Topic-Continuity and Topic-Shift Effects in Spanish Discourse: A Comparative Analysis of Referring Expressions

Iker Zulaica-Hernández
Indiana University-Purdue University Indianapolis
Indianapolis, Indiana, USA

Topic-Continuity and Topic-shift Effects in Spanish Discourse. A Comparative Analysis of Referring Expressions

Iker Zulaica-Hernández
Indiana University-Purdue University, Indianapolis
izulaica@iupui.edu

Abstract
Differences in use among referring expressions are usually explained on the basis of the cognitive accessibility of their antecedents, where antecedent accessibility has been operationalized differently in the literature; i.e. as a grammatical role, as syntactic prominence or as antecedent distance. On these grounds, it has been proposed that personal pronouns prefer topical antecedents whereas demonstratives prefer non-topical antecedents. This paper investigates the referring properties of Spanish demonstratives and direct object personal pronouns with the aim to unveil their differences and similarities. My analysis shows that these two expressions are very similar referentially when a narrow view of discourse context is considered. However, important differences show up when a broader notion of context is thrown into the picture; i.e. contexts that extend beyond the immediate previous sentence and beyond the immediate local topic of discourse. Based on my corpus evidence and on previous research on the pragmatic interpretation of referring expressions, I claim that direct object personal pronouns and demonstrative noun phrases crucially differ in the way they contribute to discourse coherence; the former playing the role of topic continuity markers and the latter focalising referents that reintroduce suspended or declining topics and marking (sub)-topic shifts in the discourse.

Keywords
salience, topic, personal pronouns, coherence, demonstratives, reference, discourse

1. Introduction

It is widely accepted that referring expressions are sensitive to the salience of their referents. In this line, a direct relationship between linguistic form and referent salience has been argued for whereby the shortest and simplest forms (i.e. unstressed pronouns) refer to most salient objects whereas more complex expressions (i.e. definite descriptions) commonly refer to less salient objects (Prince 1981a, Chafe 1987, Ariel 1988, Gundel, Hedberg and Zacharski 1993 and Grosz, Joshi and Weinstein 1995). Concepts such as salience, and related ones such as givenness, referent activation, or referent accessibility, are assumed to be essentially connected to the configurational properties of the sentence and to information packaging.[1] When the informational status of a referent is defined in cognitive terms, it is commonly assumed that the referents of demonstratives occupy a lower position in terms of cognitive accessibility (i.e. merely activated) whereas personal pronouns mark their referents as highly-accessible (i.e. in focus). From a purely syntactic point of view, the preverbal subject of a sentence is commonly assumed to be the most prominent syntactic position (i.e. more salient), whereas objects and material introduced in adjunct constituents occupy lesser prominent positions (i.e. less salient) in the sentence structure.

---

[1] Technically, a referent's givenness, accessibility or degree of activation are not identical notions but they have been extensively used as similar ways to operationalize the concept of salience.

Many studies have placed the focus on the differences between personal pronouns and demonstratives across languages (Bosch et al., 2003; Kaiser and Trueswell, 2005; Kameyama, 1999; Navarretta, 2005, 2007; Sturgeon, 2008; Vieira et al., 2002, *inter alia*). For example, in a series of papers, Bosch and Umbach (2007) and Bosch, Katz and Umbach (2008) carried out a series of quantitative and psycholinguistic experiments with the aim to compare and explain the referring properties of German demonstratives (*der*, *die*, *das*) and personal pronouns (*er*, *sie*, *es*) in discourse. Generally speaking, these studies appear to confirm the authors' hypothesis, namely, that German demonstratives clearly prefer non-subject antecedents (i.e. non-nominatives), whereas personal pronouns show a strong preference for subject antecedents (i.e. nominative NP's). These results appear to confirm the *Complementarity Hypothesis* (Bosch, Rozario and Zhao, 2003), which claims that demonstrative pronouns prefer non-topical referents whereas personal pronouns prefer referents that are discourse topics. Their results also align with data reported by Kaiser and Trueswell (2005) for Dutch and Finnish. Examples (1a-b) show the referring preferences of demonstratives and personal pronouns in German.

(1) a. $Peter_i$ wollte mit $Paul_k$ tennis spielen. Doch $er_{\{i, k\}}$ war krank.
          Peter wanted to play tennis with Paul. But he was ill.

     b. $Peter_i$ wollte mit $Paul_k$ tennis spielen. Doch der_k war krank.
          Peter wanted to play tennis with Paul. But DEM was ill.

Following previous cross-linguistic research, my hypothesis regarding the referring behavior of Spanish personal pronouns and demonstratives is that the former will show a strong preference for topical antecedents whereas the latter will prefer non-topical antecedents. However, a first analysis of some natural examples revealed that the picture does not align well with previous findings cross-linguistically. In Spanish, it is possible to refer anaphorically to a noun phrase in preverbal subject position (i.e. a highly topical and salient position) with a personal pronoun, as in (2); and the same pattern can be observed with demonstrative expressions, as in (3). Coreference is marked with a subscript to facilitate the interpretation.

(2) $[Pedro]_i$ es un gran amigo. Lo_i conocí en Barcelona hace ya unos años.
     'Pedro is a great friend of mine. I met him in Barcelona a few years ago.'

(3) $[El monarca Alfonso VI]_i$ fue el primer rey que acuño moneda propia. Este $monarca_i$ fundó una casa de la moneda en Toledo.
     'King Alfonso VI was the first king who coined his own currency. This monarch founded a mint in Toledo.'

The antecedent of a D.O. personal pronoun can be found to occupy diverse and more oblique syntactic positions within the sentence; be it as a direct, indirect object, or as an adjunct. It is reasonable to assume that the antecedent of a D.O. clitic pronoun should be a D.O. in previous discourse (an internal argument of the verb). This is the case in (4), where the personal pronoun *lo* and the D.O. *magnetismo* in the first sentence corefer. The same anaphoric pattern can be found with demonstratives, as in (5), where the D.O. *el ciclotrón* in the first sentence co-refers with the demonstrative noun phrase *este dispositivo*.

(4) La tierra emana $[magnetismo]_i$. Si le echamos cemento encima, lo_i anulamos.
     'Magnetism emanates from the Earth. If we pour concrete on it, we override it.'

(5) Dos físicos estadounidenses inventaron [el ciclotrón]$_i$. Este dispositivo$_i$ emplea un campo magnético.
'Two american physicists invented the cyclotron. This device employs a magnetic field.'

However, the D.O. clitic can also be found to corefer quite frequently with entities occupying lower syntactic positions within the verb phrase. In (6), for example, the D.O. clitic pronoun *lo* is coreferential with the stressed personal pronoun *él* ('him') in the adjunct prepositional phrase *para él*. Demonstratives can also corefer with entities in adjunct position as in (7), where the demonstrative noun phrase *este lugar* co-refers with the noun phrase within the prepositional phrase in the first sentence.

(6) Fue un golpe muy duro para [él]$_i$. Quizá el golpe que lo$_i$ mató.
'That was a very hard stroke to him. Maybe the stroke that killed him.'

(7) El calor es sofocante en [la costa sur]$_i$. Dense una vueltecita por este lugar$_i$ y lo comprobarán ustedes mismos.
'It is very hot in the Southern Coast. Take a walk around here (this place) and you will feel it yourselves.'

In this paper, I compare the referring properties of Spanish demonstratives and direct object pronouns in their role as discourse anaphors. Following previous research on the information status of referring expressions (see the work by Prince, 1981b; Ariel, 1988, 1990; Gundel et al, 1993; Hegarty et al, 2003; Poesio and Modjeska, 2005, *inter alia*), I put the main focus in checking whether significant differences can be found in the referring behavior of these expressions and whether the alleged differences, if any, may have a bearing on the information status of their referents. With this purpose, I have conducted a corpus study where I have tested two factors that will help us distinguish between the two referring expressions: *textual distance of the antecedent* and *antecedent type*. My corpus data reveals no significant differences among Spanish demonstratives (determiners and pronouns) and D.O. personal pronouns with regard to their anaphoric capability or in the information status of their antecedents; an observation that contradicts similar research cross-linguistically.

Based on these observations, my main claim is that the basic difference between Spanish D.O. clitic pronouns and demonstratives lies, in the discourse domain, in that the former play the role of topic-continuity markers, whereas the latter serve to introduce new topics or subtopics in the discourse via the focalising function of deictic demonstratives. I also argue that the topic-shifting function of demonstratives is transferred from their more basic deictic function to the 'derived' discourse anaphoric (textual) domain.

The role of D.O. clitic pronouns and demonstratives needs to be framed within a notion of dynamic discourse context; that is, an informational discourse context that extends beyond the immediate prior discourse and beyond the immediate local topic, and where the different referring behavior of these expressions as discourse anaphors can be accounted for. A dynamic discourse context is a context that evolves as the discourse unfolds, where some discourse referents lose informational prominence and are replaced by newly introduced referents, and where less salient or declined topics can be re-introduced and gain topical status. In such a dynamic context, language users establish a ranking of topicality among discourse referents, and it is only within this dynamic context that differences in the referring capabilities of demonstratives and direct object pronouns emerge.

2. Background

A good number of corpus studies have been made on the phenomenon of discourse anaphora/deixis over the last two decades, greatly contributing to our comprehension of the anaphoric processes that arise in discourse. Some of these studies have put the focus on providing an adequate annotation scheme for discourse deixis/anaphora, while others have been more focused on the quantitative part. Some studies, however, combine the two perspectives (see, for example, Poesio 2004; Poesio and Artstein 2008 for English; Dipper and Zinsmeister 2009 for German; Navarretta and Olsen 2008 for Danish and Italian; Recasens 2008 for Catalan and Spanish).

Many other studies have placed the focus on the analysis of referring expressions and information status across languages and different hierarchical scales have been proposed to account for the different distribution shown by referring expressions across languages. Prince (1981b) was the first to propose a hierarchy for discourse entities called the *Scale of Familiarity*, which is based on assumed knowledge (i.e. the assumption by the speaker of the cognitive status of the referent in the mind of the audience).

In Ariel's (1988) *Accessibility Scale*, the notion of accessibility is defined as the relative ease with which the addressee can identify the referent of a referring expression or, alternatively, the ease with which the addressee can retrieve the intended referent from memory. In Ariel's scale, demonstratives occupy an intermediate position in terms of the degree of accessibility they confer to their referents. As the accessibility scale clearly indicates, the less informative forms (gaps, PRO's, etc) occupy the highest position; that is, they are high-accessibility markers. Unstressed and cliticized pronouns also occupy a high position in the scale though not so high as phonetically null forms.

Gundel, Hedberg and Zacharski's (1993) *Givenness Hierarchy* is an implicational hierarchy of cognitive states and linguistic forms aimed to resolve the different anaphoric behavior of pronominal and non-pronominal anaphors. According to this hierarchy, the referents of demonstratives have either *activated* or *familiar* status but are never *in focus*, whereas the referent of a personal pronoun always has the status *in focus*. For a referent to be activated at a given point in the discourse, there must be a representation of the referent in short-term memory. A referent *in focus* is in short-term memory and also at the current center of attention. As these authors point out, entities *in focus* constitute a partially-ordered subset of activated entities that are more likely to be the topic in subsequent discourse.

Gundel, Hedberg and Zacharski (2005) analyzed the behavior of English demonstratives 'this' and 'that' and the unstressed pronoun 'it' in the Santa Barbara corpus of spoken American English. They observed that demonstrative anaphors were used to refer to abstract entities in 85% of the cases, whereas only 15% of the cases were anaphorically referred to with the pronoun 'it'. They claimed that the observed distributions were due to the type of material introduced in clauses: clausally-introduced entities (i.e. propositions and events, which are typical antecedents for demonstrative anaphors) are likely to be *activated*, whereas noun phrases in prominent syntactic positions are more likely to be *in focus*.

Poesio and Modjeska (2005) tried to refine linguistically the cognitive notions *in focus*, *activated* and *short-term memory* following the computational approach to anaphora resolution of Centering Theory (Grosz, Joshi and Weinstein, 1995). These authors annotated the corpus GNOME for this purpose and tested the following hypotheses regarding speakers' preferences for 'This-NPs' to refer to *in focus* entities:

- This-NPs are preferentially used to refer to entities other than the $C_b(U_i)$, the C_b of the utterance containing the This-NP.
- They are used to refer to entities other than the $C_b(U_{i-1})$, the C_b of the previous utterance.
- They are used to refer to entities other than $C_P(U_{i-1})$, the most highly ranked entity of the previous utterance.

In Centering Theory, it is assumed that new discourse entities (forward-looking centers or C_F's introduced in each utterance are ranked based on information status. The forward-looking centers of U_n only depend on the expressions that constitute that utterance; they are not constrained by features of any previous utterance in the segment. The C_b (the backward-looking center of an utterance U_n) is Centering's equivalent of the notion of topic or focus. The backward-looking center of U_i connects with one of the forward-looking centers of U_{i-1}. The $C_b(U_i)$, the backward-looking center of utterance U_i, is the highest ranked element of $C_F(U_{i-1})$ that is repeated in $C_F(U_i)$.

Poesio and Modjeska proposed a general hypothesis regarding the speaker's preference to use This-NPs for reference to *activated* (or *active* in their own terminology) discourse entities. An entity is ACTIVE if:

- It is in the visual situation; or
- It is a C_F of the previous utterance; or
- It is part of the implicit linguistic focus. They only considered as part of the implicit focus those entities that can be constructed out of the previous utterance. An entity can be constructed out of an utterance if:
  A) It is a plural object whose elements or subsets have been explicitly mentioned in that utterance; or
  B) It is an abstract entity introduced by that utterance. They consider two types of abstract entities:
  i. Propositions
  ii. Types

The distributions from their study are shown in Table 1.

Table 1. Distribution of This-NPs (Poesio & Modjeska 2005)

CLASS	n	%
Anaphora	45	40
Visual Deixis	28	25
Discourse Deixis	19	17
Type	9	8
Plurals	1	
Ellipsis	1	
Time	1	
Unsure	5	
Disagreement	3	
Total	112	

With respect to the correlation between focus and This-NPs, they found the following principal results:

- 8-11 violations to the hypothesis that a This-NP is used to refer to entities other than the $C_b(U_{i-1})$ were found, which is therefore verified by 90%-93% of This-NPs.
- The hypothesis that This-NPs are used to refer to entities other than $C_P(U_{i-1})$ is verified by 75-80% of This-NPs.
- The hypothesis that This-NPs are used to refer to entities other than $C_b(U)$ is verified by 61-65% of This-NPs.

So the hypothesis that received more empirical support is the following: 'This-NPs' are used to refer to entities which are active but not the backward-looking center of the previous utterance. Based on these results and on an in-depth study of the violation cases, they proposed the version that leads to the fewest number of violations to Grice's Maxim of Quantity (1989):

(8)   The This-NP Hypothesis: 'This-NPs' are used to refer to entities which are active in the sense specified above but not $C_b(U_{i-1})$

In a series of papers, Hegarty (2003, 2006) and Hegarty et al. (2001, 2003) studied abstract object anaphora from a semantic perspective. Generally speaking, all these studies coincide in that clausally-introduced entities are more commonly referred to with a demonstrative pronoun hence indicating that the cognitive status of these entities is *activated*. However, with regard to the cognitive statuses in the Givenness Hierarchy, Hegarty indicates that an entity will be *in focus* only if it has been mentioned by a nominal expression in a prominent syntactic argument position earlier in the utterance or in the previous utterance; a supposition which is compatible with Centering Theory and with experimental psycholinguistic data. On the other hand, peripherally introduced entities, including those introduced by less prominent nominal expressions and by clauses, will be activated upon their introduction, placed in working memory within the field of attention, but never at the center of attention.

In a similar fashion, Brennan (1998) showed that a pronoun needs to first have its referent introduced in object position before it can be made a pronoun in subject position (that is, in focus).

As we have seen so far with regard to data from English, there appears to be agreement on the information and cognitive status of the entities referred to with demonstratives and the weak personal pronoun 'it', especially when reference to abstract entities is involved. Thus, speakers would use demonstratives to refer to activated entities, which rank lower than in focus entities regarding their cognitive and information status. Unlike demonstratives, the pronoun 'it' would be strongly preferred for reference to entities in the current focus of attention; that is, in-focus.

However, cross-linguistic research indicates that these findings cannot be easily extrapolated to other languages. For example, Navarretta (2008) found language-specific results for Danish and Italian regarding the referring properties of demonstratives and personal pronouns. She found that demonstrative anaphors *det* and *dette* can be used with all antecedent types and to refer to all types of referents. Furthermore, Danish personal pronouns can also be used with clausal antecedents.

With regard to Italian, Navarretta observed that zero anaphors and personal pronouns are often used in this language in contexts where demonstrative pronouns occur in English. Also, zero anaphors are the most frequently used pronouns to refer to propositions in Italian (let us remind that the referents of zero pronouns are *in focus* in the Givenness Hierarchy). These data indicate important cross-linguistic differences in the referring behavior of referring expressions and/or the information status of abstract referents. My data from Spanish appear to point in a similar direction. I present my findings in the following paragraphs.

3. A pilot study on referring expressions and grammatical role

3.1 Demonstratives
Spanish has a tripartite system of demonstrative determiners and pronouns with proximal *este* 'this', medial *ese* 'that' and distal *aquel* 'that'. The system is shown in Table 2.

Table 2. Demonstratives of Spanish

GENDER	NUMBER	DETERMINERS AND PRONOUNS		
Masculine	Singular	Este/Ese/Aquel	(NP)	'this/that'

Feminine		Esta/Esa/Aquella	(NP)
Masculine	Plural	Estos/Esos/Aquellos	(NP)
Feminine		Estas/Esas/Aquellas	(NP)
Neuter	N/A	Esto/Eso/Aquello	

It is widely assumed that demonstratives can be used referentially in two different basic modes: deictically and (discourse) anaphorically.[2] In the deictic mode (also known as situational deixis), speakers use demonstratives to refer to physical, tangible entities in the speech situation. Like other deictic expressions, the interpretation of deictic demonstratives is bound to some contextual parameters such as the *here* and *now* of the discourse situation. When a demonstrative is used deictically in discourse, a particular communicative effect is obtained whereby the intended referent is made salient or prominent from among a set of potentially competing entities. Such effect is also known as the *joint attention effect* in the psycholinguistics literature (Diessel 2006).

In the discourse anaphoric mode of reference (see Ehlich 1982; Lyons 1975, 1979; Kleiber 1990; Maes and Noordman 1995; Cornish 2008, among others) demonstratives are coreferential with a range of discourse entities such as individuals, eventualities, event-types or propositions (Webber 1979, 1988; Asher 1993). Syntactically, the antecedents of demonstrative anaphors can be of a varied nature, i.e. noun phrases, subordinate clauses, complete sentences or even larger textual (possibly disrupted) discourse fragments. Examples of demonstratives co-referring with individuals via noun phrase antecedents were shown in examples (3), (5) and (7) above. Examples of demonstrative expressions co-referring with events are shown in (9) and (10), respectively.

(9) [Por fin Juan ha conseguido un trabajo]$_i$. Eso$_i$ me alegra.
     'Finally John got a job. That makes me happy.'

(10) Hoy celebramos el aniversario de [la Revolución Francesa]$_i$. Aquel acontecimiento$_i$ tuvo lugar en 1792.
     'Today we celebrate the anniversary of the French Revolution. That event occurred in 1792.'

3.2 Direct object personal clitic pronouns (D.O. clitic pronouns)
For the purposes of this paper I take the third person D.O. pronoun as prototypical for the class of Spanish personal pronouns. I have deliberately chosen this pronoun due in part to its ability to semantically corefer, similarly to demonstratives, with a wide range of discourse elements such as individuals, properties, events or propositions. This is, for example, the case of the neuter form *lo*. In this section, I briefly describe the anaphoric uses of the personal clitic pronoun.

Table 3. Spanish 3rd Person D.O. Pronouns

GENDER	NUMBER	3RD PERSON PRONOUNS
Masculine	Singular	**Lo** ('him')
Feminine		La ('her')
Masculine	Plural	Los ('them')
Feminine		Las ('them')
Neuter	N/A	**Lo** ('it')

---

[2] The reality about demonstratives in natural languages is much more complex than the basic distinction deictic/anaphoric may suggest. See Diessel (1999) and Dixon (2003) for detailed typological studies of demonstratives, including all parameters of variation (form, function and reference type).

According to the Nueva Gramática de la Lengua Española (2011), *lo* is a 3rd person singular accusative pronoun without grammatical distinction for masculine or neuter. Thus, *lo* is masculine whenever it refers to masculine nouns as in (11), where *lo* and the noun phrase *el niño* co-refer, and neuter when its antecedent is clausal, as in (12), where the pronoun anaphorically refers to the clause *que Juan es divertido* [i.e. an internal argument of the modal verb *creer* ('believe')]. According to the Nueva Gramática, the pronoun *lo* is also neuter when its antecedent is another neuter pronoun, as in (13); or when it refers to a noun phrase headed by the article *lo* as in (14).[3]

(11) Yo vi   al   [niño]ᵢ pero Juan no loᵢ vio.
     I   saw to-the.MASC   boy   but   John not lo saw
     'I saw the boy but John didn't see him.'

(12) María cree   que [Juan es divertido]ᵢ pero yo no loᵢ creo.
     Mary believes that John is funny   but I   not lo believe
     'Mary thinks that John is funny but I don't think so.'

(13) [Eso]ᵢ loᵢ veremos.
     That   lo   we-will-see.
     'We'll see.'

(14) [Lo que compraste]ᵢ   no loᵢ podrás   pagar.
      Lo that you-bought   not lo you-will-be-able   to-pay
     'You won't be able to pay what you bought.'

As the Nueva Gramática points out the Spanish direct object pronoun can refer, like any noun phrase, to people, things and animals. Like the stressed pronoun *ello* and other neuter pronouns such as the demonstratives *esto* ('this'), *eso* and *aquello* ('that') the pronoun *lo* can also be coreferential with clauses or verb phrases, or, more accurately, with the semantic entities that these elements denote. An example of the pronoun co-referring with a proposition was shown in (12). An example of *lo* co-referring with a verb phrase is shown in (15), where the VP *aprender inglés* ('learn English') and the pronoun in the second conjunct are coreferential.

(15) Tienes que [aprender inglés]ᵢ, y debes hacerloᵢ cuanto antes.
     'You need to learn English, and you should do that as soon as possible.'

3.3. Data and results
As I mentioned, the most widely accepted thesis, and my working hypothesis, is that antecedents of demonstratives are most commonly non-topical whereas personal pronouns commonly have topical elements as their antecedents. When topichood is assumed to be dependent on syntactic

---

[3] In Spanish, there is a *lo* that is considered a neuter definite article by some scholars (Contreras 1973), but a neuter pronoun by others (Bosque and Moreno 1990). This is the *lo* that we can find in examples like (i). For the purposes of this paper I will not consider this type of *lo* an anaphoric pronoun, hence it was not included it in my study. See also See Gutiérrez-Rexach (1999) and Leonetti (1999) for detailed studies on *lo* as a degree quantifier.
(i)   Lo que tú   digas.
      it that you say-SUBJ
      'What you say.'

configurations then highly prominent positions (i.e. subject) are topical whereas less prominent syntactic positions (i.e. object, adjunct) are non-topical. I made a pilot corpus study to check the variability in syntactic prominence of the antecedents of the direct object personal pronoun *lo* and the demonstrative determiner *ese-NP* ('that-NP'). My examples were collected from the Corpus del Español (Davies 2002-) and a total of 100 examples were analyzed in two different data sets. The examples come from both oral and written language.

In order to check the prominence of the antecedent of the referring expression I considered the grammatical role within the sentence (i.e. VP internal for objects and adjuncts and VP external for subjects). The results corresponding to the direct object personal pronoun are shown in table 4. See the Appendix for examples of some of the analyzed cases.

Table 4. Grammatical Role of NP Antecedents for D.O. Pronouns

GRAMMATICAL ROLE OF NOUN PHRASE ANTECEDENT	n	%
Subject	15	30
Direct Object	14	28
Indirect Object	0	0
Predicative in copula construction	6	12
Adjunct	5	10
Object of preposition	3	6
Left-dislocated Direct Object	7	14
Total	50	100

I found 5 cases of left-dislocated direct object, which despite being categorized differently, can be considered highly-topical constituents due to the topicalizing function of this configuration. Examples of the left-dislocated direct object and attributive constructions are shown in (16) and (17), respectively.

(16) [A Juan]$_i$ lo$_i$ vi   ayer.
    to John   him I-saw yesterday
      'I saw John yesterday.'

(17) Juan y   María eran [amigos]$_i$, pero ya   no lo$_i$ son.
    Juan and María were friends,   but already not lo they-are.
    'John and Mary used to be friends, but they are not anymore.'

The data in table 4 show a distribution with equal numbers for antecedents of the personal pronoun in topical (i.e. Subject and Left-dislocated D.O.; n = 22; 44%) and non-topical positions (i.e. D.O., Adjunct and Object of Preposition; n = 22; 44%) Cases of attribute antecedents were not accounted for because the accessibility or topical status of predicates in copula constructions is unclear. Overall, the data show that the direct object personal pronoun has a preference for antecedents having the role of direct object (i.e. D.O.'s, Left-dislocated D.O.'s and Objects of preposition combined; n = 24; 48%), hence non-topical positions, over strict preverbal subject antecedents. Nevertheless, the data become more balanced when left-dislocated direct objects are taken on syntactic prominence grounds; that is, they are treated as highly-topical elements akin to preverbal subjects. These data do not confirm the hypothesis that personal pronouns take topical elements as antecedents.

The set of data pertaining to demonstratives is given in table 5. The demonstrative analyzed is medial determiner *ese* in *ese-NP* constructions. The distribution of the demonstrative is, to a certain

extent, what we expected from our working hypothesis. The number of antecedents in non-topical positions (D.O., Adjunct and Object of preposition combined; n = 32; 64%) is higher than those in topical positions (preverbal subjects; n = 17; 36%) and these numbers seem to confirm the hypothesis that demonstratives generally take non-topical elements as antecedents. However, these figures are not clear-cut as a high number of occurrences of antecedents in preverbal subject position were found too.

Table 5. Grammatical Role of NP Antecedent for Demonstrative Determiner 'Ese NP'.

GRAMMATICAL ROLE OF NOUN PHRASE ANTECEDENT	n	%
Subject	17	34
Direct Object	16	32
Indirect Object	0	0
Predicative in copula construction	1	2
Adjunct	11	22
Object of preposition	5	10
Left-dislocated Direct Object	0	0
Total	50	100

No clear patterns can be observed from the small set of data analyzed in this pilot study. However, the qualitative analysis of the examples shows a great variability in the type of antecedent for demonstratives and direct object personal pronouns alike. As the distributions appear to indicate, the Spanish D.O. personal pronoun does not show any particular preference for highly-prominent antecedents and this would go against my starting hypothesis. On the other hand, while demonstratives show a marked tendency to refer to non-topical antecedents, the distribution is not conclusive given that a high percent of referents in subject position could be found. In my view, the syntactic position within the sentence or the grammatical role of the antecedent cannot be the only factors explaining the referring behavior of these expressions in Spanish. In Section 4, data from another corpus study are presented, where I analyze two additional factors that may constrain the referring properties of demonstratives and D.O. personal pronouns, namely, referential distance (as an alternative way of measuring antecedent salience) and type of antecedent.

4. An alternative way of measuring salience: antecedent distance
The first discourse parameter to be investigated is antecedent distance, a factor that may influence the degree of accessibility of referring expressions (see, *inter alia*, Maes and Noordman 1995 and Ariel 2001 for discussion on this topic). By antecedent distance I mean the distance between the antecedent and anaphor quantified as the number of intervening clauses between these two elements. Hegarty et al. (2001) observed that English demonstrative pronouns and adjectives (this/that-(NP) show a strong preference for their antecedents to be found in the clause next to (i.e. immediately preceding) the demonstrative expression. In quite the same line, Kirsner et al. (1988) explored the factors that affect Dutch demonstrative (deze 'this' vs die 'that') choice in written discourse. One of these determining factors was the magnitude of antecedent distance. Based on texts from different subcorpora tested on native Dutch speakers and comprising various different discourse genres, these authors showed that the Dutch proximal demonstrative deze ('this') tends to be associated with antecedent distance ≥ 1 (extrasentential retrieval of a referent) and distal die ('that') tends to be associated with antecedent distance = 0 (intrasentential retrieval of a referent). Their study showed that only 15% of NPs with deze ('this') have antecedent distance = 0, whereas a 40% of the NPs with die ('that') have antecedent distance = 0.

4.1. Data and results

For this study, I obtained my examples from the CREA corpus, a large linguistic database of over 160 million words that includes different Spanish varieties, text types and genres. The corpus search engine allows users to retrieve a discourse fragment, thus placing words in context. 50% of CREA sources are from Spain (±45 million speakers), and 50% from Latin America (±350 million speakers). 90% of the words in the corpus come from written sources, and only 10% from oral sources. This corpus is not annotated hence the annotation had to be done by hand and only for the cases analyzed. The size of the corpus and the high frequency of demonstrative expressions in Spanish made it impractical to analyze all the occurrences found. I analyzed a total number of 327 occurrences distributed as follows: 120 occurrences of the personal pronoun *lo* ('it') and 207 occurrences of demonstrative expressions.

All D.O. personal pronoun cases analyzed (n = 120) were divided into three groups corresponding to three different corpus searches: *lo entiendo* ('I understand it'), *lo necesito* ('I need it') and *lo tengo* ('I have it'), hence 40 occurrences per group were scrutinized. The reason for having analyzed these particular combinations is twofold. On the one hand, this strategy allowed me to discard other, non-referential uses of this pronoun. Also, these three groups would allow me to test not only antecedent distance but also the denotation of the antecedent and check whether it may possibly have any influence on the cognitive status and different accessibility marking shown by the personal pronoun. Thus, by using the predicates *entender* ('understand'), *necesitar* ('need') and *tener* ('have') I am forcing different semantic readings for the antecedent. As the data show, the predicate *entender* ('understand') shows a preference for higher order antecedents such as concepts or ideas rather than concrete, physical objects. Conversely, the verb *tener* in the expression *lo tengo* ('I have it') exhibits a preference for tangible referents as people have/own physical objects. The verb *necesitar* ('need') is intended to occupy an intermediate position in between the former two predicates. The aim overall was to obtain a sample ample enough to be able to draw some initial conclusions regarding the possible influence of antecedent denotation.

In order to check antecedent distance the corpus examples were segmented into constituent sentences as in (18). Sentences have been numbered from S_0 (i.e. sentence 0, that is, the sentence containing the anaphor) to $S_{\geq 4}$ (i.e. sentence 4 or higher). Sentence count is made from the position of the anaphor backwards. The anaphor is written in bold characters and the intended antecedent is shown in brackets. In this example, the antecedent to the demonstrative pronoun can be found in sentence S_3.

(18) Al fin y al cabo, si ustedes están aquí es porque
    quieren que les hable de [la Operación Ópera.]$_i$ # S_3
    ¡Me cuesta tanto volver al pasado!# S_2
    Ya comprenderán, el tiempo aquí transcurre de otra manera.# S_1
    Y todo **aquello**$_i$ sucedió en el 92.# S_0

    'After all, it is obvious that you are here because
    you want me to tell you about [the Operation Opera.]$_i$ # S_3
    It's so hard for me to go back in the past!# S_2
    As you will see, time goes by differently here.# S_1
    And all **that**$_i$ happened in 1992.#' S_0

The distribution of antecedent distance and D. O. personal pronouns is shown in table 6.

Table 6. Antecedent Distance for D.O. Personal Pronouns

	S0		S1		S2		S3		S ≥ 4	
	n	%	n	%	n	%	n	%	n	%
Lo necesito ('I need it')	3	7.5	30	75.0	4	10.0	0	0	3	7.5
Lo entiendo ('I understand it')	2	5.0	37	92.5	0	0	1	2.5	0	0
Lo tengo ('I have it')	11	27.5	26	65.0	2	5.0	1	2.5	0	0
Total	16	13.33	93	77.5	6	5.0	2	1.66	3	2.5

Next, I analyzed 207 occurrences of demonstratives (n = 207). The results of this sample are shown in tables 7 and 8. In the first place, a first sample of 50 demonstrative determiners was divided into two groups of 25 cases each: *este hecho* ('this fact') and *ese hombre* ('that man'). The reasons for having analyzed these particular noun phrases are the same that I explained for the personal pronoun in the previous paragraph. With the nouns *hecho* and *hombre* I narrow down the search space and force different denotations for the antecedent; that is, a higher-order entity and a tangible entity, respectively. A second corpus search consisted of 157 cases of demonstrative pronouns: 63 instances of the proximal demonstrative pronoun *esto* ('this'), 69 of medial *eso* ('that') and 25 of distal *aquello* ('that'). The disparity of the analyzed occurrences of demonstrative pronouns, in particular the low number of tokens for pronoun *aquello* (25), is due to the actual frequency of use of demonstratives in modern Spanish. Overall corpus figures show that demonstrative *aquello* has a very low frequency of use (6%) compared to the frequencies shown by *esto* and *eso*. Even between the latter the differences are relevant (*eso*: 60%) and (*esto*: 34%). Nevertheless, overall figures vary when the frequency of use of demonstrative determiners is considered. Demonstrative determiner *ese* has a frequency of 30% whereas determiner *este* shows a percentage as high as 61%. Again, demonstrative determiner *aquel* shows a rather low frequency of use (9%) in modern Spanish.

Table 7. Antecedent distance for Demonstrative Determiners

	S0		S1		S2		S3		S ≥ 4	
	n	%	n	%	n	%	n	%	n	%
Este hecho ('this fact')	1	4.0	21	84.0	3	12.0	0	0	0	0
Este hombre ('this man')	2	8.0	19	76.0	2	8.0	0	0	2	8.0
Total	3	6.0	40	80.0	5	10.0	0	0	2	4.0

Table 8. Antecedent Distance for Demonstrative Pronouns

	S0		S1		S2		S3		S ≥ 4	
	n	%	n	%	n	%	n	%	n	%
Esto ('this')	0	0	50	79.4	11	17.5	1	1.6	1	1.6
Eso ('that')	0	0	54	78.2	5	7.2	7	10.1	3	4.3
Aquello ('that')	0	0	19	76.0	4	16.0	0	0	2	8.0
Total	0	0	123	79.0	20	12.0	8	5.0	6	4.0

The second factor analyzed was the type of the antecedent. Two antecedent types were considered: noun phrase and sentence antecedents. The Sentence category also includes infinitival clauses and

other antecedents that expand beyond the clause (i.e. larger segments of text comprising two or more sentences). The distributions by antecedent type are shown in table 9.

Table 9. Type of Antecedent

	Noun Phrase Antecedent		Sentence Antecedent	
	n	%	n	%
Lo necesito ('I need it')	20	50.0	20	50.0
Lo entiendo ('I understand it')	13	32.5	27	67.5
Lo tengo ('I have it')	33	82.5	7	17.5
Este hecho ('this fact')	6	24.0	19	76.0
Este hombre ('this man')	25	100.0	0	0.0
Esto ('this')	6	9.5	57	90.5
Eso ('that')	16	23.2	53	76.8
Aquello ('that')	5	20.0	20	80.0
Total	124	38.0	203	62.0

In general, sentence antecedents are widely preferred over noun phrase antecedents (62% vs 38%, respectively) when all referring expressions are considered. The individual analysis of these expression types showed the following frequencies: the D.O. personal pronoun shows a slight preference for noun phrase antecedents (55%) over clausal ones (45%). Some individual differences appear to be based on the type of the predicate accompanying the personal pronoun or demonstrative determiner. For example, the expression *lo entiendo* ('I understand it'), typical of higher-order referents, shows a strong preference for sentence antecedents over NPs (27 and 13 occurrences, respectively). Conversely, the personal pronoun in the expression *lo tengo* ('I have/own it') shows a strong preference for NP over sentence antecedents (33 and 7 occurrences, respectively). Demonstrative pronouns *esto, eso* and *aquello* show a strong preference for sentence antecedents (90%, 76% and 80%, respectively), whereas demonstrative determiners show opposite preferences depending on the noun involved in each particular expression: the noun phrase *este hecho* ('this fact') shows a strong preference for sentence antecedents (76%), most likely due to the denotation of the noun, whereas the noun phrase *este hombre* ('this man') shows an even stronger preference for noun phrase antecedents (100%).

4.2. Discussion

With regard to antecedent distance, the three categories (i.e. D.O. pronouns, demonstrative determiners and demonstrative pronouns) show a strong tendency to find their antecedents in the clause next to the anaphor (S_1). Total frequencies are very similar for the three categories: 77% (personal pronouns), 80% (demonstrative determiners) and 79% (demonstrative pronouns). With only minor exceptions, a general tendency is observed that can be stated as follows:

- Concerning the use of demonstratives and D.O. personal pronouns we observed the following strong correlation: the shorter the distance between antecedent and anaphor the higher the incidence.

However, D.O. personal pronouns show a significant rate of co-occurrence with antecedents in S_0 (the clause/sentence containing the anaphor). In our corpus sample, that is due to a frequent

construction that combines a noun phrase, most likely a demonstrative pronoun, with the D.O. clitic pronoun *lo*. An example from the corpus is shown in (19).

(19) [El ojo del vertebrado tiene diseño y realiza una función]$_i$, y eso$_i$ lo$_i$ hace indistinguible de la cámara fotográfica.
'The eye of the vertebrate has a design and a function, and that makes it indistinguishable from the photographic camera.'

As I pointed out in Section 3.2, the Spanish grammar stipulates that the D.O. clitic pronoun *lo* is a neuter pronoun when its antecedent is another neuter pronoun. In this study, the cases of the pronoun in this particular configuration have been included in the S_0 group. The demonstrative pronoun *eso* refers back to the antecedent in the previous clause, or to the actual clause as in (19), and the pronoun *lo*, in turn, has the demonstrative as its antecedent. So the relevance of this construction lies in the ability of the two referring expressions to co-occur within the same clause and co-refer with the same discourse entity while having different syntactic antecedents.

In summary, my data in connection with antecedent distance show that the three anaphors show a strong preference to find their antecedent in S_1. Therefore, assuming that textual distance is an appropriate operationalization of referent salience, there appear to be no relevant or major differences between demonstratives and personal pronouns in this respect. Like the previous data with regard to grammatical role of the antecedent, the present data cannot confirm my working hypothesis that antecedents of demonstratives are less salient or accessible (i.e. *activated*) than antecedents of personal pronouns (i.e. *in focus*).

Finally, highly significant distributions were observed with regard to antecedent type. A low number of demonstrative pronouns referring to NP antecedents was observed (X^2=54.0238 (df = 2), p <0.0005). This is not surprising as demonstrative pronouns are most commonly used to refer to abstract entities in Spanish, so the distribution clearly indicates that abstract entities are most usually conveyed via clausal antecedents. Also, grouping together all demonstratives (determiners and pronouns) and personal pronoun cases, I still observed a very strong preference for clausal antecedents over NPs ones (62.4% and 37.6%, respectively); X^2=24.4165 (df = 1), p <0.0005. And similar frequencies can be observed by comparing demonstrative pronouns and personal pronouns (66.4% of clausal antecedents and 33.6% of NP antecedents, respectively) showing a highly significant distribution (X^2=43.5814 (df = 1), p <0.0005). However, D.O. personal pronouns show a slight preference for NP over clausal antecedents (55% and 45%, respectively).

Again, it looks like antecedent distance will not help us discriminate among the referring attributes of the expressions analyzed in this study. My study revealed that the preferred location of the antecedent is S_1 across referring expressions. If we consider that recency of mention is a factor having an effect on the information status of an antecedent it seems safe to conclude that no significant differences can be observed in the information status (i.e. *in focus* vs. *activated*) of the referent of personal pronouns and demonstrative expressions. It should be kept in mind that more recent antecedents (those found in S_0 and S_1) require less processing effort in terms of memory storage retrieval, hence becoming highly-accessible or in the current focus of attention.

On the other hand, although numbers indicate that demonstratives show a strong preference for clausal antecedents, my data also showed a high number of cases of clausal antecedents with the D.O. personal pronoun (55% of NP and 45% of clausal antecedents for n = 120).

The principal findings from the two pilot corpus studies can be summarized as follows:

- When the grammatical role of the antecedent is considered, the D.O. personal pronoun does not have a preference for antecedents in topical positions within the sentence (i.e. IP-external).

- When the grammatical role of the antecedent is considered, demonstratives have a preference for non-topical antecedents (i.e. IP-internal), although antecedents in topical position are also common.
- When distance of the antecedent is considered as an operationalization of salience, demonstratives and the personal pronoun alike show a strong preference to find their antecedents in the clause next to the anaphor (S_1), i.e. a position that is considered to be highly salient due to recency of mention.
- With regard to antecedent type, the denotation of the noun accompanying the anaphor matters for antecedent preference.
- Overall, the starting hypothesis that personal pronouns prefer topical (or salient, highly accessible, in focus) antecedents whereas demonstratives prefer non-topical (or less salient, less accessible, activated) antecedents could not be confirmed.

5. Demonstratives as focusing devices

As referring expressions, demonstratives encode two basic modes of reference that allow speakers refer to entities in different discourse situations: the deictic and discourse anaphoric modes. In the deictic mode, the speaker generally uses demonstratives to refer to tangible, concrete entities present in the context of utterance. In this mode, use of the demonstrative (possibly accompanied by a pointing gesture) has the communicative effect of focusing the attention of the interlocutor on a specific entity in the visual or perceptual field. By making the intended entity salient from among a set of (potentially) competing entities, both speaker and addressee focus their attention on the same entity and the desired communicative effect is accomplished. In the psycholinguistics literature this effect is known as the *joint attention effect* (Diessel 2006). In (20), the speaker points to a particular house while he utters:

(20) I bought that house last year.

In these deictic uses the gesture that accompanies the demonstrative is generally essential for the interpretation as it serves to complete the meaning of the demonstrative and to possibly resolve the reference. In this particular example, the house being pointed at could have been surrounded by other houses. In such case, use of the demonstrative without an explicit indication would not have had the desired effect and the hearer would not be able to resolve the reference.

There still exist some controversy about whether the deictic and anaphoric modes of referring should be considered totally independent modes or they should be taken as different degrees of a more general spectrum of deictic procedures available in natural languages. In this sense, and given that anaphoric reference appears to be a derivative of deixis (Lyons 1975; 1979), these referring types would not be mutually exclusive and a certain degree of overlap between them could be expected (see Ehlich 1982; Cornish 2009, on these issues and on the 'hybrid' mode of discourse reference known as *anadeixis*.)

In my view, discourse-anaphoric demonstratives somehow encode a focusing procedure, some type of deictic-like instruction derived from primary deictic uses whereby the intended referent is momentarily placed in the current focus of attention of the addressee. This procedure could be seen as an instruction from the speaker to the hearer, something like: "*Even if it is not the current discourse topic, focus on this particular entity for it will be important from now onwards*." In this regard, the only difference between the deictic and discourse anaphoric modes would lie in that the 'pointing' effect of the deictic demonstrative would be transferred from a real-world domain (i.e. exophora) into a textual domain (i.e. endophora) in the discourse-anaphoric mode. The basic communicative function would remain the same; however, the explicit deictic effect would have

evolved into a derived pragmatic function in the anaphoric use. This could help us explain the minimal differences between personal pronouns and demonstratives in Spanish shown in this paper. Thus, whereas personal pronouns would serve the primary function of marking *topic continuity* in discourse (Givón 1983) demonstratives would rather mark new topics or even old-topics that have decayed over a stretch of discourse.

For example, the (underlined) antecedent of the demonstrative (in bold characters) is introduced in the first sentence of (21) via an adjunct phrase. After processing the first sentence, it seems that the topic of this segment is 'a fortress built in the period of the Crusades'. Then, as the discourse unfolds, it appears that the topic changes to 'the military architecture' (possibly a sub-topic) for a few sentences. Next, the demonstrative brings back again, focuses on, the original topic that had been suspended.

(21) El Krak de los Caballeros (1131) en Jordania, una fortaleza construida por la Orden de los Caballeros Hospitalarios en <u>la época de las cruzadas</u>. La arquitectura militar fue una respuesta defensiva contra los avances en la tecnología militar; en todo caso, una de las estrategias más importantes seguía siendo resistir un asedio. Muchas ciudades se resguardaban dentro de una muralla fortificada y así se han conservado hasta nuestros días recintos como el de la ciudad de Ávila, en España, Aigues-Mortes y Carcasona en Francia, Chester en Inglaterra o Visby en Suecia. **Este periodo histórico** coincide con un espectacular auge de la población urbana a causa del desarrollo tecnológico....

'The Krak of the Knights (1131) in Jordan, a fortress built by the Order of the Hospitalary Knights <u>in the time of the Crusades</u>. Military architecture was a defensive response against the advances in military technology; in any case, one of the most important strategies still was how to resist a siege. Many cities would protect themselves inside a fortified wall and so enclosures like the one in the city of Avila in Spain, Aigues-Mortes and Carcassone in France, Chester in England or Visby in Sweden have been preserved until today. **This historical period** coincides with an spectacular increase in urban population due to technological advances... '

In (22), we see the interaction of the 3rd person pronoun as a marker of topic continuity and a demonstrative noun phrase as a pragmatic marker. The antecedent of both anaphors is the noun phrase *este muchacho* ('this boy') introduced in the first sentence and very likely the topic of the discourse. Notice the use of the 3rd person pronouns (i.e. *lo* and *le*) in the first sentences marking the continuity of this particular topic. Then the speaker seems to digress a bit for two-three sentences, where he mentions his thoughts about lawyers in general. Next, the demonstrative anaphor resumes the old topic 'the boy who was in prison'.

(22) Entonces, a <u>este muchacho</u> lo... pues, **lo** logré sacar de ahí con unos tres años de prisión nada más ¿ eh? **Le** reclasificaron el delito, y entonces el homicidio fue imprudencial, en realidad. Hubieron veinte mil peripecias ¿no?, a través de todo la... de toda la cosa que se ventiló. Pero quiero decirte que... el abogado debe tener una cultura muy amplia para poder saber qué cosa conviene hacer en favor de su cliente. **Este muchacho**... pues, en un... en una ocasión pensé que podría sacarse su caso por el lado psicológico.

'And then, <u>this guy</u>, well, I was able to take **him** out of there after three years in prison only, huh? **His** crime was reclassified and then the homicide was actually reckless. There were so many unforeseen events, weren't they? all through....all through this issue. But I want to tell

you that... the lawyer must possess a very broad culture so he knows what's best for his client. **This guy**... well, once.. once I thought we could win his case on the psychological side.'

Discourse (23) is similar to (22) but a bit more complex. In the first sentence, a discourse referent *un policía* 'a policeman' is introduced with an indefinite noun phrase, hence a possible new topic of the discourse. In the second sentence, a second entity is introduced with another indefinite noun phrase *un individuo* 'a guy'. Both entities are topics (i.e. a complex topic is constructed like 'policeman meets suspicious guy'). The topic continues over a few sentences and this is marked via 3rd person pronouns (in bold and subscripted to show coreference) and other means such as 3rd person possessive determiners. Then a new sentence is introduced that suspends the current topic and introduces new discourse referents (i.e. potential new competing (sub)-topics). In particular, a third discourse referent is introduced with the indefinite noun phrase *un drogadicto* ('a drug addict'), a possible new topic. However, the speaker makes the hearer focalize on the previous old topic by making salient the old referents *un policia* and *un individuo* with the demonstrative noun phrases *este policía* and *este individuo*.

(23) Imagínese usted que de pronto un día <u>un policía$_i$</u> está en el calle. ¿Está? En la calle. Ya. Y $\emptyset_i$ ve a un <u>individuo$_j$</u> que **le$_i$** infunde sospechas por **su$_j$** aspecto, por **su$_j$** forma de comportamiento.[4] $\emptyset_i$ Procede a solicitar**le$_j$** la documentación, $\emptyset_i$ procede a identificar**lo$_j$**, $\emptyset_i$ **lo$_j$** registra y $\emptyset_i$ **le$_j$** encuentra droga en los bolsillos. Una cantidad que sobrepasa lo que podría considerarse, bueno, pues el consumo de un drogadicto. **Este policía$_i$** lleva a **este individuo$_j$** a la comisaría. De ahí es puesto a disposición judicial.

'Imagine that <u>a policeman</u> is on the street one day. Where is he? On the street. Ok. And **he** sees <u>a suspicious guy</u> because of **his** appearance, **his** behavior. **He** proceeds to request **his** identification, **he** frisks **him** and he finds some drugs in **his** pockets. A quantity that exceeds what could be considered ... well ... the normal consumption of a drug addict. **This policeman** brings **this guy** to the precinct. From there, he is brought before a judge.'

Finally, in (24), from the beginning the topic would be something like 'how art can help people' or simply 'art', at least during the first two-three sentences. Then, a new discourse referent is introduced, which is of a clausal nature: *la visión de un cuadro puede levantar el espíritu*, hence denoting a proposition. This new referent is clearly connected thematically with the former topic 'art', but in our opinion it seems like a subtopic that is introduced with the demonstrative noun phrase *este tipo* 'this kind'. The basic instruction encoded in the demonstrative remains the same, namely, making salient a new discourse referent by focusing the hearer's attention on it. In this particular case, I argue that the new referent is a discourse sub-topic by virtue of being thematically related with the main topic.

(24) El arte ha de introducir una dimensión más humana en la vida moderna. También ha de reflejar las tensiones de la realidad, pero superándolas o llegando a un punto de equilibrio necesario para la buena salud de la gente. Esto son sólo ideas. Cuando trabajas, a veces piensas en ellas y otras veces no. Pero en un momento de pesimismo, <u>la visión de un cuadro puede levantar el espíritu</u>. Yo he tenido experiencias de **este tipo**, por ejemplo, ante un cuadro de Picasso. Es como si se me abriese el mundo.

---

[4] A phonetically empty expression, often referred to as *zero*, *null anaphor*, *pro* or $\emptyset$. Spanish is a null subject language, that is, a language that allows the subject of a tensed verb to be phonetically null.

'Art should introduce a more humane dimension in modern life. It should also reflect the strains of reality, but getting over them or getting to the balance point necessary for a good health. These are only ideas. When you are working, sometimes you think about them and sometimes you don't. But in a moment of despair, <u>the view of a painting can cheer you up</u>. I've had experiences of **this kind**, for example, in front of a Picasso's painting. It feels as if the world would open up to me.'

The view of demonstratives as pragmatic focusing devices presented in this paper is compatible with current research on the pragmatic functions of demonstratives. For example, Kirsner (1979, 2011) proposed a view of deixis with respect to demonstratives where he defines deixis as constituting: "an instruction to the hearer to seek out and attend to some referent." (2011: 97). Kirsner's proposal is in line with Diessel's (2006) *joint attention effect*, whereby the essential role of demonstratives is that of focusing the attention of the addressee on a particular entity in the perceptual or visual field. Whereas Diessel's proposal for demonstratives is only based on exophoric uses of demonstratives, my claim is that such joint-attention effect can also be extended to account for demonstratives in endophoric or discourse-anaphoric uses, that is, demonstratives would also get the effect of focusing the hearer's attention on a particular discourse referent with a specific communicative purpose; i.e. promoting a discourse topic (a particular referent) that has declined over a stretch of discourse into the current focus of attention of the hearer. However, whereas the joint-attention effect is generally accomplished via a pointing gesture (i.e. a pointing finger, a nod of the head, etc) accompanying the utterance of the demonstrative in deixis proper, the same attention effect would be accomplished by the contrastive use of a definite referring expression (the demonstrative) over another definite expression (i.e. a personal pronoun). Furthermore, my analysis is also compatible with the notion of contrastive meaning inherent to all demonstratives. It is widely acknowledged today that demonstratives have a contrastive meaning. Such contrastive meaning has been analyzed from different theoretical perspectives; namely, as a procedure (García-Fajardo, 2006; Scott, 2012), as a meaning component that allows us to distinguish demonstratives from other definite expressions (Hiietam and Börjars, 2002), formally as a contrast set (Roberts, 2002), or as a [+/- contrastive] feature (Ferrazzano, 2012). What all these proposals have in common is that they recognize that the use of a demonstrative (either deictically or anaphorically) implies a contrast among, at least, two elements; i.e. the contrast set. This contrastiveness meaning of demonstratives can be characterized as an entailment as in (25).[5]

(25) I was drinking out of *this* glass.
      ENTAILS: I wasn't drinking out of *that* glass.

The idea that demonstratives are commonly used to establish new topics is widespread (Linde, 1979; Ehlich, 1982; Sidner, 1983; Himmelmann, 1996, Comrie, 2000, *inter alia*). In this line, Diessel (1999: 96) points out: "anaphoric demonstratives are often used to indicate a referent that is somewhat unexpected and not currently in the focus of attention", and he goes on (1999: 99): "what all anaphoric demonstratives have in common is that they do not just continue the focus of attention; rather, they indicate that the antecedent is not the referent the hearer would expect in this context (i.e. the most topical NP)."
     As I claim for Spanish, in discourse-anaphoric uses, the contrastiveness of demonstratives should be conceived of at the informational level. Anaphoric demonstratives would thus be used to establish an implicit contrast between two discourse referents, i.e. a salient discourse referent that is

---

[5] From Ferrazzano (2013)

the topic of the preceding discourse and the intended discourse referent that is to be focused or (re)introduced. The role of demonstratives as 'topic-shift devices' has been already identified in the literature on demonstratives, and our data from Spanish appear to confirm previous research on the subject. As Diessel explains, there are two main reasons that explain why an intended referent may be less salient than other referents, namely, the referent is mentioned further back in discourse than other competing referents; or the intended referent constitutes a new topic. In the first case, the anaphoric demonstrative would shift the topic and, in the second case, the demonstrative would reactivate a previous (possibly declined) topic.

6. Conclusions
This paper shows interesting findings between the referring properties of demonstratives and direct object pronouns in Spanish. To judge from the analysis of natural examples these two referring expressions appear more similar than dissimilar regarding their referring attributes. My findings seem to go against previous research cross-linguistically. However, the similarities found among demonstratives and D.O. anaphors in Spanish are only apparent. In many respects, our data coincide with the widely accepted view that personal pronouns refer to topical antecedents and demonstratives to non-topical ones, but this interpretation only emerges when we adopt a notion of informational context that extends beyond the immediate discourse context and, therefore, beyond the most immediate local topic, i.e. the topic in the clause next to the anaphor.

Only when a narrow notion of discourse context is analyzed the two referring expressions behave in similar ways. On the other hand, when topicality is measured as the grammatical role and position of the antecedent within the sentence, D.O. clitic pronouns and demonstratives alike tend to find their antecedents in non-topical positions quite frequently. When salience is measured in purely cognitive terms (i.e. the referent is in the current focus of attention versus merely activated), and antecedent distance is taken as a plausible way of operationalizing salience, corpus data indicate that demonstratives and personal pronouns show a strong preference to find their antecedents in the clause next to the anaphor. Thus, if we consider antecedent distance as a factor having an effect on the information status of the antecedent (i.e. most recent antecedents are more accessible than antecedents located at greater distances due to processing and memory retrieval efforts), then demonstratives and personal pronouns show an almost identical behavior.

In this paper, I have claimed that in order for referential differences to emerge we need to extend our discourse context beyond the sentence that immediately precedes the anaphor. Thus, when larger discourse fragments are considered we start to see clear coincidences with previous proposals on the pragmatic interpretation and pragmatic contribution of referring expressions to the global coherence and cohesion of discourse. In my view, Spanish demonstratives and direct object pronouns basically differ in the effects they impose on discourse coherence and cohesion. I have argued that, while personal pronouns are markers of topic continuity, anaphoric demonstratives play a focalisation role that should be conceived of as an instruction on the part of the speaker for the addressee to focus on a particular discourse entity, with the precise communicative intention of making the hearer aware that a new (sub)-topic is introduced or that old-topics that have decayed informationally over a stretch of discourse have been re-introduced. In terms of cognitive or information status, anaphoric demonstratives can then be viewed as pragmatic markers used by speakers to bring entities into the current focus of attention, or 'topic-shifters' in Diessel's (1999) terminology. I also claim that this focusing property closely resembles that of demonstratives in deixis proper (i.e. use of demonstratives to point to physical entities) or nuclear pitch accent in phonological focus marking; and that this property obtains regardless of the syntactic prominence of the entity that is the current topic, the new topic or a subtopic in the discourse. Finally, such focalisation property of demonstratives in discourse-anaphoric uses is compatible with the attribute of contrastiveness inherent to the class of demonstratives.

References

Ariel, Mira. 1988. Referring and accessibility. *Journal of Linguistics* 24(1): 65—87.

Ariel, Mira. 1990. *Accessing Noun Phrase Antecedents*. London/New York: Routledge.

Ariel, Mira. 2001. Accessibility theory: an overview. In T. Sanders, J. Schilperoord and W. Spooren (eds.), *Text Representation: Linguistic and Psycholinguistic Aspects*. Amsterdam/Philadelphia: John Benjamins.

Asher, Nicholas. 1993. *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer Academic Publishers.

Bosch, Peter, Tom Rozario and Yufan Zhao. 2003. Demonstrative pronouns and personal pronouns. German *der* vs. *er*. *Proceedings of EACL2003, Workshop on the computational treatment of anaphora*.

Bosch, Peter and Carla Umbach. 2007. Reference determination for demonstrative pronouns. ZAS Papers in Linguistics 48: 39—51.

Bosch, Peter, Graham Katz and Carla Umbach. 2008. The Non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten and M. Knees (eds.), Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference. Amsterdam: John Benjamins: 145—164.

Bosque, Ignacio and Juan Carlos Moreno. 1990. Las construcciones con *lo* y la denotación del neutro. *Lingüística* 2: 5—50.

Brennan, Susan E. 1998. Centering as a psychological resource for achieving joint reference in spontaneous discourse. In M. Walker, A. K. Joshi and E. F. Prince (eds.), *Centering Theory in Discourse*, 227—249. Oxford: Clarendon.

Comrie, Bernard. 2000. Pragmatic binding: Demonstratives as anaphors in Dutch. *Berkeley Linguistics Society* 23.

Contreras, Heles. 1973. Spanish non-anaphoric *lo*. *Linguistics* 111: 5—29.

Cornish, Francis. 2008. How indexicals function in texts: Discourse, text and one neo-Gricean account of indexical reference. *Journal of Pragmatics* 40(6): 997—1018.

Cornish, Francis. 2009. Indexicality by degrees: Deixis, 'anadeixis' and (discourse) anaphora. Paper presented at the Symposium Quel sense pour la linguistique?. Université de Toulouse-Le Mirail. Downloadable at http://w3.erss.univ-tlse2.fr:8080/index.jsp?perso=cornish&subURL=web/Pap_Symp_Hon_J-Lyons_UTM.pdf

Davies, Mark. 2002. Corpus del Español: 100 million words, 1200s-1900s. Downloadable at http://www.corpusdelespanol.org.

Diessel, Holger. 1999. *Demonstratives: Form, Function and Grammaticalization*. Amsterdam/Philadelphia: John Benjamins.

Diessel, Holger. 2006. Demonstratives, joint attention and the emergence of grammar. *Cognitive Linguistics* 17(4): 463—489.

Dipper, Stefanie and Heinke Zinsmeister. 2009. Annotating discourse anaphora. In *Proceedings of the Third Linguistic Annotation Workshop (ACL-IJCNLP 2009)*, 166—169.

Dixon, Robert M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1): 61—112.

Ehlich, Konrad. 1982. Anaphora and deixis: Same, similar, or different? In R. J. Jarvella and W. Klein (eds.), *Speech, place and action. Studies in deixis and related topics*, 315—338. Chichester: Wiley.

Ferrazzano, Lisa S. 2013. Demonstratives in motion: The grammaticalization of demonstratives as a window into synchronic phenomena. PhD Dissertation, The City University of New York.

García-Fajardo, Josefina. 2006. La instrucción de contrastar en el demostrativo español. *Verba* 33: 175—186.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In T. Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1—41. Amsterdam/Philadelphia: John Benjamins.

Grice, Herbert P. 1989. Presupposition and conversational implicature. In H. P. Grice (ed.), *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.

Grosz, Barbara J., Aravind Joshi and Scott Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2): 203—225.

Gundel, Jeanette K., Nancy Hedberg and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274—307.

Gundel, Jeanette K., Nancy Hedberg and Ron Zacharski. 2005. Pronouns without NP antecedents: how do we know when a pronoun is referential? In A. Branco, T. McEnery and R. Mitkov (eds.), *Anaphora Processing: Linguistic, Cognitive and Computational Modelling*. Amsterdam/Philadelphia: John Benjamins.

Gutiérrez-Rexach, Javier. 1999. The structure and interpretation of Spanish degree neuter constructions. *Lingua*, 109, 35—63.

Hegarty, Michael, Jeanette K. Gundel and Kaja Borthen. 2001. Information structure and the accessibility of clausally introduced referents. *Theoretical Linguistics* 27: 163—186.

Hegarty, Michael. 2003. Semantic types of abstract entities. *Lingua* 113: 891—927.

Hegarty, Michael, Jeanette K. Gundel and Kaja Borthen. 2003. Cognitive status, information structure and pronominal reference to clausally introduced entities. *Journal of Logic, Language and Information* 12 (3): 281—299.

Hegarty, Michael. 2006. Type shifting of entities in discourse. In K. von Heusinger and K. Turner (eds.), *Where Semantics meets Pragmatics*, *Current Research in the Semantics/Pragmatics Interface*, *16*. Amsterdam: Elsevier.

Hiietam, Katrin and Kersti Börjars. 2002. Generative approaches to Finnic and Saami linguistics. In D. C. Nelson and S. Manninen (eds.), *Generative Approaches to Finnic and Saami Linguistics*, 1—39. Stanford: CSLI Publications.

Himmelmann, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In B. A. Fox (ed.), *Studies in Anaphora*, 205—254. Amsterdam: John Benjamins.

Kaiser, Elsi and John C. Trueswell. 2005. Investigating the interpretation of pronouns and demonstratives in Finnish: going beyond salience. In E. Gibson and N. Pearlmutter (eds.), *The Processing and Acquisition of Reference*. Cambridge, MA: MIT Press.

Kameyama, Megumi. 1999. Stressed and unstressed pronouns: complementary preferences. In P. Bosch and R. van der Sandt (eds.), *Focus, Linguistic, Cognitive and Computational Perspectives*, 306—321. Cambridge: Cambridge University Press.

Kirsner, Robert. 1979. Deixis in discourse: An explanatory quantitative study of the modern Dutch demonstrative adjectives. In Talmy Givón (ed.), *Discourse and Syntax: Syntax and Semantics Vol. 12*, 355—373. New York: Academic Press.

Kirsner, Robert and Vincent van Heuven. 1988. The significance of demonstrative position in modern Dutch. *Lingua* 76: 209-248.

Kirsner, Robert. 2011. Instructional meanings, iconicity, and l'arbitraire du signe in the analysis of the Afrikaans demonstratives. In B. de Jong and Y. Tobin (eds.), *Linguistic theory and empirical evidence*, 97—137. Amsterdam: John Benjamins.

Kleiber, Georges. 1990. Sur l'anaphore démonstrative. In M. Charolles, S. Fisher and J. Jayez (eds.), *Le Discours. Représentations et Interprétations*. Nancy: Presses Universitaires de Nancy, 243—263.

Leonetti, Manuel. 1999. El artículo. In I. Bosque and V. Demonte (eds.), *Gramática Descriptiva de la Lengua Española, Vol.1*, 787—890. Madrid: Espasa-Calpe.

Linde, Charlotte. 1979. Focus of attention and the choice of pronouns in discourse. In T. Givón (ed.), *Discourse and Syntax*. New York: Academic Press.

Lyons, John. 1975. Deixis as the source of reference. In E. L. Keenan (ed.), *Formal Semantics of Natural Language*, 61—83. Cambridge: Cambridge University Press. Reprinted as Ch.8 in Lyons, John. 1991. *Natural Language and Universal Grammar*. Cambridge: Cambridge University Press, 146—165.

Lyons, John. 1979. Deixis and anaphora. In T. Myers (ed.), *The Development of Conversation and Discourse*, 88—103. Edinburgh: Edinburgh University Press. Reprinted as Ch.9 in Lyons, John. 1991, *Natural Language and Universal Grammar*. Cambridge: Cambridge University Press, 166—178.

Maes, Alfons and Leo Noordman. 1995. Demonstrative nominal anaphors: a case of nonidentificational markedness. *Linguistics* 33: 255—282.

Navarretta, Costanza. 2005. Combining information structure and centering-based models of salience for resolving Danish intersentential pronominal anaphora. In A. Branco, T. McEnery and R. Mitkov (eds.) *Anaphora Processing. Linguistic, Cognitive and Computational Modeling*, 329—350. Amsterdam/Philadelphia: John Benjamins.

Navarretta, Costanza. 2007. A contrastive analysis of abstract anaphora in Danish, English and Italian. In A. Branco, T. McEnery, R. Mitkov and F. Silva (eds.), *Proceedings of DAARC 2007 - 6th Discourse Anaphora and Anaphora Resolution Colloquium*, 103—109.

Navarretta, Costanza. 2008. Pronominal types and abstract reference in the Danish and Italian DAD corpora. In C. Johansson (ed.), *Proceedings of the Second Workshop on Anaphora Resolution*, 63—71.

Navarretta, Costanza and Sussi Olsen. 2008. Annotating abstract pronominal anaphora in the DAD project. In *Proceedings of the Sixth International Language Resources and Evaluation* (*LREC '08*), 2046—2052.

Poesio, Massimo. 2004. The MATE/GNOME proposals for anaphoric annotation revisited. In *Proceedings of the 5th SIGDIAL Workshop*, 154—162.

Poesio, Massimo and Natalia Modjeska. 2005. Focus, activation and this-noun phrases. In A. Branco, T. McEnery and R. Mitkov (eds.), *Anaphora Processing. Linguistic, Cognitive and Computational Modeling*, 429—442. Amsterdam/Philadelphia: John Benjamins.

Poesio, Massimo and Ron Artstein. 2008. Anaphoric annotation in the ARRAU corpus. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC '08)*, 1170—1174.

Prince, Ellen F. 1981a. Toward a taxonomy of given-new information. In P. Cole (ed.), *Radical Pragmatics*, 223—256. New York: Academic Press.

Prince, Ellen F. 1981b. Topicalization, focus-movement and Yiddish-movement: A pragmatic differentiation. *Berkeley Linguistics Society* 7: 249—64.

Real Academia Española. *Corpus de Referencia del Español Actual* (CREA). Available at http://www.rae.es

Recasens, Marta. 2008. Discourse deixis and coreference: evidence from AnCora. In C. Johansson (ed.), *Proceedings of the Second Workshop on Anaphora Resolution (WAR II)*. NEALT Proceedings Series Vol. 2, 73—82.

Roberts, Craige. 2002. Demonstratives as definites. In K. van Deemter and R. Kibble (eds.), *Information sharing: reference and presupposition in language generation and interpretation*, 89—196. Stanford: CSLI Publications.

Scott, Kate. 2013. This and that: A procedural analysis. *Lingua* 131: 49—65.

Sidner, Candace L. 1983. Focusing in the comprehension of definite anaphora. In M. Brady and R. C. Berwick (eds.), *Computational Models of Discourse*, 267—330. Cambridge, MA: MIT Press.

Sturgeon, Anne. 2008. Topic and demonstrative pronouns in Czech. In G. Zybatow, L. Szuchsich, U. Junghans and R. Meyer (eds.), *Formal description of Slavic languages*. Berlin: Peter Lang.

Vieira, Renata, Sussane Salmon-Alt, Caroline Gasperin, Emmanuel Schang and Gabriel Othero. 2002. Coreference and anaphoric relations of demonstrative noun phrases in a multilingual corpus. In *Proceedings of the 4th Discourse Anaphora and Anaphor Resolution Colloquium*.

Webber, Bonnie L. 1979. *A Formal Approach to Discourse Anaphora*. New York: Garland.

Webber, Bonnie L. 1988. Discourse deixis: reference to discourse segments. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, 113—122.

APPENDIX
Some of the corpus examples analysed in this paper are presented in this appendix. The antecedent
is underlined and the anaphor marked in bold characters.

(i)   Antecedent of demonstrative. Grammatical role: SUBJECT
*La estimación del número de espectadores es el resultado de una ecuación matemática, ya que
hasta febrero de 1994 **este dato** no aparecía reflejado en los estadillos de las recaudaciones.*

(ii)   Antecedent of demonstrative. Grammatical role: DIRECT OBJECT
*Proponemos que este componente está dañado en el autismo y sugerimos que es innato y que tiene
un único sustrato cerebral. Si fuera posible señalar con precisión **este sustrato** se podría identificar
el origen biológico del autismo.*

(iii)   Antecedent of demonstrative. Grammatical role: ADJUNCT
*El entrenador italiano del Valencia, Claudio Ranieri, anunció hoy oficialmente que permanecerá
un año más dirigiendo el banquillo de **este club de fútbol**.*

(iv)   Antecedent of demonstrative. Grammatical role: OBJECT OF PREPOSITION
A: Hábleme de Gibson y de sus investigaciones.
*B: A mí me parece que **este hombre** hizo una labor notable, meritoria.*

(v)   Antecedent of demonstrative determiner. SENTENCE
*Se prohibe que desde los medios de comunicación de titularidad pública se ofrezcan
"realizaciones" del Gobierno en época preelectoral, ya que **este hecho** puede entenderse como
publicidad subliminal del partido en el poder.*

(vi)   Antecedent of demonstrative pronoun. SENTENCE
A: ¿Hay muchos mitos en esta historia?
*B: Sí, los hay; como que Watson y Crick trabajaban en una barraca o en un cobertizo para
bicicletas. ¡No es verdad! Es cierto que este grupo, o cierta parte de este grupo, ocupó la barraca
que se convirtió en laboratorio, pero **esto** ocurrió cuatro años más tarde, en el 57.*

(vii)   Antecedent of demonstrative pronoun. NOUN PHRASE
*Con la autonomía se les intenta asustar, diciéndoles que es separatismo o que significa la ruptura
de la unidad de la Patria. Y no es **eso**.*

(viii) Antecedent of 3rd person pronoun. Grammatical role: SUBJECT
*Él se levantó, ∅ fue hasta el patio y ∅ juntó la ropa seca que estaba tendida. Yo también me
levanté y **lo** ayudé.*

(ix)   Antecedent of 3rd person pronoun. Grammatical role: DIRECT OBJECT
*Traveler sacó un cigarrillo del bolsillo del piyama y **lo** encendió.*

(x)   Antecedent of 3rd person pronoun. Grammatical role: INDIRECT OBJECT
*"Esa mujer ha sido tu perdición ", le gritaba úrsula al bisnieto cuando **lo** veía entrar a la casa
como un sonámbulo.*

(xi)   Antecedent of 3rd person pronoun. Grammatical role: ADJUNCT
*Cuando terminó se limitó a meterlos en un saco de tela y, como pudo, se **lo** colocó a la espalda.*

(xii) Antecedent of 3rd person pronoun. Grammatical role: LEFT-DISLOCATED D.O.
*El primer ataque **lo** sufrió diez días antes de mi entrevista con ella.*

(xiii)Antecedent of 3rd person pronoun. Grammatical role: PREDICATIVE IN COPULA
CONSTRUCTION
A: Pero usted subraya siempre que la Historia del Arte no es una ciencia.
*B: ¡ Claro que no **lo** es!*

(xiv)Antecedent of 3rd person pronoun. SENTENCE
Hay gente que tiene sueños extraños, poco ordenados, confusos y encima dice que es feliz.
*Realmente no **lo** entiendo.*