An Implicature-Based Approach to Disjointness Effects*

Paul Marty
ZAS, Berlin

Abstract  The generation and distribution of disjointness effects raise a descriptive and explanatory challenge to linguistic theories: what are the conditions under which these effects arise and why do these effects arise in the first place? In this paper, I flesh out an account of these effects that takes up both these challenges at once in showing that these effects are derivable from current approaches to implicature-reasoning. In substance, it is argued that these effects follow from a genuine implicature-based reasoning whose outcome may, upon certain contextual pressures, result in conflicting representations, giving rise then to oddity effects. The account is shown to unify various disjointness effects under one roof and to explain the source of some of the classical conditions on co-reference and binding.

Keywords: disjointness effects, co-reference, variable binding, structural competition, scalar implicatures, relevance, oddness, blindness, exhaustivity operator

1 Disjointness Effects

This paper is concerned with a set of well-known interpretative phenomena whereby two nominal expressions, whose interpretation could a priori co-refer or co-vary, end up being interpreted as disjoint.¹ Phenomenally, these interpretative effects, known as disjointness effects, manifest themselves in two noticeable ways. First, in run-of-the-mill contexts, we observe in examples like (1) that certain identity relations between different nominals, occurring or not in the same sentence, are spontaneously ruled out by speakers.

---

* This paper includes and extends some of the ideas and materials originally presented in Chapter 5 of my doctoral dissertation, Implicatures in the DP domain. I am indebted to Irene Heim, Danny Fox and David Pesetsky for sharing with me their invaluable insights on this topic over the years. I am also thankful to my colleagues of the Semantics and Pragmatics group (FB IV) at ZAS, Itai Bassi, Louise McNally, Sophie Moracchini, Despina Oikonomou, Philippe Schlenker, the reviewers and audiences at SALT 28 for useful discussion and feedback at earlier stages of this work. This work was supported by Bundesministerium für Bildung und Forschung (Grant Nr. 01UG1411).

¹ We will say that a nominal expression $\alpha$ is interpreted as disjoint from a nominal expression $\beta$ if the interpretation of $\alpha$ does not — exhaustively or partially — co-refer with or co-vary with that of $\beta$.
Disjointness Effects as Implicatures

(1) a. The husband of John’s childhood sweetheart arrived.
   $\leadsto$ John is not the husband of John’s childhood sweetheart

   b. John’s wife talked to Peter’s boss.
   $\leadsto$ John and Peter’s boss are distinct individuals

   c. She criticized Sue’s work.
   $\leadsto$ Sue and she are distinct individuals

   d. What is Sue gonna do?
   $\leadsto$ Sue is neither I (the speaker), nor you (the addressee)

Second, in cases where such semantic relationships are known to hold, the sentences in (1) are perceived as deviant, as exemplified in (2).² Hence, a sentence like (2c) cannot felicitously be used to express the thought that Sue criticized her own work, although the $\phi$-features of she match those of Sue. Similarly, the sentence in (2a) is unacceptable if it is known to the speakers that John married his childhood sweetheart, although nothing in the meaning of the description the husband of John’s childhood sweetheart forces its denotation to be distinct from that of the description John, as demonstrated by the fact that we can felicitously assert identity statements such as The husband of John’s childhood sweetheart is John (himself).

(2) a. Context: John married his childhood sweetheart
   # The husband of John’s childhood sweetheart arrived.

   b. Context: John is Peter’s boss
   # John’s wife talked to Peter’s boss.

   c. Context: someone asked ‘What did Sue do?’
   # She$_Sue$ criticized Sue’s work.

   d. Context: Sue is the addressee of the speaker
   # What is Sue gonna do?

While these examples only constitute a fragment of the data that a fully-fledged theory of disjointness effects has to account for, they allow us to intuit the main challenges that any such theory has to take up. The first one pertains to the description of the general conditions under which disjointness effects arise. In the past literature, these effects have usually been broken down into distinct groups and subsumed under different generalizations, e.g. the principles of the Binding Theory (Chomsky 1981, 1993; Lasnik 1989b), the so-called Crossover Constraints (Postal 1971; Wasow 1972; Lasnik & Stowell 1991), Grodzinsky & Reinhart’s (1993) Rule I,

² Here and throughout this paper, I use the symbol ‘#’ to indicate that a sentence is semantically deviant in that it is perceived as odd by speakers. This characterization will be shown to be in line with the analysis developed in this paper where such sentences are not ungrammatical per se, but infelicitous due to the presence of an implicature conflicting with common knowledge.
Fox’s (2000) Rule H, Vergnaud’s (1974) Disjunction Condition later extended to the ‘i-within-i’ Condition (Chomsky 1981), Schlenker’s (2005) Minimize Restrictors! Yet it has also been noticed that some of these generalizations overlap in terms of empirical coverage and can be either reduced to, or unified with one another (a.o., Reinhart 1983; Grodzinsky & Reinhart 1993; Fox 2000; Büiring 2005). Hence, can we go one step further and propose a unified theory of these effects, one from which all the valuable generalizations mentioned above would follow? Next, beyond the descriptive level, the very existence of these effects also raises an explanatory challenge, namely that of their raison d’être: why is it that certain identity relations between nominals are systematically rejected by speakers, although those relations are logically possible and pragmatically plausible?

In this paper, I offer an account of disjointness effects that takes up both these challenges at once in deriving these effects from current approaches to implicature-reasoning. The basic intuition underlying this account is that the core phenomenology of disjointness effects is similar in essence to that of scalar implicatures (SIs). First, in run-of-the-mill contexts, sentences like (3) are enriched by adding to their logical representations the negation of one of their relevant, excludable alternatives, ruling out in effect certain logical possibilities left open by their plain meaning.

(3)  
   a. John gave an A to some of his students.  
       SI: ¬(John gave an A to all of his students)  
   b. John or Peter arrived.  
       SI: ¬(John and Peter both arrived)  
   c. John arrived.  
       SI: ¬(Peter arrived)

Second, in cases where the truth of the target alternatives is contextually known to hold, these same sentences are now perceived as deviant by speakers, (4).

(4)  
   a. Context: John gave the same grade to all his students  
       # John gave an A to some of his students.  
   b. Context: John and Peter are conjoined (Siamese) twins  
       # John or Peter arrived.  
   c. Context: John and Peter are conjoined (Siamese) twins  
       # John arrived.

To account for these and related observations, it has been proposed that (a) the set of excludable alternatives entering the computation of SIs includes not just stronger alternatives, but more generally non-weaker ones (a.o., Fox 2007; Chemla & Spector 2011; Marty 2017), and that (b) the oddity effects in (4) result from the presence of conflicting SIs (a.o., Magri 2009, 2011; Schlenker 2012; Marty 2017).
Disjointness Effects as Implicatures

In a nutshell, the sentences in (4) sound odd in context because their corresponding alternatives are relevant and therefore the SIs associated with them are still computed, thus resulting in contextually contradictory representations (e.g., *John gave an A to some, but not all of his students*). Current theories of SIs featuring both (a) and (b) end up with something along the lines of (5) as a general felicity condition:

(5) **Felicity condition imposed by implicature reasoning**

A sentence $\phi$ is felicitous at a context $C$ only if there is no sentence $\psi$ s.t.

a. $\phi$’s plain meaning entails $\psi$ in $C$, but

b. $\phi$’s strengthened meaning entails $\neg\psi$, i.e. if

i. $\psi$ is a formal alternative to $\phi$, and

ii. $\psi$ is logically non-weaker than $\phi$, and

iii. $\psi$ is relevant in $C$.

In the following, I argue that, once supplemented with suitable mechanisms for substituting and interpreting nominals in alternatives, theories of SIs provide an explanatory account of a wide range of disjointness effects, subsuming under (5) three core results from the literature on binding and co-reference: (I) reference to the same individual via distinct but presupposedly coreferential descriptions is dispreferred (*Postal 1970; Heim 1982; Aloni 2001*), (II) semantic binding is preferred over co-reference (*Grodzinsky & Reinhart 1993*), and (III) pronoun binding seeks the closest antecedent (*Fox 2000; Kehler 1993*). The resulting account will be shown to do justice to various insights from the past literature and in particular to encompass the following familiar ideas: apparent syntactic restrictions on co-reference and binding are in fact symptoms of semantic and pragmatic restrictions (*Bolinger 1977, 1979; Reinhart 1983; Levinson 1987, 2000; Schlenker 2005*); the generation of disjointness effects often relies on the availability of more minimal alternatives and thus involves some form of manner considerations (*Bolinger 1979; Levinson 1987; Schlenker 2005*); finally, identity relations are established by means of co-dependent readings whenever possible (*Evans 1980; Safir 2004, 2005*).

2 **Overview of the proposal**

The proposal I will put forward is that disjointness effects follow from a genuine implicature-based reasoning. To illustrate this proposal, let us consider the sentence in (1a)/(2a). This sentence, repeated below in (6a), has (among others) the sentence in (6b) as a competing alternative; one can verify that this alternative is derivable from (6a) by substituting the whole subject DP with one of its sub-constituent DPs and that it is logically independent from (6a).
Marty

(6)  
a. The husband of John’s childhood sweetheart arrived.
    \[\alpha \text{ the husband of } [\gamma [\beta \text{ (the John)’s child. sweet.}]] \text{ arrived}\]

b. John arrived.
    \[\beta \text{ (the John) arrived}\]

I argue that the comparison of (6a) to (6b) leads us to infer that the speaker doesn’t believe that the husband of John’s childhood sweetheart and John are the same individuals, regardless of whether (6b) is assumed to be relevant or irrelevant. To explain this, let us consider both possibilities in turn. Suppose first that (6a) is uttered in a context in which (6b) is relevant as well. Since (6b) is an excludable alternative of (6a), the strengthened meaning of (6a) will entail the negation of (6b). The resulting representation, together with the axioms of logic, gives us then the disjointness inference we were interested in: if the husband of John’s childhood sweetheart arrived but John didn’t, then both descriptions must denote distinct individuals on the speaker’s beliefs, as shown below.3

\[
\begin{align*}
\lbrack \text{EXH}_R \lbrack \text{the husband of John’s childhood sweetheart arrived}\rbrack \rbrack^w & \iff \lbrack \text{the husband of John’s child. sweet. arrived} \rbrack^w \land \lbrack \text{John arrived} \rbrack^w \\
\therefore \lbrack \text{the husband of John’s childhood sweetheart} \rbrack^w & \neq \lbrack \text{John} \rbrack^w
\end{align*}
\]

The presence of this implicature not only explains why this sentence is understood as involving distinct individuals in run-of-the-mill contexts, but also why it is perceived as infelicitous in cases where it is known that John married his own childhood sweetheart, as in (2a). In such cases, (6a) and (6b) are contextually equivalent; since relevance is closed under contextual equivalence, we have it that, if (6a) is relevant, then so is (6b). The computation of the SI associated with (6b) becomes then mandatory, resulting in a representation that contradicts speakers’ common assumptions, hence the oddness, consistent with the generalization in (5).

Turning to the second possibility, suppose now that we take (6b) to be irrelevant. In this case, the implicature we just described cannot go through. But relevance considerations still apply. Specifically, if (6a) is relevant but (6b) isn’t, then (6a) and (6b) must not be contextually equivalent. This is possible only if (6a) and (6b) are

3 Here and throughout this paper, I adopt a grammatical view on SI-computation on which the SIs of a sentence are computed recursively with the compositional side of meaning by applying a covert but syntactically realized operator called the exhaustion operator and notated EXH (a.o., Chierchia 2006; Fox & Hackl 2006; Fox 2007; Chierchia, Fox & Spector 2008; Fox & Spector 2009; Fox & Katzir 2009; Magri 2011; Chierchia, Fox & Spector 2012; Romoli 2012; Spector 2014; Magri 2014; Marty 2017). Just like any other quantifier, the domain of quantification of EXH is hypothesized to be restricted by a contextually assigned relevance predicate \(R\), hence the notation EXH\(_R\). The specifics of this operator will be discussed and refined in the following section. For now, we can consider that applying EXH\(_R\) to a proposition \(\phi\) (i.e., its prejacent) outputs the truth of \(\phi\) together with the falsity of every excludable, structural alternative to \(\phi\) that is relevant given \(\phi\) (i.e., \(\psi \in R\)).
Disjointness Effects as Implicatures

not true in all the same worlds compatible with the speakers’ common assumptions, i.e. if it is not presupposed that the husband of John’s childhood sweetheart and John refer to the same individual, as shown below.

\[(8) \quad \llbracket \text{the husband of John’s child. sweet. arrived} \rrbracket \not\in C \llbracket \text{John arrived} \rrbracket \]
\[
\therefore \exists w \in C : \llbracket \text{the husband of John’s childhood sweetheart} \rrbracket^w \neq \llbracket \text{John} \rrbracket^w
\]

Crucially, note that, on this alternative scenario, (6a) also gives rise to a disjointness inference, albeit a weaker one, namely that the speaker(s) do(es) not mutually believe that the husband of John’s childhood sweetheart is John. Consequently, (6a) should be felicitous only in contexts where the interlocutors believe that this identity relation doesn’t hold or are uncertain as to whether it does, as in (9).⁴

\[(9) \quad \text{Context: A and B don’t know yet that John married his childhood sweetheart}
A: Who did John’s childhood sweetheart marry?
B: I don’t know: maybe John himself, maybe someone else.
A: Well, I guess we’ll find out soon enough: the husband of John’s childhood sweetheart just arrived.
\]

In sum, whenever (6b) is relevant, (6a) gives rise to a (possibly mismatching) implicature which leads us to infer that identity is believed to not hold; whenever (6b) is irrelevant, this implicature disappears but we are then lead to infer that identity is not believed to hold; whichever is the case, it follows that (6b) can never felicitously express in context the meaning literally expressed by its alternative (6a). In the next section, I lay down the specifics of this proposal by making explicit certain assumptions about the structural complexity of nominals and then by discussing in further detail the mechanisms whereby the set of alternatives is determined and meaning strengthening performed. The empirical scope of the resulting account will be discussed in Section 4. It will be shown that the SI-reasoning sketched out in this overview applies to all the cases in (1)/(2).

Cases of overt disagreement like (i) can be seen as special cases of uncertainty, where uncertainty arises at the inter-subjective level (i.e., interlocutors have conflicting opinions) creating a similar situation as in (9), i.e. the identity relation under discussion is not mutually accepted. I notice however that, in the absence of such specific contexts of disagreement or mutual ignorance, which actively depends upon the addressee’s epistemic state, the target of our inference tends to be narrowed down from common knowledge to the speaker’s epistemic state, i.e. from ‘it is not common ground that identity holds’ to ‘the speaker does not believe that identity holds’. This form of epistemic narrowing has been discussed in Chemla 2008 and more recently in Rouillard & Schwarz 2017.

(i) \quad \text{Context: A and B are disagreeing about the identity of John’s wife}
A: So John married his childhood sweetheart then?
B: No, he didn’t! He married Sue’s daughter!
A: Well, you can verify by yourself: the husband of John’s childhood sweetheart just arrived.
3 Specifics of the proposal

3.1 Structural alternatives in the nominal domain

The present account relies on the idea that disjointness effects originate in the comparison of a given nominal with one of its alternatives, where the intended notion of alternatives is that of *structurally-defined alternatives* developed in Katzir 2007 (see also Katzir 2008; Fox & Katzir 2011).\(^5\) When this comparison involves fully-fledged definite descriptions as in (6), the structural relation is usually readable from the surface forms: one can easily see that (6b) is a structural alternative to (6a) while the reverse is not true, i.e. \((6b) \preceq (6a)\) but \((6a) \npreceq (6b)\). The example in (1b)/(2b), repeated below, offers further instances of transparent structural relations between alternative descriptions of identical or lesser syntactic complexity.

\[(1b)\] John’s wife talked to \([\alpha [\beta (the) Peter)]’s boss]\)

a. John’s wife talked to \([\alpha [\beta (the) John)]’s boss\] by substitution within \(\beta\)

b. John’s wife talked to \([\beta (the) John]\) by substituting further \(\alpha\) with \(\beta\)

When it comes to pronouns, however, what appears on the surface is of little help as a given phonological form may correspond to different structural realities. Hence, to see how pronominal forms relate to other nominals and to each other, we need to scratch the surface and look further into the details of their internal composition. For space reasons, I will focus here on the case of 3rd person pronouns (1st/2nd person pronouns will be discussed in \(4.1\)). As a baseline for our metrics, we can consider the basic structure of DPs to be as shown in (10): a genuine DP consists of a determiner D and an NP made up of a nominal stem and a certain number of meaningful nominal features (e.g., number, gender).

\[(10)\] **Skeleton of genuine DPs:** \([DP \text{ Determiner } [NP [N F_1] \ldots [N F_n] [N \text{ Stem}]]]\)

e.g., every student, some girls, a man, (the) John, these people

In line with the previous literature, I will assume that the structure of pronouns varies depending on whether they are referential or bound (a.o., von Stechow 2002; Reinhart 2006; Heim 2008; Johnson 2012). On the one hand, I will follow the view that (3rd person) referential pronouns are definite descriptions. This idea can be implemented in at least two ways: one is to treat these descriptions as involving at their core a variable, in the guise of a numerical index (a.o., Heim & Kratzer\(^6\)).

---

\(^5\) In a nutshell, the set of structural alternatives for a syntactic structure \(\phi\), noted \(\text{ALT}_{\text{str}}(\phi)\), is the set of syntactic structures that are derivable from \(\phi\) by a finite series of substitutions in \(\phi\) — e.g., substitution of one terminal element for another terminal element, substitution of one constituent with one of its sub-constituents — and that are therefore *structurally at most as complex as \(\phi\)*, i.e. \(\text{ALT}_{\text{str}}(\phi) := \{\psi : \psi \preceq \phi\}\). See Katzir (2007: (18)-(20)) for formal definitions and discussion.
Disjointness Effects as Implicatures

1998; Büring 2005; Heim 2005, 2007a,b; Adger 2011); another one is to treat them as genuine DPs converted into pro-forms via nominal ellipsis, in the guise of N’-ellipsis (a.o., Postal 1966; Elbourne 2001, 2005). For the time being, I will favor this second option, restricting the use of indices to represent in the grammar the relation of variable-binding, much in the spirit of Reinhart 1983, 2006. On this view, a pronoun like she can be analyzed on its referential reading as in (11). Number and descriptive features have their familiar lexical denotations. The feature [D def] is interpreted like the definite determiner the in English, i.e. it maps a property of individuals to the maximal individual that has that property (a.o., Link 1983).

(11) The student/she\textsubscript{the student} arrived.
   a. Syntax: $\text{DP} \{\text{def}\} \{\text{NP} \{\text{sg}\}[\text{fem}]\{\text{student}\}\}\text{Nomina}tive$
   b. Meaning: the individual x s.t. x is an atom & x is female & x is student
   c. Plain vs. Phonologically reduced descriptions:
      i. $\{\text{def}\} \{\text{sg}, \text{fem}, \text{student}\} \rightarrow \text{‘the student’/\_ Nom}inative$
      ii. $\{\text{def}, \text{sg}, \text{fem}\}\text{student} \rightarrow \text{‘she’/\_ Nom}inative$

On the other hand, we will take bound pronouns to be ‘minimal pronouns’ in the sense of Kratzer: a bound pronoun consists at LF of a bare index, which upon semantic binding, inherits feature values from the DP that binds it (a.o., Kratzer 1998; von Stechow 2002, 2003; Schlenker 2003; Heim 2008; Kratzer 2009; Johnson 2012). On this view, the features appearing on bound pronouns only contribute to how these pronouns are spelled out. For concreteness, I will implement this idea by following the proposal in Johnson 2012 and treat indices as nominal features (e.g., $\{N_1\}$, $\{N_2\}$, $\{N_3\}$), hereby letting them access the features of their binders under agreement. These indices are treated at LF as variables and are interpreted as usual via variable assignments (i.e., $\{[n]\}^g = g(n)$). Note that, in this system, a derivation can converge at the interfaces only if every index that occurs in it is bound.

On these assumptions, the intended reading of a sentence like (12) can in principle obtain from two distinct representations, depending on whether the identity relation between her and the subject DP is established by co-reference, (12a), or by variable-binding, (12b). In the first case, the pronoun is referential: it is syntactically construed as a genuine description and undergoes N’-ellipsis at PF; in the second case, it is bound: it comes solely with an index and receives from its binder the features it needs to be pronounced at PF.

6 NP-ellipsis (or rather N’-ellipsis) is implicitly assumed to be subject to certain pragmatic conditions. In particular, in our cases, the result of NP-ellipsis should comply with the general constraints that rule the felicitous use of pro-forms. For instance, the use of she in (11) is felicitous only if the context allows the addressee to recover enough of the elided descriptive content to establish reference. This recoverability-based restriction can be thought of as an equivalent of Heim & Kratzer’s (1998) Appropriateness Condition on the alternative treatment of referential pronouns as free variables.
(12) The student criticized her work.

a. (Co)-Referential representation
   [SUBJECT]^{7\{sg, fem\}} \ldots [vP t_7 \ldots [DP [def] [NP [sg][fem][student]]']s work]

b. Bound variable representation
   [SUBJECT]^{7\{sg, fem\}} \ldots [vP t_7 \ldots [NP [7]']s work]

It is crucial to observe that the above construals are structurally related to each other in a non-symmetric fashion: the bound-variable construal in (12b) is a structural alternative to the co-referential construal in (12a), while the reverse is not true since (12a) is more complex than (12b). More generally, on this analysis, one can observe that (i) a referential expression occurring in an environment in which it could be semantically bound always compete with a simpler bound-variable alternative, and that (ii) by our metrics, a bound pronoun (i.e., an index) can only have other bound pronouns (i.e., other indices) as structural alternatives. We shall later see that it is essentially these two properties that drive the competition which result in Rule I and Rule H effects. One question then remains to be addressed: what indices do we consider when deriving the alternatives of a sentence? As the present system does not allow indices to have free occurrences at the global level, we shall minimally assume that their introduction is guided by the rule in (13).

(13) **Substitution rule for indices**: a numerical index $i$ is in the substitution source of a structure $\phi$ only if $i$ already occurs in $\phi$ (i.e., $i$ occurs on a potential binder or else on a trace left by the movement of a potential binder).

We shall now explain how such competitions between nominals interact with meaning strengthening and contribute to determine the range of contexts in which a sentence like (12) can be used felicitously.

3.2 Excludability, relevance and exhaustification

As previewed in fn. 3, I adopt the view that meaning strengthening is performed in the grammar via the application of a syntactic operator called the *exhaustivity operator*, notated $EXH_{\phi}$. The conceptualization of $EXH_{\phi}$ I will follow is essentially that proposed in Magri 2009, 2011, 2014, which relies on three axioms. The first one is that an occurrence of $EXH_{\phi}$ is present in syntax at every scope site (recursive exhaustification is only optional). For instance, in (14) and (15), $EXH_{\phi}$ occurs at the two scope sites available, ① and ②, corresponding roughly to the TP and vP level. In line with our analysis, note that these sentences have bound-variable alternatives in common, namely those alternatives containing the vP $[t_7 v criticize [7]'s work]]$. 

374
The student criticized Sue’s work.

\[
\text{EXH}_{\phi} \begin{array}{c}
\text{the student}\end{array}^7 \ T_{\text{PAST}} \ \text{EXH}_{\phi} \begin{array}{c}
t_7 \text{ v criticize } [(\text{the) Sue}’s \text{ work}]]
\end{array} \\
\text{ALT}_{str} (\begin{array}{c}2\end{array}): \begin{array}{c}t_7 \text{ v criticize [7]’s \text{ work}]}
\end{array}
\]

(15) (#) She criticizes Sue’s work.

\[
\text{EXH}_{\phi} \begin{array}{c}
\text{she } \begin{array}{c}
\text{Sue}\end{array}\end{array}^7 \ T_{\text{PAST}} \ \text{EXH}_{\phi} \begin{array}{c}
t_7 \text{ v criticize } [(\text{the) Sue}’s \text{ work}]]
\end{array} \\
\text{ALT}_{str} (\begin{array}{c}2\end{array}): \begin{array}{c}t_7 \text{ v criticize [7]’s \text{ work}]}
\end{array}
\]

The second axiom is that the notion of entailment relevant for determining the set of excludable alternatives is that of logical entailment (see also Fox & Hackl 2006; Chierchia et al. 2008, 2012). For our purposes, we will take the excludable alternatives to a sentence \( \phi \) to be those structural alternatives to \( \phi \) that are logically non-weaker than \( \phi \) (i.e., not entailed by \( \phi \)). Since \( \text{EXH}_{\phi} \) can occur at embedded levels, we also need to specify how excludability is computed for representations involving indices that are not semantically bound locally, as above. My proposal is that the computation of excludable alternatives is blind to the actual values of all contextual parameters of evaluation: a structural alternative \( \psi \) to a sentence \( \phi \) is excludable if there is a possible valuation of the parameters of evaluation (i.e., world of evaluation, variable assignment, context of evaluation) that makes the falsity of \( \psi \) consistent with the truth of \( \phi \). This notion of extended blindness can be implemented as in (16a). As usual, the set of excludable alternatives associated with a sentence \( \phi \), \( \text{EXCL}(\phi) \), is further restricted to the subset of ‘innocently excludable’ alternatives (a.o., Fox 2007), \( \text{IE}(\phi) \). The semantics for \( \text{EXH}_{\phi} \) is given in (16c).

16. **Extended blindness, innocent exclusion and exhaustification**

   a. \( \text{EXCL}(\phi) := \{ \psi \in \text{ALT}_{str}(\phi) : \exists w \exists g \exists c ([\phi]^{w,g,c} \wedge \neg [\psi]^{w,g,c}) \} \)

   b. \( \text{IE}(\phi) := \{ \phi' : \text{EXCL}'(\phi') \subseteq \text{EXCL}(\phi) \text{ and EXCL}'(\phi) \text{ is a max. subset of } \text{EXCL}(\phi) \} \)

   c. \([\text{EXH}_{\phi} \phi]^{w,g,c} := [\phi]^{w,g,c} \wedge \forall \psi \in \text{IE}(\phi) : (-[\psi]^{w,g,c} \vee [\psi]^{w,g,c} \notin R)\]

Crucially, this characterization of \( \text{EXH}_{\phi} \) makes it applicable to propositions involving free individual variables at embedded levels. For instance, we can verify that the bound-variable alternatives to (14)/(15) are locally excludable: that \( g(7) \text{ criticized Sue’s work in } w \) does not logically entail that \( g(7) \text{ criticized } g(7)’s \text{ work in } w \) for there is a possible world \( w \) and assignment \( g \) that assigns to index 7 a value distinct from the denotation of \( \text{Sue} \) in \( w \). Strengthening the meaning of these sentences at embedded levels delivers the following result:

17. \([\text{EXH}_{\phi} [\begin{array}{c}
t_7 \text{ v criticize } [(\text{the) Sue}’s \text{ work}]]\end{array}]^{w,g,c} \text{ is true iff}

   a. \([t_7 \text{ v criticize } [(\text{the) Sue}’s \text{ work}]]]^{w,g,c} \text{ is true, and } \text{(prejacent)}

   b. \([t_7 \text{ v criticize [7]’s \text{ work}]]^{w,g,c} \text{ is false or irrelevant } \text{(strengthening)}

\)
The third and last axiom concerns the process whereby we evaluate the relevance of an alternative to determine whether the disjunctive inference obtained upon exhaustification is strengthened to what we properly call a scalar implicature. Following Magri’s proposal, we will assume that, by default, the prejacent \( \phi \) of \( \text{EXH}_R \) is always relevant simply by virtue of being asserted, i.e. \( \phi \in \mathcal{R} \). Since relevance is closed under contextual entailment, any alternative \( \psi \) to \( \phi \) that makes the same contribution as \( \phi \) in context is also relevant, i.e. if \( \phi \in \mathcal{R} \) and \( \psi \Leftrightarrow_C \phi \), then \( \psi \in \mathcal{R} \). In such cases, the SI associated to that alternative is mandatory. It is expected to be so in examples like (17) whenever the value that \( g(7) \) receives by composition ends up contextually indistinguishable from the denotation of Sue.\(^7\)

As in the case discussed in Section 2, the interpretation of (17) leaves us with two possible scenarios, both of which result in the generation of a disjointness inference.

i. **Strong disjointness:** If (17b) is relevant, then the target SI obtains, i.e. *The student/Sue didn’t criticize her (own) work*, and disjointness follows. In the example in (15), since (17a)\(\Leftrightarrow_C(17b)\), this SI is mandatory and results in a contradictory representation. An utterance of (15) is thus predicted to be trivially false in context, hence infelicitous (i.e., (15) \( \cap C = \emptyset \)).

(15) **Context:** Sue had to comment on her essay for a class.

  # She\(_{Sue}\) criticized Sue’s work.

  **Conflicting SI:** # She\(_{Sue}\) didn’t criticize her own work

ii. **Weak disjointness:** If (17b) is irrelevant, then (17b)\(\not\Leftrightarrow_C(17a)\). This alternative scenario is possible in (14). In this case, it must not be presupposed that *the student* and Sue refer to the same individual. Consequently, a sentence like (14) is predicted to be felicitous only in contexts where some uncertainty remains as to whether both descriptions refer to the same person, as in (18) for instance where identity is questioned rather than presupposed.

(18) A: Is the student over there Sue?

  B: I believe so: ✓ This student/She\(_{the\ student}\) is wearing Sue’s coat.

In the next section, I show that this implicature-based approach to disjointness effects applies in full generality and offers an explanatory account of (some of) the classical descriptive generalizations proposed in the previous literature.

\(^7\) For our purposes, the procedure for determining whether a sentence is contextually equivalent to one of its alternatives at embedded levels can be stated as follows. Let \([\ldots \text{EXH}_R [\psi] \ldots]\) be any sentence where \( \psi \) is a constituent in \( \phi \) with a propositional type. An alternative \( \chi \) to \( \psi \) is contextually equivalent to \( \psi \) in \( \phi \) if and only if their plain contextual contributions are equivalent, i.e. iff \([\ldots [\chi] \ldots]\) is contextually equivalent to \([\ldots [\psi] \ldots]\). One can verify by this procedure that (15) is contextually equivalent to its bound-variable alternative at the embedded level, \( \exists \), hence its infelicity.
4 Empirical Coverage

4.1 Don’t multiply guises: ‘i-within-i’ and related effects

It has long been observed that a nominal expression cannot co-refer with any another description that contains it, as stated and exemplified in (19) (a.o., Vergnaud 1974; Higginbotham & May 1981; Brody 1982; Jullens 1983; Williams 1982; Chomsky 1981; Hoeksema & Napoli 1990; Chomsky 1993; Jacobson 1994). 8

(19) The ‘i-within-i’ Condition: *[α . . . [β]j . . . ]
No description may co-refer with a description that contains it.
b. # [[John]j’s financial advisor’s coach]j is a former Marine.
c. # Mary met [the famous author of [D.Brown]j’s books]j.
d. # John is dating [the mother of the son of [this woman]j].

In the theories of Chomsky 1981, Lasnik 1989a and Webelhuth 1995, the ‘i-within-i’ Condition is stipulated and conceived as a primitive of the grammar. It shall be clear by now that the present theory offers a rationale for the source of these effects (see Section 2): ‘i-within-i’ sentences all have, by definition, simpler structural alternatives which are contextually equivalent to them, yet not logically entailed by them. The representations resulting from their strengthened meanings are thus always contradictory with common knowledge, hence their oddness.

Crucially, our account predicts that the logic underlying the generation of these effects shall not be specific to the ‘i-within-i’ environment so to speak. Specifically, we expect such oddity effects to be found whenever a given referential expression presupposedly co-refers with one of its (simpler) structural alternatives. The examples in (20) and (21) show that this expectation is indeed borne out. Note that, in these examples, structural complexity breaks symmetry: the (b)-sentences are structural alternatives to the (a)-sentences while the reverse is not true. I refer the reader to Marty (2017: chapter 5) for a discussion of the generality of this finding.

(20) Context: John is Mary’s assistant
a. # [[John]j’s boss] gave [[Mary]j’s assistant]j a raise
b. [[John]j’s boss] gave [John]j a raise

8 Subscript letters (j,k,l,...) are used informally in the following to represent presupposed coreference and thus to clarify the state of the common ground. Presupposed co-reference is understood as follows: any two referential expressions α and β are presupposed to be co-referential at a context C just in case α and β pick up the same individual in each world in C (a.o., Heim 1982, 1998, 2007a).
(21)  **Context:** Sue is married, and she and her husband have a son.

a. # [[Sue]’s boss] talked to [the father of [her] son]
   (ok if divorced)
b.  [[Sue]’s boss] talked to [[her] husband]

Pursuing this idea further, I propose that the disjointness inferences triggered by the use of 3rd person referential expressions relative to the speaker and addressee of the utterance context (see (1d)/(2d)) follow from a similar kind of competition, one involving 1st and 2nd person pronouns. Following Kratzer’s (2009) proposal, 1st and 2nd referential pronouns can be analyzed as in (22), where nominal 1st and 2nd person features are treated as irreducibly referential. On this analysis, 1st/2nd referential pronouns are thus, in terms of structural complexity, between 3rd person referential expressions (more complex) and bound indices (least complex). A brief comparison of (11) vs. (22) allows us to verify that 1st/2nd referential pronouns asymmetrically compete with 3rd person referential expressions.

(22)  a. Syntax: [NP [sg][1st]NOMINATIVE arrived
   b. Meaning: the atomic speaker in c arrived
   c. PF: [{sg}, {1st}] → ‘I’/ NOMINATIVE

Recall that, on our view, the computation of excludable alternative is blind to all contextual parameters of evaluation, including to the information pertaining to the ‘objective’ utterance context (e.g., identity of the interlocutors, utterance time and place). A sentence like (23) is thus predicted to give rise to a variety of disjointness inferences, preventing both John and Mary from being denotationally identical to the speaker or addressee and therefore giving rise to oddity effects in contexts where it would be assumed to be otherwise.

(23)  John is looking at Mary

a. EXHₐ [φ [(the) {sg, masc} John]ₗ [tₗ is looking at [(the) {sg, fem} Mary]]
b. ALTₐₗ [φ]: [{sg, 1st}]ₗ [tₗ am looking at 7],
   [{sg, 1st}]ₗ [tₗ am looking at [{sg, 2nd}]],
   [{sg, 2nd}]ₗ [tₗ are looking at [{sg, 1st}]],
   [{sg, 2nd}]ₗ [tₗ are looking at 7]]

Implicatures: I am not looking at myself & I am not looking at you & you are not looking at yourself & you are not looking at me.

---

9 See Siewierska 2004 for evidence of the special status of 1st/2nd person referential pronouns in many pronominal languages. As Kratzer (2009: p.220) emphasized, one indication of this special status is that *speakers and listeners are normally only referred to by person markers, whereas reference to any other things or individuals can, and in some languages must, be achieved via demonstratives or full-fledged definite descriptions.*
Our theory thus derives the ‘i-within-i’ Condition and, more generally, the constraint that reference to the same object via two distinct guises is possible only as long as the speaker(s) treat(s) it as an open question whether the same object is behind these two guises (a.o., Postal 1970; Heim 1982; Aloni 2001). Consistent with the scope of our proposal, we have provided evidence that this constraint applies in full generality, beyond the restricted area of the ‘i-within-i’ environment.

4.2 Binding over co-reference: Rule I and Binding Principles

Several generalizations have been proposed to capture the observation that identity relations between nominals is preferentially expressed by means of co-dependent rather than co-referential readings. This preference is remotely articulated in the Binding Theory where Principle C prevents a referential expression from co-referring with an expression that ccommands it and interacts with Principle A and B to favor in such cases the use of more dependent forms, i.e. reflexive forms in local environments and pronouns elsewhere. It is explicitly stated in Grodzinsky & Reinhart’s (1993) Rule I, where it takes the form of a competition at LF between co-referential vs. bound-variable construals (see Büring (2005) for discussion).

(24) Rule I: An NP $\beta$ cannot corefer with an NP $\alpha$ if replacing $\beta$ with $\gamma$, $\gamma$ a variable semantically bound by $\alpha$, yields an indistinguishable interpretation.

a. Co-reference: $[\alpha]^7_\ldots [t^7_\ldots [\beta]_j \ldots ]$

b. Variable-binding: $[\alpha]^7_\ldots [t^7_\ldots [\var^7_\ldots ] ]$

The present account derives this generalization and explains its raison d’être. Co-referential construals of the form in (24a) structurally compete with simpler bound-variable construals of the form in (24b) that are contextually equivalent to them and yet locally excludable (see 3.1 and 3.2). On our account, such construals generate a conflicting implicature, one that falsifies an identity relation that is contextually known to hold, and are thus predicted to be semantically deviant, (25), consistent with the felicity condition in (5). By contrast, variable-binding offers a more minimal structural way to express an identity relation between two nominals, while preventing it from being defeated upon implicature reasoning.

(25) # $[\alpha]^7_\ldots \text{EXH}_{\mathcal{R}} [\phi \ldots t^7_\ldots [\beta]_j \ldots ]$

a. $[\psi \ldots t^7_\ldots ]$ is a simpler structural alternative to $\phi$, and

b. $\psi$ is innocently excludable, and

c. $\psi$ and $\phi$ are contextually equivalent, hence $\psi \in \mathcal{R}$
Marty

More generally, in a similar vein as Rule I, the present account predicts that a sentence $\phi$ is felicitous at a context $C$ only if, for any two nominals $\alpha$ and $\beta$ in $\phi$, if $\alpha$ could but does not semantically bind $\beta$, then the alternative to $\phi$ that obtains by replacing $\beta$ with an index semantically bound by $\alpha$ makes a contribution in $C$ distinct from that of $\phi$. Whether it is so ultimately depends on whether $\alpha$ and $\beta$ denote the same individual on speakers’ conversational assumptions (a.o., Postal 1970; Heim 1982; Higginbotham 1980) as well as on additional considerations related to disambiguation strategies (see Schlenker 2005 for discussion).\textsuperscript{10}

This theorem subsequently captures the effects subsumed under Principle C and can further capture those subsumed under Principle A and B on certain assumptions. On our view, one way to think of the contrast in (26) is to assume that the morphophonology of English (and of many other languages) distinguishes referential from bound pro-forms in certain environments. In other words, speakers know from the phonology that me/you/him/her/it are not possible spell-outs of an index whose $\phi$-features are valued via agreement with a local DP, and thus know just from the surface forms that those pro-forms can only be construed as free referential expressions in sentences like (26a), hence the resulting disjointness effect.

(26) a. # [John]$_j$ praised [him$_{sg, \text{masc}}$ John]$_j$ \hspace{1cm} \checkmark \text{Free, } \ast \text{Local Agree}

b. [John]$_j$ praised [himself$_7$] \hspace{1cm} \checkmark \text{Free, } \ast \text{Local Agree}

Now, our account also makes certain predictions that are beyond the scope of Rule I and the Binding Principles. Consider for instance the sentence in (27). On its intended reading, this sentence sounds odd because it suggests that Mary and her brother have different mothers (it is ok then if Mary and her brother are step/half-siblings). Crucially, neither Rule I, nor Principle C capture this disjointness effect: in this sentence, neither of the descriptions that could be semantically bound by the subject [Mary’s brother] co-refers with it. It is explained however on our account: the referential expression [[her Mary] mother] competes with a simpler bound-variable alternative, e.g. [[7]’s mother], which ends up denoting the same individual as [[her Mary] mother] once bound by the subject [Mary’s brother].

(27) # [Mary]$_k$’s brother$_j$ talked to [[her]$_k$ mother]$_l$.

a. LF: \[TP [Mary’s brother]$_7$ EXH$\,\checkmark$ \vspace{.2cm} [vP t$_7$ talked to [her Mary] mother]]

\textsuperscript{10} For instance, a speaker may prefer to repeat a fully-fledged description (e.g., John told Bill that Luke talked to John’s mother) instead of using a bound pronoun (i.e., John told Bill that Luke talked to his mother) to avoid some referential ambiguity that may have arisen from using that pronoun (e.g., multiple potential antecedents: John, Bill and Luke). These disambiguation strategies can be integrated to our account if we assume that the conditions imposed on the felicitous use of pronouns are taken into account when evaluating whether an alternative would make or not a distinct contextual contribution. Avoiding a referential ambiguity in context seems like a good tie-breaker.
Disjointness Effects as Implicatures

b. \( \text{ALT}_{str}(\text{vP}): [\text{vP} t_7 \text{ talked to [7]'s mother}] \)

Conflicting SI: # Mary’s brother didn’t talk to his own mother

Another important prediction of our theory is that the preference for semantic binding should always be relativized to structural complexity. Hence, a bound nominal expression may still compete with a referential one if it is structurally simpler. The example in (28) shows that this prediction is borne out: in this sentence, the co-dependent expression [7’s ex-boyfriend] competes with a simpler alternative, namely [John], which happens to denote the same individual as [7’s ex-boyfriend] upon semantic binding by the subject [John’s ex-girlfriend]. Once again, the oddness of (28) is left unaccounted for by Rule I and the Binding Theory.

(28) # [[John]’s ex-girlfriend] \_ \_ \_ j bumped into [her ex-boyfriend] \_ \_ \_ k
  a. LF: [TP [John’s ex-girlfriend] □ \_ \_ \_ EXH \_ \_ \_ [vP t_7 bumped into [7’s ex-boyfriend]]
  b. \( \text{ALT}_{str}(\text{vP}): [t_7 \text{ bumped into [John]}] \)

Conflicting SI: # John’s ex-girlfriend didn’t bump into John

In sum, our theory derives the preference to have binding in environments where binding is possible, and offers a rationale for the competition between co-referential vs. bound-variable construals at the core of Rule I. In addition, it offers fine-grained predictions regarding the competition between these construals in environments where neither Rule I, nor the Binding Principles apply.

4.3 Transitive binding over co-binding: Rule H and Crossover Effects

Another important generalization from the past literature is that co-binding is generally dispreferred to transitive binding: if an NP is construed as a bound variable, then it must be bound to the closest antecedent possible. A well-known formulation of this locality condition on variable binding is Fox’s (2000) Rule H, (29), which has been invoked for instance to account for a number of otherwise puzzling facts about ellipsis resolution, including ‘Dahl’s puzzle’ (see Fox 2000: chapter 4).11

(29) Rule H: A pronoun \( \alpha \) can be bound by an antecedent \( \beta \) only if there is no closer antecedent \( \gamma \) such that it is possible to bind \( \alpha \) to \( \gamma \) and get the same semantic interpretation.

a. Co-binding: \( *[\alpha]_7 . . . [[\text{var}_7]_8 . . . [\text{var}_7]_8 . . .] \)

b. Transitive binding: \( \checkmark [\alpha]_7 . . . [[\text{var}_7]_8 . . . [\text{var}_8]_8 . . .] \)

11 A linking condition very similar to Rule H is also proposed in Kehler (1993: (26)): A referential element is linked to the most immediate coreferential element that c-commands it in the syntax.
To illustrate, consider the sentence in (30). Among the conceivable LFs that could express a reading on which this sentence means ‘The boy said that the boy criticized the boy’s mother’, Rule I only allows (30a) and (30b), and Rule H further rules out (30a) in favor of (30b).

\[(30)\] The boy_i said that he_i criticized his_i mother.

\[\begin{align*}
\text{a. } & \text{∗[the boy}\ _7\text{ said that [he}\ _7\text{][φ } _8\text{ t} _8\text{ criticized his}_7\text{ mother]]} \\
\text{b. } & \text{✓[the boy}\ _7\text{ said that [he}\ _7\text{][ψ } _8\text{ t} _8\text{ criticized his}_8\text{ mother]]}
\end{align*}\]

Our account also explains the *raison d’être* of this competition. In (30a), the embedded sentence \(\phi\) has the sentence \(\psi\) in (30b) as a structural alternative, which obtains from \(\phi\) by replacing a (bound) index with another (bound) index available in \(\phi\)’s substitution source (by contrast, note that \(\phi\) is not derivable from \(\psi\) since 7 is not in \(\psi\)’s substitution source). At all levels, this alternative is contextually equivalent to \(\psi\) and, since it is locally excludable, it must be excluded. Upon meaning strengthening, the SI associated with this alternative is thus computed, resulting in a contradictory representation, i.e. *The boy said that he criticized his own mother but not his own mother.* On our view then, (i) binding is the privileged way to express and preserve identity relations in the grammar (Rule I), and (ii) transitive binding is the optimal way to preserve these relations throughout the derivation (Rule H).

\[(31)\] # [α]_7...[7][EXH_7φ t8...[7]...]

\[\begin{align*}
\text{a. } [\psi \ t8...[8]... ] \text{ is a structural alternative to } \phi, \text{ and} \\
\text{b. } \psi \text{ is innocently excludable, and} \\
\text{c. } \psi \text{ and } \phi \text{ are contextually equivalent, hence } \psi \in R
\end{align*}\]

Rule H

This line of explanation extends to the Strong Crossover (SCO) effects. The SCO construal, just like the co-binding construal above, is one in which a conflicting implicature is predicted to systematically arise (since contextual equivalence always obtains). This result is exemplified in (32) where the relevant vP-alternative has the same structure as the vP in *Every boy criticized his mother* on its co-varying interpretation, namely *every boy* subject to meaning strengthening, the SI associated with this alternative is thus computed, resulting in a contradictory representation, i.e. *The boy said that he criticized his own mother but not his own mother.*

\[(32)\] #He_i criticized every boy_i’s mother.

\[\begin{align*}
\text{a. LF: [every boy]}_7[he}_7[φ} _8\text{ t} _8\text{ criticized t} _7\text{’s mother}] \\
\text{b. ALT}_{str}(vP): [t} _8\text{ criticized [8]’s mother]}
\end{align*}\]

Conflicting SI: # Every boy didn’t criticize his own mother
Disjointness Effects as Implicatures

Interestingly, I notice that this account does not extend to the Weak Crossover (WCO) construals, (33), which are treated on a par with sentences like 

\[ \text{His}_i \text{ mother criticized John}_i \] (Possible SI: But she didn’t criticize herself) or like \[ \text{Every boy}_i \text{'s mother criticized him}_i \] (Possible SI: But every boy’s mother didn’t criticize herself).

Crucially, these sentences may give rise to implicatures, but those implicatures are not mandatory and do not result in contradictory representations.

(33) His\(_i\) (own) mother criticized every boy\(_i\).

\[ \begin{align*}
\text{a. LF: } & \left[ \text{every boy}\right] ^7 \left[ \text{his}_7 \text{ mother}\right] ^8 \text{ EXH}_P \text{ [vP t}_8 \text{ criticized t}_7 \right] \\
\text{b. ALT}_{str}(vP): & \left[ \psi t_8 \text{ criticized [8]} \right] \\
\text{Possible SI: } & \text{Every boy’s mother didn’t criticize herself}
\end{align*} \]

In sum, our theory derives Rule H through the requirement to have local binding and offers a rationale for the competition between co-binding vs. transitive binding. This theory further accounts for the unavailability of variable-binding in SCO construals and, in that respect, draws a clear line between those construals and the WCO construals. The merit of this last result is left to the reader.

5 Conclusion

The present approach offers a rationale for the existence of disjointness effects and, by uncovering their underlying logic, permits to subsume under one roof several descriptive generalizations from the previous literature. This result is achieved at minimal costs by elaborating on two ideas which exist inchoately in the literature. The first one is that nominals compete with each other on the basis of their structural complexity and that these competitions lay the groundwork for implicature-reasoning (e.g., referential expressions vs. bound pronouns, complex co-dependent vs. simpler referential expressions, more complex vs. simpler definite description). The second concerns the depart between contextual and logical information in computing implicatures: the relevance of an alternative depends upon contextual considerations, including speakers’ extra-linguistic knowledge of co-referential relations, whereas its excludability is determined on the sole basis of logical considerations, with only access to those identity relations that are encoded in the grammar proper. To the best of my knowledge, the resulting theory achieves an empirical coverage of the disjointness effects never reached before. Further investigations are still needed however to fully evaluate the empirical scope of this theory and in particular to assess which effects of a similar kind it can account for and which ones are to be attributed to other grammatical constraints or other forms of pragmatic reasoning. An in-depth study of cross-sentential disjointness effects and of their possible treatment on the present view is also left for future work.
References

Disjointness Effects as Implicatures

Heim, Irene. 2007a. Forks in the road to Rule I. In *Proceedings of the 38th annual meeting of the north east linguistic society*, NELS.
Heim, Irene. 2007b. Person and number on bound and partially bound pronouns. *ms. Massachusetts Institute of Technology*.


Magri, Giorgio. 2014. An account for the homogeneity effects triggered by plural definites and conjunction based on double strengthening. *Pragmatics, Semantics and the Case of Scalar Implicatures* 99–145.


Schlenker, Philippe. 2003. Indexicality, logophoricity, and plural pronouns. *Ams-
Disjointness Effects as Implicatures

terdam studies in the theory and history of linguistic science series 4 409–428.

Paul Marty
Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS)
Schützenstr. 18, D-10117 Berlin
marty@leibniz-zas.de