashgh-Abad on Nowruz, the Persian New Year. In the South, everyone was celebrating *New Year*; to the North, *it* was a regular day.

(14) a. The rate of increase of *the December 2000 CPI in entire Spain* stayed at 2.9 per cent . . . Regarding Catalonia, *the CPI* stays at 3.5 per cent.
      b. The Barcelona Chamber of Commerce has marked *the Catalan GDP growth* during last year in 3.7 per cent . . . Regarding the growth of the economy during last year's last three months, *the GDP growth figures* reached 3.9 per cent, three tenths over *that obtained in the previous months*.

4.2.2. Multifaceted entities

Another frequent target of near-identity relations are DEs that belong to more than one taxonomy, which Pustejovsky (1995) formalizes by concatenating the two taxonomic features as a *dot object*. For example, information-bearing objects such as newspapers are physical objects (they have pages, a weight, etc.) but they also bear information. A person can play a kinship role but also a professional one. A company produces a product, is headquartered in a location, employs a president, etc. Similarly, a country is a physical place, a political organization, its population, the ruling government, etc. The fact that the different facets can be referred using the same noun is the reason why they are often seen as coreferent, although they can truly only be near-identical.

Two mentions can invoke two different aspects of an entity but still be presented as the *same* entity (i.e., neutralization), or they might be kept separate and distinct (i.e., refocusing). In (15), the actor and father pertain to two different roles of the same individual Gassman. The opposition expressed in the citation (note the use of *but*) pertains to the typical activities of Gassman (actor-like actions versus father-like ones) and so causes a complete value replacement (Fig. 3). In communicative contexts where no contrast is drawn beetween the different facets (16), there is no need for fine-grained distinctions such that establishing coreference through neutralization keeps the discourse coherent (Fig. 4): the entity that was criticized and accused is the same. For cases like (16), Poesio et al. (2006) posit that a "super-interpretation" is made available that is underspecified but good enough for the corresponding discourse. From a strict point of view, the two mentions are not identical, but coreference is possible thanks to the super-interpretation.

(15) "Your father was the greatest, but he was also one of us," commented an anonymous old lady while she was shaking Alessandro's hand—*Gassman*'s best known son. "I will miss *the actor*, but I will be lacking *my father* especially," he said.

(16) Yugoslav opposition leaders sharply criticized both the United States and *Russia* today as a general strike against President Slobodan Milosevic gained momentum . . . Kostunica accused *the Russian government* of indecision.

In order for a near-identity relation to hold, the two mentions must refer to two entities/facets that are identically related in features other than the name. Two roles of a person, for instance, share not only the individual's name, but the actual agent. In contrast, a producer and its product might be accessible through the same name, but there is no identity between them. In (17), the company is a completely different entity from the tires, and so neither coreference nor near-coreference can be established.

(17) *Firestone* chairman John Lampe, on a telephone conference call with reporters this afternoon . . . I see the concern in people's faces. And they're very apprehensive about
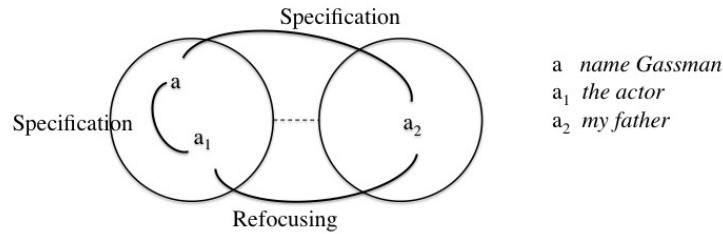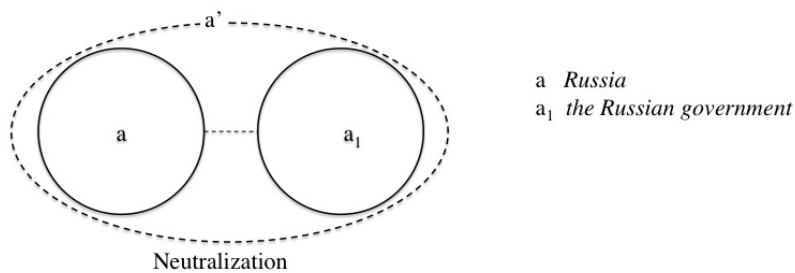
14

Figure 3: Mental space configurations of (15).

a   *name Gassman*
a_1 *the actor*
a_2 *my father*

Figure 4: Mental space configuration of (16).

a   *Russia*
a_1 *the Russian government*

purchasing *Firestones*.

Two kinds of multifaceted entities whose facets are linked by a form of identity are representational objects (e.g., pictures, statues, toy replicas, characters, maps, etc.) and informational objects (e.g., story, book, review, etc.). The former generate, for an entity X, two mental spaces containing respectively Real-X and Image-X (or more than one Image-X) and the same name gives access to all of them. For Image-X to be a representation of Real-X, Jackendoff (1983, p. 221) points out two preference rules: (i) dubbing, by which the creator of the image has stipulated the entity in question as an Image-X, and (ii) resemblance, by which Image-X must somehow look like Real-X.

The example in (18) is especially interesting because, like (15), it splits an individual (Halle Berry) into two roles (mother and actress), but it also splits her into the real and a representation (the cover of Bazaar magazine), as made clear in the first sentence. The refocusing split becomes fully relevant in the context because Berry's daughter is accessing her as her mother while the other lady in the supermarket is only accessing her from the actress role. On top of this, the lady, unlike the daughter, does not recognize the relation between the image on the cover and the real woman standing next to her. The story finishes with Berry revealing all connections. It constitutes an excellent example of refocusing.

In (19), the representation is of a more abstract kind: the people's mental conceptualization of Dalí, which might not necessarily be consistent with the real Dalí.

(18)   [Halle Berry speaking] *I* am in the supermarket and *I* was just on the cover of Bazaar magazine. At an early age my daughter would recognize *me* in the photo ... she's over my shoulder and I hear her say *"Mama, mama"*, and I knew "Oh, she saw that cover,

15

that's cute." And this woman behind her was sort of cooing with her, and I heard the woman say "Oh, no, honey, *that*'s not your mama, *that*'s Halle Berry."

(19)   Montse Aguer claimed that there is an image of *Dalí*, which is the easiest one: *the provocative Dalí*, whose most popular works are known.

In the case of informational objects, an individual (with a unique name) can be split into different tokens each of which has a different value for the format feature, specifying the format in which the information is presented or manifested (e.g., book, movie, speech, etc.). The tokens can also differ in the ending, location, number of characters, etc., but the fact that they are all accessed with the same name and that they share the core content makes them near-identical. Refocusing underlies the near-identity relations in (20) and (21) where the books and the films of the same name are being compared.

(20)   Patricia Ferreira makes progress making thriller films with *her second feature film, "The Impatient Alchemist," presented yesterday in the competition section of the Spanish Film Festival*. The film, based on *the novel of the same title by Lorenzo Silva*, is a thriller.

(21)   *A beloved American holiday story* comes to the big screen in *a Universal Pictures comic fantasy starring Jim Carey*. Alan Silverman has a look at the first feature film adaptation of *Dr. Seuss's "How the Grinch Stole Christmas"* . . . Director Ron Howard set out to film *the fantasy* not as a cartoon, but with actors in costumes and settings in the spirit of *the book*.

4.2.3. Sets

The Jews example above (6), very similar to Versley's (2008) example reproduced in (22), already showed the tendency to neutralize DEs that denote two overlapping (possibly unbounded) sets. The members of the set represented by *several hundred disabled people* are not identical with the members referred to by *us* or *the congregated*: not all the demonstrators were probably disabled, not all of them carried a poster, and *us* is used generically to refer to the larger group of disabled people represented by the smaller group of demonstrators. Nonetheless, discourse participants intuitively neutralize these values as near-identical. The collection usually consists of repeated, closely similar, members, that perform a similar function. This is also illustrated in (23), where the two mentions can be coreferred only if neutralization is assumed: clearly, one of the supporters might say that s/he does care that Nader has no chance to become President, but this makes no difference to the interpretation of this discourse. In (24), there are two notions of *team*, one that includes the coach (*we*) and one that does not (*his seniors; they*).

(22)   For a "barrier-free Bremen," *several hundred disabled people* went onto the streets yesterday—and demonstrated for "Equality, not Barriers." . . . "Why always *us*" *the congregated* asked on the posters.

(23)   Nader condemns corporations, drug companies, pesticide manufacturers, banks, landlords, the media. *His supporters* say *they* don't care that he has no chance to become President.

(24)   Meanwhile, at the Sun Ball in El Paso Texas, the University of Wisconsin Badgers held off *the University of California at Los Angeles* 21-20. The Badger's coach Barry Averett says that *his seniors* showed leadership in making their last game one of their best. *We* were soft after that first drop. Sometimes when it comes too easy you can get soft but I

liked the way *they* responded.

The complexity of the interpretation is increased when the temporal dimension is added, as in (25). Neutralizing the set members not only individually but also temporally accounts for the interpretation of the two mentions as referring to the same population around Popo, although the first mention, unlike the second one, includes people who have already died. Talking of near-identity instead of identity is thus more appropriate.

(25)   For centuries here, *the people* have had almost a mystical relationship with Popo, believing the volcano is a god. Tonight, *they* fear it will turn vengeful.

Another kind of near-identity relation in which sets often participate is the one taking place between a mention of the set and a mention of the sum of its members (26). This is allowed by the fact that a set can be accessed either as an individual or as its members. The set, however, is more than the sum of its members. *The cartel* and *the team* in (26) are not merely the conjunction of *Sotheby's and Christie's* or of *18 soccer players*, but have further implications, although the two are perceived to be nearly identical in this context. This is extensible to near-identity relations that are pragmatically established between a whole and a functionally differentiated part of this whole. In (27), President Clinton is seen as a functioning part of the entire US government. By neutralizing them we drop those features of Clinton that make him a person and keep only those that make him a government functionary. This example reflects that it is possible for principles of social organization such as legitimate representation to underlie near-identity.

(26)   a.   The gigantic international auction house Sotheby's pleaded guilty to price-fixing with Christie's—its only real competition in an industry that does $4 billion in business every year ... *The cartel* consisted of Sotheby's and Christie's ... *Sotheby's and Christie's* are all about money.

         b.   The former Real Madrid player is the only novelty in the list, comprised of *18 soccer players*. Eto'o said that *the team* should not be too confident because of the result of the first leg of Copa del Rey.

(27)   Bangladesh Prime Minister Hasina and *President Clinton* expressed the hope that this trend will continue ... Both *the US government* and American businesses welcomed the willingness of Bangladesh to embrace innovative approaches toward sustainable economic growth.

4.3. The role of context

By *context* we mean the linguistic material, the speaker's communicative intention, and the mutual knowledge between the conversants. In the preceding analysis, we have already alluded to context to explain some categorization choices. In order to illustrate the impact of context on establishing near-identity relations, compare (15) above with (28). The context presented here does not make the Gassman-the-actor alteration relevant but simply adds more detail to the first mention, without suggesting any contrast and without providing inconsistent material either. As a result, there is no need for either refocusing or neutralization. Only the first mental space of Fig. 3 is needed. Gassman and his role (actor) can be in the same mental space since they are only related by specification. The difference between the contexts of (15) and (28) is the reason why different mental space configurations are set up in each case.

17

(28)   Hollywood beckoned and *Gassman* was put under contract at MGM but the studio didn't know how best to exploit *the actor*'s capabilities.

Barker (2010) finds support for his claim that nouns do not have criteria of identity as part of their meaning but that pragmatic variability is a very influential factor in that, under the right pragmatic circumstances, typical individual-level nouns like *person* (8-b) can give a stage-level reading (29-a), and conversely, typical stage-level nouns like *passenger* (8-a) are capable of an individual-level interpretation (29-b).

(29)   a.   Newton has a new, state-of-the-art, award-winning Library which served 602,951 *people* in 1993.
         b.   Half of the *passengers* were women.

The pragmatic needs and discourse determine the granularity level at which DEs should be individuated. A contrast between a context leading to neutralization and one leading to refocusing has been exemplified by (12) and (13), respectively. Thus, language users follow contextual clues to conceptualize the world under the grain size that is required, e.g., to stop at the level of Postville and ignore differences at a more granular level, or to distinguish between the old Postville and the new Postville, or between Postville in 1960, Postville in 1970, Postville in 1980, etc., if relevant. Some linguistic variations are just a stylistic phenomenon of no further consequence, like (28). Our cognitive-linguistic system will naturally perform neutralization or refocusing if needed for the discourse to be meaningful and coherent. In fact, both can take place and the output of one can serve as the input to another, as in the case of (30), borrowed from Fauconnier and Turner (2002).

(30)   In France, *the president* is elected for a term of seven years, while in the United States *he* is elected for a term of four years.

First, all the fillers of *the president* role in France are neutralized across time as one unique individual, and so are all the fillers of *the president* role in the United States. Then, the two different locations together with the contrastive connector *while* require a refocusing operation that separates the French role from the U.S. role.

4.4. Degrees of near-identity

By their very nature, neutralization places emphasis on the identical features between two near-identical entities, whereas refocusing places emphasis on the non-identical features. It follows that DEs whose differences are inhibited by neutralization are perceived as more identical than DEs split by refocusing. Notice that we are referring to DEs as it is upon these that conceptualization mechanisms operate. We do not consider the degree of near-identity between the actual environmental inputs, but between the entities as constructed by the discourse. In (5), the language user is required to distinguish between the old and the new Postville as two different entities, although they are near-identical in that the latter is just a changed, later version of the former. In contrast, the discourse in (6) is built in such a way that one does not even perceive the actual distinctive features that exist between *Jews*, *we*, and *the crowd*. These will only be revealed upon close inspection, like through annotation.

This first factor is directly related with the second one, i.e., the number of shared features. The larger number of features two DEs have in common, the more identical they are. By removing the incompatible features, neutralization maximizes the number of shared features, while the

reverse is true of refocusing. Identity becomes harder as more granular distinctions are made. But the number of shared features can also make a difference within refocusing or neutralization. In (15), Gassman-the-actor and Gassman-the-father, which only differ in terms of their role, are more identical than the film and the book in (20), which differ not only in their medium but probably also in the number of characters, ending, etc. The double refocusing in (18) accounts for its being less identical than (15). At the outer limit, we find DEs that are only identical in their name (17), which is not enough for coreference to take place.

Similarly, with respect to neutralization, we find differences in degree of near-identity depending on the number of shared features. Singular entities tend to be more identical than plural entities, as the latter provide more room for differences. The different flu instantiations in (11) are more identical than the different instantiations of the set of Nader's supporters in (23). The near-identity degree decreases when neutralization occurs at more than one level, as illustrated by (25).

It appears relevant, in this respect, that Tobin (2010) finds a group of exceptions to Sweetser's (1997) examples in the case of NPs referring to artistic or literary works, what we called "informational objects." Unlike (7-b), it is not straightforward whether (31) involves an "individual" or a "role" interpretation. According to Tobin (2010), what licenses the use of a monolexemic change predicate and so an "individual reading" of *the poem* despite the existence of multiple, non-identical documents is that it "rests on a cultural pre-compression that is highly conventionalized." Thus, although two manifestations of the same literary work share similar content, and this makes them more identical than other near-identical DEs, culture also influences what we see as identical to a larger or lesser extent.

(31)   *The poem* initially expands from thirteen lines ... to thirty-eight lines ... but then shrinks to a lonely three.

4.5. Form and function

Let us finally turn to the correlations between form and function. Proper nouns usually designate the most individuated entities, those that are directly identifiable, whereas common nouns lie at the opposite extreme. Given that we have greater knowledge of DEs expressed by a proper noun, they are more prone to undergo a refocusing operation than DEs which can only be named by means of a common noun. Therefore, it makes more sense to talk about the old and the new Postville than, say, two temporal slices of a random glass.

The examples above have shown a wide variety of referring expressions participating in near-identity relations, but two main tendencies emerge. On the one hand, referring expressions that are classified higher on Ariel's (1988) Accessibility Scale or on Gundel et al.'s (1993) Givenness Hierarchy, i.e., expressions that refer to entities with a high degree of accessibility or that are in focus, are more frequent in neutralization operations. In particular, pronouns: *we* (6) (24), *it* (10) (11) (12), *they* (10) (23) (25), and so forth. On the other hand, expressions that are classified lower on these scales, i.e., expressions that refer to entities with a low degree of accessibility or that are not in focus, are more frequent in refocusing operations. This especially applies to definite NPs and proper nouns: *the old Postville* (5), *the GDP growth figures* (14-b), *the actor*, *my father* (15), and so forth.

This prompts us to suggest that the speaker picks the expression that will more likely lead the hearer to construct the discourse model that s/he means. By using typically anaphoric expressions (e.g., pronouns), the hearer will probably try to interpret the two mentions as (near-)coreferent

and, if they are not truly identical, carry out neutralization. Conversely, by using potentially appropriate expressions for disjoint reference (e.g., modified definite NPs), the hearer will probably perceive the feature value clash and construct two distinct DEs. This relates to the point above about entities undergoing neutralization being perceived as more identical than entities undergoing refocusing, as the hearer needs to use the expression that will contribute maximally to conveying the right interpretation. But these are just tendencies, there is no strict form/function correlation, as the hearer might be led to perform refocusing or neutralization by other phrases in the discourse (e.g., temporal expressions), or selectional preferences, or simply by the pragmatic needs.

5. Conclusion

We discussed the shortcomings of a categorical understanding of coreference as it is too limited to take into account the role of cognitive processes in the dynamic interpretation of discourse, and hence leads to contradictory analyses and annotation. It fails when confronted with the full range of natural language phenomena, as pointed out by Versley (2008) and Poesio et al. (2006). The complexity of coreference becomes apparent once we reject the naive view of linguistic expressions as mere pointers to a unique objective world, and acknowledge that the categories and concepts of our mental apparatus rely on a projected world. Discourse constructs its own model with its own entities, which language users conceptualize at a coarser- or finer-grained granularity depending on the linguistic and pragmatic context. This granularity level underlies the identity criteria according to which coreference relations are established by discourse participants, in a fashion very different from mathematical or logical identity. Accordingly, we argued for a continuum approach to coreference that contemplates middle-ground relations of near-identity, which can be accommodated in the framework of Fauconnier's (1985; 1997) mental space theory. Near-identity appears to be key to describe connections between elements of different spaces that share most but not all feature values.

Our near-identity approach is grounded in two major cognitive operations of categorization that imply a shift toward greater or lesser granularity, and toward greater or lesser identity, depending on whether there is a complete value replacement (refocusing shift) or a loss of a distinctive value (neutralization shift). We illustrated these operations by explaining a wide range of examples found in real data that are hard to account for from an either-or coreference definition, and we identified several types of features that typically result in near-identity relations when their value differs, e.g., time, location, role, set members, etc. The degree of near-identity was modeled as a function of the operation involved, and the number of shared features. Regarding the form/function correlation, we distinguished a preference for pronominal expressions to appear in neutralizing contexts, and a preference for fuller expressions to appear in refocusing contexts. The model we have presented is best viewed as a set of directions and tendencies that help interpret how coreference phenomena occur in discourse under the understanding that there are no absolute and universal rules.

As future work, we plan to refine our model by conducting both annotation and psycholinguistic experiments that provide new insights into the nature of near-identity relations. Additionally, future work will formalize the operations of refocusing and neutralization in more objective and measurable terms with a view to computational linguistics. For this purpose, typed feature structures such as dot objects (Pustejovsky, 1995) might be a promising line to pursue, as Versley (2008) puts forward.

Acknowledgements

References

Ariel, M., 1988. Referring and accessibility. Journal of Linguistics 24 (1), 65–87.

Baker, M. C., 2003. Lexical Categories. Cambridge University Press, Cambridge.

Barker, C., 2010. Nominals don't provide criteria of identity. In: Rathert, M., Alexiadou, A. (Eds.), The Semantics of Nominalizations across Languages and Frameworks. Mouton de Gruyter, Berlin, pp. 9–24.

Charolles, M., Schnedecker, C., 1993. Coréférence et identité: le problème des référents évolutifs. Langages 112, 106–126.

Doddington, G., Mitchell, A., Przybocki, M., Ramshaw, L., Strassel, S., Weischedel, R., 2004. The Automatic Content Extraction (ACE) Program – Tasks, data, and evaluation. In: Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004). Lisbon, pp. 837–840.

Fauconnier, G., 1985. Mental Spaces: Aspects of Meaning Construction in Natural Language. MIT Press, Cambridge.

Fauconnier, G., 1997. Mappings in Thought and Language. Cambridge University Press, Cambridge.

Fauconnier, G., Turner, M., 2002. The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. Basic Books, New York.

Frege, G., 1892. On sense and reference. In: Geach, P., Black, M. (Eds.), Translations from the Philosophical Writings of Gottlob Frege. Basil Blackwell (1952), Oxford, pp. 56–78.

Geach, P., 1962. Reference and Generality. Cornell University Press, Ithaca.

Geach, P., 1967. Identity. Review of Metaphysics 21, 3–12.

Gundel, J., Hedberg, N., Zacharski, R., 1993. Cognitive status and the form of referring expressions in discourse. Language 69 (2), 274–307.

Hobbs, J., 1985. Granularity. In: Proceedings of the 9th International Joint Conference on Artificial Intelligence (IJCAI 1985). Los Angeles, pp. 432–435.

Jackendoff, R., 1983. Semantics and Cognition. MIT Press, Cambridge.

Jackendoff, R., 2002. Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford University Press, Oxford.

Karttunen, L., 1976. Discourse referents. In: McCawley, J. (Ed.), Syntax and Semantics. Vol. 7. Academic Press, New York, pp. 363–386.

Lasersohn, P., 2000. *Same*, models and representation. In: Jackson, B., Mathews, T. (Eds.), Proceedings of Semantics and Linguistic Theory 10. CLC Publications, Cornell, pp. 83–97.

Nunberg, G., 1984. Individuation in context. In: Proceedings of the 2nd West Coast Conference on Formal Linguistics (WCCFL 2). Stanford Linguistics Association, Stanford, pp. 203–217.

Poesio, M., Artstein, R., 2005. The reliability of anaphoric annotation, reconsidered: Taking ambiguity into account. In: Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II. Ann Arbor, pp. 76–83.

Poesio, M., Sturt, P., Artstein, R., Filik, R., 2006. Underspecification and anaphora: Theoretical issues and preliminary evidence. Discourse Processes: a multidisciplinary journal 42, 157–175.

Pradhan, S. S., Hovy, E., Marcus, M., Palmer, M., Ramshaw, L., Weischedel, R., 2007. Ontonotes: A unified relational semantic representation. In: Proceedings of the 1st International Conference on Semantic Computing (ICSC 2007). Washington, pp. 517–526.

Pustejovsky, J., 1995. The Generative Lexicon. MIT Press, Cambridge.

Recasens, M., Hovy, E., Martí, M. A., 2010. A typology of near-identity relations for coreference (NIDENT). In: Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010). Valletta, Malta, pp. 149–156.

Recasens, M., Martí, M. A., 2010. AnCora-CO: Coreferentially annotated corpora for Spanish and Catalan. Language Resources and Evaluation 44 (4), 315–345.

Russell, B., 1905. On denoting. Mind 15, 479–493.

Sweetser, E., 1997. Role and individual readings of change predicates. In: Nuyts, J., Pederson, E. (Eds.), Language and Conceptualization. Oxford University Press, Oxford, pp. 116–136.

Talmy, L., 1978. The relation of grammar to cognition: a synopsis. In: Proceedings of the 1978 Workshop on Theoretical Issues in Natural Language Processing. Urbana-Champaign, Illinois, pp. 14–24.

Tobin, V., 2010. Grammatical and rhetorical consequences of entrenchment in conceptual blending: Compressions involving change. In: Parrill, F., Tobin, V., Turner, M. (Eds.), Meaning, Form, and Body. CSLI Publications, Stanford, CA, pp. 329–347.

Van Deemter, K., Kibble, R., 2000. On coreferring: Coreference in MUC and related annotation schemes. Computational Linguistics 26 (4), 629–637.

Versley, Y., 2008. Vagueness and referential ambiguity in a large-scale annotated corpus. Research on Language and Computation 6, 333–353.

Webber, B. L., 1979. A Formal Approach to Discourse Anaphora. Garland Press, New York.

Wittgenstein, L., 1953. Philosophical Investigations. Blackwell, Oxford.