Fake Features and Valuation From Context

by

Itai Bassi

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author . . . . . .
Department of Linguistics and Philosophy
September 9, 2021

Certified by . . . . . .
Danny Fox
Anshen-Chomsky Professor of Language and Thought
Thesis Supervisor

Certified by . . . . . .
Irene Heim
Professor of Linguistics
Thesis Supervisor

Certified by . . . . . .
Roger Schwarzschild
Professor of Linguistics
Thesis Supervisor

Certified by . . . . . .
Patrick D. Elliott
Doctor of Linguistics
Thesis Supervisor

Accepted by . . . . . .
Danny Fox
Anshen-Chomsky Professor of Language & Thought
Department Head

Fake Features and Valuation From Context

by

Itai Bassi

Submitted to the Department of Linguistics and Philosophy
on September 9, 2021, in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

Abstract

This thesis offers a new account of a persisting puzzle in the theory of ellipsis and association with focus: the fact that ϕ-featural content on full DPs and on bound pronouns can sometimes be ignored in focus alternatives and in calculating identity for ellipsis ('**Fake Features**'). I present new data about gender and number mismatch in ellipsis which proves difficult to model on existing approaches to fake features (e.g. Sauerland 2013; Sudo & Spathas 2020). The heart of the proposal is a derivational theory of contentful ϕ-features: they do not, as usually assumed, enter a derivation from the lexicon with listed meanings (presuppositions) that constrain the denotation of their host DP; rather, they are inserted late in the derivation towards PF by a process called "Valuation from Context" (Kučerová 2018): the features are inserted based on the meaning of the DP in the (local or global) context of evaluation. Ellipsis identity and focus alternatives are computed off of the feature-less representation. The theory assumes that the construction of local contexts for embedded constituents (Schlenker 2009) is blind to information encoded in focus alternatives. The account supports an architecture of grammar in which representations that are submitted to semantic interpretation (meaning-in-context) feed morphological valuation processes. It also implies that there is no substantial difference between "interpreted" and "uninterpreted" ϕ-features; in a sense, both are uninterpreted, the distinction being whether they are valued from context or from pieces in the structure.

Thesis Supervisor: Patrick D. Elliott
Title: Doctor of Linguistics

Thesis Supervisor: Danny Fox
Title: Anshen-Chomsky Professor of Language and Thought

Thesis Supervisor: Irene Heim
Title: Professor of Linguistics

Thesis Supervisor: Roger Schwarzschild
Title: Professor of Linguistics

Acknowledgments

A long and exciting journey as a Ph.D student must come to an end now. The years at MIT have shaped me in ways that are hard to express in words – as a linguist, but also generally as a person, by interacting with so many people with different backgrounds than mine. I know I will one day look back at these years in envy.

First, thanks to my committee: Danny Fox, Irene Heim, Roger Schwarzschild, and Patrick Elliott for all their invaluable thoughts, support, guidance, and willingness to listen even to the craziest ideas. From my early days as a B.A. student, Danny has taught me most of what I know about syntax, semantics, and pragmatics. I am always struck by his ability to be so analytical about a puzzle presented to him and to immediately unfold all the possible ways forward, their merits and challenges, before I had the chance to understand what I myself was saying. Irene's rare (and known) skill to communicate complicated thought-processes in exceptional clarity is inspiring. I think that anybody who is familiar with some of her work would be able to trace her influence on many pages in this text, if not in content then in style. Roger is a great mentor, always incredibly helpful both in finding paths through difficult empirical terrains, and for his insistence on precision whenever I (mistakenly) thought it wasn't needed. I admire his ever-willingness to be immersed in pondering about unfamiliar data or half-baked ideas. Patrick joined late to my committee and I benefited a lot from my conversations with him about this work and other projects. His vast knowledge about compositional techniques and dynamic semantics made central contribution to this thesis (even if it won't seem so to him at first glance).

There's a great many excellent people with whom I had the fortune to cross paths during my Ph.D and to talk linguistics (and other things!). Naturally, the MIT community as a whole deserves huge credit: the students, faculty, visitors and occasional invited speakers. I've learned some from all of you and a lot from some of you. Due to failure of memory, I fear that an elaborate list would leave out too many of you. So if you and I ever exchanged ideas about material even remotely related to this thesis, or to other projects, consider yourself mentioned here.

Special thanks to Ling-15, my cohort-mates at MIT: Rafael Abramovitz, Justin Colley, Colin

Davis, Suzana Fong, Verena Hehl, Maša Močnik, and Mitya Privoznov, for great moments together. They made the MIT experience, in particular the first year, fun and less stressful than it could have been.

I thank my family in Israel for being there for me: my dad Yonatan, my mom Nomi, my brothers and sisters: Noga, Shaul, Assaf, Yiska, Orna and all their families. I miss you a lot, and am sorry for being distant and not as communicative as I wanted to. I think that will change for the better now.

I don't know how this dissertation would have seen the light of day if it wasn't for the love and emotional support of my wife, Debbie and our son, Nadav at various stages of the journey. When I try to express my gratitude for their existence in my life, I feel the limits of language. The dissertation is dedicated to them.

Contents

1

Overview of dissertation

Propositions can represent the whole reality, but they cannot represent what they must have in common with reality in order to be able to represent it—the logical form.

– Ludwig Wittgenstein, *Tractatus Logico-Philosophicus, 4.12*

1.1 Main claims

The topic of this dissertation is nominal ϕ-features – number, gender and person on pronouns and nouns – at the interface between morpho-syntax and meaning. The study of ϕ-features constitutes one of the most important domains of inquiry for theories of syntax. Cross-linguistically, ϕ-features are the main exponents of grammatical dependencies such as predicate-argument agreement, and the interest in studying them lies, among other reasons, in their being useful for diagnosing and uncovering structure. But how do ϕ-features enter the syntactic engine in the first place, and how do they interface with the semantic component of grammar? That is the central question that this dissertation deals with.

Intuitively, as far as meaning goes, ϕ-features on nominals have a quite different status from their appearance on agreeing verbs and adjectives; most contemporary theories capitalize on the intuition that the semantic locus of ϕ is in the nominal domain, and their appearance outside of it,

e.g. on verbs and adjectives, is a reflex of agreement with the interpreted features; ϕ-features on verbs and adjectives are not themselves associated with interpretation. In the standard model in Generative Grammar, this means that ϕ-features that are interpreted are base-generated, whereas those on adjectives and verbs are perhaps derived in the syntax.

Semanticists who attend to ϕ-features usually assign them presuppositional meanings (Cooper 1979; Heim 2008; Sudo 2012; Charnavel 2017, a.o.). On this view, from the perspective of the connection between syntax and semantics, nominal and pronominal ϕ-features have 'lexical semantics', not qualitatively different from normal lexical items like *apple* and *drive*: they are drawn into the syntactic derivation from the lexicon with encoded piece of meaning.

This dissertation offers a different organization of the grammar of ϕ-features. The main theses I will argue for are the following:

(1) *Main claims in the dissertation*:

- **Feature-free base generation**: A syntactic derivation starts without specified values for (some) ϕ-features on D-heads. Number/gender/person that surface in pronouns and NPs do not (always) constrain the interpretation of the pronoun/NP.
- **Valuation from Context**: ϕ-features are then valued on their host NP/DP in a late stage of the derivation, based on the properties of the referent of their DP host.
- **'Interpreted' features as uninterpreted**: ϕ-features that are valued from context are not interpreted in the traditional sense, and hence do not affect the semantic value of sentences that host them. After being inserted to the syntax, they are interpreted only by PF (mediated by the morphology). The only difference between the customary labels 'interpretable' and 'uninterpretable' is whether the feature is valued from the context or from an agreement operation with some other element already present in the structure.

The empirical domain through which I will motivate these claims and explore their consequences is the phenomenon of **Fake Nominal Features** in ellipsis constructions and in association-

with-focus. This is the observation that, in various configurations, content from ϕ-features in the nominal domain—but not ordinary lexical content—can be systematically ignored in the computation of focus alternatives and in the calculation of identity for ellipsis (Heim 2008; Kratzer 2009; Sauerland 2013; McKillen 2016; Sudo & Spathas 2020, a.o.).

I will present new data about gender and number mismatch in ellipsis which proves difficult to model on the most popular contemporary existing approaches to fake features (e.g. Sauerland 2013; Bruening 2019; Sudo & Spathas 2020), but are accounted for on the current proposal. Examples of the relevant data are given (2)-(3), which many consulted speakers accept.

(2) *Context: Noam has a **female** head teacher. Tal has a **male** one.*

Noam$_1$ ohev et ha-mexanex**-et** Sel-o$_1$. Tal$_2$ lo △. (*Hebrew*)
Noam likes ACC the-head.teacher**-FEM.SG** of-him$_1$. Tal$_2$ not △.
*'Noam likes his head teacher, who is **female**, but Tal doesn't △.'*

Possible ellipsis: △ = *'like his$_2$ **male** teacher'*

(3) *Context: Becca got just **one** marble. Anna got **two**.*

Becca$_2$ hid her **marble**. So did Anna$_1$ [$_{VP}$ △].

Possible ellipsis: △ ='hid her$_1$ **marbles**'

The proposal, in a nutshell, is that the offending features that surface in the antecedent of the ellipsis (gender in (2), number in (3)) are not present at the relevant level for calculating identity for ellipsis. They are inserted into morphosyntax only later, based on the meaning (semantic value) of the host DP in its context.

The thesis will also analyze analogous fake features in association-with-focus constructions. The facts to be covered also include the well-known observation that ϕ-features on bound pronouns can be ignored in ellipsis (Ross 1967) and association-with-focus, but without requiring a rule of syntactic agreement to achieve that:

(4) Only I did **my** homework.

$\rightsquigarrow$ *sloppy reading:* No one other than me did their own homework.

Valuation-from-context requires an organization of grammar according to which certain pieces of semantic/conceptual representation can intrude into syntactic representation. If the main idea in this thesis is on the right track, it supports models of grammar that are like 'Generative Semantics' (Lakoff 1971) or close in spirit to it (e.g. Sauerland & Alexiadou 2020).

1.2 Structure of the thesis

Chapter 2 introduces the puzzling empirical landscape of Fake Features and the theoretical background assumptions necessary for describing the puzzles. In this chapter I will also critically evaluate the most prominent existing theory of the phenomenon, namely the Weak Presupposition Projection theory of ϕ-features (Sauerland 2013; McKillen 2016; Bruening 2019). I will argue that that account does not account for the full array of facts, and also fails to provide an explanation as to what is special about the presuppositions of features specifically.

Chapters 3-4 form the heart of the thesis. There I motivate, develop and implement the approach I outlined in (1) in light of the data. In chapter 3 I lay my assumptions about the Valuation-from-Context framework and how it works to account for a portion of the puzzling data. In chapter 4 I extend the basic system to deal with the bulk of the puzzles from section 2. The extension requires building a notion of **local (modified) context**, since it requires valuating features on nominals that dominate bound variables (as in (2) and (3)). The main upshot of that chapter is that the feature valuation rules look not at the global context of the utterance, but one that has been locally enriched with referents that have been recorded up to the point of evaluating the DP in question. The resulting theory will derive as a theorem that a bound variable in focus always agrees with its antecedent (while not necessarily "agreeing" with it across the focus alternatives), without a dedicated rule of agreement. The theory I end up with also requires assuming that focus-alternatives are systematically counted off from the computation of local contexts.

Finally, Chapter 5 zooms in on fake features in relative clauses, such as *I am the only one who likes **my** parents* (Kratzer 2009; Wurmbrand 2017a), a construction that is known to raise various

challenges to a unified theory of the phenomenon of fake features. The main puzzle concerns the syntactic relationship between the element that hosts the fake feature and its antecedent (the matrix subject), which is more distant than a theory of the syntax-semantics interface allows. I develop and argue for a new analysis of this construction. The main novelty towards overcoming the challenges is the claim that the underlying LF-syntax is not the way it appears at first glance. The Valuation-from-Context framework developed in previous chapters also plays an important role as an ingredient in the analysis. I will argue for my novel analysis based on independent diagnostics for structure (NPI licensing).

THIS PAGE INTENTIONALLY LEFT BLANK

2

The Puzzles of Fake Features in Focus and Ellipsis

Language disguises the thought; so that from the external form of the clothes one cannot infer the form of the thought they clothe, because the external form of the clothes is constructed with quite another object than to let the form of the body be recognized.

– Ludwig Wittgenstein, *Tractatus Logico-Philosophicus, 4.003*

Abstract

This chapter introduces the puzzling empirical landscape of Fake Features and the theoretical background assumptions necessary for describing the puzzles. I also critically evaluate the most prominent existing theory of the phenomenon, the Weak Presupposition Projection theory of ϕ-features.

2.1 Introduction: Focus Alternatives and their role in grammar

This section provides necessary background for a large portion of the phenomena that the thesis is about. It will consist mostly of introduction to the central concepts in the theory of focus alternatives, and will review the applications of the theory to the empirical domains that will figure in later chapters, such as association with focus, ellipsis and binding. An informed reader might wish to skim it quickly. The next section, 2.2, will present the phenomenon about Fake Features in association with focus.

2.1.1 Focus and Discourse

The meaning contribution of focus, on standard analyses, is to invoke alternatives. Those alternatives affect truth conditions when interacting with focus-sensitive operators, and they also feed pragmatic rules for establishing discourse coherence. To be more concrete, consider first classical examples of the discourse effects of focussing in (5)-(6). Throughout, capitalization represent high prosodic prominence which I take to reflect the presence of contrastive focus.

(5) A: Max bought the red dress.

a. B: No, ILYA bought the red dress.

b. # B': No, Ilya bought the RED dress.

(6) A: Ilya bought the yellow dress.

a. # B: No, ILYA bought the red dress.

b. B': No, Ilya bought the RED dress.

(5) shows that if one wants to claim that *Ilya bought the red dress* as an opposition or correction to the claim that *Max brought the red dress*, one must place prosodic prominence on *Ilya*, but not on e.g. *red*. Whereas for (6), where the same proposition is used to oppose the proposition *Ilya*

bought the yellow dress, the pattern is the reverse. Quite generally, there seems to be a systematic relationship between focus placement in a sentence and a contextually-salient antecedent in the discourse: positions marked by focus differ from parallel position in the antecedent, while non-focused positions have to match material in the antecedent.

These facts (and a host of similar ones related to discourse congruence) are commonly accounted for by principles that make reference to prosody, meaning, and a syntax that links them. A syntactic feature, F(ocus), is postulated (Jackendoff 1972) which can attach to constituents, as represented in (7). F-marking has consequences both for semantic interpretation and for the phonology.

(7) a. Syntax of (5a): [Ilya]$_F$ bought the red dress

b. Syntax of (5b): Ilya bought the [red]$_F$ dress

On the phonological side, F-marking translates to some kind of high prosodic prominence within a certain domain, which (in simple cases) is primarily perceived as increased pitch movements.[1] On the semantic side, as mentioned, F-marking triggers alternatives. Rooth's (1985) Alternative Semantics models alternatives as sets of ordinary denotations.[2] Every expression X is assigned two semantic values: the ordinary one, marked $[\![X]\!]$ as standard, plus a focus semantic value, notated as $[\![X]\!]_f$, which represents the dimension of focus alternatives. $[\![X]\!]_f$ produces a (possibly singleton) set of denotations; if an expression is not F-marked, its focus semantic value is just a singleton set consisting of its ordinary value. If it is F-marked, its focus semantic value is a (non-singleton) set of elements of the same semantic type.

(8) a. $[\![\text{Ilya}]\!]_f = \{[\![\text{Ilya}]\!]\} = \{\mathit{Ilya}\}$

b. $[\![\text{Ilya}_F]\!]_f = \{x \mid x \text{ is an individual}\} = \{[\![\text{Ilya}]\!], [\![\text{Max}]\!], [\![\text{Anna}]\!], ...\}$

[1] I will not go into the intricacies of the phonology of focus in any detail. Rather, I will be assuming that our raw intuition about stress will suffice to determine where F-marking lies in the examples I discuss. See also the last paragraphs of this subsection.

[2] The work done in Alternative Semantics using sets of denotations is replaced in other theories by entailment relationships between propositions with existentially closed variables (e.g. Schwarzschild 1999; Beaver & Velleman 2011). The idea is the same and my points in this dissertation apply equally to those implementations, as well as to the Structured Meaning approach to focus (Krifka 2006).

(9) a. $[\![\text{red}]\!]_f = \{[\![\text{red}]\!]\} = \{\lambda y.\ y \textit{ is red}\}$

b. $[\![\text{red}_F]\!]_f = \{P \mid P \text{ is a property}\} = \{[\![\text{red}]\!], [\![\text{yellow}]\!], [\![\text{green}]\!], ...\}$

In the version of Alternative Semantics I am adopting, focus values are semantic (e.g. model-theoretic) objects, but for now the points are the same on syntactic accounts of alternatives (Fox & Katzir 2011). In fact, I will sometimes use syntactic rather than semantic representations, and sometimes even mix between representations (trusting that no confusion arises), as this often simplifies the exposition.

Focus-semantic values for larger expressions that contain F-marking are computed from their parts by appropriate rules of composition (basically a point-wise version of the ordinary rules). Glossing over the details, the general recipe for deriving alternatives boils down to (10).

(10) **Focus Alternatives** (informal): $[\![\text{S}]\!]_f$, the alternatives of some expression S, is the set of objects which are just like S except that any F-marked element in S is replaced with a different element of the same semantic type.

Given (10), the sets of focus alternatives derived from the structures in (7) are given in (11):

(11) a. $[\![\text{Ilya}_F \text{ bought the red dress}]\!]_f$

$= \{x \textit{ bought the red dress} \mid x \text{ is an alternative to } \textit{Ilya}\}$

$= \{\textbf{\textit{Max}} \textit{ bought the red dress}, \textbf{\textit{Anna}} \textit{ bought the red dress}, ...\}$

b. $[\![\text{Ilya bought the red}_F \text{ dress}]\!]_f$

$= \{\textit{Ilya bought the } P \textit{ dress} \mid P \text{ is an alternative to } \textit{red}\}$

$= \{\textit{Ilya bought the } \textbf{\textit{yellow}} \textit{ dress}, \textit{Ilya bought the } \textbf{\textit{green}} \textit{ dress}, ...\}$

These objects allow us to formulate rules of discourse coherence, relating a focus-containing utterance to its antecedent. Here I adopt from Rooth a general principle of Appropriate Contrast (Rooth 1992a,b).

(12) Appropriate Contrast: A constituent S *contrasts appropriately* with some discourse entity α (an individual, property, or proposition made salient in the context) iff:

a. $\alpha \neq [\![S]\!]$,[3] and

b. $\alpha \in [\![S]\!]_f$.

The facts in (5)-(6) are now accounted for, assuming these discourses must obey Appropriate Contrast. In (5) the antecedent *Max bought the red dress* is a member of (11a) but not a member of (11b), so (5a) but not (5b) contrasts appropriately with it. In (6) the pattern is reversed because the antecedent is different.[4]

Besides the F-marked material, all other material in the felicitous discourses above is phonologically reduced (destressed). This corresponds to the intuition that reduced material must be given in the discourse. As a useful descriptive generalization, we can say that material that is phonologically reduced cannot be F-marked.[5]

This is an opportunity to say a word of caution. In this dissertation I will be assuming that F-marking is the only information-theoretic feature that links intonation, meaning and discourse in English. This is likely a huge oversimplification. According to recent literature, predicting the full range of data that intuitively go under the term 'focus effects' requires dividing the labour between at least two notions (e.g. Féry & Samek-Lodovici 2006; Beaver & Velleman 2011; Büring 2016; Schwarzschild 2019; Kratzer & Selkirk 2020). The precise identity of the information-structural factors relevant for accent placement and how they interact in grammar are currently actively debated questions. Schwarzschild (2019) and Kratzer & Selkirk (2020), for example, each in their own and different way, distinguish between Givenness and Focus, both being encoded in grammar; 'Givenness' is roughly about whether material is discourse-old (given) or discourse-new, and 'focus' is reserved for material that receives exhaustive or contrastive interpretation and

[3]This condition might have to be strengthened in light of the discussions in Wagner (2012), Katzir (2013), Büring (2016:115-116) and Büring (2019), but in a way that doesn't bear on present concerns.

[4]The rule in (12) should be taken as a necessary but not sufficient condition for an intuitively felicitous use of focusing. Another condition is AvoidF! (Schwarzschild 1999) which requires minimizing F-marking as much as possible. This will prevent e.g. *ILYA bought the RED dress* from being an appropriate response to either (5) or (6). Kratzer & Selkirk (2020:35) effectively bake AvoidF into the rule of Appropriate Contrast. For us (12) will do.

[5]By Destressing I generally mean lack of pitch accent. Second-Occurrence Foci (von Fintel 1994) are the exception, being deaccented although not entirely stress-less relative to the surrounding environment (see e.g. Beaver *et al.* 2007).

associates with sentence-internal operators.[6] On those theories the analyses of facts like (5)-(6) and others below are more complicated than I make them out to be.

I will try to steer away from these complications altogether because I don't see that they have direct bearings on the issue I am concerned with here which is the mechanism for generating alternatives. Whatever the right theory of focus, Givenness and their influence on accent turns out to be, it is as far as I know universally agreed that—at least in the examples I discuss here—*some* information-theoretical property is responsible for highlighting alternatives to the accented phrase, and that the procedure for generating alternatives is by replacement along the lines of (10). This much is sufficient to appreciate the puzzles in the center of this thesis. I thus keep to just F-marking and a simple notion of Appropriate Contrast in (12), and will use examples in which the mapping from intonation to the placement of the F-marks is quite straightforward and hopefully uncontroversial on any theory.[7]

2.1.2 Association with focus

As mentioned, alternatives also affect truth-conditional content, when interacting with Focus-sensitive operators like *only*, *even*, *always*. In (13)-(14), the propositions that *only* can negate (exclude) are constrained by the focus structure of its scope. The scope of *only* is called its PREJACENT.

(13) only ILYA bought the red dress

a. *propositions in* $[\![Ilya_F \text{ bought the red dress}]\!]_f$ *are possible exclusions* (cf. 11a)

b. *propositions in* $[\![Ilya \text{ bought the red}_F \text{ dress}]\!]_f$ *are* ***not*** *possible exclusions* (cf. 11b)

(14) Ilya only bought the RED dress

[6] Facts that have been taken to motivate a two-factor system and figure in this debate include: (i) second-occurrence focus (Krifka 2004; Beaver *et al.* 2007, a.o.), (ii) cases where material that is discourse-new and also contrasted is acoustically different from material that is new but not contrasted (Katz & Selkirk 2011; Kratzer & Selkirk 2020), and (iii) cases where a primary accent can or cannot structurally intervene between an operator and its associated focus (Schwarzschild 2019).

[7] My use of F-marking corresponds roughly to Kratzer & Selkirk's (2020) FoC-marking, and to Schwarzschild's (2019) N-marking as well as F-marking.

a. *propositions in* $[\![Ilya_F$ *bought the red dress*$]\!]_f$ *are* ***not*** *possible exclusions*

b. *propositions in* $[\![Ilya$ *bought the red*$_F$ *dress*$]\!]_f$ *are possible exclusions*

Only thus encodes sensitivity to the focus structure of its prejacent (Beaver & Clark 2008). Its denotation, informally and glossing over irrelevant details, is stated in (15). Note that I assume a syntax where focus-sensitive adverbs always take propositional scope at LF (Hirsch 2017), as shown in (16).

(15) [[only S]] is true iff [[S]] is true and all contextually relevant alternatives in $[\![S]\!]_f$ are false.

(16) a. Logical Form of (13): only [Ilya$_F$ bought the red dress]

b. Logical Form of (14): only [Ilya bought the red$_F$ dress][8]

The syntax of focus association

For the ensuing discussion, it will become important that the LF of sentences with pre-subject *only* indeed always looks like (16a) and not like (17) in which *only+subject* is a constituent to the exclusion of post-subject material.

(17) [only Ilya$_{(F)}$] [bought the red dress]

On the surface it does seem that (13) has the syntax of (17), and in general that so-called 'Constituent *only*' forms a constituent with its surface sister (Wagner 2006, a.o.). To wit, the string *only+subject* behaves like a constituent for the purpose of movement and other syntactic processes such as Negative Inversion in English and the V2 effect in Germanic. However, this is only evidence that the *surface* syntax looks like (17). I assume, relying on the arguments in Hirsch (2017); Hirsch & Wagner (2019) (see also references cited therein) that the semantically active *only* is covertly attached at a clausal level as in (16a). Constituent *only* is the overt realization of the interpreted one.

[8] This representation could result from reconstruction of the subject to its *vP*-internal position; Hirsch 2017.

I'll provide here one piece of independent evidence that the correct LF representation is as in (16a). Jacobson (2012) remarked that pre-subject *only* can associate with focus inside the VP, if it also associates with the subject (or something within the subject):

(18) *Scenario: I ask you whether anyone will help bring anything to the dinner. You answer:*
I hate to tell you, but **only SUE will bring SALAD**. (Jacobson 2012:ex.47)
$\rightsquigarrow$ *No one else will bring anything else*

(19) a. only [$_S$ Sue$_F$ will bring salad$_F$]

b. $[\![S]\!]_f$ = {x will bring y | x, y are alternatives to *Sue*, *Salad* respectively}

Getting the correct entailment in (18) is straightforward if pre-subject *only* takes propositional scope and doubly-associates with the focused positions, cf. (19a). It is not as easy to explain if the scope of *only* is limited to the subject, as in (17).[9]

I will assume that when constituent *only* is in subject position, the (covert) interpreted operator it realizes is adjoined at least as high as TP, whereas for constituent *only* in lower positions it is adjoined to *v*P (Hirsch & Wagner 2019). I further assume that attaching the interpreted operator at TP or higher is possible only when some F-marking is placed above *v*P; this accounts for why subject-*only* (more generally, sentence-initial *only*) can't associate with focus in lower positions *unless* it also associates with something in its surface sister, cf. (18).[10] For the rest of the dissertation I abstract away from these details and simply represent *only* as taking clausal scope.

2.1.3 Ellipsis and Rebinding

Focus structures and alternatives are important ingredients in the theory of ellipsis as well. Ellipsis typically requires semantic identity with an antecedent (Keenan 1971; Sag 1976; Hankamer & Sag 1976; Williams 1977; Merchant 2019, a.o.). But as is well known, the licensing condition on ellipsis cannot involve something as simple as merely checking for identity at the level of the

[9] Another relevant data point is that children often make mistakes of associating a pre-subject *only* just with object positions, as documented extensively by Crain *et al.* (1994) and Sugawara (2016). This suggests that children, more freely than adults, allow a pre-subject *only* to take scope above the object.

[10] And in light of footnote 9, this last restriction might be absent in children's grammar.

elided constituent itself; the theory must make reference to Appropriate Contrast of possibly larger constituents (**Focus Parallelism**; see Tancredi 1992; Rooth 1992a; Jacobson 1992; Fiengo & May 1994; Heim 1997; Takahashi & Fox 2005; Charlow 2019 among many). I focus here on ellipsis of VPs, but I assume that (20) reflects a general condition on ellipsis.

(20) Focus Parallelism Condition on Ellipsis: Ellipsis of VP_E is licensed only if VP_E is embedded in a constituent, S, and there is another constituent in the surrounding discourse, A, and S contrasts appropriately with $[\![A]\!]$.[11]

This entails that the focus alternatives of S need to contain $[\![A]\!]$ for ellipsis to be successful. The argument for something like (20) over a simpler identity condition that doesn't refer to Appropriate Contrast, comes from the interaction of ellipsis and variable binding. The relevant cases, which often go by the name Rebinding, are structures where a variable in the ellipsis is bound from outside of it. Variable binding is going to take a central place in this thesis (chapter 4), so I will spell out the issue in some detail now. Consider first standard cases of strict-sloppy ambiguities in VP ellipsis (Ross 1967).

(21) Mary saw her parents, and ANA did [$_{VP}$ $\triangle$] too.

a. $\rightsquigarrow$ *sloppy-identity reading:* Mary saw Mary's parents, Ana saw Ana's parents.

b. $\rightsquigarrow$ *strict-identity reading:* Mary saw x's parents, Ana saw x's parents. (x = Mary or someone else contextually salient)

Sloppy readings are represented using variable binding as usual. The representation in (22) presupposes familiarity with the basic concepts and notations in Heim & Kratzer (1998) about the syntax-semantics of binding.

(22) LF of (21a): Mary $\underbrace{\lambda_1\ t_1 \text{ saw her}_1 \text{ parents}}$. Ana_F did $\underbrace{\lambda_2\ t_2 \text{ saw her}_2 \text{ parents}}_{VP_E}$, too.

[11] Parallelism as defined in (20) requires an *overt* antecedent (*S*) with the exact same focus structure as the sentence that hosts the ellipsis site. Rooth (1992a) showed that this requirement is too strict and must be weakened to allow for certain cases where an antecedent is accommodated/implied rather than overtly expressed ('Implicational Bridging'). Fox (1999) further proposed a restriction on the allowable accommodations that Rooth permitted. I will not discuss the interesting complications arising from Implicational Bridging and will work with (20).

The underscored constituents are semantically identical, as they express the property of seeing one's parents. So a simple semantic-identity condition would correctly license ellipsis.

Now consider Rebinding. (23) are modeled after examples in Jacobson 1992.

(23) a. Mary thinks you will [$_{VP}$ call her], while SUE thinks JOHN will [$_{VP}$ $\triangle$].
'Mary$_1$ thinks you will call her$_1$, while Sue$_2$ thinks John will call her$_2$'

b. I know which pills$_1$ George has [$_{VP}$ taken t_1], but not which pills MARY has [$_{VP}$ $\triangle$].

Representations that would yield the intended interpretations of (23) are in (24).[12]

(24) a. Mary λ_1 t_1 thinks you will [$_{VP}$ call her$_1$], while
Sue$_F$ λ_2 t_2 thinks John$_F$ will [$_{VP}$ call her$_2$].

b. I know which pills λ_1 George has [$_{VP}$ taken t_1],
but not which pills λ_2 Mary$_F$ has [$_{VP}$ taken t_2].

In both (24a)-(24b) the VP in the ellipsis clause contains an occurrence of a different variable from the one in the antecedent VP. Now, it is persistently tricky to state a notion of semantic identity that would apply to VPs in these configurations, without running into problems elsewhere (Sag 1976; Williams 1977; Heim 1997; see more in the excursus below), as the two VPs contain occurrences of different variables (this problem does not arise in (22) because even though the ellipsis site and the antecedent site include occurrences of different variables name, they also include their respective λ-binder).

Focus Parallelism in (20), however, is adequate. In both (24a)-(24b) the whole second clause appropriately contrasts with the first. In particular, the following holds:

(25) a. *Mary thinks you will call Mary* $\in$ {x thinks y will call x | x, y are individuals}

b. *I know which pills George has taken* $\in$ {I know which pills x has taken | x is an individual}

[12] Binding is known to create technical compositionality problems for classical alternative semantics in configurations like (24). The problems are fixed on other versions including the one in which my official proposal will be implemented later, so we needn't worry about that.

In general, given the problems of stating an adequate identity condition for unbound variables, we can conclude that a legitimate Parallelism domain for an elided VP that hosts a bound variable must include the binder of that variable.

And since ellipsis is sensitive to Focus Parallelism, it could thus be used as a tool to investigate the properties of Parallelism domains. In particular, ellipsis will be used here to probe questions about Fake ϕ-features in alternatives.

2.1.4 Excursus: indexations and the identity condition on ellipsis

Here I'd like to make a small excursus and review a known (and somewhat pedantic) issue regarding indexation, ellipsis and binding. It concerns the question fo whether modeling sloppy readings for variables necessitate syntactic representations in which the variables are counter-indexed (Heim 1997; Charlow 2019). This excursus is not strictly necessary to follow the main plot of the thesis and may be skipped, but its conclusion anticipates the ideas I will explore later as part of my theory of fake features.

Looking again at rebinding in (23)-(24), one might wonder whether focus parallelism is really called for in a theory of ellipsis licensing. What if instead of counter-indexing in (24) we reused the variable name in the ellipsis site, making it the same as in the antecedent? then we could have kept to a local-identity condition on VP-ellipsis that requires just the VPs themselves to be semantically identical – if 'identity' means sameness of value relative to **every** variable-assignment function. The representation for (23a) that would feed ellipsis on that attempt would be:

(26) Mary λ_1 t_1 thinks you will [$_{\text{VP}}$ call her$_1$], while
Sue$_F$ λ_1 t_1 thinks John$_F$ will [$_{\text{VP}}$ call her$_1$].

Indeed, it seems like this would remove any need for focus parallelism for the purpose of ellipsis. However, such a removal would create problems elsewhere; it would allow ellipsis licensing when the antecedent VP and the elided VP each contain an occurrence of the same variable which aren't bound from parallel positions. (27) is a case in point, where the VPs are identical although

one occurrence of the variable 1 is free and the other is bound by the subject.[13] The VPs have the same meaning under every assignment to the variable 1, but ellipsis is in fact not possible (the ellipsis site cannot mean "met Bill's parents", when the VP in the antecedent doesn't).

(27) Anastasia thinks I have [$_{VP}$ met his$_1$ parents]. (where $his_1 \neq Bill$)

*BILL λ_1 t_1 KNOWS I have ~~[$_{VP}$ met his$_1$ parents]~~.

Focus Parallelism thus seems necessary for ellipsis licensing. If we examine Parallelism at the level of the whole clauses, it is easy to see that the two clauses don't stand in the Appropriate Contrast relation. *Anastasia thinks I have met his$_1$ parents*, whichever boy *his*$_1$ refers to, is not an element of {*x P* I have met *x*'s parents | x is an entity, P an element of the type of attitude verbs}.

This is not the end of the story, however. There is still a problem lurking. Focus Parallelism merely instructs us to find *one* constituent that dominates the ellipsis site and satisfies Appropriate Contrast, so we must rule out all possible candidates in (27). Consider as a candidate a domain which is below the binder of the variable, in particular the constituent [*knows$_F$ I have met his*$_1$ *parents*]. Does this phrase contrasts appropriately with anything? we can't actually tell, because it contains a variable free, and we haven't been explicit in (20) about how Appropriate Contrast handles assignment-sensitivity. But if we followed the idea entertained above and defined Contrast by universal quantification over variable assignments, we would run into a problem again. That is, if we were to define that S appropriately contrasts with some α if it appropriately contrasts with it (in the old sense) under every assignment g, we would incorrectly permit ellipsis in (27). This is because [*knows$_F$ Mary met his*$_1$ *parents*] would appropriately contrast (in the new sense) with an antecedent, namely with [*thinks I have met his*$_1$ *parents*]: under every assignment g, the latter is a member of the focus value of the former. Simply put, the two occurrences of *his* carry the same referential index, and therefore have the same interpretation under any g.

Considerations of this nature led Heim (1997) to do away with the complications that bound variables create by adopting a syntactic well-formedness condition on the use of variables. Heim

[13] Nothing crucial hinges on the first *his* being free in (27); the point will be the same if it is bound by e.g. *John* from higher, as long the two *his*'s aren't bound from parallel positions. That would just make the exposition more complicated.

(1997) formulated a No Meaningless Coindexation (NMC) convention, which simply stipulated that if one occurrence of a variable is bound by a λ-node, all other occurrences (in the discourse) of the same variable must be bound by that λ-node. NMC bans representations like (27) from being legitimate LFs from the start, since one occurrence of his_1 is λ-bound and the other is free. It also blocks the derivation in (26) (the two occurrences of her_1 aren't bound by the same node), leaving only counter-indexing as in (24) as a possible representation of the intended meaning in rebinding structures. NMC immediately eliminates the headaches caused by variables for a theory of ellipsis licensing, but at the price of a syntactic stipulation whose intended effect should, intuitively, follow independently from the semantics.

Recently, Charlow (2019) showed how it is possible to have NMC's effects follow from the semantics and thus dispense with it in a theory of ellipsis, as long as we are careful in stating how Appropriate Contrast works. As Charlow observed, the problem that forced Heim to stipulate NMC stems from the implicit assumption that the way to define identity/Appropriate Contrast for constituents that contain unbound variables is by universally closing off the variable assignment that interprets those constituents. But it is arguably more natural to check Appropriate Contrast against the **local assignment (local context)**: those assignments that are 'relevant' for evaluating the semantic value of the constituents in question. Informally, even though the two his_1's in (27) carry the same referential index, their semantic value differ locally because the assignment that is used to interpreted them is different. In slightly more detail (but still informally), in the ellipsis clause *Bill* binds occurrences of variable 1 (via λ_1), and this modifies the variable assignment that evaluates everything that follows *Bill* by adding the information that the value of index 1 is now Bill. Against that modified assignment, then, his_1 refers to Bill. The occurrence of his_1 in the antecedent clause, on the other hand, doesn't refer to Bill in *its* own context, which happens to be the global context of the discourse.

This difference in local semantic value, Charlow (2019) argues, is enough to disrupt Appropriate Contrast and thus to correctly block ellipsis in (27). No constituent in the second clause of (27) will Appropriately Contrast with anything in the first clause, once we make sure to evaluate the in-

terpretation of a constituent—both its ordinary semantic value and, importantly, its focus semantic value—against its own local context. In particular:

(28) In (27):

$[\![\text{thinks Mary met his}_1 \text{ parents}]\!]^{g}$ (where $g(1) \neq \textit{Bill}$)

$\notin \left\{ [\![\text{knows}_F \text{ Mary met his}_1 \text{ parents}]\!]_f^{g^{1 \to \textit{Bill}}} \right\}$

So even with exact-identity of indexation at the VP level, a local-assignment-sensitive formulation of Appropriate Contrast rules out ellipsis in (27). No NMC convention is needed. Importantly, ellipsis in the felicitous rebinding structures is still predicted to be ok, even with exact identity of indexation as in (26), because a Parallelism domain that includes the binder in both positions satisfies Appropriate Contrast.

Be that as it may, on either Heim's NMC (which requires counter-indexing to represent sloppy readings in ellipsis) or Charlow's local-assignment approach (which doesn't), there is no escape from having a condition on ellipsis that looks at focus values of possibly larger constituents than the one that is targeted for ellipsis.

But Charlow's reduction of NMC to the working of semantics-in-context constitutes, in my view, a step forward, as it allows for a simplification of the theory of ellipsis licensing. More broadly, a lesson we can draw from the discussion is that when dealing with a complicated phenomenon that involves bound variables, it is useful to think seriously about the mechanism of context-change by binding (assignment modification) and to embed that mechanism within the general theory of that phenomenon.

The approach to Fake Features in ellipsis and focus alternatives defended in this dissertation will also take this point seriously. I will argue that taking local contexts in binding configurations into consideration will help coming up with a unified solution to the puzzling data we will tackle in the next sections.

End of Excursus.

2.2 The puzzles of Fake Features in Focus Alternatives

In the previous section I reviewed a number of phenomena that make reference to focus alternatives. The theory of all of them relies on the rather innocent assumption that alternatives are arrived at by a simple replacement algorithm which is given in (29) (basically a restatement of (10)). (29) conforms to the age-old intuition that the very purpose of focus placement is to indicate the shape of alternative(s) that the speaker contrasts her utterance with. Everything that isn't F-marked '**strongly projects**' – it must stay constant across all the alternatives (29b).

(29) **Deriving Alternatives** (simple algorithm): To derive $[\![S]\!]_f$, the focus-alternative set of S, do the following:

a. Replace F-marked constituents in S with an element of the same type.

b. Leave everything else the same.

'Weak Projection' names the observation that things are not as simple as that (Sauerland 2003; McKillen 2016; Bruening 2019, a.o.). Specifically, (29b) is too restrictive. The facts I present in this section motivate a weakening of (29), as follows:

(30) **Deriving Alternatives** (more adequate algorithm): To derive $[\![S]\!]_f$, the focus-alternative set of S, do the following:

a. Replace F-marked constituents in S with an element of the same type.

b. Leave everything else the same **except for information encoded by grammatical ϕ-features (number, gender, person). ϕ-features don't have to contribute their expected semantics to alternatives**.

By this I mean that ϕ-featural information that constrains the interpretation of a DP (pronouns and definite nominals) do not have to constrain its interpretation across focus alternatives. I will call those **Fake Features**. The phenomena include the well-known fact that ϕ-features on bound-variable pronouns are ignored for the in bound-variable readings in association with focus.[14]

[14]The authors cited above group under 'Weak Projection' various other facts that lie outside the scope of (30). In

This section presents facts that motivate a switch from (29) to (30). It will also present the only theory available to date that unifies the phenomena: the Weak *Presupposition* Projection theory (Sauerland 2013). The next section, 2.4, will present a new empirical challenge for the Weak Presupposition Projection theory, based on certain acceptable mismatches of features in ellipsis which that theory undergenerates. I also discuss the explanatory adequacy of the generalization in (30), and that discussion will serve as a segue to my alternative proposal in the following chapters.

The structure of this section is as follows. First I discuss Fake number features on definite nominals, based on an observation by von Heusinger (2007) (discussed further in Sauerland 2013) and others (2.2.1). Then I show (2.2.2) that parallel facts hold for gender features on nominals. Finally I review the problem of ϕ-features on bound-variable pronouns (2.2.3). For each of these, I also show how it is handled by the Weak Presupposition Projection theory (Sauerland 2013).

2.2.1 Number features on nouns

von Heusinger (2007) brought to attention an example similar to (31).

(31) Mary only talked to the GERMAN professor (based on von Heusinger 2007)

a. *Felicitous in this context*: There is one German professor, 2 Japanese professors, and 5 American professors at the party, as well as students. For all we know Mary might have talked to some students.

(31), with focus on *German* and deaccenting *professor*, intuitively entails that Mary didn't talk to any professor except for the German one. Both the prosody and the context suggest the focus structure in (32a),[15] and the simple algorithm for generating alternatives predicts (32b).

this thesis I am limiting my discussion to ϕ-features, though I will get back to one more case that I analyze, the case of strict readings of reflexives (in section 4.6.2).

[15]That focus is crucially on the adjective and not on the NP/DP layer is suggested both by established generalizations about the nature of focus projection from adjectives (Büring 2016; Büring 2019 and references therein), and by noticing that if F-marking were on the whole DP or the NP part, we would get a wrong reading: we would expect excluded alternatives to vary at the place of *professor* and would therefore predict (31) to entail that Mary didn't talk to students either, contrary to context.

Having said that, it is possible that an appropriately sophisticated theory of discourse, focus and accent (outside my hypothesis space) could eventually predict the correct entailments in (31) while assuming that the associated focus is

(32) a. LF: Only [$_S$ Mary talked to the [German]$_F$ professor]

b. Predicted $[\![S]\!]_f$: {*Mary talked to the x professor* | *x* is an alternative to *German*}

The crucial observation is that the singularity (uniqueness) inference associated with *the German$_F$ professor*—that there is just one such professor—needs to be ignored at the level of alternatives that *only* excludes. If the uniqueness inference projected to all alternatives, and assuming that *x* in (32b) ranges over nationalities, then (31) will suffer from reference failure and shouldn't be felicitous in the context; there are two Japanese professors.[16]

What would help is if the alternatives that *only* excludes had the option of encoding different number information from the prejacent, i.e. if the relevant alternatives were denoted by sentences in which singular were replaced by plural: {*Mary talked to the Japanese professor***s**, *Mary talked to the American professor***s**}. But it is not obvious how to achieve this in a principled way, given that number marking in nominal phrases is uncontroversially located at the NP/DP level, outside the adjective and therefore outside the scope of F-marking.

The facts are parallel when we consider the non-uniqueness inference of plural DPs. (33) entails that I didn't talk to the German professor, so the plurality inference of *the Japanese professors* needs to have the option of being suspended in the alternatives.

(33) (*Same context*:) Mary only talked to the JAPANESE professors.

a. $\rightsquigarrow$ *Mary didn't talk to the German professor*

b. LF: Only [$_S$ Mary talked to the [Japanese]$_F$ professors]

c. Predicted $[\![S]\!]_f$ = {*Mary talked to the x professors* | *x* is an alternative to *Japanese*}

placed on the whole DP even though *professor* is deaccented. I do not know what such a theory would look like, and thus I continue to assume the LF in (32a), but I point out that even if we had one for cases like (31), the puzzle remains for examples that involve pronoun binding that I discuss below (cf. 36)-(37)).

[16]Roger Schwarzschild (p.c.) asks what if *the German$_F$ professor* could have alternatives that are fine-grained enough to distinguish between the professors and guarantee unique reference, e.g. the alternatives *the tall Japanese professor* and *the short Japanese professor*. The context given in (31) obviously doesn't provide such information; all we know about the Japanese professors is that there are two of them (we would moreover need to play the same trick for 5 American professors). Assuming that the alternatives to *German* that figure into the propositions excluded by *only* must be properties that are salient in the context, or at least can be safely accommodated without risk of reference failure, it is not clear how such a view could work. I should also point out that the theory of formal alternatives in Fox & Katzir 2011 disallows *the tall Japanese professor* from being a legitimate alternative to *the German professor*.

Here I point out that it is the morphosemantics of the number feature that spells the trouble. If a lexical numeral is present, it must add its contribution to alternatives, as predicted by the simple algorithm. The minimally different (34) (pronounced with deaccented *two*, to make sure it isn't F-marked) is infelicitous in the above context; it can only exclude alternatives that are "about" two N professors, for some nationality N.

(34) Mary only talked to the **two** JAPANESE professors. (# *in the context of (33)*)

Sauerland 2013

Sauerland (2013) proposes to explain (31) by invoking two main assumptions. One is that the uniqueness inference of singular DPs is the result of a singular-number feature imposing a presuppositional restriction to atomic (singular) reference, and the other is that this feature has the option of not contributing its presupposition to alternatives.

I summarize the gist of Sauerland's (2013) proposal in (35), which is going to be useful for my purposes. His account is embedded within a theory of the syntax-semantics of definite DPs on which the DP layer's denotation is compatible with both atomic and non-atomic reference (cf. (35b); there will be a precise semantics in the next chapter); the number feature adjoins on top of DP and adds the atomicity presupposition (Sauerland 2003). For ease of illustration, here I employ a syntactic representation for alternatives (see 35c), and I cash out the idea that singular number is ignored in alternatives using a deletion operation. The SING feature is present in the prejacent but gets deleted at the level of its alternatives, and therefore the alternatives are compatible with plural reference.[17] (35d) is the semantic representation of the alternatives and is to be taken as the set that includes, for each S' in (35c), the interpretation of S'. I notate it $[\![S]\!]_{f_{WK}}$ (WK for 'weak') to distinguish it from the standard $[\![S]\!]_f$.

[17] Feature deletion is not in fact embraced in Sauerland 2013, as he opted for a semantic framework for alternatives. A deletion operation does not make sense on a semantic (i.e. model-theoretic) conception of alternatives. Rather Sauerland proposed that the SING feature (as well as other ϕ-features, see below) can have a vacuous denotation in its focus value. Either way, the end result is that SING does not contribute meaning to alternatives.

(35) *Sauerland (2013): The singular feature on a full DP does not (have to) contribute its meaning at the level of alternatives*

a. LF of (31):
Only $\left[_S\right.$ Mary talked to [$_{\text{NUMP}}$ SING [$_{\text{DP}}$ the [German$_F$ professor]]]$\left.\right]$

b. $[\![[_{\text{DP}}$ the $[_{\text{NP}}$ ... $]]\!]$ = the maximal contextually-salient individual that satisfies $[\![$NP$]\!]$, where $[\![$NP$]\!]$ is defined for both atomic and non-atomic (sum-)individuals.

c. Alternatives of the prejacent S:

i. Mary talked to [$_{\text{NUMP}}$ ~~SING~~ [$_{\text{DP}}$ the [Japanese professor(s)]]] ,

ii. Mary talked to [$_{\text{NUMP}}$ ~~SING~~ [$_{\text{DP}}$ the [American professor(s)]]] , ...

d. $[\![\text{S}]\!]_{f_{\text{WK}}}$ = {*Mary talked to the x professor or professors* | *x* is an alternative to *German*}

For the plural case (33), not previously discussed as far as I know, the story could be analogous: a feature PL contributes a non-atomicity presupposition but not necessarily in alternatives. Another possibility is that PL is semantically vacuous from the get-go and receives the non-atomicity inference indirectly, through competition with SING (Sauerland 2003, 2008). On the latter option deletion in alternatives might not be necessary to get the desired entailments.[18]

Sauerland (2013) suggests that the ability of singular to be ignored in alternatives is a property of a larger class of items referred to as PURE-PRESUPPOSITION TRIGGERS. I argue against this claim in section 2.4, but for now what is relevant is that on his theory the simple (restrictive) algorithm of generating alternatives must be abandoned. This is because it is committed to the existence of operations that manipulate the shape/interpretation of alternatives beyond a simple replacement of F-marked material.

[18]Proponents of the weak theory of plurality, on which plural is semantically vacuous, usually derive the non-atomicity inference of plurals either as a scalar implicature (e.g. Zweig 2009; Spector 2007) or as an implicated presupposition, through the principle MAXIMIZE PRESUPPOSITION! (Sauerland 2003, 2008). For some challenges to the weak theory of plurals see de Swart & Farkas 2010; Grimm 2013; Martí 2020. I will briefly return in chapter 3 to some of the considerations that militate for or against the weak theory of plurals in the context of definite DPs.

Fake Features that dominate a bound variable

Let us expand our data set a bit. Abstractly, the problem that emerges for the simple theory of alternatives comes about whenever a semantically-active element X is not dominated by an F-marked constituent yet is observed not to be semantically-active at the level of alternatives (in (35) X=SING). This is what motivated postulating the deletion in (35c). In the LF in (35a) F-marking is in the scope of X, but that is not necessary to bring about the puzzle; it arises for instance also for structures in which an F-marked phrase binds a variable inside the scope of X. Consider (36)-(37).

(36) *Context: Becca has 3 advisors, Eleanor has 2 advisors, Kelly has 1 advisor.*

a. Only KELLY$_1$ will meet with [$_{DP}$ her$_1$ advisor] tomorrow.

b. $\rightsquigarrow$ $\forall x \in \{Becca, Eleanor\}$: *x won't meet with x's advisors tomorrow.*

(37) *Context: Officer Mary arrested 1 guy, Officer Kelly arrested 2 guys and Officer Becca arrested 3 guys. Which officers were humane and which ones cruel during the arrests?*

a. Only MARY$_1$ treated [$_{DP}$ the guy she$_1$ arrested] with respect.

b. $\rightsquigarrow$ $\forall x \in \{Kelly, Becca\}$: *x didn't treat the guys x arrested with respect.*

On the intended meanings, the interpretation of the focused phrase co-varies with a pronoun's reference (her_1) across the focus alternatives. I assume this obtains when the focused phrase binds the pronoun, as in (38a). The point is that the singularity inference of the singular definite/possessive DP that contains the pronoun does not have to project to the alternatives; in the context of (37), the alternatives to Mary arrested more than one guy. Using Sauerland's idea, the remedy could be the same: the number restriction can disappear from alternatives (38b), so [*the guy...*] is compatible with non-atomic reference, (38c).

(38) *Sauerland 2013-style analysis of (37a)*

a. LF:

Only $\left[_S\ \text{Mary}_F\ \lambda_1\ t_1\ \text{treated}\ [_{\text{NUMP}}\ \text{SING}\ [_{DP}\ \text{the guy she}_1\ \text{arrested}]]\ \text{with respect}\right]$

b. Alternatives of *S*:

i. Kelly λ_1 t_1 treated [$_{\text{NUMP}}$ ~~SING~~ [$_{\text{DP}}$ the guy she$_1$ arrested]] with respect ,

ii. Becca λ_1 t_1 treated [$_{\text{NUMP}}$ ~~SING~~ [$_{\text{DP}}$ the guy she$_1$ arrested]] with respect , ...

c. $[\![S]\!]_{f_{\text{WK}}}$ = {*x treated the guy or guys x arrested with respect* | *x* is an individual}

I have used only definite DPs when illustrating weak number projection, and I will continue to do so. One would expect, at least as a null hypothesis, that the points could be illustrated with indefinite DPs as well. This is however not easy to do, at least when indefintes' behavior is probed using association with *only*, which negates alternatives. Indefinites in episodic contexts (in argument positions) carry existential force both in their singular ('a guy') and plural ('guys') occurrences, and when negated they support virtually the same entailments, whether singular or plural. So it is not obvious how to test whether SING/PL on indefinites needs to disappear from alternatives.[19]

2.2.2 Gender features on nouns

I show here that there are cases of weak gender projection (Fake Gender) in focus in cases structurally similar to von Heusinger's (2007) example that we saw in (31). (39) is from German.[20]

(39) *German Fake Gender*

a. *Context: At the conference I talked to some students, but to almost no professor.*

b. Ich habe nur mit [der DEUTSCHEN Professor-**in**] gesprochen
I have only with [the GERMAN professor-FEM] spoken
'I only talked to the GERMAN professor, who is female'

c. $\rightsquigarrow \forall P, P \neq$ *German*: *I didn't talk to the P professor.*

[19] I also point out that the analysis of features on indefinites should not obviously be assumed to be the same as the one for features on definites, syntactically and semantically. The standard presuppositional account of ϕ-features on pronouns (Cooper 1979) works well for pronouns and definite DPs but it is debated whether number (and gender) features on indefinites are presuppositional in the first place, and if they are, how to integrate them into the composition of indefinites (for a discussion of some of the complications see Sauerland 2003). This depends on the proper analysis of indefinites independently of features, e.g. whether they are quantificational, referential, or something else, and how indefinites project presuppositions. Because the situation is not clear with number on indefinites both empirically and theoretically, I will stick on the bulk of the dissertation to features on definite (semantic type *e*) elements.

[20] I thank Verena Hehl for providing a judgment for (39). In this thesis I generally restrict my attention to cases where gender has semantic consequences, and don't analyze non-semantic ('idiosyncratic') gender.

On a prominent reading, (39) does not just say that I didn't talk to female-identified professors except for the German one, but also to no male-identified ones. The gender feature on 'professor' does not project to alternatives.[21] (40) from Hebrew makes the same point.

(40) *Hebrew Fake Gender*

a. *Context: A singing competition is on TV. There are 10 contestants and only 5 qualify to the next round. Alex, Benny and myself are watching the show, and each of us has a favorite contestant; mine is **female**, the other two are **male**. At the end of the show when the judges announce their decision, it turned out that...*

b. rak [ha-mitmoded-**et** še ANI ahav-ti] alt-a šalav
only [the-contestant-**FEM**.SG that $I_{(F)}$ liked-1SG] went.up-FEM.3.SG level
*'Only the contestant that I_F liked, who is **female**, qualified'*

c. $\rightsquigarrow \forall x \in \{Alex, Benny\}$: *the contestant that x liked didn't qualify.*

By now it is easy to see why these examples are problematic for the simple algorithm of alternative generation. Take (39) as a representative, and assume that its syntactic analysis is analogous to number from (35): FEM attaches on top of DPs (41a) and restricts the denotation to individuals identified as females. The simple algorithm produces the too-restrictive alternatives in (41b), whereas the desired alternatives look something like (41c), either by deletion of the feature or setting it to a different value.

(41) a. LF: Only $\left[_S$ Maria talked to $[_{GENDP}$ **FEM** $[_{DP}$ the [German$_F$ professor]]]$\right]$

b. $[\![S]\!]_f$ = {*Mary talked to the P professor, who is **female*** | *P* is alt' to *German*}

c. $[\![S]\!]_{f_{WK}}$ = {*Mary talked to the (**male or female**) P professor* | *P* is alt' to *German*}

In (42) I give an example with variable binding into the scope of FEM, and its analysis is schematized in (43). It has the same logical structure as we had in (36)-(37), this time with gender.

[21] Yatsushiro & Sauerland (2006) notice that in German and other languages, a phrase like *the most popular politician.***FEM** can convey that the politician is more popular than male politicians too. That is, FEM doesn't restrict the comparison class to female politicians. This looks suspiciously similar to (39)-(40). Maybe the same mechanism in charge of ignoring feature contributions in focus alternatives is behind this case too, given that it has been argued that fixing the comparison class for superlative quantification is constrained by focus (Rooth 1985; Heim 1999; Tomaszewicz 2015, a.o.). I won't discuss superlatives further, though.

This sentence is fine even if the salient alternatives to Yosi had **male** head teachers in elementary school.[22]

(42) *Hebrew*

rak YOSI$_1$ histader im [ha-mexanex-**et** šel-o$_1$ ba-yesodi].
only YOSI got.along with [the-head.teacher-**FEM.SG** of-his$_1$ in-elementary.school]
'Only Yosi got along with his head teacher in elementary school (his was female).'

(43) a. LF: only $\left[_S \text{ Yosi}_F\ \lambda_1\ t_1 \text{ got along with } [_{\text{GENDP}} \textbf{FEM}\ [_{\text{DP}} \text{the [head-teacher of his}_1]]]\right]$

b. $[\![S]\!]_{f_{WK}}$ = {*x got along with x's (**male or female**) head-teacher* | *x* is an individual}

2.2.3 Features on bound pronouns

The final case I discuss in this section is ϕ-features on bound-variable ('sloppy') pronouns. This is perhaps the most famous case of fake features, and it can be described as another instance of the puzzle of weak projection (although it is not always recognized as such). ϕ-features on pronouns bound by a focused phrase can be ignored in association with *only* and other focus-sensitive operators (Kratzer 1998; Heim 2008, a.o.).[23] I concentrate on person and gender features.

(44) Only I did **my** homework

⇝ *sloppy reading:* No one other than me did their own homework.

⇝ *strict reading:* No one other than me did my homework.

(45) Only Mary did **her** homework.

⇝ *sloppy reading:* No one other than Mary, **including non-females**, did their homework.

⇝ *strict reading:* No one other than Mary did *x*'s homework.

[22] Giorgos Spathas (p.c.) tells me that similar examples to (42) also work fine in Greek, with the noun *jatro* 'doctor' (see also section 2.3.4).

[23] There are other constructions that have been grouped under the umbrella of fake features on bound pronouns. One particularly interesting configuration is fake indexicals in relative clauses (Partee 1989), which is the topic of chapter 5. For now I concentrate on focus constructions like (44). Another relevant construction is dependent plurals (plurals on pronouns bound by distributed quantifiers), as in *the boys each thought **they** were the only boy in the room* (see Rullmann 2004; Heim 2008). For a theory about these compatible with my general claims, see Sudo (2014).

On common assumptions about the Logical Form of pronouns, this is the same phenomenon as in the previous cases of weak projection: A variable in the scope of a ϕ-feature is bound from outside by the focused phrase and the feature is observed not to project to focus alternatives.

(46) a. LF for the sloppy reading of (45):

Only [Mary$_F$ λ_1 t_1 did [FEM pro_1]'s homework]

b. Desired alternatives:

{ John λ_1 t_1 did [~~FEM~~ pro_1]'s homework,

Bill λ_1 t_1 did [~~FEM~~ pro_1]'s homework, ... }

c. $[\![S]\!]_{f_{WK}}$ = {*x did x's homework* | *x* is an individual}

(47) a. LF for the sloppy reading of (44):

Only [I$_F$ λ_1 t_1 did [**1SG** pro_1]'s homework]

b. Desired alternatives:

{ John λ_1 t_1 did [~~**1SG**~~ pro_1]'s homework,

Mary λ_1 t_1 did [~~**1SG**~~ pro_1]'s homework, ... }

c. $[\![S]\!]_{f_{WK}}$ = {*x did x's homework* | *x* is an individual}

The Minimal Pronoun+Feature Transmission Approach

Although many authors indeed take Fake Features on bound pronouns to be explained in terms of disappearance of an active feature from alternative (Spathas 2010; Sudo 2012; Jacobson 2012; Sauerland 2013; Bassi & Longenbaugh 2018; Bruening 2019; Sudo & Spathas 2020, a.o.), there is also a quite different, syntactic account of these cases. This is the MINIMAL PRONOUN approach (Kratzer 1998, 2009; von Stechow 2003; Heim 2008; Wurmbrand 2017a, a.o.). Minimal Pronouns conjectures that fake features on bound pronouns are not interpreted but are mere agreement markers, reflecting post-syntactic agreement between binders and bindees. On Minimal Pronouns, what is marked FEM and 1SG in (46)-(47) are not present at LF at all, neither in the alternatives nor in the generated prejacent itself.

In this thesis I will develop an account that takes an intermediate position between Weak Projection and Minimal Pronouns, combining insights of both approaches. I agree with Minimal Pronoun theories that variables are featureless at LF and that they acquire their features late in the syntactic derivation (so no 'generation + deletion' is involved). However, I deny that the process by which they do so is mediated by agreement. Instead I will propose that *all* variables – referential and bound alike – are generated without ϕ-features, and those are acquired late based on what the variable refers to in context.

There are (at least) two persisting problems with Minimal Pronouns, on top of not being able to extend to the cases of fakeness on full DPs discussed in the previous two subsections (which did not concern features *on* bound pronouns at all). One problem is its reliance on a process of agreement. If a morphosyntactic agreement process is involved here, it is a quite atypical one. Agreement relationships are usually very local, but the agreement between binders and bindees can span arbitrary long distances (48a), is not constrained by c-command (48b), and arguably not even LF C-command (48c).[24]

(48) a. Only *I* bought a gift for the person who hosted **me** for thanksgiving. (✓ sloppy *me*)

b. Only MY mother will pick **me** up.
(✓ sloppy; Sudo 2012:p.153, attributed to Irene Heim)

c. Only if *SUE* has trouble in school I would help **her**.
(✓ sloppy, mixed gender; based on Tomioka 1999:238)

Second, the realization of standard agreement dependencies is subject to significant cross-linguistic variation, but it is my impression that the obligatoriness of feature matching in binder-bindee dependencies like the above is much more invariant. I am not aware of a language that has configurations of focus-association similar to (46) and (47), but uses e.g. a 3rd person pronoun

[24] In (48c) the antecedent of the bound pronoun is embedded in an *if*-clause and the bound pronoun is in the main clause, so if the former were to raise at LF to a position that c-command the latter, it would have to do so while violating established syntactic restrictions on LF movement. This is what is meant in saying that (48c) shows that binder-bindee agreement is arguably not even constrained by LF c-command. See McKillen (2016) and Bassi & Longenbaugh (2018) for more discussion on this kind of data, including parallel examples with person features.

instead of a 1st person when its binder is 1st person.[25]

2.2.4 Taking Stock

Taking stock, we have seen evidence that suggests that a simple substitution-algorithm for deriving focus alternatives is too strong, and that ϕ-features are systematically discounted from the requirement that anything that isn't F-marked must project to all alternatives. We now have an unnatural-looking generalization about how focus alternatives are derived, repeated from (30).

(49) **Deriving Alternatives** (repeated from (30)): To derive $[\![S]\!]_{f_{WK}}$, the focus-alternative set of S, do the following:

a. Replace F-marked constituents in S with an element of the same type.

b. Leave everything else the same except for information encoded by grammatical ϕ-features (number, gender, person). ϕ-features don't have to contribute their expected semantics to alternatives.

[25] Excluding possible but irrelevant languages which would always use a dedicated form of a pronoun to express bound readings.

2.3 Fake Features in Ellipsis

Until now I have used Sauerland's (2013) theory of Weak Presupposition Projection (pursued also in McKillen 2016; Bruening 2019; Sudo & Spathas 2020) as a way to describe the puzzle. It asserts that ϕ-features do not necessarily contribute their meaning to focus alternatives. In this section I raise challenges for this theory. Specifically, I will present data about acceptable cases of Fake Features in ellipsis which the Weak Presupposition Approach undergenerates.

2.3.1 Pronouns

As a background, I first present well-known (and still unproblematic) data regarding features on bound pronouns. Bound pronouns, for many speakers, allow mismatch in features:

(50) *ϕ-mismatches on bound pronouns in ellipsis*:

a. I did **my** homework, but you didn't [$_{VP_\triangle}$ ~~do **your** homework~~].

b. Mary did **her** homework, but Jack didn't [$_{VP_\triangle}$ ~~do **his** homework~~].

These cases are predicted by the Weak Projection theory of features. Recall that the Focus Parallelism condition on ellipsis requires that an elided VP be contained in a constituent S that appropriately contrasts with an antecedent A, i.e such that $[\![\text{A}]\!] \in [\![\text{S}]\!]_f$ and $[\![\text{A}]\!] \neq [\![\text{S}]\!]$ (see section 2.1.3). Suppose we update Approriate Contrast to refer to the weaker $[\![\text{S}]\!]_{f_{\text{WK}}}$ rather than $[\![\text{S}]\!]_f$. Then, taking S to be the whole second clause of (54a) and A to be the whole first clause, the two satisfy Appropriate Contrast:

(51) *In (50a)*:

a. $[\![_{\text{A}}\ \text{I}\ [\lambda_i\ \text{t}_i\ [_{\text{VP}}\ \text{did}\ \textit{my}_i\ \text{homework}]]]\!] \in [\![_{\text{S}}\ \text{YOU}_F\ [\lambda_i\ \text{t}_i\ \text{did}\ [_{\text{VP}}\ \text{do}\ \textit{your}_i\text{'s homework}]]]\!]_{f_{\text{WK}}}$

b. $[\![\text{A}]\!]$ = *the speaker did the speaker's homework*

c. $[\![\text{S}]\!]_{f_{\text{WK}}}$ = {*x did x's homework* | *x* is an individual}

Even though *your* in the ellipsis clause carries the presupposition that its semantic value is the

addressee, the hypothesis that this presupposition can be absent in alternatives will derive the possibility of ellipsis.[26]

2.3.2 A note about the empirical picture

I make here a note about the data which will be relevant throughout this section. Since the earliest works on sloppy readings under ellipsis, it has been known that pronominal mismatches have a somewhat intermediate grammaticality status. Not all speakers accept them, as mentioned already by Grinder & Postal (1971) and confirmed by Lakoff (1973); Dahl (1973); Sag (1976); Kitagawa (1991); Hestvik (1995), among others. Grinder & Postal (1971:p.30), for instance, report that while virtually all speakers accept a sloppy reading in (52a), many speakers don't in (52b).

(52) Grinder & Postal (1971:p.30)

a. John painted his house, and so did his father. (✓sloppy)

b. John painted his house, and so did his mother. (% sloppy)

Based on such facts, the authors cited above converge on the conclusion that there are (at least) two populations with respect to accepting ellipsis mismatches in pronominal form.

This complicates the empirical picture, but what should we infer from that? the fact that a subset of speakers, call them Population 1, accept form mismatches under sloppy ellipsis suggests that grammar should in principle generate them. The fact that many speakers, call them Population 2, do not tolerate them, or at least highly disprefer them, could be interpreted in several ways. Following Lakoff (1973), I suggest that those speakers have a stricter identity conditions for ellipsis; they require the semantic-identity condition (Focus Parallelism), but on top of that also require (or highly prefer) phonological identity between the antecedent clause and the ellipsis clause had there not been ellipsis.

[26]There is another route to deriving ellipsis here which does not even require activating the Weak Presupposition Projection hypothesis. Namely, to suppose that the bound pronoun which is marked *your* in the ellipsis site is not generated with features at all and is a bare variable at LF (we cannot tell, because the VP is elided). In that case, ⟦A⟧ will be a member of $⟦S⟧_f$, not only $⟦S⟧_{f_{wk}}$. This move works but will be of little for the other cases I discuss presently.

In what follows, I will mostly concentrate on the population which accepts mismatches (although it is important to keep the judgment split in mind); as I show next, there are speakers who accept mismatches in cases that goes beyond the descriptive power of Weak Presupposition Projection theory. The two problematic cases involve mismatches on full nominals in argument positions. I first discuss number (section 2.3.3), then gender (section 2.3.4).

2.3.3 Number in full DPs (in argument positions)

I have found that for many speakers, VP-Ellipsis allows number mismatches on full (definite) DPs in argument positions, in a rather parallel fashion to the facts from association with focus in section 2.2.1. Specifically, when an elided VP contains a variable in a definite DP bound from outside the ellipsis, there can be number mismatches with the antecedent. This holds, for most speakers consulted, whether the antecedent contains plural and the ellipsis contains singular, or vice versa:

(53) *Context: Becca has 3 advisors, Eleanor has 2 advisors, Kelly has 1 advisor.* (=(36))

a. Becca$_2$ will meet with her$_2$ **advisors** tomorrow, and KELly$_1$ will [$_{VP}$ △] too.

△ = *meet with her$_1$* ***advisor*** *tomorrow*

b. Kelly$_1$ will meet with her$_1$ **advisor** tomorrow, and BECca$_1$ will [$_{VP}$ △] too.

△ = *meet with her$_1$* ***advisors*** *tomorrow*

(54) *Context: Officer Mary arrested 1 guy, Officer Kelly arrested 2 guys and Officer Becca arrested 3 guys. Which officers were humane during the arrests?* (=(37))

a. Becca$_2$ treated the **guys** she arrested with respect, and so did MAry$_1$ [$_{VP}$ △].

△ = *treat the* ***guy*** *she$_1$ arrested with respect*

b. Mary$_1$ treated the **guy** she arrested with respect, and so did BEcca$_2$ [$_{VP}$ △].

△ = *treat the* ***guys*** *she$_2$ arrested with respect*[27]

[27] One speaker reported to me that he accepts the mismatch in (54a) but not in (54b). He did not perceive such a contrast in (53), accepting both versions there. Perhaps he belongs to a variety of Population 2 (see above, section 2.3.2), although I have not systematically investigated whether there is a correlation with his judgments for form-mismatches on bound pronouns.

Moreover, just like in association with focus (cf. ex. (34)), such mismatches seem to be confined to information coming from the morphological feature; it is not possible to ignore the contribution of lexical numeral. For all speakers, (55) on the sloppy reading is only felicitous in contexts where the subject of the ellipsis clause has three advisors.[28]

(55) *Context as in (53)*

Becca will meet with her **three** advisors tomorrow, and KELly/Eleanor will [$_{VP}$ △] too.

Can all these facts be explained by Sauerland's (2013) analysis of weak number projection into alternatives? not without further stipulations. Consider first (54a), the easier case among the pair in (54). The Parallelism condition requires that an elided VP be contained in a constituent S that appropriately contrasts with an antecedent A, i.e such that $[\![A]\!] \in [\![S]\!]_f$ and $[\![A]\!] \neq [\![S]\!]$. Take S to be the whole second clause of (54a), and our A is the first clause. These do satisfy Appropriate Contrast, if: (i) Appropriate Contrast is defined over $[\![S]\!]_{f_{WK}}$ instead of $[\![S]\!]_f$, as already suggested above (section 2.3.1) ; and (ii) PL is semantically vacuous. To verify this, examine (56) which shows the relevant LF and semantic representations of (54a).

(56) a. $\Big[_A$ Becca λ_2 t_2 treated [$_{\text{NUMP}}$ **PL** [$_{\text{DP}}$ the guys she$_2$ arrested]] with respect$\Big]$, and
$\Big[_S$ Mary$_F$ λ_1 t_1 treated [$_{\text{NUMP}}$ **SING** [$_{\text{DP}}$ the guy she$_1$ arrested]] with respect$\Big]$

b. $[\![A]\!]$ = *Becca treated the guys Becca arrested with respect*

c. $[\![S]\!]_{f_{WK}}$ = {*x treated the guy or guys x arrested with respect* | *x* is an individual} (=(38c))

On the weak-projection hypothesis SING has no contribution in focus alternatives and therefore DP in S is compatible with plural reference across its alternatives. If PL does not hard-code a semantic presupposition of non-atomic reference, then indeed $[\![A]\!] \in [\![S]\!]_{f_{WK}}$, and so S appropriately contrasts with A and ellipsis is correctly licensed (if PL *does* contribute a semantic presupposition, then $[\![A]\!] \notin [\![S]\!]_{f_{WK}}$ because the proposition $[\![A]\!]$ now carries a presupposition that is absent in the relevant member of $[\![S]\!]_{f_{WK}}$).

[28]The contrast between (53a) and (55) have been confirmed by five speakers in an informal survey. Two speakers have reported that (53a) is not impeccable, but even they found it acceptable and perceive a clear contrast with (55). Again, I suspect this variability is related to the variability about form mismatches on pronouns.

Obviously, the issue is dissolved by taking PL to be semantically-vacuous. However, this is of little help for examples like (54b) in which SING and PL switch places. In (54b) it is SING that is in the A-clause. SING is assumed to impose a semantic atomicity-presupposition in its **ordinary** semantic value, see (57b); but on Sauerland's (2013) account there is no path into having this atomicity presupposition be present in the propositions in $[\![S]\!]_{f_{WK}}$, cf. (57c). We are driven to conclude that $[\![A]\!] \notin [\![S]\!]_{f_{WK}}$ here, and consequently ellipsis shouldn't be licensed, contrary to fact.

(57) a. $\Big[_A$ Marry λ_1 t_1 treated $[_{\text{NUMP}}$ SG $[_{\text{DP}}$ the guy she$_1$ arrested]] with respect$\Big]$, and
$\Big[_S$ Becca$_F$ λ_2 t_2 treated $[_{\text{NUMP}}$ PL $[_{\text{DP}}$ the guys she$_2$ arrested]] with respect$\Big]$

b. $[\![A]\!]$ = *Becca treated the guy or guys Becca arrested with respect*
Presupposition: Becca arrested exactly one guy

c. $[\![S]\!]_{f_{WK}}$ = {*x treated the guy or guys x arrested with respect* | *x* is an individual} (=(56c))

There are a number of imaginable fixes, but they go beyond the core of the Weak Projection hypothesis and require casting the theory quite differently. One option is to assume that in generating alternatives we can somehow replace feature values, rather than deleting them (or rendering them vacuous in some other way). I will not develop this possible emendation further. Rather, I suggest that these data indicate that the problem—and therefore the fix—is not with the theory of focus alternatives specifically, but concerns the syntax-semantics of the relevant feature more generally. My goal in the next chapters is to develop an account which overcomes the problem by claiming that neither SG nor PL are semantically interpreted in the usual technical sense at all, and therefore aren't imposing presuppositional restrictions at the level that serves as input to the calculation of focus alternatives and Parallelism. There is no semantic presupposition in (57b) that disrupts ellipsis licensing in the first place.[29]

[29]Modeling Population 2 will still require a phonological identity condition on top of Focus Parallelism.

2.3.4 Gender

Turning to gender now, there are analogous puzzles. In Hebrew, post-subject ellipsis allows gender mismatch with definite nominals. In (58), a pronoun is bound by the subject and is dominated by a DP with FEM marking. But ellipsis is fine in (58) even though it is entailed by context that the antecedent DP is male.[30]

(58) *Context: Yosi has a **female** head-teacher, Avi has a **male** one.*

Yosi$_1$ histader im [ha-mexanex-**et** šel-o$_1$]. Avi lo △. (*Hebrew*)
Yosi got.along with [the-head.teacher-**FEM.SG** of-his$_1$]. Avi not △.
⇝ △ : *got along with his head teacher, who is **male**.*

In German, at least for some speakers, there are parallel facts, although the effect seems weaker. One consultant reports she finds the ellipsis in (59) acceptable (albeit 'confusing').[31]

(59) *Established context: Paula has a* FEM *English teacher; Maria has a* MASC *English teacher.*

Paula mag ihr-e Englischlehrer-**in**, aber Maria nicht △. (*German*)
Paula likes her-FEM English.teacher-**FEM**, but Maria not △.
⇝/? △ : *likes her head teacher, who is **male**.*

The behavior of the speakers who accept these examples is again not captured by Weak Projection theory. On that theory the focus semantic value of (58) is in (60a) (either because MASC in the ellipsis site is semantically vacuous, or because it is deleted from alternatives). The antecedent clause's meaning in (60a), however, should carry a presupposition that Yosi's head teacher is identified as female, which then should violate Appropriate Contrast and should disrupt ellipsis.

(60) In (58):

a. $[\![S]\!]_{f_{WK}}$ = {*x got along with x's (**male/female**) head-teacher* | *x* an individual} (= 43b))

[30] This judgment has been confirmed by five out of the six people I consulted with. One speaker found it marginal. The judgment perhaps gets easier if what is stated as the 'context' part in (58) is uttered by the speaker, as part of one discourse, immediately before the target sentence. I reemphasize my conjecture that the marginality of form-mismatches in ellipsis for some speakers is a general phenomenon, section 2.3.2.

[31] Another speaker did not accept (59). Perhaps this speaker belongs to Population 2, or this particular noun in German disfavors a mismatch quite generally (see below for a discussion of different classes of nouns w.r.t gender mismatch).

b. ⟦A⟧ = *Yosi got along with his head teacher*

Presupposition: Yosi's head teacher is female.

Note that, in a fashion similar to the number mismatch cases, Weak Projection theory does account for cases of mismatches in the other direction (which are also possible for many speakers), i.e. where the antecedent clause hosts MASC, on the assumption that MASC is semantically vacuous (imposes no presuppositions).

2.3.5 Other cases of fake gender features?

In this subsection I will take a little detour to review and evaluate other potential cases of fake gender on nominals discussed in recent literature. A reader who is only interested in the main plot is invited to skip to section 2.4.

There is by now a growing consensus that some nominals in some languages, in some configurations, allow gender mismatches in ellipsis (Bobaljik & Zocca 2011; Merchant 2014; Sudo & Spathas 2016, 2020 a.o.). For instance, Bobaljik & Zocca (2011:156, ex:28) present examples from Brazilian Portuguese where the targeted noun is in a predicative position.

(61) *Gender Mismatch in Predicative Positions* Bobaljik & Zocca (2011:156, ex:28)

A Marta é médic-**a** e o Pedro também é ~~medic-**o**~~. (*Brazilian Portuguese*)
The Marta is doctor-FEM and the Pedro also is ~~doctor-MASC~~
'Marta is a doctor, and Pedro is too.'

At the same time, not all gendered nominals are alike. As documented in Bobaljik & Zocca (2011); Merchant (2014); Sudo & Spathas (2020), languages generally have three classes of nouns with respect to their behavior in ellipsis: one class allows gender mismatches in either direction (like Brazilian Portuguese *médic-o/a*, Greek *jatro*, and Hebrew *mexanex/-et*), another allows them in neither direction (e.g. *prince-princess* in English), and a third allows mismatches when the noun is male-denoting in the antecedent and female-denoting in the ellipsis site but not vice versa. Furthermore, the classes contain different sets of nouns cross linguistically, and, to the best of my

knowledge, even within a language there aren't reliable predictors (semantic or morphological) about which noun belongs to which class (with few exceptions, e.g. nobility terms apparently invariably belong to the second class; see the authors above for discussion). The most interesting class for my purposes, and the one I focus on, is the first class which allows mismatch in either direction.

Gender mismatches in predicative positions, as in (61) and most of the reported data in the literature I am aware of about acceptable gender mismatches, does not however present a puzzle on the same scale as the fake gender cases discussed in the earlier subsection (such as (58)). (61) is amenable to an analysis entirely compatible with standard assumptions about the syntax-semantics of ellipsis and/or alternatives. Namely, gender assignment on the predicate here could result from post-syntactic agreement with the subject, under natural assumptions about locality of agreement. In that case the two nouns could be semantically unspecified for gender at the point relevant for the identity condition on ellipsis; this is in fact the view adopted in Bobaljik & Zocca (2011) to explain these cases.

We should therefore concentrate on the behavior in ellipsis of these nouns in argument positions, which is what was done in the previous subsection from Hebrew and German. Those require a different mechanism, on the assumption that—as opposed to (61)—no relevant syntactic agreement is established between the offending DP and a local target.[32]

Sudo & Spathas (2020) do present data from Greek that appears more relevant. They give (62), as case of nominal ellipsis in Greek. The noun *jatro* 'doctor' can be interpreted as male-referring in the antecedent and female-referring in the ellipsis (62a), or vice versa (62b).[33]

(62) *Gender mismatch in Greek nominal ellipsis (Sudo & Spathas 2020:12)*

[32] From a preliminary inquiry, two speakers of Brazilian Portuguese do not accept gender mismatch with definite *médic-o/a* analogous to (58). This could suggest that they are both in Population 2, a suggestion which gains some support by the observation that at least one of them also does not accept the mismatch in the other way (MASC in the antecedent, FEM in the ellipsis), and that she finds mismatches of form on bound pronouns somewhat marginal as well. A more thorough investigation is thus required.

[33] Sudo & Spathas (2020) report that the judgments regarding these were not very stable across speakers. This is in line with the general theme that feature-mismatches in ellipsis have an intermediate status in grammar, or show some population split.

a. O Petros episkefthike **enan** jatro stin veria, ke **mia** ~~jatro~~ stin katerini
the Petros visited one.M doctor in Veria, and one.F ~~doctor~~ in Katerini
'Petros visited a male doctor of his in Veria, and a female doctor in Katerini.'

b. O Petros episkefthike **mia** jatro stin veria, ke **enan** ~~jatro~~ stin katerini
the Petros visited one.F doctor in Veria, and one.M ~~doctor~~ in Katerini
'Petros visited a female doctor of his in Veria, and a male doctor in Katerini.'

Here too, however, a local-agreement analysis could on certain assumptions also explain these data. Since this is nominal ellipsis, it is possible that the semantic locus of the feature is at the DP level, outside the ellipsis site. If so, and if *jatro* 'doctor' starts a derivation without underlying gender, then these cases are accounted for without assuming something special about ellipsis and/or focus alternatives.[34]

It is therefore better to inspect structures in which a larger syntactic chunk than just the noun undergoes ellipsis. As a first piece of evidence, Giorgos Spathas (p.c.) reports that a case structurally parallel to (58) (post-subject ellipsis) in Greek with *jatro* does make the gender mismatch possible, but more empirical work is required.[35]

2.4 Why only ϕ?

In this final section of the current chapter, I address the question of whether the puzzling behavior of ϕ-features thus far reviewed is shared by other elements that have a presuppositional semantic status.

As mentioned earlier, Sauerland suggested that the behavior of ϕ-features is just a special case

[34] Sudo & Spathas (2020) note that all nouns with the behavior of (62) in Greek are Epicene, i.e. morphologically identical whether referring to female or male doctors, and the grammatical reflex of the presence of gender is only on the agreeing adjective.

[35] One additional note is in order. In presenting the challenging ellipsis data in section 2.3.4, I have only discussed gender in *definite* nominals (just like for the case of number). For other types of nominals in argument positions, in particular indefinites, the picture is even more murky empirically, as well as unclear theoretically (see also fn. 19). Whether gender features on indefinites are presuppositional in the first place, and whether they are born on definites in the same way as gender on definites, are open questions (for a relevant discussion see section 5.2.2 in Sudo & Spathas 2020). I will not deal with these issues in this thesis, and the analysis I put forth in the next chapters will therefore only deal with gender in/on referential phrases (type *e*). A more serious research (both empirically and theoretically) regarding gender on indefinites and its behavior in ellipsis and focus will have to be left for a future occasion.

of so-called PURE-PRESUPPOSITION TRIGGERS: elements that contribute nothing more than a presupposition to the semantics (in particular, nothing to the assertive dimension). The hypothesis is stated in (63).

(63) The pure-triggers hypothesis (Sauerland 2013; McKillen 2016; Bruening 2019):
Pure-presupposition triggers do not have to contribute their presuppositions to focus alternatives.

In this section I criticize (63), in the service of motivating my own proposal in the next chapter. The criticism is that there are elements that arguably fall under the rubric of 'pure-presupposition triggers', yet they always strongly project to alternatives; I discuss *again*, *also* and *already*. The result of the investigation will imply that the Weak Projection analysis—even if it were to account for the ellipsis data from the previous section (2.3) by some other means—must apparently remain confined to ϕ-features specifically, and this raises the question why that should be the case.

Before I continue, let me be a bit more precise about what 'pure-presupposition' trigger means. It means a restricted-identity function. The entries below illustrate, and are the ones that are interpretable in the LFs I have provided so far.[36] A feature takes an individual-denoting argument, checks it satisfies the presupposition of the feature, and if it does returns that individual.

(64) *ϕ-features as pure-presupposition triggers*

a. $[\![\text{1ST}]\!]^c = \lambda x : x \text{ includes the speaker in } c.\ x$

b. $[\![\text{SG}]\!] = \lambda x : x \text{ is an atomic individual.}\ x$

c. $[\![\text{FEM}]\!] = \lambda x : x \text{ is identified as female.}\ x$

Somewhat informally, an item is purely presuppositional if removing it from the structure would not affect the assertive component of the meaning of its LF sister.

[36] Notation: The part between the colon and the dot defines the domain of the function described by the λ-term and is meant to represent presuppositional information.

2.4.1 Pure triggers that strongly project

Again

McKillen (2016) already noticed that *again* raises a challenge to the hypothesis in (63). *Again* is a pure trigger, at least this seems so based on raw intuition: *again* only contributes a presupposition, and does not affect the assertive component of its scope. Both s *John is/isn't here again* presuppose that John was here before, but are felt to assert exactly what they would assert without *again*. Thus a suitable lexical entry is given below:[37]

(65) a. 'S again' uttered at time t presupposes that 'S' was true before t. Asserts: 'S'.

b. $[\![\textbf{again}]\!]^{g,t} = \lambda S_{\langle i,t\rangle}$: there is a time $t' < t$ such that $S(t')$ is true. S

But *again* has to project to all alternatives, as shown below. In (66b) I provide a context for which focus-association with *again* would be predicted felicitous if *again* showed weak projection. The fact that the result is infelicitous suggests that *again* strongly projects.

(66) *The pure-presupposition trigger **again** projects its meaning to all focus alternatives:*

a. (Among Mary, John and Sue,) only Mary_F is talking to Ed **again**.

$\rightsquigarrow$ *All (salient) alternatives to Mary talked to Ed before.*

b. Context suitable for weak projection: Mary talked to Ed before and she is talking to him now. John and Sue aren't talking to Ed now, and we don't know whether they did before.

only Mary_F is talking to Ed again.

Deleting *again* from ellipsis sites is also not tolerated, as (67) shows.

(67) *Context: John has been late to the morning meeting before. Bill never has (and knows he hasn't), but today he knows he'll be late.*

John_1 said that he_1 will be late **again** this morning.

*Bill_2 did ~~say that he_2 will be late to the meeting this morning~~, too.

[37] Bacovcin *et al.* (2018) conclude based on experimental evidence that "the content of the presupposition of *again* is entirely backgrounded".

Also, too

Additive particles like *also* and *too* are pure-presupposition triggers, inducing the presupposition that their the focus value of their scope contain a (contextually salient; Kripke 2009) true alternative, and otherwise contributes nothing to assertion. But, *also/too* must project strongly to alternatives. Again I provide in (68c) a context that on a weak-projection behavior would make the sentence felicitous, contrary to facts.

(68) *The pure-presupposition trigger* ***also*** *projects its meaning to all focus alternatives:*

a. $[\![\textbf{Also}]\!]^g = \lambda S_{\langle t \rangle}$: there is some S' in the focus alt' of S and S' is true. S

b. (Among Mary, John and Sue,) only MARY will **also** talk to a GERMAN professor

(**also** associates with GERMAN)

$\rightsquigarrow$ *All alternatives to Mary will talk to some professor other than a German one.*

c. Context suitable for weak projection: Mary will talk to a Japanese professor and also to a German professor. We know for sure that Sue and John will not talk to a German professor, but for all we know they might or might not talk to some other professor. So...

\# Only MARY will **also** talk to a GERMAN professor.

The oddness of (68c) is explained if the sentence is predicted to presuppose that both Sue and John will talk to some non-German professor; but the context explicitly conveyed ignorance about this matter.

Already

Already could be considered a pure presupposition trigger as well (Löbner 1989; Ippolito 2007):

(69) John has started cooking **already**.

a. Presupposition: John wasn't (necessarily) expected to be cooking by now.

b. Assertion: John has started cooking.

c. $[\![\textbf{already}]\!]^{g,t} = \lambda S_{\langle i,t\rangle}$: it wasn't (necessarily) expected that $S(t)$ is true. S [38]

But *already* strongly projects its meaning to all focus alternatives:

(70) *The pure-presupposition trigger **already** projects its meaning to all focus alternatives:*

a. (Among Mary, John and Sue,) only John$_F$ has **already** started cooking.

$\rightsquigarrow$ *All (salient) alternatives to John are such that they weren't (necessarily) expected to be cooking by now.*

b. Context suitable for weak projection: Bill was expected to be cooking by now. John wasn't, but he started nevertheless.

only John$_F$ has **already** started cooking.

[38] This toy denotation is a massive simplification, which I'm hoping does not affect the argument. See Ippolito 2007 for extensive discussion and details.

THIS PAGE INTENTIONALLY LEFT BLANK

3

Valuation-from-Context theory of fake features

> *The proposition communicates to us a state of affairs, therefore it must be* essentially *connected with the state of affairs. And the connexion is, in fact, that it is its logical picture.*
>
> – Ludwig Wittgenstein, *Tractatus Logico-Philosophicus, 4.03*

Abstract

In this chapter I start developing the Valuation-from-Context framework. I show how it works to account for a portion of the puzzling data from the previous chapter, without having to complicate the mechanism that generates focus-alternatives. The key idea is that ϕ-features are inserted into the syntax rather than being base-generated.

3.1 A fresh perspective on fakeness

In the previous chapter we accrued challenges to the pure-presupposition theory of Weak Projection. That theory stipulated that ϕ-features are special (pure-)presupposition triggers: ones whose contribution can be absent in focus alternatives. Empirically, it ran into difficulties accounting for some of the ellipsis facts I presented in section 2.3; it was insufficient to explain mismatches where the antecedent has a strong gender/number feature and the ellipsis a weak one. In light of the results in section 2.4.1, it also requires the stipulation to be specifically about ϕ-features and not other presupposition triggers, a split which ideally should be reduced to something more principled once we have an insight on what is special about ϕ-featural information. Is there a way to understand the phenomenon without having to assume different classes of presuppositions with respect to their projection behavior in focus alternatives? In this thesis I defend a 'yes' answer to this question.

The proposal I advance is that ϕ-features on pronouns and definite DPs are not presupposition triggers to begin with, hence they don't have a 'projection' property at all. They do come to be associated with backgrounded information, very much like real presupposition triggers, but they do so via a grammatical process fundamentally different from imposing genuine presuppositions. In fact, I'll propose, such ϕ-feature do not technically have semantic denotations at all. Rather, they are inserted into the syntax of a pronominal or full DP based on the properties of the DP referent, by an operation I call VALUATION FROM CONTEXT.[1] The purpose of this chapter is to start developing this idea and show its merits in accounting for some of the puzzling facts. The next chapter will extend the theory to cover more data.

Before I present the details of the proposal, I would like to spend the rest of this introduction section to explicate the intuition that underlies it. That intuition has been partly expressed already in Irene Heim's paper on features on bound pronouns (Heim 2008). Heim remarked that ϕ-features on pronouns aren't really needed for interpretation, given their pure-presuppositional semantics, but they *are* required to appear in the structure, being "essentially a morphological requirement"

[1] I borrow the term 'Valuation from Context' from Kučerová (2018), who uses it in a sense not unlike the one intended here (though for an unrelated set of data and with a different implementation).

(Heim 2008:p.48). She used that idea to motivate an agreement rule that is responsible for spelling out the binder's features on any featureless variable it binds.

I do not adopt an agreement rule since it does not extend to fake features on full DPs (and for other reasons I will elaborate on later), but I adopt the idea that the purely formal status of ϕ-features may be relevant to why they can be fake. Namely, even though **semantically** ϕ-features appear to share with adverbials like *again*, *also* and *already* the property of encoding presuppositional information, ϕ-features cross-linguistically are part of a purely formal, **morpho-syntactic** system which excludes those adverbials. In many languages number, gender and person on pronouns and nouns determine agreement on verbs and adjectives, whereas agreement in e.g. 'again' is as far as I know unheard of; at least it is very rare. ϕ-features can also be distinguished from other lexical content morphologically: it is quite common for ϕ-features, cross-linguistically, to be expressed as affixes rather than free morphemes, to participate in productive inflectional paradigms, and to undergo processes such as allomorphy; whereas e.g. *again*, *also*, and other lexical content is far more morphologically invariant. Certain kinds of semantic information—including ϕ-information—drive morphological and morphosyntactic processes cross-linguistically, while others do not (*again*, *also*).[2]

But why should all this matter for fakeness in ellipsis and focus? building on Heim's intuition, I propose that the reason that (semantically-contentful) ϕ-features specifically can be ignored in focus alternatives and in ellipsis is that they can be derived in the syntax rather than be base-generated. ϕ-features don't have to be interpreted in the first place; but even if they aren't interpreted, morphosyntax independently requires ϕ-features to be expressed on pronouns and nominals whenever the relevant properties (number, gender, person) of the referent are known. Fakeness is thus not a result of an element that has been base-generated with encoded meaning and somehow

[2]This of course begs the question of why it is that only some conceptual information seem to be relevant to grammar. This is an age-old question in the theory of syntactic features and I don't have much to offer by way of answer (for relevant discussion see e.g. Adger & Svenonius 2011; Panagiotidis 2021). One would hope that the perspective expressed here would predict that other semantically grounded information that traffics in morphology and/or syntax, e.g. Tense, could be ignored in ellipsis and in focus association in the same manner to nominal features. However, there are independent confounds that have to do with the theory of the semantics of tense that create difficulties for probing the question of whether tense behaves the same. I leave a more thorough investigation for future work.

managed to have its contribution discharged in focus alternatives; quite the opposite: it is a result of an element in the structure that has not been base-generated with meaning in the first place but has been (therefore) forced to appear there because morphosyntax needs to realize it at all costs. The need can be satisfied by having the feature be inserted into the syntactic derivation at a late stage, bases on the semantic interpretation of the host DP, but without itself imposing a presuppositional requirement. The process by which they are inserted to syntax in this way is called **Valuation from Context**; and Context, as I propose later, is oblivious to focus alternatives.

In section 3.2 I lay out my assumption about the the syntax of ϕ-features on pronouns and nominals. Section 3.3 presents the theory of valuation-from-context and how it explains a (small) subset of the puzzling data from chapter 2. In section 3.4 I discuss the architectural consequences of the proposal, which necessitate the semantic module of grammar to feed insertion of syntactic and morphological material.

3.2 Assumptions about the representation of ϕ-features

Since my account takes seriously the featural status of ϕ information, it is necessary to be more explicit about how it is represented in syntax, which is what I turn to now.

Nominal and pronominal ϕ-features will both be taken here to be D(eterminer)-features; i.e. their syntactic locus is on D heads whether they are exponed on pronouns or in NPs. Some morpho-syntactic operation which I will neglect to specify is responsible for realizing D-features lower in the NP (see more on this below). I employ a [*attribute*:**value**] format for features.[3]

An illustration is provided in (71). An English DP like *the dog(s)* has its D head specified for number (71a); A German DP is also specified for gender (71b). For pronouns (71c,d), I assume as standard that they host a numerical index ('referential address') that represents the individual picked out by the pronouns; the index could be thought of as the root of the pronoun.

[3]My account is as far as I can see also compatible with other systems for modeling feature systems, in particular with the idea that features are privative and are organized in feature geometries (Harley & Ritter 2002). I also note that the choice of placing features on D is made for convenience; I could also assume that they are on a dedicated ϕ-node right above D (Sauerland 2003).

(71) *ϕ-features in DPs*

a. $[_{DP}\ [D_{[num:\mathbf{pl}]}\ \text{the}\ [_{NP}\ \ldots\ \sqrt{\text{DOG}}\ ]\ ]\ \rightarrow\ \left/\textit{the dogs}\right/$

b. $[_{DP}\ [D_{[num:\mathbf{sg},\ gen:\mathbf{f}]}\ \text{the}\ [_{NP}\ \ldots\ \sqrt{\text{PROFESSOR}}\ ]\ ]\ \rightarrow\ \left/\textit{die professor-in}\right/$[4] (German)

c. $[_{DP}\ [D_{[num:\mathbf{sg},\ gen:\mathbf{f},\ pers:\_]}\ 7\ ]\ \rightarrow\ \left/\textit{she/her}_7\right/$

d. $[_{DP}\ [D_{[num:\mathbf{sg},\ (gen:\mathbf{f/m/\_}),\ pers:\mathbf{1}]}\ 8\ ]\ \rightarrow\ \left/\textit{I/me/my}_8\right/$

That category D is where ϕ features sit in DPs is an oversimplification. Much recent literature has argued based on cross-linguistic evidence that number and gender in DPs are (or can be) introduced in dedicated functional heads along the nominal spine, and also that there is possibly more than one functional head for one and the same exponed feature (on these assumptions for number see Ritter 1995, Alexiadou 2011, Ouwayda 2014, Landau 2016b, a.o.; for gender see Steriopolo & Wiltschko 2010, Matushansky 2013, Pesetsky 2013, Rappaport 2013, Kučerová 2018, Sigurðsson 2019, a.o.). But there are indirect reasons to assume that all ϕ-features must anyway be represented on D, even if they start their life lower in the structure, so that this head will be 'ϕ-complete' for the purposes of syntactic dependencies between a DP and DP-external elements (see Danon 2011).[5] I thus continue to assume that (semantically-active) ϕ-features live on D, and won't bother with possible intermediate projections.[6]

In (71c) I chose to model a 3rd person pronoun in terms of absence of person value, in line with the common idea that 3rd person is the unmarked or 'elsewhere' person value. This claim should be understood as one regarding *semantic* unmarkedness, namely that 3rd person elements are com-

[4] I concentrate on semantically predictable gender features. 'Idiosyncratic' (non-semantic) gender plausibly comes lower in the structure, closer to the root (Kramer 2016a). It is worth noting that existing proposals which argue for more than one locus of gender inside DPs (see discussion immediately below) almost universally assume that the semantic gender is located higher than the idiosyncratic one, and many of them place the semantic gender on D as is done here (see Kramer 2016a:672-3 for a literature overview). Also for number, a similar structural asymmetry has been argued to exist where semantically-predictable number is higher than idiosyncratic number, see e.g. Alexiadou 2011, Ouwayda 2014, Landau 2016b, Kramer 2016b.

[5] Danon (2011) suggested a feature sharing operation (Pesetsky & Torrego 2007) which copies features onto D from below, as a remedy for the tension between the position of gender and number inside DP and their need to end up on D. In my proposal immediately below I adopt the idea that D has unvalued slots for ϕ, although I claim that they (at least sometimes) get valued from the extra-syntactic context (Kučerová 2018; Sigurðsson 2019).

[6] In order for the system I develop to work, really what is important is that all the features are located on a head whose maximal projection denotes an individual (semantic type *e*). So the account could in principle be fitted with the aforementioned works (without the assumption that features are always in D), if the semantics is adjusted appropriately.

patible with mixed-reference situations in quantificational contexts and in unknown-identity scenarios (Sauerland 2005 and references therein). In the domain of (animate) gender, a similar status of unmarked should be endowed to masculine, in those languages in which masculine is compatible with female-identified referents. In contemporary English, on the other hand, a *he*-pronoun in many dialects is restricted to male-identified referents; but there is still an unmarked option for gender which is spelled-out as a *they* pronoun (Bjorkman 2017). As for number on nouns and pronouns, it is not entirely obvious that there is a (semantically) unmarked value, but I will assume following the work of Sauerland (2003, 2005 et. seq.) and as mentioned previously that the plural is semantically unmarked.[7]

3.3 Valuation-from-Context theory of weak projection

I am now ready to outline the central proposal of this thesis. It is summarized as follows:

- **Feature-free base generation**: A syntactic derivation starts without specified values for ϕ-features on D-heads. Number/gender/person values do not constrain the interpretation of the DP in question.[8]
- **'Weak Projection' as an epiphenomenon**: Since a DP is not (semantically) constrained by ϕ information, it isn't across its focus alternatives either; there is nothing to project to begin with, and the puzzle of 'weak projection' to alternatives disappears.
- **Valuation from Context**: But ϕ-features on D-heads nevertheless have to be checked for value, because the morphosyntax of the language in question requires it. D-features are then valued by properties of the referent of the DP, i.e. according to rules that refer to the ordinary semantic value of the DP in the context of evaluation (ignoring focus semantic values).

[7] I return briefly to this at the end of 3.3.1. I wish to emphasize that issues of markedness are somewhat orthogonal to my main proposal, so I don't worry if this or that detail is wrong about how a feature structure is organized.

[8] I must qualify this claim. The general form of the arguments here merely allow me to make the weaker assertion that a derivation *can* start without specified D-features, and that D-features don't *have to* be semantically interpreted. But it is more instructive to pursue as a working assumption the stronger claim (which is consistent with my results) that D-features never start a derivation with (presuppositional) meanings. This is why the language in the text is the way it is.

- **'Interpreted' features as uninterpreted**: ϕ-features that are valued from context are not interpreted in the traditional sense, and hence do not affect the semantic value of sentences that host them. After being inserted to the syntax, they are interpreted only by PF (mediated by the morphology). The only difference between the customary labels 'interpretable' and 'uninterpretable' is whether the feature is valued from the context or from an agreement operation with some other element already present in the structure.

Valuation from Context requires an architecture of grammar in which properties of our conceptual representation determine pieces of a syntactic derivation and morphological realization. This is in contrast to mainstream Generative Grammar (Chomsky 1965 et seq.) but in line with other approaches to various phenomena at the interface between syntax and meaning, e.g. Bobaljik & Wurmbrand 2012; Jackendoff 1997; Seuren 1984. More specifically for the domain of ϕ-features, the claim that a semantic-pragmatic representation feeds rules for introducing features into syntax has been recently argued for by Matushansky 2013, Kučerová 2018, Sigurðsson 2019 (and to some extent Ackema & Neeleman 2018), who were each motivated by different data. Here I develop this idea for the data this thesis is concerned with. I also try to be more precise than previous work about what exactly is the notion of 'context' at play when we talk about valuation from context.

This chapter builds a partial account of the puzzling data from chapter (2), one which only allow us to capture Fake Features in focus alternatives in those structures that do not involve variable binding, as we will see. In the next chapter (4) I extend the model to formally deal with cases that involve fake features in variable binding as well (in both ellipsis and focus alternatives).

In section 3.3.1 I present how the system works for number, showing how it works to solve some of the puzzles of weak projection in the previous chapter. Section 3.3.2 does the same for gender. I discuss the architectural consequences of the proposal in section 3.4.

3.3.1 Number 'weak projection' explained

I repeat the example from 2.2.1, which showed that the restriction to atomic reference of *the German professor* isn't satisfied at the level of alternatives.

(72) Mary only talked to the GERMAN professor (based on von Heusinger 2007)

a. *Felicitous in this context*: there's one German professor, 2 Japanese professors, 5 American professors at the party. Mary didn't talk to any professor except for the German one (she might have talked to non-professors).

b. *The alternatives negated by (72)*: {Mary talked to the (two) Japanese professor**s**, Mary talked to the (five) American professor**s**}

My proposal is that singular number here is not base-generated and (therefore) isn't semantically interpreted in either the prejacent or its alternatives, but rather gets inserted into the DP based on the meaning of the DP, at a stage after semantic interpretation takes place. Once valued, features are interpreted only at PF. Here is how it works.

LF syntax and semantics

A syntactic derivation proceeds to yield the structure schematized in (73a), and I zoom in the full structure of the DP in (73b). The number feature on D is crucially not valued at this point.

(73) Generate LF:

a. only $\left[_{\text{TP}}\right.$ Mary talked to $[_{\text{DP}}$ the German$_F$ professor(s)$]$ $\left.\right]$

b. $[_{\text{DP}}$ $[\text{D}_{[num:\_\_]}$ the $[_{\text{NP}}$ $\sqrt{\text{GERMAN}}_F$ ... $\sqrt{\text{PROFESSOR}}$ $]$ $]$

As for the semantics of this structure, I borrow from Sauerland (2013) the idea that the denotation of DPs underdetermines whether the reference is atomic or not, see (74).

(74) $[\![_{\text{DP}}$ the German professor(s)$]\!]^C$ = the maximal contextually-salient individual containing one or more German professors

The semantic composition that yields (74) is given in detail in (75). NP predicates invariably range over both atomic and non-atomic individuals, in (75a). The determiner *the* maps an NP predicate to the maximal contextually salient individual (atomic or plural) that satisfies the predicate, presupposing there is one (cf. Link 1983):

(75) a. $[\![\text{professor(s)}]\!]^C = \lambda X.\ professor(X)$.[9]

b. $[\![\text{german professor(s)}]\!]^C = \lambda X.\ german(X) \wedge professor(X)$.

c. $[\![\text{the}]\!]^C = \lambda f \in D_{\langle e,t\rangle} : \textsc{max}_C(f)$ is defined. $\textsc{max}_C(f)$.

d. $\textsc{max}_C(f) =$ $\iota X[f(X)=1 \wedge X$ is salient in $C \wedge \forall Y$, if $f(Y)=1$ and Y is salient in C, then $Y \leq X]$;

Defined only if $\exists X[f(X)=1 \wedge X$ is salient in C $\wedge\ \forall Y$, if $f(Y)=1$ and Y is salient in C, then $Y \leq X]$.[10]

($\textsc{max}_C(f)$ is the maximal contextually-salient individual that makes f true, assuming there is one)

Since *German* is F-marked, the alternatives that the sister of *only* (the TP) in (73a) triggers are in (76). They are arrived at by a boring replacement procedure. At the moment, it does not actually matter if alternatives are syntactic objects as in (76) or the corresponding semantic objects.

(76) *Alternatives of the prejacent of only in (73a):*

a. Mary talked to [$_{\text{DP}}$ the German professor(s)],

b. Mary talked to [$_{\text{DP}}$ the Japanese professor(s)],

c. Mary talked to [$_{\text{DP}}$ the American professor(s)], ...

The hypothesis that NPs—whether 'plural' or 'singular'—uniformly have both atomic and non-atomic individuals in their extension is the reason that reference to non-atomic individuals is possible across the focus alternatives in (76). Indeed, in our context there is just one German professor but more than one Japanese and American ones.

Only conveys the negation of the alternatives in (76b-c), and these alternatives contain reference to non-atomic individuals. I assume that the negation of propositions with reference to non-atomic individuals obeys the Homogeneity property of definite descriptions (Schwarzschild 1994; Löbner

[9] $professor(X) = 1$ iff $\forall x \leq X, professor(x) = 1$ ($\leq$ is the part-whole relation).

[10] The part that says '$Y \leq X$' in (75d) can be replaced with '$Y = X$' with an equivalent result, as long as we buy the assumption that if a plural-individual is considered 'salient' in some context C, no (atomic or plural) proper sub-part of it is considered salient in that C. I think this assumption is defensible. I am making this remark because I anticipate a small technical problem that (75d) gives rise to in combination with a much later part of my proposal (defining local contexts for variables bound inside the NP argument of *the*), a problem that is fixed by the change to '=' that I was now describing. But since I won't go into detail about that technical problem, I will stay with (75d).

2000), so negating a plural definite ends up distributing the negation for each atomic individual. This accounts for the intuition in (72) that Mary didn't talk to any of the non-German professors. As von Heusinger (2007:501) observes, (72) is not compatible with a situation in which Mary talked to some of the American professors. In line with this observation, Bar-Lev (2020) provides evidence that when a plural definite is trapped in the scope of negation, it virtually always gives rise to this strong reading (¬∃, rather than the weaker ¬∀).

Number valuation

Now for the valuation part. I assume that a D-feature must be checked for a value, and it can get one based on the meaning-in-context of the DP.[11] I propose the valuation rule in (77) which is operative (in languages that mark a number distinction) when a structure is sent to a morpholosyntactc level of representation (see section 3.4 for the architecture).

(77) **Number valuation from context**:

Let DP be headed by $D_{[num:\_\_]}$, in a syntactic derivation S in context C.

a. If $[\![DP]\!]^C$ is an atomic individual, then $D_{[num:\underline{\textbf{sg}}]}$.

(=If the semantic value of DP in C is an atomic individual, then its head D is assigned the value **sg** for its number feature.)

b. Else, $D_{[num:\underline{\textbf{pl}}]}$.

Note that by treating plural as the elsewhere element, (77) comes down to adopting the hypothesis that plural is the semantically unmarked option (Sauerland *et al.* 2008; Alexiadou 2019, a.o.; see 3.3.1 below for more on this).

The context parameter C plays a key role. I understand C to be an object that represents all sorts of beliefs and perceptions that the speaker has about the context of utterance (in Heim's

[11] 'Idiosyncratic' values that come with the noun, I assume, can also value D-features. This is what is likely responsible for plural number in pluralia tantum structures, e.g. *these scissors are...*, and maybe also for singular number in mass nouns as in *this furniture is...*. The same goes for idiosyncratic gender. But in the semantically-predictable cases, valuation is from meaning-in-context.

(2007) terms, C stands for a *subjective* rather than *objective* context[12]). Valuation from Context rules are therefore instructions for the speaker to use information from his conception of the context as source for feature valuation.

Among other things, C records information about the salient individuals in discourse. Suppose C is such that the salient professors in it are: [Uli, Ayaka, Yasu, David, Elizabeth, Seth, Lisa, Beth]. Suppose Uli is the only German among them, Ayaka and Yasu are Japanese and the rest are American. For this context the condition in the first clause of (77) is satisfied—the maximal individual containing one or more German professor is one in number. So, the number feature on D gets valued for **sg**. The whole process is summarized in (78).

(78) a. Base-generation (=73b): $[_{DP}\ [D_{[num:\_\_]}$ the $[_{NP}\ \sqrt{\text{GERMAN}}_F$ ... $\sqrt{\text{PROFESSOR}}\ ]]\ ]$

b. Information from context: *There is exactly one German professor* (cf. 72)

c. After Valuation from Context: $[_{DP}\ [D_{[num:\underline{\mathbf{sg}}]}$ the $[_{NP}\ \sqrt{\text{GERMAN}}_F$ ... $\sqrt{\text{PROFESSOR}}\ ]]\ ]$

d. Phonology: /the GERman professor/[13]

If, in the same context, the structure that was realized was *the Japanese professor(s)*, it would be spelled out as plural because the context makes salient two Japanese professors, (79). But the focus alternatives still include the proposition about the (single) German professor.

(79) a. Base-generation: $[_{DP}\ [D_{[num:\_\_]}$ the $[_{NP}\ \sqrt{\text{JAPANESE}}_F$ ... $\sqrt{\text{PROFESSOR}}\ ]]\ ]$

b. Information from context: *There are two Japanese professors*

c. After Valuation from Context: $[_{DP}\ [D_{[num:\underline{\mathbf{pl}}]}$ the $[_{NP}\ \sqrt{\text{JAPANESE}}_F$ ... $\sqrt{\text{PROFESSOR}}\ ]]\ ]$

[12]The objective-subjective distinction is needed in order to model instances of uncertainty or partial knowledge about the identity of referents. Following Heim 2007, we cash it out by modelling C (the subjective notion) as a set of potential objective contexts, i.e. a set of worlds *w* compatible with how C is depicted in the speaker's mind. Then the rules would look like in (i) (and the semantic composition rules would correspondingly refer to *w* as well):

(i) **Number valuation from context**:
Let DP be headed by $D_{[num:\_\_]}$, in a syntactic derivation S in context C.
a. If for every *w* in C, $[\![DP]\!]^{w,C}$ is an atomic individual, then $D_{[num:\underline{\mathbf{sg}}]}$.
b. Else, $D_{[num:\underline{\mathbf{pl}}]}$.

I will not dwell so much on partial knowledge, but see some discussion at the end of 3.3.1.

[13]As earlier mentioned, there are intermediate steps between step (c) and (d) that I am not representing here, which are eventually responsible for realizing the number feature in D on the noun.

d. Phonology: /the JapaNESE professors/

The valuation rule in (77) makes reference to the semantic value of DP in the context C. It should be clear that by this I mean the *ordinary* semantic value of the DP; Crucially, focus-semantic values are not counted in deciding how the feature should be valued. Presumably, this is because they don't correspond to a structure that gets overtly realized.

Binding and fake number

As we saw in 2.2.1, there are fake features in structures involving variable binding. For example, when an elided VP contains a variable in a definite DP bound from outside the ellipsis, there can be number mismatches with the antecedent. This holds whether the antecedent contains plural and the ellipsis contains singular, or vice versa.

(80) *Context: Becca has 3 advisors, Eleanor has 2 advisors, Kelly has 1 advisor.*

a. Becca$_2$ will meet with her$_2$ advisors tomorrow, and KELly$_1$ will [$_{VP}$ △] too.
△ = *meet with her$_1$ advisor tomorrow*

b. Kelly$_1$ will meet with her$_1$ advisor tomorrow, and BECca$_1$ will [$_{VP}$ △] too.
△ = *meet with her$_1$ advisors tomorrow*

I am not ready yet for a complete analysis of these cases, because it requires incorporating binding and its affect on context, which will only be tackled in chapter 4. But the idea is basically the same and can already be appreciated. The LF of the antecedent sentence doesn't prejudge the number on the DP *her advisors*. Ellipsis licensing requires focus parallelism: the semantic value of the antecedent sentence needs to be a member of the focus-semantic value of the ellipsis-containing sentence. This is met because at the level of semantic values there is no number information. The DP gets valued only on the way to PF, but that is irrelevant for checking the licensing condition on Ellipsis. Valuation will see that the meaning of the antecedent 'her$_1$ ADVISOR(S)' is Kelly's advisor, and since this entity is atomic in the context, the DP will be valued **sing**.

Partial knowledge and semantic unmarkedness in number

I'd like to make a small digression here about semantic unmarkedness. The valuation rules in (77) raise interesting questions about cases of referential uncertainty. What happens when the speaker is not sure about the cardinality of the referent? The rules lead us to expect that the plural should be inserted, at least if "else" in (77b) also includes scenarios where the cardinality is undecided (see footnote 12 for a way to formally explicate this). Specifically for definite DPs, this prediction faces challenges. Mayr (2015) claimed that definite DPs cannot be used in either singular or plural in cases of uncertainty:

(81) Mayr 2015:211.

Context: It is common belief that Paul either wrote exactly one song or several ones.

a. #The song that Paul wrote are good.

b. #The songs that Paul wrote are good.

However, The empirical picture is not so conclusive.[14] While I agree with Mayr that (81b) is not perfectly natural in the scenario, it is still acceptable, and this has been confirmed by three speakers I have consulted with. Moreover, these speakers judge the following minimal variant as totally acceptable:

(82) *Context: Every time the band worked on a new album, Paul contributes one or more songs. There's a new album coming up, and this time is no exception.*

a. # Probably, the song that Paul wrote are horrible.

b. Probably, the songs that Paul wrote are horrible.

What is the difference? I believe that (81b) is degraded for irrelevant reasons, namely because it requires accommodating a rather unusual state of affairs – one in which we know that a certain

[14]Mayr's obligatory-implicature account of these facts runs as follows. Number is introduced at the NP level below the determiner, and a non-atomic inference of **pl** in definite DPs is obligatorily arrived at by exhaustifying over the stronger **sg** alternative (computed locally, below the determiner). I cannot adopt Mayr's theory (and I don't have to, given the text), since for the current proposal the NP level must crucially be compatible with both atomic and non-atomic predication for the purpose of weak projection.

number of songs are (each) good without even knowing if their cardinality is one or more. (82) removes the degradedness because the context together with the sentence's assertion implies that Paul usually writes horrible songs, no matter their cardinality given a particular album. And we see that when this is established, the plural version is acceptable.

With pronouns the judgments are even sharper. One can use the plural (but not the singular) in cases where the cardinality of the referent of the pronoun is undetermined (the evidence for this in English needs to come from non-human referring pronouns because of the phenomenon of singular *they* for human referents):

(83) One song of mine, or maybe two, will be played at the event. They/#It will be performed by the lead singer.

Kanazawa (2001) shows the same effect with donkey pronouns. The plural pronoun can be used in a (local) context in which the pronoun's reference ranges over atoms, (84a). And (84b) shows parallel data with full DPs.

(84) a. Every farmer who owns [one or more donkeys]$_1$ beats them$_1$. (Kanazawa 2001:385)

b. Every farmer who owns one or more donkeys beats the donkeys/#donkey he owns.

In sum, I maintain that plural is the semantically unmarked element and is used in elsewhere contexts, as predicted by the valuation rule in (77) (cf. footnote 12).

3.3.2 Gender 'weak projection' explained

Moving on to gender, I repeat from chapter 1 an example of weak projection in German. (85a) does not necessarily restrict the negated alternatives to be about female-identified professors.

(85) *Context: At the conference I talked to some students, but to almost no professor.*

a. Ich habe nur mit [der DEUTSCHEN Professor-**in**] gesprochen (German)
I have only with [the GERMAN professor-**FEM**] spoken
'I only talked to the GERMAN professor, who is female'

b. $\rightsquigarrow \forall P, P \neq German$: *I didn't talk to the P professor.*

The proposal works virtually the same as with number, so I can be succinct. A derivation starts without a valued gender (and number) D-feature:

(86) Generate LF:

a. only $\left[_{\text{TP}}\right.$ I have spoken to $[_{\text{DP}}$ the German$_F$ professor(s)$]$ $\left.\right]$

b. $[_{\text{DP}}$ $[\text{D}_{[num:\_\_,\ gen:\_\_]}$ the $[_{\text{NP}}$ $\sqrt{\text{GERMAN}}_F$ ... $\sqrt{\text{PROFESSOR}}$ $]$ $]$

The DP doesn't restrict the reference to individuals identified as female, and therefore neither do the alternatives in (88).

(87) $[\![_{\text{DP}}$ the German professor(s)$]\!]^C$ = the maximal contextually-salient individual containing one or more German professors

(88) *Alternatives of the prejacent of only in (86):*

a. I have spoken to $[_{\text{DP}}$ the German professor(s)$]$,

b. I have spoken to $[_{\text{DP}}$ the Japanese professor(s)$]$,

c. I have spoken to $[_{\text{DP}}$ the American professor(s)$]$, ...

Gender valuation

Once again, "interpreted" D-features are those that get values from the context, and this applies to gender in German.[15] The relevant rule is given in (89) and assumes that what is called 'masculine' is the lack of gender value in German. It says that gender remains unvalued unless all individuals referred to are identified as female.

(89) **Gender valuation from context** (in German, Hebrew, Greek, ...):

Let DP be headed by $\text{D}_{[gen:\_\_\ ,\ ...]}$ in a syntactic derivation S in context C.

[15] Like in the case of Number, I assume that idiosyncratic gender that comes with the noun can also value D-features (cf. fn. 11). Moreover, there seems to be a preference for valuation by noun over valuation by context, since in languages that have both idiosyncratic and semantic gender, it is the idionsyncratic value that wins for the purpose of verb agreement and morphology of determiners; e.g. German *Das/*Die Mädchen* 'the.SG.NEU/*FEM girl.NEU'. The phenomenon of hybrid agreement, where idiosyncratic and semantic gender can both be active for agreement at the same time (see e.g. Corbett 1991; Pesetsky 2013; Rappaport 2013; Landau 2016a; Kramer 2016a) further complicates the picture. I won't discuss those complications and how they might affect the theory; see Kučerová 2018 for a concrete proposal about the dominance of idiosyncratic over semantic gender in terms of timing of valuation in the derivation.

a. If $\forall x \leq [\![\text{DP}]\!]^C$, x is identified as female, then $\text{D}_{[gen:\underline{\textbf{fem}}\ ,\ ...]}$.

b. Else, $\text{D}_{[gen:\_\_\ ,\ ...]}$.

If the context contains just one German professor and she is female, D will be valued **sg** (as before) and **fem** (because of 89a). The derivation is summarized in (90).

(90) a. Base-generation (=86b): $[_{\text{DP}}\ [\text{D}_{[num:\_\_,\ gen:\_\_]}$ the $[_{\text{NP}}\ \sqrt{\text{GERMAN}}_F$ ... $\sqrt{\text{PROF}}$']]]

b. Information from context: *There is exactly one German professor, she's female*

c. After Valuation from Context: $[_{\text{DP}}\ [\text{D}_{[num:\underline{\textbf{sg}},\ gen:\underline{\textbf{fem}}]}$ the $[_{\text{NP}}\ \sqrt{\text{GERMAN}}_F$ ... $\sqrt{\text{PROF}}$']]]

d. Phonology: /the GERman professor-FEM/

If the context contained more than one German professor and they are all identified as females, the derivation would end in values **pl** and **fem**. And if the group is mixed with respect to gender, the result is the value **pl** for number and the lack-of-value __ for gender; this predicts correctly (and unsurprisingly) that the 'masculine' plural form (*Die Professor-en* 'the professor-PL') can denote a referent consisting of both female and male professors.

In all these different scenarios, the *semantics* as well as the derivation of focus alternatives is one and the same, with just the context of utterance responsible for the different morphological outcome.

English pronouns and markedness

For many English speakers, as mentioned in section 3.2, the masculine *he* pronoun is restricted to refer to male-identified individuals. A *they* pronoun seems to be the elsewhere element (for human/animate-referring pronouns). For these speakers, an appropriate rule is given below, with the assumption that the morphology spells out an unvalued gender D-feature as a *they* pronoun (gender on English pronouns will be relevant for the next chapter).

(91) **Gender valuation from context** (English pronouns):

Let DP be headed by $\text{D}_{[num:\textbf{sg},\ gen:\_\_]}$, in a syntactic derivation S in context C.

a. If $[\![\text{DP}]\!]^C$ is identified as female, then $\text{D}_{[num:\textbf{sg},\ gen:\underline{\textbf{fem}}]}$.

b. If $[\![\text{DP}]\!]^C$ is identified as male, then $\text{D}_{[num:\textbf{sg},\ gen:\underline{\textbf{masc}}]}$.

c. Else, $\text{D}_{[gen:\ \_\_]}$.[16]

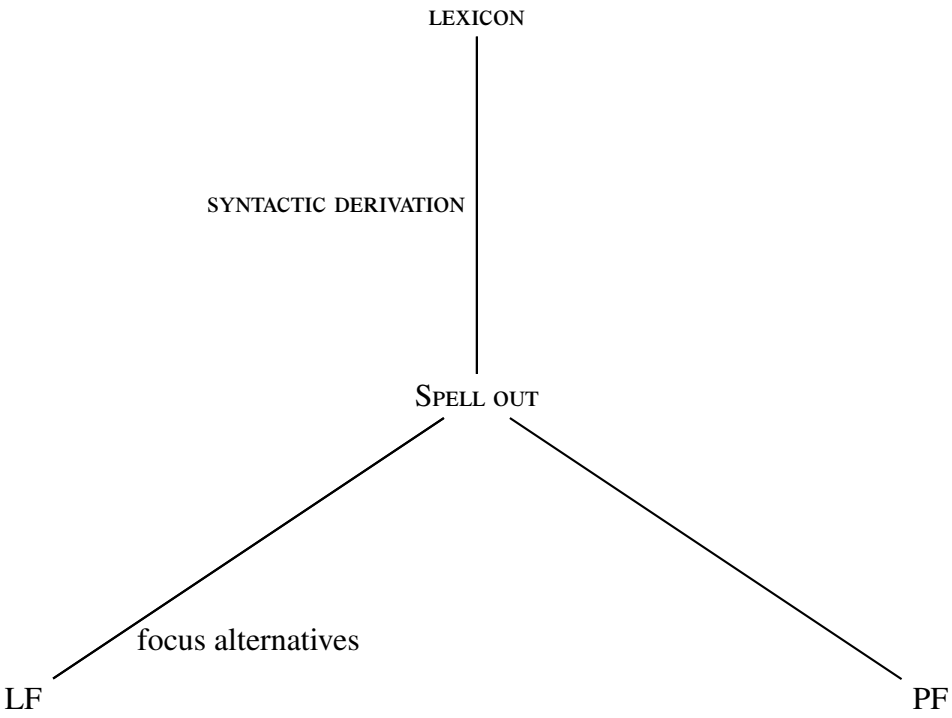
The rules in (91) are conditioned to apply only when D is valued **sg**, because there are no gender distinctions in English plural pronouns. This entails that the procedure for gender valuation follows number valuation. Alternatively, we could state the (a) and (b) clauses of (91) to be in principle compatible with plural reference, i.e. with conditions like that in (89): *if every atom in* $[\![DP]\!]^C$ *is female, then* $D_{[num:\ldots\ gen:\underline{\textbf{\textit{fem}}}]}$. In that case no hierarchy of number-over-gender valuation is necessary, but the condition would only have a visible effect with atomic reference. That would be somewhat unattractive for English, although it would be more appropriate for languages in which there are gender distinctions in the plural.

3.4 Consequences for the architecture

Before moving on, I discuss the architectural consequences of the proposed framework. The idea that ϕ-feature values are not generated at the base is **incompatible** with the prominent architecture of grammar in current generative theory (since its inception), the Y model:

(92) *The Y model* (e.g. Chomsky 1965; Chomsky & Lasnik 1977)

[16] These rules only apply to human/animate pronouns; non-human/animate referents are referred to by *it* of course. Assuming that non-humanness/non-aminacy is also a gender feature, it could be easily modeled by adding another sub-rule in (91). To keep things as simple as possible, I won't discuss this category further.



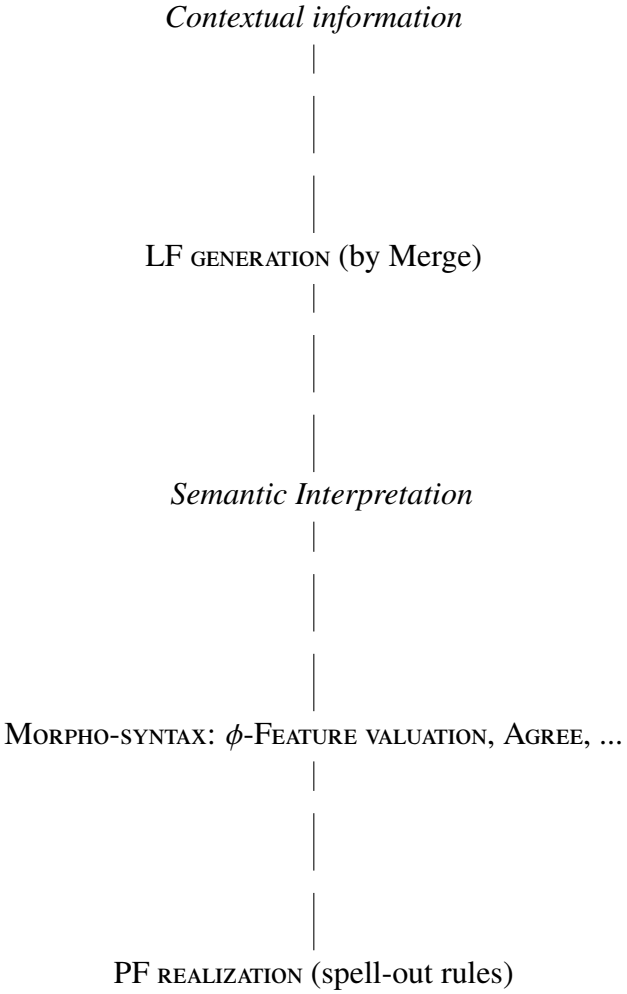
Here is why the Y-model in (92) is ill-suited to do the required work. In that model, semantic-pragmatic information can never feed morphosyntax. The link between semantics (the LF side) and phonological realization (PF) is mediated by a narrow-syntactic component that feeds into both of them. Crucially, every piece that is associated with both a semantic interpretation and phonological realization enters the syntactic derivation from the lexicon, at the beginning of the structure-building operations, and at the end it reaches the interfaces – in the case of LF, it will receive a semantic interpretation.

Now, on the Y model, focus alternatives undoubtedly live somewhere at the LF interface (either in a syntactic or post-syntactic level). Therefore, if the generalization I motivated is correct, this model requires that ϕ-features start their life with interpretation from the lexicon (this is by definition of the model) and then be deleted specifically when it comes to computing focus alternatives, which is what the the weak presupposition projection theory of Sauerland 2013; Bruening 2019 says. This is not a very economical derivation. More problematically, it doesn't answer what is it about features specifically that allow them to disappear from alternatives. Why can't *again*, *also*, *already* be deleted in a similar fashion? The proposal here, which says that ϕ-features do not have

semantics in the first place, provides an answer. But it crucially requires a different model, where a representation that gets submitted to interpretation, as well as a representation of context, can feed the morphosyntactic engine. For concreteness, I assume the model depicted in (93), which is closer in spirit to the older Generative Semantics framework (e.g. Lakoff (1971); Seuren (1984)).[17]

[17] Other models, such as the recent 'Meaning First' approach outlined in Sauerland & Alexiadou 2020, might fit the bill too.

(93) *Proposed Model*:



The place of the computation of focus alternative, crucially, is somehwere on the spine between LF generation and Interpretation. LFs are interpreted (against a context), and some of them are shipped to morphology where DPs will have to be equipped with features. This stage is fed by a representation that already has access to interpretation (and context).

3.5 Presuppositional inferences not as presuppositions

I have denied that ϕ-features (always) enter the syntax from the lexicon with hard-coded semantics, which means that they don't technically have the status of a presupposition trigger in the grammar. Nevertheless, the current analysis duplicates a basic result of the standard presuppositional-trigger account, namely the feeling of oddness or infelicity when the wrong feature value is used. Let's take gender on English pronouns. Recall the rule:

(94) **Gender valuation from context** (English pronouns):

Let DP be headed by $D_{[\textit{num}:\textbf{sg},\ \textit{gen}:\_\_]}$, in a syntactic derivation S in context C.

a. If $[\![\text{DP}]\!]^C$ is identified as female, then $D_{[\textit{num}:\textbf{sg},\ \textit{gen}:\underline{\textbf{fem}}]}$.

b. If $[\![\text{DP}]\!]^C$ is identified as male, then $D_{[\textit{num}:\textbf{sg},\ \textit{gen}:\underline{\textbf{masc}}]}$.

c. Else, $D_{[\textit{gen}:\ \_\_]}$.

If the context makes clear that the individual intended to be referred to is identified as a female individual, then using a 'he'-pronoun violates (94a). This violation will be responsible for the feeling of oddness and will justify a hearer to protest, or—if the hearer is charitable—they will accommodate the mistake and keep the conversation going on the assumption that either the speaker or they themselves have wrong beliefs about the gender identity of the referent. What was before encoded directly as a semantic presupposition is now arrived at indirectly by reasoning on the part of the hearer over what possible contexts the speaker considers when they referred to some individual.

3.6 Valuation on referential pronouns

Throughout I make two important though unremarkable assumptions about pronouns. One is that a pronoun contains at LF a referential index (a 'pointer'), which represents the semantic value of the pronoun;[18] Second is that the context of utterance records values for indices ('discourse referents').

[18] My analysis could be cast within variable-free semantics (Jacobson 1999), with only technical and insignificant differences.

With this, the rules can make a well-formed pronoun out of a bare index, as illustrated in (95) for a 3rd-fem-sg pronoun.

(95) her_7

a. LF: $[_{D[num:\_\_,\ gend:\_\_,\ pers:\_\_]}\ 7]$

b. Information from context: the index 7 is mapped to a female atomic individual

c. After valuation $[_{D[num:\ \mathbf{sg},\ gend:\ \mathbf{fem},\ pers:\_\_]}\ 7]$

The same analysis extends to 1st-sg pronoun:

(96) my_1

a. LF: $[_{D[num:\_\_,\ gend:\_\_,\ pers:\_\_]}\ 1]$

b. Information from context: the index 1 is mapped to the speaker

c. After valuation $[_{D[num:\ \mathbf{sg},\ pers:\ \mathbf{1st}]}\ 1]$[19]

Given present assumptions, it is predicted correctly that it is not possible to shift the reference of a referential pronouns across the alternatives. In a sentence like *I only gave her a BOOK* it is not possible to vary the reference of *her* across alternatives. This is because, while the features are inserted based on the prejacent alone, the index itself doesn't change values across alternatives, when it is not bound

Things change when the index is itself bound by a focused or focus-containing phrase; in that case the reference of the index shifts across alternatives along with its binder. Modelling these cases requires a more sophisticated notion of valuation from context, which will be provided in the next chapter.

[19]Strictly speaking a rule for gender is also invoked in (96), assigning e.g. *fem* when the speaker is feminine. But in English and other languages in which gender is not distinguished on 1st/2nd pronoun, this will not show its effect at PF. One common way to model this and similar cases where morphosyntactic features are neutralized with respect to their PF exponents is with 'Impoverishment rules' which can delete features at the interface between morphosyntax and phonology (Bonet 1995; Halle 2000; Harley & Noyer 2014, a.o.). In another implementation of this, the valuation rules for gender don't apply at all if person has been valued to 1st/2nd. Since I don't deal with the morphology-phonology interface, I can remain neutral on the exact type of mechanism that achieves this and I simply avoid representing the relevant stage in the derivation.

4

Valuation from *Modified* Context: incorporating variable binding

Propositions cannot represent the logical form: this mirrors itself in the propositions.

That which mirrors itself in language, language cannot represent.

– Ludwig Wittgenstein, *Tractatus, 4.121*

Abstract

In this chapter I extend the system of valuation-from-context from the previous chapter to deal with features on bound pronouns and (more generally) instances where a variable is bound in the scope of a D-feature. The extension requires a notion of *modified* context: the valuation rules look not at the global context of the utterance, but one that has been locally enriched with referents recorded by λ-operators. The resulting theory will derive as a theorem that a bound variable in focus always agrees with its antecedent (while not necessarily "agreeing" with it across the focus alternatives), without a dedicated rule of agreement. The enrichment requires assuming that focus-alternatives is systematically counted off from the computation of local contexts.

4.1 Local Contexts for bound anaphora: Overview of proposal

4.1.1 The issue

In the previous chapter I proposed a theory of valuation of ('interpreted') features from context: a pronominal or full DP doesn't have to start the derivation with features; but DPs *must* be equipped with the right values for these features in the course of derivation to PF. The rules of morphological valuation are sensitive to the semantic value (reference) and certain information stored in contextual parameters. This view and the valuation rules are repeated here for convenience. I add rules for valuation of person:

(97) D-heads (pronouns and heads of full DPs) start the derivation unvalued for ϕ-features: $D_{[num:\_\_,\ pers:\_\_,\ gend:\_\_]}$. Those get valued in the course of derivation to PF as follows:

a. **Number valuation from context**:

i. If $[\![DP]\!]^C$ is an atomic individual, then $D_{[num:\underline{\textbf{sg}}]}$.[1]

ii. Else, $D_{[num:\underline{\textbf{pl}}]}$.

b. **Gender valuation from context** (English pronouns):

i. If $[\![DP]\!]^C$ is identified as female atomic individual, then $D_{[num:\textbf{sg},\ gen:\underline{\textbf{fem}}]}$.

ii. If $[\![DP]\!]^C$ is identified as male atomic individual, then $D_{[num:\textbf{sg},\ gen:\underline{\textbf{masc}}]}$.

iii. Else, $D_{[gen:\_\_]}$.

c. **Person valuation from context** (English pronouns):

i. If $[\![DP]\!]^C$ is or includes the speaker in C, then $D_{[\ldots\ pers:\textbf{1}]}$. Else,

ii. If $[\![DP]\!]^C$ is or includes a conversational participant in C, then $D_{[\ldots\ pers:\textbf{2}]}$.

iii. Else, $D_{[\ldots,\ pers:\_\_]}$.

We saw cases of weak projection to focus alternatives that this perspective explains – number and gender on full DPs that contain focus. There is however a substantial portion of the data that

[1] Again, this is a shorthand for 'if for every world $w \in C$, $[\![DP]\!]^{w,C}$ is..., then ...'. The rules refer to the value of DP *throughout* C. For convenience, the main text incorporates worlds only when necessary.

the system cannot yet handle, namely fake features in structures involving variable binding. In chapter 2 I presented old and new data from both association with *only* and ellipsis in which the features on a bound pronoun, or on a full DP that dominates a bound pronoun, are ignored for the relevant purposes. Three such examples are repeated below: person/gender on bound pronouns (98a-i-98a-ii), number in full DPs (98a-iii) and gender in full DP (98a-iv).

(98) *Fake Features with bound variables*

a. *In Association with Focus*:

i. Only I did **my** homework.

$\rightsquigarrow$ *sloppy: No one other than me did* ***their*** *own homework.*

ii. Only MARY$_1$ did **her**$_1$ homework.

$\rightsquigarrow$ *sloppy: No one other than Mary,* ***including non-females****, did their homework.*

iii. Only MARY$_1$ treated [$_{DP}$ the **guy** she$_1$ arrested] with respect.

$\rightsquigarrow$ $\forall x \neq$ *Mary*: *x didn't treat the* ***guy or guys*** *x arrested with respect.*

iv. rak YOSI$_1$ histader im [ha-mexanex-**et** šel-o$_1$]. (*Hebrew*)
only YOSI got.along with [the-head.teacher-FEM.SG of-his$_1$]
$\rightsquigarrow$ $\forall x \neq$ *Yosi: x didn't get along with x's head teacher,* ***male or female****.*

b. *In Ellipsis*:

i. I did **my** homework, but you didn't [$_{VP_\triangle}$ ~~do **your** homework~~].

ii. Mary did **her** homework, but Jack didn't [$_{VP_\triangle}$ ~~do **his** homework~~].

iii. Mary$_1$ has treated [$_{DP}$ the **guy** she$_1$ arrested] with respect, but

ANN$_2$ hasn't [$_{VP_\triangle}$ ~~treat [$_{DP}$ the **guys** she$_2$ arrested] with respect~~].

iv. Yosi$_1$ histader im [ha-mexanex-**et** šel-o$_1$]. Avi lo $\triangle$. (*Hebrew*)
Yosi got.along with [the-head.teacher-FEM.SG of-his$_1$]. Avi not $\triangle$.
$\rightsquigarrow$ $\triangle$: *got along with his head teacher, who is* ***male****.*

The reason that the system can't capture these is that it is not clear what notion of context feeds rules of valuation here. Let's see why. I illustrate with cases of fakeness on pronouns in association with focus, but the morals extend to the cases with full DP, and to the ellipsis examples.

In order to derive the possible entailments of (98) with the simple and mechanical algorithm for generating alternative sets, the variable marked with subscript 1 needs to be bound by the focused phrase but also unconstrained by featural information. My conjecture about fake features from before carries over: pronominal variables are (at least in these cases) 'minimal' at LF, having just an index and no interpreted features. The surface features on the pronoun pro_i, just like referential pronouns, receive it for the purpose of morphology based on the semantic value of the DP in context.

(99) a. Sloppy-identity LF of (98a-i) (simplified): Only [I_F $\boldsymbol{\lambda_i}$ t_i did $\boldsymbol{pro_i}$ homework]

b. Sloppy-identity LF of (98a-ii) (simplified): Only [Mary$_F$ $\boldsymbol{\lambda_i}$ t_i did $\boldsymbol{pro_i}$ homework]

c. Alternatives of the sister of *only*:

{ Avi λ_i t_i did $\boldsymbol{pro_i}$'s homework,

Yosi λ_i t_i did $\boldsymbol{pro_i}$'s homework, ... }

d. $\boldsymbol{pro_i} \equiv [\mathrm{D}_{[num:\_\_,\ gend:\_\_,\ pers:\_\_]}\ i]$[2]

But, how does pro_i end up surfacing as *my* in (98a-i) and *her* in (98a-ii)? the valuation from context rules formulated above presuppose an intuitive notion of 'reference', and a simple notion of 'context'; the problem is that bound variables do not have reference directly from context. Therefore the rules currently fail to apply to such examples. At best, the system would output the elsewhere element *their* as the spell out of the bound pronoun, contrary to fact. The same goes for the full DP cases, in (98a-iii) and (98a-iv). And things are not better off for the ellipsis counterparts of all these examples, in (98b).

In fact, if ϕ-features are *always* valued from context and never start a derivation with what in the standard theory is semantic content (presupposition), it is not possible to handle even the most run-of-the-mill patterns of bound anaphora in quantification, e.g. in (100a); those pronouns, too, don't have a semantic value simpilicier from context. (100b) strengthens this point by showing

[2]Strictly speaking also the subject *I* in (99a) is (or can be) a minimal pronoun and has the exact same structure at LF as the variable it binds (no interpreted features). But I don't represent it as such, for easier exposition because this detail doesn't have interesting consequences for the main concerns of this chapter. Featureless referential pronouns are simply assigned features from the global context (cf. section **??**).

fake gender on bound pronouns in ellipsis environments.

(100) a. No girl$_i$ said that she$_i$ was busy.

b. No girl$_i$ said that she$_i$ was busy. No BOY$_j$ did ~~say that he$_j$ was busy~~, either.

4.1.2 Solution: valuation from *local* contexts

But while bound variables indeed don't have a semantic value with respect to the global context, they do have a value relative to a **Local Context** (Schlenker 2009; Singh 2011; Charlow 2019), and the natural step to take is to suppose that the morphological valuation rules are checked against this more sophisticated notion of context. Intuitively, since the subject phrase in (99) (and in the other examples in (98a)-(98b)) binds, it creates a new context relative to the original one, and in that context the variable has a semantic value—one that corresponds to its binder (see details below).[3] If Local Contexts are defined in just the right way, and if the valuation rules are reformulated to respect sensitivity to them, the facts will be accounted for. One of the main aims of this chapter is to formally explicate a notion of Local Contexts that would do the desired job as just informally described, and explore its empirical and theoretical consequences. Parts of this chapter will be slightly more technical than the previous ones.

To my knowledge, local contexts are used chiefly in the theory of presupposition (e.g. Heim 1983; Schlenker 2008), to model presupposition projection or how context update by linguistic expressions affects the felicity conditions on the use of presupposition triggers. Because it is standard in the semantic literature to treat ϕ-features as presupposition triggers, such theories of local contexts automatically make predictions about projection of ϕ-features. But we saw that ϕ-features do not show the projection behavior of other triggers in one specific environment: focus alternatives, and I proposed that they are therefore not presupposition triggers at all. However, we can still harness the insights and techniques of the literature on local contexts to do a job for us, namely create a method allowing us to talk about the 'referent' of a bound variable.

[3] I remind the reader the excursus on ellipsis from section 2.1.4. Charlow's (2019)'s work on the relevance of local contexts for the theory of ellipsis licensing formed part of my motivation to explain fake features by recourse to local contexts.

Let me give a birds-eye view of the proposal, as well as the questions that it must address, before diving into the nitty-gritty details which will be laid out in section 4.2. Drawing on Schlenker 2009's insight (though I depart from him in various respects), the local context for some expression α embedded in an LF can be identified with the smallest context one may restrict attention to when evaluating the contribution of α to the overall semantic value of the whole LF, given the global context of the LF plus information encoded by material up to the point where α is evaluated. For the purpose of modeling bound-variable anaphora, local contexts can be identified with sets of assignment functions.[4] In accordance, I define the Local Context (**LOCOs**) of α as the smallest set of assignments that expand the global context (the starting assignment) with referents recorded by λ-operators that c-command α. Once we have LOCOs at our disposal, it is possible to talk about the semantic value of a bound variable – with respect to its LOCO. This informal description will be made precise soon.

The system I build will produce as one of its consequences that a pronominal variable bound by an expression of type *e* will 'refer' across its local context to the same individual that its binder refers to. This will be the case no matter whether the binder is F-marked (as in (99a)-(99b)) or not (see right below on why F-marking shouldn't matter). Therefore, in (99a) the pronominal variable *i* will refer in its local context to the speaker, and in (99b) to Mary. With this, the morphological rules can successfully assign the correct feature-values to the pronouns to produce the surface form in (98a-i) and (98a-ii), respectively. Binder-bindee 'agreement' falls out from the system of interpretation and valuation, and the other cases mentioned in the beginning of this section are accounted for, too. We just need to update valuation-from-context rules to refer to LOCOs, as follows. The relevant changes are underlined ('g' ranges over assignments):

(101) D-heads start the derivation unvalued for ϕ-features: $D_{[num:\_\_,\ pers:\_\_,\ gend:\_\_]}$.
Let C' be the Local Context of DP (given a global context).

a. **Number valuation from context**:

[4] Or sets of assignment-world pairs, if we don't abstract away from intensionality and matters that have to do with presupposition satisfaction. For most of this section I will ignore worlds in the main text (but see footnotes) because I try to keep the system as simple as possible. But worlds will eventually become important.

i. If $\forall g' \in C'$, $[\![\text{DP}]\!]^{g'}$ is an atomic individual, then $\text{D}_{[num:\underline{\textbf{sg}}]}$.

ii. Else, $\text{D}_{[num:\underline{\textbf{pl}}]}$.

b. **Gender valuation from context** (English pronouns):

i. If $\forall g' \in C'$, $[\![\text{DP}]\!]^{g'}$ is identified as female atomic individual, then $\text{D}_{[num:\textbf{sg},\ gen:\underline{\textbf{fem}}]}$.

ii. If $\forall g' \in C'$, $[\![\text{DP}]\!]^{g'}$ is identified as male atomic individual, then $\text{D}_{[num:\textbf{sg},\ gen:\underline{\textbf{masc}}]}$.

iii. Else, $\text{D}_{[gen:\_\_]}$.

c. **Person valuation from context** (English pronouns):

i. If $\forall g' \in C'$, $[\![\text{DP}]\!]^{g'}$ is or includes the speaker in C', then $\text{D}_{[\ldots\ pers:\textbf{1}]}$. Else,

ii. If $\forall g' \in C'$, $[\![\text{DP}]\!]^{g'}$ is or includes a conversational participant in C', then $\text{D}_{[\ldots\ pers:\textbf{2}]}$.

iii. Else, $\text{D}_{[\ldots,\ pers:\_\_]}$.

One advantage of this system is that it unifies feature-valuation on both referential and bound-variable pronouns (and DPs that contain bound variables). In fact, valuation on referential DPs and referential pronouns are just a special case of valuation from local context, if a global context is thought of as one particular local context (i.e. the one where no sentence-internal binding has taken place).

4.1.3 Local contexts ignore focus alternatives

What I just described will already allow us to explain, in sections 4.3-4.4, the ellipsis facts that proved difficult for the Weak Presupposition Projection approach. To explain the focus-association cases, one additional assumption is required. As hinted above and as careful readers may have wondered, there is a question about the role of focus alternatives in the process of computing local contexts as just sketched. Crucially, the construction of LOCOs must be blind to information coming from focus alternatives. Otherwise we would end up deriving the wrong local context for the variables in (99a)-(99b), and consequently the wrong feature realization.

To appreciate the issue it is useful to contemplate on the striking difference (Heim 2008) between pronouns bound by 'regular' quantifiers and those bound by focused phrases. *No one but*

X [...] is truth-conditionally equivalent to *Only* X_F *[...]*, yet variables bound by *No one but X* are realized differently from those bound by *only* X_F:

(102) *On a bound-variable reading:*

a. Only I did **my**/*his/*their homework.

b. No one but me did his/their/***my*** homework.

(103) *On a bound reading with* ***mixed-gendered*** *domain of quantification:*

a. Only MARY did **her**/*their homework.

b. No student but Mary did their/***her*** homework.

The generalization is that in quantification over focus alternatives (*only* X_F *[...]*), a bound pronoun is spelled out with the features of X alone, although (or even if) the alternatives to the binder aren't compatible with those features; whereas in ordinary quantification (*no one but X [...]*), the spell-out takes into account all the individuals in the domain of quantification. When the domain of quantification is mixed with respect to gender, the spell-out is according to the elsewhere element. For person, elsewhere is 3rd person and therefore the choice in (102b) must be 3rd: 'his', 'her' or 'their' (depending on the members in the domain of quantification); for gender the elsewhere (for most speakers) is a 'they' pronoun which is why in (103b), in light of the domain being mixed-gendered, a 'they' pronoun is obligatory.

As is by now evident, I suggest that the difference in realization is a by-product of the fact that grammar encodes 'normal'-quantification differently from focus-quantification. I propose that only the ordinary semantic value of a phrase counts for deriving LOCOs for subsequent phrases. A focus-alternative of that phrase doesn't count. Focus alternatives can manipulate variable-assignments for the purpose of the narrow semantics alone, namely for the purpose of assigning a semantic value to bound variables in their scope; but this manipulation is done internally in each alternative,[5] at a level of representation that does not have access to the construction of lo-

[5]More precisely, in each alternative in the scope of the focus-sensitive operator that binds the focus; see details in section 4.3.2.

cal contexts. Only those assignments manipulated in the ordinary semantic value count for Local Contexts. Let me call this the CONTEXT ENCAPSULATION HYPOTHESIS, stated succinctly in (104):

(104) **Context Encapsulation from Alternatives Hypothesis**: Information encoded (merely) in focus alternatives does not have an effect on the construction of embedded contexts.

Encoding the Context Encapsulation hypothesis in the system will be technically easy (given the rest of the system) and will be done in section 4.3.2, but it raises a *why* question. Why should (104) hold? For now it will have to remain a rock-bottom assumption, although I will make a suggestion in section 4.5.6 regarding what can ground (104). I will suggest there that (104) should follow from a more general theory of context update by salient information, and that it is ultimately a result of the asymmetry in salience status between information encoded in prejacents (ordinary meanings) and information encoded in focus alternatives.

Section 4.5.5 will show that the hypothesis in (104) does not imply that information encoded in **presupposition triggers**, e.g. *again*, merely have to be satisfied by the prejacent. Focus-sensitive operators like *only* project presuppositions universally from their **restrictor** position, which is the set of alternatives they operate over. Since presupposition triggers like *again*, unlike ϕ-features, are base generated and interpreted as standard, when they are in the scope of focus they will be present across all alternatives and therefore project universally from the restrictor (even though their local context doesn't contain information from focus alternatives, per (104)).

4.1.4 Summary of desiderata and a Look-ahead

The rest of this chapter is devoted to making what I said until now more precise. Here is a summary of the desiderata from the system:

(105) **Desiderata from a system of Local Contexts for bound anaphora and focus**:

a. i. When a variable, i, is bound by a type-e element, the local context (set of assignments) of the scope of X maps i to the meaning of its binder (cf, 102a)-(103a)).

ii. When a variable, i, is bound by a quantifier (type-$\langle et, t\rangle$), the local context of the scope of X maps i to the meaning of some individual in the domain of the binder (cf, 102b)-(103b)).

b. i. Focus Alternatives don't alter the picture; only information coming from the prejacent is relevant for the rules that construct local contexts.

ii. Standard presuppositions (*again, also, both*) project universally in the scope of focus because they project from the restrictor position of the focus-sensitive operators that interprets the focus.

Except for their projection properties into focus alternatives, gender, number and person features in ordinary quantificational contexts seem to show the same behavior as standard presuppositions (e.g. Sudo 2012 for gender). This means that the system to be constructed should produce the same results as the standard theory, which cashes out the fact in terms of presupposition projection/satisfaction in (local) context.[6]

4.2 Local Contexts (basic cases, without focus alternatives)

I now start developing a basic system of local contexts for anaphora, inspired somewhat by Schlenker 2009's work on local contexts for presuppositions, but the details will be different. In this section I completely ignore the contribution of focus alternatives and focus-sensitive operators like *only*, and concentrate on deriving local contexts for constructions that don't involve them. Focus and alternatives will be incorporated in section 4.3.2, and I will build it in such a way that the addition will have affect only on the semantics but not on the system of local contexts, as discussed above.

[6] Although see Yanovich 2010; Stokke 2020 for interesting problematizations of the idea that gender behave like classical presupposition triggers (independently of the issues under consideration here). Yanovich 2010 observed that gender features on a pronoun in consequents of counterfactual conditionals must relate to the 'real-world' gender of the antecedent of the pronoun, and not to its gender in counterfactual worlds introduced by *if*-clauses (**If John*$_1$ *were female, she*$_1$ *would...*). Whereas standard presupposition triggers are sensitive to counterfactual worlds and thus can be filtered out in analogous constructions (*If you had smoked, you would have stopped by now*). Yanovich also noticed that when the gender of the referent is left open in the actual world, gender again behaves as expected from presuppositions (e.g. *if the spy*$_i$ *we are after is female, she*$_i$*/??they*$_i$ *will enter this bathroom and not the other one*).

There are many possible routes to defining local contexts, depending on one's goals. Since I am not aware of an off-the-shelf theory of local contexts for bound anaphora which both suits my needs and is relatively simple, I'll now create one. The system I present has a mixture of static and dynamic properties. The semantics will be static, with denotation gaps: presuppositions will be encoded as definedness conditions on domains of semantic values. The spirit of how to construct local contexts will be dynamic—guided by the question: how linguistic material affect the context (assignment) for evaluating succeeding linguistic material. But the semantics itself is static. Finally, the technique I use for deriving the local contexts will follow Schlenker 2009 in using variables over contexts in the object language.[7]

One last introductory remark. In the following I concentrate on accounting for a very small fragment of bound-variable anaphora, essentially just the basics described earlier: binding by individual-denoting phrases, binding by quantifiers, and ultimately the (non-)interaction with focus. Therefore I allow myself to develop a quite primitive system, without a treatment of some of the outstanding problems in the theory of anaphora. In particular I limit myself to intra-sentential, in-scope binding which static semantics can handle easily, so the system is not suitable for cross-sentential anaphora or donkey pronouns.[8] I therefore won't have much to offer about the distinction in anaphoric potential between indefinites and definites, on the one hand, and universal quantifiers on the other, for example. I also ignore basic issues that have to do with left-right asymmetries and the effects of linear order on presupposition satisfaction (e.g. how presuppositions in a right conjunct are satisfied based on material introduced in a left conjunct). Adequately addressing all these issues will require an enrichment beyond the scope of my interests here, but in any case my main results should carry over to any richer theory of anaphora and context update.

I will proceed as follows. First I lay out the syntax and semantics of binding which I'll be

[7] I do not wish to claim that this is the most desirable mixture of properties in a general theory of local context, it is just one that I find easy to work with in a way that allows us to systematically talk about the local semantic value of a bound variable. The main take home point from this exercise is that whatever system of local contexts one favors, focus quantification should ultimately be encapsulated from it.

[8] For this reason, the system I provide will stop short of accounting for Fake Features on donkey pronouns in focus (Bassi & Longenbaugh 2018), which require going full-fledged dynamic in the semantics. If the account offered here for classical binding is appreciated, it seems to me that implementing the switch that donkey fake features necessitate in the semantics will be not a lot more than a technical exercise. Which I leave for another occasion.

assuming (part 4.2.1)). These will based on completely standard truth-conditional semantics—importantly, with definedness conditions on denotations—and a fairly simple pragmatics (global context). Then, I will formulate Local Contexts for embedded constituents, based on the global context and the semantics (part 4.2.2).

4.2.1 Static truth-conditional semantics

Logical Forms

I assume a Heim & Kratzer 1998-style logical forms for variable binding constructions. LFs are generated in which λ-operators are optionally inserted and can bind indexed variables in their syntactic scope. Two types of sentences and their LFs with which I will illustrate the system are given in (106) and (107). As usual, pro_i in (106a) and (107a) is really a shorthand for the more elaborate structure with unvalued features, (106b).

(106) Mary$_i$ did her$_i$ homework

a. LF: Mary [λ_i [$_{vP}$ t$_i$ did *pro*$_i$'s homework]]

b. *pro*$_i$: [$_{D[num:\_\_,\ gend:\_\_,\ pers:\_\_]}$ *i*]

(107) Every(/no) student$_i$ did her$_i$/his$_i$/their$_i$ homework

a. LF: Every(/no) student [λ_i [$_{vP}$ t$_i$ did *pro*$_i$'s homework]] [9]

b. *pro*$_i$: [$_{D[num:\_\_,\ gend:\_\_,\ pers:\_\_]}$ *i*]

The λ_i-operator in the above is the binder of occurrences of the variable *i* in its scope. As in common practice, I will also sometimes use the term 'binder of a variable *i*' in a derivative sense, to refer to the sister DP of the predicate-abstract headed by the λ-operator that binds (occurrences of) *i*: *Mary* in (106) is a binder, *every/no student* in (107) is a binder.

[9] I am mostly interested here in the realization of overt pronouns and phrases, but I assume that the trace position of the quantifier is assigned features too, which is reflected on verb agreement (in number, and—in the relevant language—in gender). I am representing the trace position as an impoverished index, but it could also be a full copy with NP content, in line with the theory of Trace Conversion or something to that effect (Fox & Johnson 2016).

Global context and assignment modification

Logical forms are semantically interpreted relative to contextual parameters, the relevant one being an assignment function to variables. Every global context furnishes one.[10] Pairing of an LF and a context is subject to the appropriateness rule in (108).

(108) **Approporiate Context** (cf. Heim & Kratzer 1998:243):

A context C is appropriate for an LF only if C determines a variable assignment g whose domain includes every index which has a free occurrence in the LF.

An assignment g records referents by pairing them with numerical indices. Assignments can be partial. It will be easiest to illustrate the ensuing system with an empty global context, i.e. one in which no discourse referent is recorded (i.e. one which is only appropriate for an LF with no free variables). This is simply the empty set $\emptyset$. Assignments can be enriched (modified) with discourse referents in the course of computation by sentence-internal λ-operators. I give the definition of an assignment extension in (109). Standard notation: $g^{[i \to x]}$ is an extension of g with the pair <i, x>.

(109) g' is an **Assignment Extension** of g iff g' is a (proper or improper) superset of g.

I.e., g' is an Extension of g if it is just like g but with zero or more entries (zero or more <index, individual> pairs added to g).

Note that this definition implies that a g' is *not* an extension of g if they both have some index i in their domain but disagree on its value. Extensions cannot rewrite the value given to some variable.

Semantics

Still as standard, these logical forms are interpreted according to normal interpretation procedures in a truth-conditional setting (Heim & Kratzer 1998). They are assigned semantic value using the interpretation function $[\![\ ]\!]$, relativized to an assignment function, g. (110) is the rule for interpreting variables – both free and bound.

[10] Or a *set* of assignments/ assignment-world pairs. I represent contexts as sets of assignment-world pairs only when necessary (see footnotes).

(110) $[\![pro_i]\!]^g = g(i)$, if i is in the domain of g; undefined otherwise.

Two familiar rules of composition that will be most relevant are below – Functional Application (111) and Predicate Abstraction (112). I am being explicit about the definedness conditions of the function and the argument and how they project, as definedness is going to play a role in defining local contexts. As usual, (112) is the rule that interprets binding, by utilizing the operation of assignment modification.

(111) **Functional Application (FA)**:

$[\![\text{A B}]\!]^g$ is defined iff $[\![\text{A}]\!]^g$ is defined and $[\![\text{B}]\!]^g$ is defined and $[\![\text{A}]\!]^g$ is in the domain of $[\![\text{B}]\!]^g$ (or vice versa). If defined, $[\![\text{A B}]\!]^g = [\![\text{B}]\!]^g\left([\![\text{A}]\!]^g\right)$ (or vice versa).

(112) **Predicate Abstraction (PA)**:

$[\![\lambda_i\ \text{S}]\!]^g = \lambda y :\ [\![\text{S}]\!]^{g[i\to y]}$ is defined. $[\![\text{S}]\!]^{g[i\to y]}$. (cf. Heim & Kratzer 1998:p.125)[11]

Finally, and importantly, I assume that quantifiers project definedness universally: when composing with predicates, the result is defined only if the predicate is defined for **every** individual in the restrictor of the quantifier ('universal projection').

(113) **Universal Projection from quantifiers**:

a. $[\![\text{Every student}]\!]^g = \lambda P_{\langle e,t\rangle} : \forall y \in \textsc{student}[y \in dom(P)].\ \forall y \in \textsc{student}[P(y) = 1]$

b. $[\![\text{No student}]\!]^g = \lambda P_{\langle e,t\rangle} : \forall y \in \textsc{student}[y \in dom(P)].\ \forall y \in \textsc{student}[P(y) = 0]$[12]

[11] A predicate abstract is always defined. But for the next node up above λ_i, Functional Application requires the predicate abstract to be defined for its argument.

[12] I assume the same universal projection for indefinites (*a student*), which is famously a problematic assumption in the literature on presupposition projection (Heim 1982; Schlenker 2008; Sudo 2012, a.o.). Sometimes universal projection from the scope of indefinites is indeed felt to be present—and with certain triggers even obligatorily so (Charlow 2009), but in many cases the projection is weaker. As far as I am aware, there is no agreement about how to handle the puzzling projection behavior of indefinites and other existential quantifiers. In the present system, assuming (possibly restricted) universal projection is necessary because of the unique way local contexts are compositionally built (see right below); encoding merely existential projection will jam the process of deriving local contexts for the scope of indefinites. For concreteness, I assume that indefinites indeed encode universal projection from their scope, but can sometimes flexibly allow that domain to be narrowed down even when other quantifiers cannot (Schlenker 2008 makes that claim too).

4.2.2 Local Contexts

Local Contexts for embedded constituents will be now derived from global contexts (assignments) and semantic values of the LF ingredients. Schlenker's (2009) insight is that the local context for some embedded constituent ψ (for the purpose of presupposition satisfaction and projection) is the most minimal object one may restrict attention to when evaluating the contribution of ψ to the overall truth conditions of the root sentence, given information up to the point where ψ is evaluated. I will do something similar, though not quite the same, as I will not talk about the minimal contribution to *truth* conditions but to *definedness* conditions. I will translate Schlenker's insight to the present system as follows: the local context of ψ is the smallest set of assignments that extend the global context (i.e. the starting assignment) with individuals recorded by λ-operators that c-command ψ.

Let me give the intuition before the details. Look again at the LF in (106a): *Mary* λ_i [$_{vP}$ t_i *did her*$_i$ *homework]*, and consider the following question: If we evaluate this whole LF against a contextual assignment g, how does information encoded by the binder *Mary* minimally affect this g so as to produce a modified context (assignment) with which to evaluate *v*P (and everything inside *v*P)? The answer to that will be the Local Context of *v*P. And the intuition is that, if we start, say, with the empty assignment $\emptyset$, the minimally required change to the initial assignment is one that produces the assignment $\emptyset^{[i \to Mary]}$, i.e. the one which is just like the starting assignment except modified so as to map i to *Mary*. The intuition (to be technically implemented presently) is summarized in (114).

(114) In (106a) (to be derived below):

a. Starting assignment: $\emptyset$

b. Local Context of *vP* (and anything in it): $\{ \emptyset^{[i \to Mary]} \}$

Similarly, consider the same question for the LF of a quantified sentence as in (107a): *Every/no student* λ_i [$_{vP}$ t_i *did her*$_i$ *homework]*. How does information encoded by the binder *every/no girl* minimally affect the initial context so as to produce assignments that can evaluate *v*P? Since a

quantifier by hypothesis imposes a universal demand on its scope (the scope must be defined for every individual quantified over), we are required to minimally modify g to a *set* of assignments, a set that contains one modification per individual in the domain of quantification—one for each student:

(115) In (107a) (to be derived below):

a. Starting assignment: $\emptyset$

b. Local Context of *vP* (and anything in it): $\{ \emptyset^{[i \to y]} \mid y \text{ is a student} \}$

We will now derive these desiderata compositionally. Readers who are not interested in the technical details and trust that the system works may want to skip to section 4.2.4.

LC Definition

We follow Schlenker 2009 in making use of a special syntactic variable over contexts, which can be embedded inside LFs and which I call LC ('Local Context') variable. A LC variable ranges over sets of assignment functions. I assume that there is only one such variable in the object language—one is all we will ever need—and it will be named C'. C' can attach to any phrase ψ. When it does, I notate it like so: $[^{C'}\psi]$. The purpose of this device is to allow us to target one specific constituent in an LF and extract all the necessary information for deriving its local context. To keep the system well-behaved, I assume that no more than one phrase can be adjoined with C' in one and the same LF.

Since technically C' is a variable, it itself needs an assignment for interpretation. I employ a special assignment function which I mark as $[C' \to G]$. It stands for 'G is the value of C' '. So now our interpretation function is relative to both g and $[C' \to G]$.[13] Here is how C' is interpreted:

(116) **LC Variable Interpretation**

For any assignment g' and a context G (set of assignments) that C' denotes:

$[\![^{C'}\psi]\!]^{g', [C' \to G]}$ is defined iff $\boldsymbol{g' \in G}$ and $[\![\psi]\!]^{g', [C' \to G]}$ is defined. If defined, $= [\![\psi]\!]^{g', [C' \to G]}$

[13] And a world parameter, but for now worlds don't play a role (see footnotes).

In words, (116) imposes a definedness condition on the relationship between the designated value of C' when it attaches to some ϕ in an LF, and the active assignment function that interprets that ψ: the latter must be a member of the former. Put differently, the contribution of the local context variable is to record the (ordinary) assignment function which the constituent in question is evaluated against.[14]

Apart from (116), the new parameter $[C'\rightarrow G]$ is idle, though we need to update the rest of the system to respect sensitivity to this parameter. Lexical denotations and the composition rules are cast in a rather boring way in terms of $[C'\rightarrow G]$. Predicate Abstraction, for example, is defined as follows:

(117) **Predicate Abstraction (PA)** (LC version):

$$[\![\lambda_i\ \mathrm{S}]\!]^{g,\ \mathbf{[C'\rightarrow G]}} = \lambda x :\ [\![\mathrm{S}]\!]^{g[i\rightarrow y],\ \mathbf{[C'\rightarrow G]}} \text{ is defined. } [\![\mathrm{S}]\!]^{g[i\rightarrow y],\ \mathbf{[C'\rightarrow G]}}. \qquad \text{(cf. 112)}$$

With this much in place, I now define Local Contexts for embedded constitutes, in (118). Notation: when $\mathcal{P}$ is a set of sets with a unique smallest element, $\textsc{min}(\mathcal{P})$ is the smallest set in $\mathcal{P}$.

(118) **Local Contexts (LOCO)**. In an LF [$_{\mathrm{S}}$ ... ψ ...] evaluated against global assignment g, the *Local Context* of ψ is notated $\mathrm{LOCO}_{g,\mathrm{S}}(\psi)$ and is defined as follows:

$$\mathrm{LOCO}_{g,\mathrm{S}}(\psi) := \textsc{min}\Big(\big\{G \ \big|\ [\![\mathrm{S}\ldots\ [^{C'}\psi]\ldots]\!]^{g,\ C'\rightarrow G} \text{ is defined}\big\}\Big)$$

That is, to determine the local context of some embedded constituent ψ, append the LC variable to ψ and find the smallest value for it that will make the whole LF defined w.r.t to the global assignment.[15] In particular the LC Variable Interpretation rule in (116) makes sure that $\mathrm{LOCO}_{g,\mathrm{S}}(\psi)$ includes one assignment extension of g for every individual that 'binds' a variable in ψ.[16]

[14] World-sensitive version of (116):

(i) LC Variable Interpretation (with worlds)
For any $< w', g' >$ *and a context G that C' denotes:*
$[\![^{C'}\psi]\!]^{w',g',\ [C'\rightarrow G]}$ is defined iff $\mathbf{< w', g' >\in G}$ and $[\![\psi]\!]^{w',g',\ [C'\rightarrow G]}$ is defined. If defined, $= [\![\psi]\!]^{w',g',\ [C'\rightarrow G]}$.

[15] Given our set up, in particular given universal projection of quantifiers and given that assignment modifications always extend and never rewrite information in the original assignment (cf. (109)), there will always be a smallest value for G, no matter what is the starting assignment.

[16] The definition of LOCO in (118) takes a single assignment as the starting context. As mentioned in the footnotes, it is more appropriate to represent an initial context C as a set of world-assignment pairs, in order to model instances

4.2.3 Illustration

Let us run a couple of examples to illustrate the workings of the system, one with binding by individual-denoting phrases (119a), one with binding by quantifiers (120a).

(119) Mary$_i$ did her$_i$ homework

a. LF: $\left[_{\text{S}}\ \text{Mary}\ [\lambda_i\ [_{vP}\ \text{t}_i\ \text{did}\ \textit{pro}_i\text{'s homework}]]\ \right]$

(120) No student$_i$ did their$_i$ homework

a. LF: $\left[_{\text{S}}\ \text{No student}\ [\lambda_i\ [_{vP}\ \text{t}_i\ \text{did}\ \textit{pro}_i\text{'s homework}]]\ \right]$

Binding by individual-denoting phrases

I start with (119a). Assume the initial (global) assignment is the empty one $\emptyset$ and suppose we want to determine $\text{LOCO}_{\emptyset,\text{S}}(vP)$, the local context of vP. (118) instructs us then to attach C' to vP and find the smallest value for it, G, such that: $[\![\text{Mary}\ [\lambda_i\ \boldsymbol{C'}\ [_{vP}\ \text{t}_i\ \text{did}\ \textit{pro}_i\text{'s HW}]]]\!]^{\emptyset,\ C'\rightarrow G}$ is defined.

We can prove that every choice of G that will make the above defined will include the assignment $\emptyset^{i\rightarrow Mary}$. To see this, choose an arbitrary G and compute the definedness conditions:

(121) $[\![\text{Mary}\ \lambda_i\ [\boldsymbol{C'}\ _{vP}]]\!]^{\emptyset,\ [C'\rightarrow G]}$ is defined, iff (by FA)

$Mary \in dom\Big([\![\lambda_i\ [\boldsymbol{C'}\ _{vP}]]\!]^{\emptyset,\ [C'\rightarrow G]}\Big)$, iff (by PA in (117))

$Mary \in dom\Big(\lambda y : [\![[\boldsymbol{C'}\ _{vP}]]\!]^{\emptyset^{i\rightarrow y},\ [C'\rightarrow G]}$ is defined. $[\![[\boldsymbol{C'}\ _{vP}]]\!]^{\emptyset^{i\rightarrow y},\ [C'\rightarrow G]}\Big)$, iff

$[\![[\boldsymbol{C'}\ _{vP}]]\!]^{\emptyset^{i\rightarrow mary},\ [C'\rightarrow G]}$ is defined, iff (by LC interpretation (116))

$\emptyset^{i\rightarrow mary} \in G$.[17]

of uncertainty of referents and presuppositions. Once we do that, and after we incorporate worlds into the composition rules and the semantic denotations of the pieces (see especially fn.14), the definition of LOCO could be updated in the following way, where LOCOs will now be sets of assignment-world pairs instead of sets of assignments:

(i) $\text{LOCO}_{C,\text{S}}(\psi) := \bigcup_{<w,g>\in C} \text{MIN}\Big(\big\{C' \mid [\![_{\text{S}}\ldots\ [^{\text{C'}}\psi]\ldots]\!]^{w,g,\ \text{C'}\rightarrow C'}\ \text{is defined.}\big\}\Big)$

This says that to get the LOCO of an embedded constituent, do for each $< w, g >$ in the starting context what (118) originally instructed, and then collect the result. This gives back a set of world-assignment pairs. In the main text I will mostly keep to the simpler version unless worlds are necessary.

[17] Actually, there's a missing conjunct in the last line in (121); it should read (addition underlined):
$\emptyset^{i\rightarrow mary} \in G$ and $[\![\text{vP}]\!]^{\emptyset^{i\rightarrow Mary},\ [C'\rightarrow G]}$ is defined. But it is harmless to ignore this conjunct; its satisfaction will

If every G must contain $\emptyset^{i \to Mary}$, the smallest G is just $\{\emptyset^{i \to Mary}\}$. In other words, the smallest set of assignments that makes $[\![\text{Mary } \lambda_i\ [^{\boldsymbol{C'}}{}_{vP}]]\!]^{\emptyset,\ C' \to G}$ defined is just $\{\emptyset^{i \to Mary}\}$. So that is the LOCO of vP in this LF, under an initial empty assignment (the conclusion is not affected by a different choice of an initial assignment, see below).

(122) In (119a):

a. Starting assignment: $\emptyset$

b. $\text{LOCO}_{\emptyset,S}(vP)$: $\{\emptyset^{[i \to Mary]}\}$

Technical remarks

A couple of remarks are worth mentioning. First, given the notation for interpreting the LC variable as a definedness condition (cf. 116), and the nature of the composition rules which project definedness conditions up, more than one positions for placing C' in the syntax could result in one and the same LOCO. In particular, if we adjoined C' to any other position inside the vP in (119a), for example on the DP pro_i, the end result would be the same.

(123) Also in (119a):

a. Starting assignment: $\emptyset$

b. $\text{LOCO}_{\emptyset,S}([\textbf{DP } \boldsymbol{pro_i}])$: $\{\emptyset^{[i \to Mary]}\}$

In general, two positions for C' differ in terms of the resulting LOCO only if there is a λ-binder in between them.

Second, had we started in (121) with a more informative global assignment, i.e. one in which some referents have been introduced, e.g. $\emptyset^{k \to Sue}$, the resulting LOCO would of course be different, but not in any devastating way. In that case the LOCO of our vP can be proven to be $\{\emptyset^{[k \to Sue,\ i \to Mary]}\}$. As long as assignment modification does not allow an index to be rewritten

not constrain or change the possible values for G further. Recall that by hypothesis the VC variable C' is only ever embedded once in an LF, on the constituents whose LOCO we are after; so $[C' \to G]$ is going to be idle for constituents inside vP, hence G isn't going to record any more assignments. The point generalizes to other possible positions for placing C', though proving this is tedious.

(as ensured by the definition of assignment extension, cf. (109)), there will always be a minimal modification that satisfies definedness, and therefore there will always be a LOCO.

Binding by quantificational phrases

Next I illustrate the system for the scope of quantifiers, (120a), w.r.t initial $\emptyset$. Following the recipe in (118), we again attach C' to *v*P (or anything within it) and examine what is the smallest set of assignments G such that: $[\![$No student $[\lambda_i$ **C'** $[_{vP}$ t$_i$ did *pro*$_i$'s HW$]]]\!]^{\emptyset,\ C'\rightarrow G}$ is defined.

Choosing an arbitrary G, we can be convinced that it has to include all modifications of $\emptyset$ of the form $\emptyset^{i\rightarrow y}$, for any student y. This is ultimately because of universal projection of quantifiers. Here is the computation:

(124) $[\![$No student λ_i $[$**C'** $_{vP}]]\!]^{\emptyset,\ [C'\rightarrow G]}$ is defined, iff (by FA+ quantifiers (113))

$\forall y \in$ STUDENT : $y \in dom\Big([\![\lambda_i$ $[$**C'** $_{vP}]]\!]^{\emptyset,\ [C'\rightarrow G]}\Big)$, iff (by PA)

$\forall y \in$ STUDENT : $[\![[$**C'** $_{vP}]]\!]^{\emptyset^{i\rightarrow y},\ [C'\rightarrow G]}$ is defined, iff (by LC notation (116))

$\forall y \in$ STUDENT : $\emptyset^{i\rightarrow y} \in G$.[18]

Furthermore, the *smallest* such G is one that contains *only* those extensions that map i to some student; all other assignments can be dispensed with for computing definedness.

(125) In (107a):

a. Starting assignment: $\emptyset$

b. Local Context of *vP* (and anything in it): $\{\ \emptyset^{[i\rightarrow y]} \mid y$ is a student$\}$

More than one binder

The system also derives the intuitively correct LOCOs for more than one λ-binder in an LF. Here I present two examples and state the end result (the reader is invited to prove that the claims are correct).

[18] Once again I am harmlessly dropping a conjunct (see the previous footnote). The fourth line in (124) should strictly read (addition underlined): $\forall y \in$ STUDENT : $\emptyset^{i\rightarrow y} \in G$ and $[\![$vP$]\!]^{\emptyset^{i\rightarrow y},\ [C'\rightarrow G]}$ is defined. But *v*P does not embed C', so G is never going to record further assignments beyond those that the students record.

(126) Every girl$_i$ told Bill$_j$ that she$_i$ likes him$_j$.

a. LF: Every girl λ_i Bill λ_j [t_i told t_j that pro_i likes pro_j].

b. LOCO of constituents below λ_j (on initial $\emptyset$): $\left\{\emptyset^{[i\to y, j\to Bill]} \;\middle|\; y \text{ is a girl}\right\}$

(127) Every girl$_i$ told every boy$_j$ that she$_i$ likes him$_j$.

a. LF: Every girl λ_i every boy λ_j [t_i told t_j that pro_i likes pro_j].

b. LOCO of constituents below λ_j (on initial $\emptyset$): $\left\{\emptyset^{[i\to y, j\to z]} \;\middle|\; y \text{ is a girl, z is a boy}\right\}$

The next three sections are devoted to making what I said until now more precise. Here is a summary of the desiderata from the system:

4.2.4 Recap

As an intermediate summary, we have constructed a system for computing Local Contexts in (simple cases of) bound anaphora, where some of our desiderata from before are met:

(128) **Desiderata from a system of Local Contexts for bound anaphora** (cf. 105):

a. i. When a pronominal variable i is bound by a type-e element, the local context (set of assignments) of the scope of X maps i to the meaning of its binder.

ii. When a pronominal variable i is bound by a quantifier (type-et, t), the local context of the scope of X maps i to an individual in the domain of the quantifier.

As I remarked in the beginning of this section, this is an impoverished notion of a local context but one that will do for us to get a system going. To faithfully model presupposition projection/satisfaction in this system (Schlenker 2009; Singh 2011), it is necessary to add worlds and talk about initial contexts as set of world-assignment pairs (see footnotes 14,16 and section 4.5).

4.3 Feature Valuation from Local Context

4.3.1 The valuation rules

Now that we have the more sophisticated Local Contexts (LOCOs) at our disposal, we can generalize the notion of 'valuation from context' to refer to this object. I repeat from (101):

(129) D-heads start the derivation unvalued for ϕ-features: $\text{D}_{[num:\_\_,\ pers:\_\_,\ gend:\_\_]}$.
Let $\mathbf{LOCO_{g,S}(DP)}$ be the Local Context of DP (given g and the whole LF, S).

a. **Number valuation from context**:

i. If $\forall g' \in \text{LOCO}_{g,S}(DP)$: $[\![\text{DP}]\!]^{g'}$ is an atomic individual, then $\text{D}_{[num:\underline{\textbf{sg}}]}$.[19]

ii. Else, $\text{D}_{[num:\underline{\textbf{pl}}]}$.

b. **Gender valuation from context** (English pronouns):

i. If $\forall g' \in \text{LOCO}_{g,S}(DP)$: $[\![\text{DP}]\!]^{g'}$ is identified as female atomic individual, then $\text{D}_{[num:\textbf{sg},\ gen:\underline{\textbf{fem}}]}$.

ii. If $\forall g' \in \text{LOCO}_{g,S}(DP)$: $[\![\text{DP}]\!]^{g'}$ is identified as male atomic individual, then $\text{D}_{[num:\textbf{sg},\ gen:\underline{\textbf{masc}}]}$.

iii. Else, $\text{D}_{[gen:\_\_]}$.

c. **Person valuation from context**:

i. If $\forall g' \in \text{LOCO}_{g,S}(DP)$: $[\![\text{DP}]\!]^{g'}$ includes the speaker, then $\text{D}_{[...\ pers:\underline{\textbf{1}}]}$. Else,

ii. If $\forall g' \in \text{LOCO}_{g,S}(DP)$: $[\![\text{DP}]\!]^{g'}$ includes a conversational participant, then $\text{D}_{[...\ pers:\underline{\textbf{2}}]}$.

iii. Else, $\text{D}_{[...,\ pers:\_\_]}$.

[19] Versions of the rules in (129) that don't ignore worlds (see previous footnotes, especially fn. 16), look as follows:

(i) If for every pair $< w', g' > \in \text{LOCO}_{g,S}(DP)$, $[\![\text{DP}]\!]^{w',g'}$ is ..., then ....

'

4.3.2 Checking the predictions of the system

Gender assignment on pronouns

The system now produces the result that when a feature-less pronoun is bound by an individual-denoting phrase like 'Mary', referring to a female individual, a pronoun bound by it will be valued FEM for its gender feature. This is summarized in (130).

(130) *Mary submitted her paper*

a. LF: $\left[_{S}\text{ Mary }[\lambda_i\ [\text{t}_i\text{ submitted }\textit{pro}_i\text{'s paper}]]\ \right]$

b. $\underline{\textit{pro}_i}$: $[_{D_{[num:\_\_,\ gend:\_\_,\ pers:\_\_]}}\ i]$

c. $\underline{\mathbf{LOCO}_{g,S}(\textit{pro}_i)}$: $\{g^{[i\rightarrow Mary]}\}$ (=122)

d. After Valuation from LOCO: $[_{D_{[num:\underline{\mathbf{sg}},\ gend:\underline{\mathbf{fem}},\ pers:\_\_]}}\ i] \Rightarrow \left/\text{her}\right/$ (by (129b))

If the binder is a quantifier, e.g. 'every student', valuation will depend on the identity of the individuals in the domain of quantification. If (it is known that) the domain consists just of female-identified individuals, then again we'll get FEM on a bound pronoun.

To cash this out we'll need to restrict the domain of quantification to only those that are contextually salient (as in any theory of domain restriction of quantifiers). I assign a subscript *RD* ('restricted domain') on the determiner to do the job.

(131) *Every student submitted her paper*

a. LF: $\left[_{S}\text{ Every}_{RD}\text{ student }[\lambda_i\ [\text{t}_i\text{ submitted }\textit{pro}_i\text{'s paper}]]\ \right]$

b. $\underline{\textit{pro}_i}$: $[_{D_{[num:\_\_,\ gend:\_\_,\ pers:\_\_]}}\ i]$

c. $\underline{\mathbf{LOCO}_{g,S}(\textit{pro}_i)}$: $\{\ g^{[i\rightarrow y]}\ \big|\ y \text{ is a student in } [\![\text{RD}]\!]^g\}$ (cf. 125)

d. After Valuation from LOCO:

If every student y in $[\![\text{RD}]\!]^g$ is female, then $[_{D_{[num:\underline{\mathbf{sg}},\ gend:\underline{\mathbf{fem}},\ pers:\_\_]}}\ i] \Rightarrow \left/\text{her}\right/$.

If every student y in $[\![\text{RD}]\!]^g$ is male, then $[_{D_{[num:\underline{\mathbf{sg}},\ gend:\underline{\mathbf{masc}},\ pers:\_\_]}}\ i] \Rightarrow \left/\text{his}\right/$.

So, the (global) context fixes the domain of quantification, and depending on its identity, one

and the same LF can be spelled out with different features. If it is known that the domain consists just of male-identified individuals the value will be MASC.

Person assignment on pronouns

For person, the story is trivially analogous. If a binder is 'I' (itself possibly a feature-less index, but one whose local context is not modified from the global one)—variables bound by it will be equipped with value 1ST for their person feature:

(132) *I submitted my paper*

a. LF: $\left[_S \text{ I } [\lambda_i \text{ [t}_i \text{ submitted } pro_i\text{'s paper]]} \right]$

b. $\mathbf{LOCO_{g,S}(pro_i)}$: $\{g^{[i \rightarrow \textit{the speaker}]}\}$

c. After Valuation from LOCO: $[\text{D}_{[num:\underline{\mathbf{sg}},\ gend:\_\ pers:\underline{\mathbf{1}}]}\ i] \Rightarrow \left/\text{my}\right/$ (by (129c))

Number on Full DPs

The more interesting cases is when a variable is bound into a full DP in need of features. But the story is not qualitatively different, as I now show.

(133) *Mary treated the* ***guy*** *she arrested with respect.*

As a reminder from chapter 3, the semantics of definite DPs doesn't prejudge number; it merely encodes maximality, presupposing maximality is satisfied (hence an existence presupposition of *the*). I repeat from earlier (because of the presupposition of *the*, it is necessary to go with the world-sensitive version of the system, see fns. 16 and 19):

(134) a. For any C and $< w, g > \in C$,

$[\![\text{the}]\!]^{w,g,C} = \lambda f \in D_{\langle e,t\rangle} : \text{MAX}_C(f)$ is defined. $\text{MAX}_C(f)$.

b. $\text{MAX}_C(f) = \quad \iota X[f(X)=1 \wedge X$ is salient in $C \wedge \forall Y$, if $f(Y)=1$ and Y is salient in C, then $Y \leq X]$;

Defined only if $\exists X[f(X)=1 \wedge X$ is salient in C $\wedge \ \forall Y$, if $f(Y)=1$ and Y is salient in

C, then $Y \leq X$].

($\text{MAX}_C(f)$ is the maximal contextually-salient individual that makes f true, assuming there is one)

The maximality presupposition must be satisfied in the (global) context for the LF to be defined and the sentence felicitous. This means that Mary must have arrested at least one (salient) guy in every world in C.

The process of valuation is laid out in (135). The LF representation (135a) serves as the input. Because the subject *Mary* binds the variable i into the DP, the local context of DP maps i (throughout the worlds) to Mary, (135b). Assume it is known in the context, i.e. throughout C, that Mary arrested just one guy (this was indeed established in the mixed-number scenarios in ellipsis we were considering earlier). Then, the semantic value of DP **against its LOCO** is an atomic individual throughout its LOCO. So the DP will end up with **sing** value for number, (135d).

(135) a. LF: $\left[_{\text{S}}$ Mary [λ_i [t_i treated [$_{\text{DP}}$ the GUY pro_i arrested] with respect]] $\right]$

b. $\underline{\mathbf{LOCO}_{C,\text{S}}(DP)} = \left\{ < w, g^{[i \to Mary]} > \;\middle|\; < w, g > \in C \right\}$ (cf. fn. 16)

c. Information from (global) context: Mary arrested exactly one guy. Therefore:

for every $< w', g' > \in \text{LOCO}_{C,\text{S}}(DP)$, $[\![\text{DP}]\!]^{w',g'}$ is an atomic individual.

(because: the semantic value of the DP in each $< w, g^{[i \to Mary]} >$ in (135b) is the maximal individual that satisfies the predicate $[\![$GUY that pro_i arrested$]\!]^{w,g^{[i \to Mary]}}$)

d. After Valuation from LOCO:

$[\text{D}_{[num:\underline{\text{sg}}]}$ the GUY pro_i arrested$]$ $\Rightarrow$ /the guy.../

Gender on full DPs

For gender on full DPs as in (136) the derivation is parallel, (137). The subject binds into a DP whose semantic value in its local context is female, so valuation results in FEM.

(136) Yosi$_1$ histader im [ha-mexanex-**et** šel-o$_1$]. (*Hebrew*)
Yosi got.along with [the-head.teacher-**FEM.SG** of-his$_1$].
⇝ *'Yosi got along with his head teacher, who is **female**'*

(137) a. LF: $\left[_{\text{S}}$ Yosi [λ_i [t_i got.along with [$_{\text{DP}}$ the TEACHER-**FEM.SG** of-pro_i]]] $\right]$

b. $\underline{\mathbf{LOCO}_{C,\mathrm{S}}(DP)} = \left\{ < w, g^{[i \to Yosi]} > \;\middle|\; < w, g > \in C \right\}$ (cf. 16)

c. Information from (global) context: Yosi's head teacher is female. Therefore:

for every $< w', g' > \in \mathrm{LOCO}_{C,\mathrm{S}}(DP)$, $[\![\mathrm{DP}]\!]^{w',g'}$ is (atomic) female.

(because: the semantic value of the DP in each $< w, g^{[i \to Yosi]} >$ in (135b) is the (maximal) individual that satisfies the predicate $[\![\textsc{teacher}\ \text{of-}pro_i]\!]^{w,g^{[i \to Yosi]}}$)

d. <u>After Valuation from LOCO</u>:

$[\mathrm{D}_{[num:\mathbf{sg},\ gend:\underline{\mathbf{fem}}]}\ \text{the}\ \textsc{teacher}\ \text{of-}pro_i] \Rightarrow$ /the teacher-**FEM**.SG of-his/

4.4 Explaining Fake Features in Ellipsis

After so much investment in technical details, we can finally cash out some of the empirical coins. We have all the tools to account for ellipsis mismatches with bound variables:

(138) *Feature Mismatches In Ellipsis* (repeated from (98b)):

a. I did **my** homework, and you did [$_{VP_\triangle}$ ~~do **your** homework~~], too.

b. Mary$_1$ has treated [$_{\mathrm{DP}}$ the **guy** she$_1$ arrested] with respect, but

ANN$_2$ hasn't [$_{VP_\triangle}$ ~~treat [$_{\mathrm{DP}}$ the **guys** she$_2$ arrested] with respect~~].

c. Yosi$_1$ histader im [ha-mexanex-**et** šel-o$_1$]. Avi lo △. (*Hebrew*)
Yosi got.along with [the-head.teacher-**FEM.SG** of-his$_1$]. Avi not △.
$\rightsquigarrow$ △ : *got along with his head teacher, who is* ***male***.

Given the build-up, the account is quite straightforward. Ellipsis requires identity of **meaning (semantic value)** with an antecedent – either at the level of the elided constituent itself, or, in the cases here, via a focus parallelism condition (cf. section 2.1.3). Crucially, as we just saw, the DP/pronouns with the offending features do not constrain the semantic value proper of the structure, but is added late. Because of that, the ellipsis clause and the antecedent clause respect identity at the level relevant for ellipsis licensing: the elided VP, for example in (138a), is semantically identical to its antecedent.

The relevant fact that makes Focus Parallelism satisfied is in (139). Focus Parallelism is stated in (140) (see section 2.1.3).

(139) $[\![_{\text{A}}\ \text{I}\ [\lambda_i\ \text{t}_i\ [_{\text{VP}}\ \text{did}\ \textit{pro}_i\text{'s homework}]]]\!]^g \in [\![_{\text{S}}\ \text{YOU}_F\ [\lambda_i\ \text{t}_i\ \text{did}\ [_{\text{VP}}\ \text{do}\ \textit{pro}_i\text{'s homework}]]]\!]^g_f$

(140) Focus Parallelism Condition on Ellipsis: Ellipsis of VP_E is licensed only if VP_E is embedded in a constituent, S, and there is another constituent in the surrounding discourse, A, such that $[\![\text{A}]\!]^g \in [\![\text{S}]\!]^g_f$ (and $[\![\text{A}]\!]^g \neq [\![\text{S}]\!]^g$).

This particular case is also accounted for by Feature Transmission approaches to fake features, as well as the Weak Presupposition Projection approach to features. But the current theory also explains (138b)-(138c) with the same mechanism, whereas the competitors do not. On the current approach, the singular-plural mismatch and the feminine-masculine mismatch are merely present at PF and do not constrain interpretation. But for the speaker who accept form-mismatches, identity of interpretation is all that ellipsis is sensitive to.[20]

4.5 Adding focus and alternatives

I now turn to fake features in association with focus alternatives. In 4.1.3 we saw that what needs to be ensured is that the addition of focus into the syntactic and semantic representations basically keeps intact the system of valuation we've had: a binder creates the same embedded context for valuation, no matter whether it is focused or not. Focus should matter for truth conditions, of course, as well as for other phenomena such as Congruence and Contrast, but I suggest that it is encapsulated from the process that builds local contexts for embedded expressions. It is as though the LF representation that local context construction reads is such that F-marking has been removed from it. This hypothesis is stated again in (141).

(141) **Context Encapsulation from Alternatives Hypothesis** (repeated from (104)):

[20] Recall that there are two populations w.r.t. form-mismatches (section 2.3 and in particular 2.3.2). I continue to assume that the other population requires on top of Parallelism a phonological-identity condition.

Information encoded (merely) in focus alternatives does not have an effect on the construction of embedded contexts.

A useful way to describe the desideratum that (141) is supposed to deliever is in (142), taking *only* as a representative for any focus-sensitive operator.

(142) **Desideratum**:

In an LF of the form $[...[only_{\text{ALT}}\ [...Z_F\ \phi\ ...]...]]$, where Z is focused and ϕ its sister, the LOCO of any constituent in ϕ (given any starting context) is the same LOCO of that constituent in $[...Z\ \phi\ ...]$ (given that starting context)

The way I will cash this out, once we have a system for generating focus alternatives, is to define that the C' variable—our local-context-producer—can be idle at the level where alternatives are computed. This will be done in subsection 4.5.2. As I have remarked, this part of the account is something that must remain a rock-bottom assumption, but in section 4.5.6 I will speculate on what it should follow from, given some truths about context update by linguistic expressions mroe generally. But it will be useful before doing that to first see how the system formally works.

The structure of this section is as follows: I first incorporate focus semantics, so that binding across focus alternatives can be compositionally captured (subsection 4.5.1). This will serve as the platform on which (in subsection 4.5.2) I encode (141) in the system of local contexts. Then I show how the resulting framework accounts for the interaction between focus, variable binding and assignment of features based on local contexts (subsection 4.5.4). In subsection 4.5.5 I will show that the system does not pose a threat to the behavior of genuine presupposition triggers in the scope of focus.

4.5.1 Focus semantics with distinguished focus-variables

The semantics for focus I adopt is based on the implementation in Büring (2016:ch.10) of the system developed in Kratzer 1991 and Wold 1996 (used also in Schwarzschild 1999; Beck 2016). It makes use of distinguished variables to represent focus in LFs and special assignment functions

that interpret those variables. Schematically, LFs for focus association look as in (143a), and the paraphrase that guides the composition (detailed below) is in (143b).[21]

(143) Only BOB works.

a. LF: $\left[\text{only } [_S \text{ Bob}_{F_7} \text{ works}]\right]$

b. Informal paraphrase: ⟦(143a)⟧ is true iff S is true on a focus-assignment that does not see the variable F_7 and false on all other (relevant) focus-assignments to F_7.

Foci as variables and Focus-assignment functions

Focused constituents are represented with a sub-scripted indexed F-mark. A focus assignment function h is a function from a numerical (F-)index to a denotation (not to be confused with 'ordinary' assignment functions g). Every constituent is semantically interpreted relative to h. h is only active when a constituent is focused. The interpretation scheme is given in (144).

(144) a. If α doesn't bear a F-index, $[\![\alpha]\!]^{w,g,h} = [\![\alpha]\!]^{w,g}$.

b. $[\![\alpha_{F_7}]\!]^{w,g,h} = \begin{cases} h(7), & \text{if } 7 \in dom(h); \\ [\![\alpha]\!]^{w,g,h}, & \text{otherwise} \end{cases}$

In words: for constituents that aren't F-marked, h is idle (144a). An F-marked constituent is assigned a value by h, provided that the F-index is in the domain of h; otherwise, the F-marking is 'switched off' and we resort to the normal semantic value, (144b).[22]

I reserve '$h_\emptyset$' to designate the empty focus-assignment function: one with an empty domain. By the definition in (144), $h_\emptyset$ has the property that it ignores any F-marking in the phrase that it interprets and always outputs the ordinary meaning of that phrase.

[21] The primary reason for going with the indexed-foci framework is that it allows for a simple way to integrate variable-binding in focus semantics, which obviously is crucial for the purposes here. In particular, defining the rule of λ-abstraction in more standard theories of the semantics of focus (Rooth 1985, 1992b) is known to create technical challenges (see e.g. Romero & Novel 2013) which are avoided here. Another advantages is that it allows for selective focus association (see fn. 23 later).

[22] F-indices are actually pairs $\langle n, \tau \rangle$ of a number and a semantic type, and the domain of a focus-assignment function is a set of such pairs. For convenience, however, we will represent only the number part of the index because it is easy to read the semantic type off of the expression it attaches to.

Each time a root LF is processed, the initial (i.e. global) focus-assignment function is $h_\emptyset$ (cf. Kratzer 1991; Wold 1996). In other words, the h resets every time a new sentence is evaluated:

(145) A root LF is only interpreted against $h_\emptyset$ (for any context C and any $< w, g > \in C$).

Composition

Composition rules work as usual to assign an (h-sensitive) interpretation to larger constituents; two of the rules are given below. They are the usual ones, except h has been added as a parameter.

(146) **Functional Application (FA)**:

$[\![\text{A B}]\!]^{w,g,h}$ is defined iff $[\![\text{A}]\!]^{w,g,h}$ is defined and $[\![\text{B}]\!]^{w,g,h}$ is defined and $[\![\text{A}]\!]^{w,g,h}$ is in the domain of $[\![\text{B}]\!]^{w,g,h}$ (or vice versa). If defined, $[\![\text{A B}]\!]^{w,g,h} = [\![\text{B}]\!]^{w,g,h}\left([\![\text{A}]\!]^{w,g,h}\right)$ (or vice versa).

(147) **Predicate Abstraction (PA)**:

$[\![\lambda_i\ \text{S}]\!]^{w,g,h} = \lambda y :\ [\![\text{S}]\!]^{w,g^{[i\to y]},h}$ is defined. $[\![\text{S}]\!]^{w,g^{[i\to y]},h}$.

In (148) is an illustration for a simple example computed using FA.

(148) $[\![\text{Bob}_{F_7}\ \text{works}]\!]^{w,g,h} = [\![\text{works}]\!]^{w,g,h}\left([\![\text{Bob}_{F_7}]\!]^{w,g,h}\right) = \begin{cases} h(7)\ works_w, & \text{if } 7 \in dom(h); \\ Bob\ works_w, & \text{otherwise} \end{cases}$

And in (149) is a computation of a pronoun-binding example (I leave out the definedness conditions for perspicuity).

(149) $[\![\text{Mary}_{F_7}\ [\lambda_i\ [_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s homework}]]]\!]^{w,g,h} =$ (by FA and PA)

$[\![_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s homework}]\!]^{w,g^{[i\to [\![\text{Mary}_{F_7}]\!]^{w,g,h}]},h} =$ (by (144))

$\begin{cases} h(7)\ did_w\ h(7)'s\ homework, & \text{if } 7 \in dom(h); \\ Mary\ did_w\ Mary's\ homework, & \text{otherwise} \end{cases}$

Note how the interpretation of the (ordinary) variable i in the scope of λ depends upon the identity of the h that the whole LF is computed against. This will produce binding across alternatives.

Constructing alternatives

The above supplied interpretations relative to a particular focus-assignment. To build focus alternatives (alternative sets), we abstract over possible h's. To do that, we first for convenience define the index-sensitive modifications of h, in (150). The notation defines what it means for an assignment to differ from another just in the value for a particular index.

(150) For any focus-assignment function h and F-index n,

$H^h_n := \{h' \mid h' \text{ is just like } h \text{ except for the value it assigns to } n.\}$

$= \{h' \mid h' = h^{[n \to x]}, \text{ for some } x \text{ (of the relevnt type)}\}.$

With this, we now define focus alternatives in (151). The definition makes reference to the notation in (150).

(151) **Variable-sensitive Focus Alternatives**:

For any constituent α, assignments g, h and F-index i:

$\boldsymbol{ALT^{g,h}_n(\alpha)} := \{\lambda w : [\![\alpha]\!]^{w,g,h'} \text{ is defined. } [\![\alpha]\!]^{w,g,h'} \mid h' \in H^h_n\}$

Alternative are thus partial propositions (as standard). I write '$\lambda w.[\![\alpha]\!]^{w,g,h'}$' as a shorthand for what comes before $\mid$ in (151). To illustrate:

(152) $\boldsymbol{ALT^{g,h}_7(\textbf{Bob}_{F_7}\ \textbf{works})} = \{\lambda w.[\![\text{Bob}_{F_7}\ \text{works}]\!]^{w,g,h'} \mid h' \in H^h_7\} = \{\lambda w.\ h^1(7)\ works_w,$
$\lambda w.\ h^2(7)\ works_w,$
$\lambda w.\ h^3(7)\ works_w, \ldots\}$

Depending on the identity of the relevant h's, this set could be resolved to, for example: {*Ann works*, *Jack works*, *Melissa works*}.

(153) concisely shows a computation of the focus alternatives of variable-binding structures, against $h_\emptyset$ as the focus-assignment function.

(153) $\boldsymbol{ALT^{g,h_\emptyset}_7(\textbf{Mary}_{F_7}\ [\lambda_i\ [_{vP}\ \textbf{t}_i\ \textbf{did}\ \textit{pro}_i\textbf{'s HW}]])} =$ (by (151))

$\{\lambda w.[\![\text{Mary}_{F_7}\ [\lambda_i\ [_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s HW}]]]\!]^{w,g,h'} \mid h' \in H^{h_\emptyset}_7\} =$ (by convention (150))

$$\left\{\lambda w.[\![\text{Mary}_{F_7}\ [\lambda_i\ [_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s HW}]]]\!]^{w,g,h^{[7\to x]}_{\emptyset}} \;\middle|\; x \text{ is an individual} \right\} = \text{(by (149))}$$

$$\left\{\ \lambda w : x\ has_w\ HW.\ x\ did_w\ x's\ HW \;\middle|\; x \text{ is an individual} \right\}$$

Focus-sensitive operators

The last piece needed for the semantics is the interpretation of focus-sensitive operators, which in this framework are a kind of binders of focus variables. Entries for *only* and Rooth's (1992b) '∼'-operator are provided below. They each come with a superscript (to their left), which marks what F-index (or indices) they manipulate.[23] They also come, as all operators, with a domain restriction variable, which is notated with a subscript 'A'. As before, I am being explicit about definedness.

(154) $[\![^{n}\!\sim_{A}\ \alpha]\!]^{w,g,h}$

is defined iff $[\![\alpha]\!]^{w,g,h}$ is defined $\wedge$ $A \subseteq ALT_n^{g,h}(\alpha)$ $\wedge$ $\forall\beta \in A\ [\beta(w)$ is defined].

If defined, $[\![^{n}\!\sim_{A}\ \alpha]\!]^{w,g,h} = [\![\alpha]\!]^{w,g,h}$.

(155) $[\![^{n}\textbf{only}_{A}\ \alpha]\!]^{w,g,h}$

is defined iff $[\![\alpha]\!]^{w,g,h}=1$[24] $\wedge$ $A \subseteq ALT_n^{g,h}(\alpha)$ $\wedge$ $\forall\beta \in A\ [\beta(w)$ is defined].

If defined, $[\![^{n}\textbf{only}_{A}\ \alpha]\!]^{w,g,h} = 1$ iff $\forall\beta \in A\ [\beta \neq [\![\alpha]\!]^{g,h} \rightarrow \beta(w)=0]$.

(In words: *'only α'* presupposes that the presuppositions of all the (relevant) focus alternatives to α are met and that α's ordinary semantic value is true. Asserts that the alternatives are false.)

A couple of things to pay attention to in these entries. First, after Rooth (1992b) (et seq.), the value of the domain variable A is constrained to be a subset of the focus structure of the prejacent (this is the second conjunct in the definedness conditions). Second, every proposition quantified over must be defined for the input world, therefore it must have its presuppositions met in the context; this corresponds to the intuition that presuppositions project universally from alternatives (see below, section 4.5.5).

[23] A focus-sensitive operator can selectively bind foci in its scope, in configurations of the kind discussed by Krifka (1991); Wold (1996); Beck (2016); Büring (2016). This system can model selective association because the operators indicate which F-index they manipulate. If an operator associates with more than one focus, this is done by letting the operator carry more than one index.

[24] This requires $[\![\alpha]\!]^{w,g,h}$ to be defined.

The semantics now produces the right interpretation for the LF in (156a) evaluated against $h_\emptyset$. An informal paraphrase of the end result is in (156b).

(156) Only MARY did her homework.

a. LF: $^{7}\text{only}_A$ [$_s$ Mary_{F_7} [λ_i [$_{vP}$ t_i did pro_i's homework]]]

b. *Mary did her homework, and every contextually-relevant alternative proposition in (153) is false.*

4.5.2 Local Contexts and Encapsulation from alternatives

In LF (156a), pro_i is of course an unrestricted variable; this is what in principle allows its value across negated alternatives to be resolved to male referents as well. But variables must be equipped with morphological features, and they do so based on their meaning in their local context.

We now add local context computation into what has been achieved thus far. In the current framework, computing local contexts for some embedded constituent involves adjoining the LC variable C' to that constituent and using this to retrieve the smallest possible object that ensures the result is defined. Suppose we minimally upgrade the earlier definitions of LC variable interpretation and LOCO to be sensitive to the computation of focus, as follows:[25]

(157) $\underline{\text{LOCO}_{C,S}(\psi)} := \bigcup_{<w,g>\in C} \text{MIN}\Big(\big\{C' \mid [\![S \ldots [^{C'}\psi] \ldots]\!]^{w,g,\,[C'\to C'],\boldsymbol{h_\emptyset}} \text{ is defined.}\big\}\Big)$ (cf. fn. 16)

(Take a world-assignment pair $<w, g>$ in the initial context C and compute the smallest set of world-assignment pairs that would make the whole LF defined against $<w, g>$. Do the same for all other assignment-world pairs in C, and collect the result.)

(158) LC Variable Interpretation, first version (based on fn. 14; **to be replaced with** (160))

For any $\boldsymbol{h}$, $< w', g' >$, *and a context* C' *that C' denotes:*

$[\![^{C'}\psi]\!]^{w',g',\,[C'\to C'],h}$ is defined iff $[\![\psi]\!]^{w',g',\,[C'\to C'],h}$ is defined and $< w', g' > \in C'$.

If defined, $= [\![\psi]\!]^{w',g',\,[C'\to C'],h}$.

[25] (157)-(158) are based on the definitions from the footnotes because here we will be careful to respect world-sensitivity. Also, in (157) the h parameter is $h_\emptyset$ because of (145): evaluating a root LF is always against $h_\emptyset$.

Let us assume for simplicity, but without loss of generality, that there is only one $<w, g>$ in the initial context C. The question then boils down to: what is the smallest value of C' that would make (159) below defined against $< w, g >$? (as before, it does not make a difference where exactly C' is placed in the scope of λ_i.)

(159) $^{7}\text{only}_{A}$ [$_{s}$ Mary_{F_7} [λ_i [$_{vP}$ t_i did $^{C'}pro_i$'s homework]]]

The definition in (158) would incorrectly produce that the LOCO for the scope of λ_i in (159) (and hence for pro_i) would have one (ordinary-)assignment per each relevant focus alternative to *Mary*, in addition to one for *Mary*. Rather, the LOCO for the scope of λ_i in (159) should only include assignments (read: world-assignment pairs) in which i's value is *Mary*.

As previewed earlier—and this is the central hypothesis of the current section—I propose that only information from prejacents (ordinary semantic values) are required for computing local contexts (cf. 141). I encode this by postulating that C', our Local Context variable, is idle under any focus-assignment except for $h_\emptyset$, the one that doesn't read F-marking (cf. the interpretation rules in (144)). This is achieved by replacing (158) with (160):

(160) **LOCOs ignore information encoded in focus alternatives** (replaces (158))

For any h, $< w', g' >$, and a context C' that C' denotes:

$[\![^{C'}\psi]\!]^{w',g',[C'\to C'],h}$ is defined iff $[\![\psi]\!]^{w',g',[C'\to C'],h}$ is defined, **and $< w', g' > \in C'$ if $h = h_\emptyset$**.

If defined, $= [\![\psi]\!]^{w',g',[C'\to C'],h}$.

The underlined part says that possible values for C' do not record ordinary-assignments that evaluate the constituent in question (ψ), when that constituent is computed against a focus assignment h which isn't the empty one $h_\emptyset$.[26] Since $h_\emptyset$ does not interestingly interpret foci at all and only sees prejacent information, the rule in (160) essentially dictates that only prejacent information can ever determine the local context for embedded constituents. This derives:

(161) In (159), $\mathbf{LOCO}_{C,s}(pro_i) = \left\{ < w, g^{[i \to Mary]} > \;\middle|\; < w, g> \in C \right\}$

[26] Actually (160) says that possible values for C' do not **all** record ordinary-assignments under non-empty h's; it doesn't exclude that some do. But because we eventually pick the *smallest* value for C' to compute the LOCO, this does not make much difference; superfluous values for C' will be thrown out. I thank Roger Schwarzschild (p.c.) for suggesting me to formulate (160) this way.

Illustration

To see how this works formally, we need to compute the definedness conditions of (159) against w, g, an arbitrary C' and $h_\emptyset$ (and then picking out the smallest such C'). Instead of going over the whole computation which is tedious, it will be enough to concentrate just on two pieces of the definedess condition of (159), namely the requirement that the scope of *only* be defined w.r.t the mentioned parameters, and that the alternatives quantified over, $ALT_7^{g,h_\emptyset}$, be defined as well.

(162) $[\![(159)]\!]^{w,g,\ C'\to C',h_\emptyset}$ is defined only if: (by (155))

a. $[\![\text{Mary}_{F_7}\ [\lambda_i\ ^{C'}[_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s hw}]]]\!]^{w,g,\ C'\to C',h_\emptyset}$ is defined, and

b. $[\![\text{Mary}_{F_7}\ [\lambda_i\ ^{C'}[_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s hw}]]]\!]^{w,g,\ C'\to C',h'}$ is defined for each relevant $h' \in H_7^{h_\emptyset}$.

We already essentially computed (162a) earlier, in (149). With some innocuous shortcuts, we have:

(163) $[\![\text{Mary}_{F_7}\ [\lambda_i\ ^{C'}[_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s homework}]]]\!]^{w,g,\ C'\to C',\ h_\emptyset}$ is defined iff (by FA, PA)

$[\![^{C'}[_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s homework}]]\!]^{w,g^{[i\to[\![\text{Mary}_{F_7}]\!]^{w,g,h_\emptyset}]},\ C'\to C',\ h_\emptyset}$ iff (by (144))

$[\![^{C'}[_{vP}\ \text{t}_i\ \text{did}\ pro_i\text{'s homework}]]\!]^{w,g^{[i\to Mary]},\ C'\to C',\ h_\emptyset}$ iff (by (160))

$< w, \boldsymbol{g^{[i\to Mary]}} > \in C'$.

As for (162b), the only difference from (162a) is that h has been extended to include some value for F_7. But that difference is critical because (160) does not (require to) record the resulting $< w', g' >$ in C'. For any $h \neq h_\emptyset$, $[\![^{C'}\psi]\!]^{w',g',\ [C'\to C'],h}$ is defined if $[\![\psi]\!]^{w',g',\ [C'\to C'],h}$ is defined. Therefore, the smallest value for C' that makes (159) defined is $\{< w, \boldsymbol{g^{[i\to Mary]}} >\}$, and we derive (161).

4.5.3 Recap

We have now obtained a system which achieves the central desiderata set out in the beginning of this chapter, repeated in (164):

(164) **Desiderata from a system of Local Contexts for bound anaphora and focus** (cf. (164)):

a. i. When a variable, i, is bound by a type-e element, the local context (set of assignments) of the scope of X maps i to the meaning of its binder.

ii. When a variable, i, is bound by a quantifier (type-et, t), the local context of the scope of X maps i to the meaning of some individual in the domain of the binder.

b. i. Focus Alternatives don't alter the picture; only information coming from the prejacent is relevant for the rules that construct local contexts.

4.5.4 Explaining fake features in association with focus

With this, the morphological valuation rules now successfully assign the right feature-values to variables and (more generally) variable-dominating DPs in the scope of focus.

On bound pronouns

We saw in (161) that when pro_i is bound by a focused *Mary*, its semantic value in its LOCO is Mary (because its LOCO is $\{< w, \boldsymbol{g}^{[i \rightarrow \textit{Mary}]} > \mid < w, g> \in C\}$). Therefore, the rules of valuation from before assign FEM to its gender feature. The derivation is transparently parallel to the one detailed in section 4.3.2, example (130), so there is no need to repeat it.

By the same reasoning, we get a fake indexical surfacing when its binder is 1st person pronoun:

(165) *Only I did my homework*

a. LF: $^{7}\text{only}_A$ [$_s$ I_{F_7} [λ_i [$_{vP}$ t_i did $^{(C')}pro_i$'s homework]]]

b. In (165a), $\underline{\mathbf{LOCO}_{C,s}(\boldsymbol{pro_i})} = \left\{< w, \boldsymbol{g}^{[i \rightarrow \textit{the speaker}]} > \mid < w, g> \in C\right\}$

The reader can verify that when the binder is a quantifier, features on a pronoun bound by it will be valued only based on information encoded by the ordinary semantic value of the quantifier:

(166) Only every FIRST year student completed **her** homework.

a. LF: $^{7}\text{only}_{A}$ [$_{s}$ [Every [first]$_{F_7}$ year student] [λ_i [$_{vP}$ t_i did $^{(C')}pro_i$'s homework]]]

b. $\underline{\mathbf{LOCO}_{C,S}(pro_i)} = \left\{ < w, g^{[i \to x]} > \mid < w, g > \in C, x \text{ is a first-year student} \right\}$

This, correctly, does not require every second/third/... year student to be female. It only predicts (correctly) that every first year student must be female.

On full DPs

Finally, the theory covers the examples in which a bound pronoun is bound into a DP in need of features. They are repeated from the beginning of the chapter (98a):

(167) **Fake Number**:

a. Only MARY$_1$ treated [$_{DP}$ the **guy** she$_1$ arrested] with respect.

$\rightsquigarrow \forall x \neq Mary$: *x didn't treat the* ***guy or guys*** *x arrested with respect.*

b. LF: Only $\left[_{S}\ \text{Mary}_{F_7}\ \lambda_1\ t_1 \text{ treated } ^{(C')}[_{DP} \text{ the GUY } pro_1 \text{ arrested}] \text{ with respect}\right]$

The DP in (167) is number-less at the base and gets valued by context. $\mathbf{LOCO}_{C,S}(DP) = \left\{ < w, g^{[1 \to Mary]} > \mid < w, g > \in C \right\}$; if the context entails that Mary has arrested exactly one guy, number is then valued for **sing**.[27]

(168) **Fake Gender**:

a. rak YOSI$_1$ histader im [$_{DP}$ ha-mexanex-**et** šel-o$_1$]. (*Hebrew*)
only YOSI got.along with [$_{DP}$ the-head.teacher-**FEM.SG** of-his$_1$]
'Only Yosi got along with his head (his was female).'

b. $\rightsquigarrow \forall x \neq Yosi$: *x didn't get along with x's head teacher,* ***male or female***.

[27] *the NP* presupposes that there exists NP (given the definition of MAX from earlier). The account here predicts that at the local context of DP this existence presupposition is met, when the global context entails that. This is because the set of worlds that comprise the local context of DP is a subset of the set of worlds that comprise the global context. If in each world in the global context Mary arrested just one guy, so she did in each world in the local context.

c. LF: Only $\left[_S\ \text{Yosi}_{F_7}\ \lambda_1\ t_1\ \text{got along with}\ {}^{(C')}[_{\text{DP}}\ \text{the teacher of}\ pro_1]\right]$

In (168), $\underline{\mathbf{LOCO_{C,S}(DP)}} = \left\{< w, \boldsymbol{g}^{[\mathbf{1}\rightarrow \textit{Yosi}]} > \ \middle|\ < w, g> \ \in C\right\}$; if the context entails that Yosi's head teacher is female, gender is valued for **fem**.

4.5.5 Presupposition triggers and universal projection

While deriving the basic desideratum, the hypothesis that information in alternatives does not affect local contexts (160) does not mean that *presuppositions* will show 'weak projection' in association-with-focus constructions. As desired, presuppositions will still universally project, essentially because their triggers are base-generated and therefore present in each alternative (unlike ϕ-features). They simply project from the **restrictor** (A) of *only* ϕ, which is constrained by the focus-semantic value of the prejacent. Recall that an LF of the form $only_A\ \phi$ presupposes that all the alternatives in A must be defined for the input world (cf. the definedndess conditions of *only* ϕ in (155)); this means, in the general case, that they must be met in the local context **of** $\boldsymbol{only_A}\ \phi$, which in simple unembedded cases is just the global context.

(169) Only MARY is talking to Ed **again**.

a. LF: $\text{Only}_A\ \left[\ \text{again}\ [_S\ \text{Mary}_{F_7}\ \text{is talking to Ed}]\ \right]$

b. $\text{A} \subseteq \left\{\lambda w : x\ \textbf{talked to Ed before}\ .x \text{ is talking to Ed now} \ \middle|\ x \text{ is an individual}\right\}$

c. LOCO of the whole LF: $\{w : w \in C\ (\text{the initial context})\}$[28]

d. Requirement of $Only_A$: $\forall w \in (169c), \forall \beta \in A\ [\beta(w)\ \text{is defined}]$.

That (169d) holds, and therefore universal projection follows, is the result of any theory that adopts something like (170):

(170) Local Satisfaction of Presuppositions (Heim 1983; Schlenker 2009; Singh 2011):

The presuppositions of each constituent must be entailed by its LOCO.

For each world w in the LOCO of XP, $[\![XP]\!](w)$ *must be defined.*

[28] Here I don't represent assignments because they are irrelevant.

Indeed, if we only looked at *again*'s very local context (or the local context of anything in the scope of *only*), then the information available will merely require that Mary talked to Ed before, given that the local context of the scope of *only* looks just at prejacent information. But the presuppositions ultimately project to the restrictor of *only*, and from there—due to (170)—they must project to the (global) context universally.

4.5.6 Context-encapsulation: a more general pragmatic condition?

The Context-Encapsulation hypothesis in (141), as encoded in (160), is a officially a stipulation, and will remain so; I have thus far provided no independent reason to go for (160) rather than (158). Before leaving this section, I'd like to make a remark about what I believe (160) could ultimately follow from, although I'll keep the discussion informal and speculative.

If one inspects the logic of what (160) does, one can see that its sole effect is to create an **asymmetry** between information encoded by overt material—the information that comes from an F-marked material as if it wasn't F-marked—and information encoded by its focus alternatives. Specifically, only the former ends up contributing to the construction of embedded contexts.

I believe that this effect should at the end of the day follow from a better understanding of the pragmatic conditions on context update, independently of focus and more generally than just for *local* contexts. Specifically, it should at least partly reduce to the fact that information encoded in an uttered expression is (arguably) much more **salient** than information encoded in its focus alternatives—at least it has been *made* to be highly salient right after it has been uttered.

In general, independently of focus, there is a correlation between the degree of saliency of some piece of information in a discourse context and its potential to provide antecedents for anaphoric devices—pronouns, definite descriptions, elided material. If the conversation was about Anna two minutes ago, and is now about Sue, then an utterance of "she" made in the conversation now is much more likely to refer to Sue than to Anna.

It is also known that some anaphoric devices even go as far as demanding, or highly preferring, that their antecedent be overtly expressed—for instance, VP ellipsis normally requires an overt

antecedent (Hankamer & Sag 1976); merely implied information is not enough for ellipsis of VP to be licensed. Anaphoric (prosodically weak) pronouns are also known to tolerate overt antecedents much more readily than merely implied antecedents; maybe this is what underlies Heim's (1982) famous *marbles* examples (attributed to Barbara Partee) and others like it:

(171) a. I dropped ten marbles and found all except for one. It is probably under the sofa.

b. I dropped ten marbles and found only nine of them. ??It is probably under the sofa.

(172) a. John has a wife. I'll meet*'er* tomorrow.

b. John is married. ??I'll meet*'er* tomorrow.

The asymmetry between the prejacent and its alternatives that (160) encodes, then, may thus be the result of the asymmetry in their salience status. Put differently, the suggestion is that the high degree of saliency of content expressed by an overt expression compared to its focus alternatives explains why only the former has the privilege to affect context update for subsequent constituents.

At the moment this suggestion is not even a beginning of a theory. To explore it seriously would require explicating 'saliency', the way it affects context update, and the relationship of all this to focus – something which is far beyond the scope of this thesis.[29] Nevertheless, once we have a good grasp on these issues and other pragmatic conditions on context update, I contend that (160) should eventually be reduced to those conditions.

Be that as it may, I continue to just assume (160).

[29]The fact that the system here is static and only models *local* context update is also of little help towards a reduction of (160).

4.6 offshoot

The main goal of this chapter has been achieved. Before I leave it and turn my attention to Fake Feature in relative clauses (chapter 5), I briefly go over two issues that, I argue, the current proposal can shed new light on. One issue has to do with split-binding (section 4.6.1), the other with the theory of reflexive pronouns and principle A of the binding theory (section 4.6.2).

4.6.1 Split-bound pronuons (Rullmann 2004; Heim 2008)

The system developed here allows us to talk about the local semantic value of a pronoun for the purpose of feature realization. The phenomenon of Split Binding (Partee 1989, Rullmann 2004, Heim 2008) exhibits a particularly nice illustration of the usefulness of this perspective.

'Split Binding' refers to cases where a pronoun has two separate antecedents, as in (173). It can be modeled by letting a pronoun be a spell out of two (or more) indices at LF that are summed together, (173b).

(173) Every girl told John that **they** should get together.

a. split-bound reading: Every girl x told John that ***x+John*** should get together

b. LF: Every girl λ_7 t_7 told john that $\underbrace{\textbf{pro}_{[7+8]}}_{\textbf{they}}$ should get together $(g(8) = john)$

As Rullmann (2004) and Heim (2008) showed, 1st (and 2nd) person pronouns can also be split-bound in assocation with focus constructions. For example, (174) uttered by one of John's ex-wives to John's other ex-wives has the reading in (174a). Once more, this reading can be captured with a complex structure for the pronoun, as in the LF in (174b) where one part is bound by the focused I_F.

(174) ("All of us wanted to separate from John on peaceful terms, but)...
Only I hoped **we** would eventually reunite" (based on Heim 2008:52)

a. Split-bound reading: No ex-wife x besides me hoped ***x+*****john** would reunite

b. LF: only [I_F λ_7 t_7 hoped $\underbrace{\textbf{pro}_{[7+8]}}_{\textbf{we}}$ would reunite] (where $g(8) = \textit{john}$)

Observe that the relevant reading in (174) *requires* the split-bound pronoun to surface as 1ST-person; pronouncing *you* or *they* instead of *we* loses the split-bound reading. Obviously this has to do with the fact that the binder of one part of the pronoun is a 1st person pronoun. More generally, it can be shown using parallel examples with different antecedents that the way that split-bound pronouns are morphologically realized follows the generalization in (175):

(175) *Generalization about the morphological realization of split-bound pronouns:*

a. If one part of a split-bound pronoun **refers to, or is bound by,** a 1ST person element, the split-bound pronoun is spelled out as 1ST-PL. (e.g. in 174)

b. Else, if one part of the split-bound pronoun **refers to, or is bound by,** a 2ND person element, the split-bound pronoun is spelled out as 2ND-PL.

c. Else, the split-bound pronoun is spelled out as 3RD-PL. (e.g. in 173b)

The theory offered here predicts exactly this generalization. In the local context of the split-bound pronoun, both parts of the pronoun have semantic values, and the overall meaning (the sum) feeds valuation. That number is always plural on split-bound pronoun is because '+' denotes the sum formation, and sum formation yields a non-atomic individual.[30] As for the person value of the pronoun, it just tracks the valuation rules that we have: 1ST if the value includes the speaker, else 2ND if it includes an addressee, else 3RD. In 174) the value of index 7 in its local context is the speaker so the spell-out rules yield 1ST; in (173b) the value of 7 across its local context is some individual in the domain of the binding quantifier (and the other part of the pronoun refers to non-speaker, non-addressee), so the value has to be 3rd.

[30]Except the pathological case where an individual is summed with itself, which we can rule out by a separate condition.

4.6.2 Strict reflexives and condition A: a proposal

In this section I offer a solution within the framework developed here to a stubborn problem in binding theory, namely the problem of strict readings of reflexive anaphors (Fiengo & May 1994; Hestvik 1995; McKillen 2016).

Strictness and condition A

It has been observed that reflexive anaphors (*-self* forms) in English give rise to strict-sloppy ambiguities in ellipsis and association with focus.

(176) a. Only JOHN voted for himself. (*strict, sloppy*)

b. Mary defended herself, and John did too. (*strict, sloppy*)[31]

We can describe the challenge in terms of the proper formulation of condition A of the Binding theory—the condition that is supposed to predict the grammaticality of *self*-forms in English. How to state condition A in a way that would predict possible both a sloppy reading (the less problematic case) and a strict reading for a reflexive? Consider the two relevant LF representations that express the two readings. I'll illustrate with association with focus, but the major points carry over to ellipsis.

(177) a. Sloppy LF: Only [John$_F$ λ_1 t_1 voted for himself$_1$]

b. Strict LF: Only [John$_F$ voted for himself$_1$] $(g(1) = John)$

Condition A requires a local relationship between a reflexive pronoun and its antecedent.[32] But how exactly to state the relationship? The simplest generalization seems to go as follows:

[31] The literature contains some conflict of judgment on whether strict reflexives are possible and in which configurations (see Hestvik 1995 vs. McKillen 2016; Büring 2019). I find them very natural, especially when context supports. McKillen (2016) provided experimental evidence that naive speakers generally accept strict reflexives, in both association-with-focus and ellipsis contexts.With McKillen (2016), I will proceed on the assumption that they are good across the board.

[32] We're ignoring the so-called 'exempt anaphora' cases here (e.g. Reinhart and Reuland 1993) and cases of perspectival anaphora.

(178) **Generalization**: a reflexive pronoun must have the same semantic value as a local antecedent, but this requirement is suspended in focus alternatives.

Sauerland 2013, followed by McKillen 2016; Bruening 2019, proposed to derive this generalization within their Weak Presupposition Projection theory of alternatives. They assume that a *self*-morpheme is: (i) pure-presupposition trigger—requiring the co-arguments of the nearest predicate to have the same (ordinary) denotation, and (ii) the requirement can be lifted from alternatives, because *self* is a pure-presupposition trigger.

(179) *Pure-presupposition theory:* self *as a presupposition trigger*

$[\![\text{self}]\!] = \lambda R_{\langle e,et\rangle}\lambda x\lambda y : x = y.\ R(x)(y)$[33]

Thus a representation of the focus alternatives in *only MARY defended herself* that would adhere to the above generalization and express the strict reading is given in (180).

(180) *Pure-presupposition theory: co-valuation without binding (strict reading)*

a. Prejacent:

Mary defended herself$_x$ ($x = Mary$; no λ-binding by subject)

b. Focus Alternative:

i. **Sue** defended her~~self~~$_x$ ($x = Mary$)

ii. **Jill** defended her~~self~~$_x$ ($x = Mary$)

This theory does capture strict readings (sloppy readings are captured by a variable binding representation). It also goes against the spirit of the proposal in this thesis, which aims to eliminate that idea that there is sucha thing as weak presupposition projection.

But there is a more empirical difficulty with this proposal, which has to do with its reliance on the assumption that *self* is presuppositional. The presupposition that it purportedly imposes is about the co-arguments of a predicate. As has been noted by the above authors, this analysis therefore does not easily extend to configurations of Exceptional Case Marking (ECM), where *self*

[33] This entry implies that *self* attaches at LF to the predicate. There could be other formulation on which it is in situ (McKillen 2016).

and its antecedent are not even co-arguments of the same predicate. But *self* has a strict reading in ECM:

(181) Only BIDEN expects himself to win.
strict: Trump did not expect Biden to win.

This fact requires exceptional maneuvers to explain within the Pure Presupposition theory. This is not to say that it cannot technically be done (see Sauerland 2013), just that it makes the claim for a presupposition less appealing.

Proposal

I propose instead to view *self*-morphemes as syntactic D-features, much like ϕ-features, instead of presuppositoin-imposers. The guiding idea is that *self* is a purely morphological realization of a reflexivity feature that gets valued whenever the right syntacto-semantic conditions that Principle A describes are met. More specifically, reflexivity is a feature on a pronominal index that gets valued +, and thus realized morphologically, if it has a co-valued antecedent in its local context. If it doesn't, we resort to the non-reflexive pronominal form: the 'elsewhere' realization of the reflexivity feature. Instead of having a licensing condition on the cases where a reflexive pronoun can appear, we will have a derivational theory of reflexivity that works off of the valuation-from-context framework. *self* is inserted into the structure if the condition on its insertion—which refers to the meaning-in-context of the host pronominal index—is met.

The theory I've developed so far is rich enough to cash this out; importantly, both strict and sloppy readings are predicted to be realized with a reflexive. Whether a pronominal index is bound as in (177a) or not as in (177b), its semantic value in its local context will determine the shape it'll take.

Pronouns, including reflexives, enter the derivation unvalued for REFL feature. A representation that corresponds to the strict reading is generated as in (182) (*only* takes sentential scope as usual):

(182) only [$_{vP}$ Mary$_F$ defended pro$_{i_{[\text{REFL:}__]}}$]

This LF, and the information from context, serves as input to the Condition A rule. To state condition A, it would be useful to define two terms first. (183) defines what it means to be a local domain for the purpose of binding theory; (184) defines what it means for two DPs to be (semantically) co-valued given a local context.

(183) Local Binding Domain (for concreteness): The local binding domain of a DP is the smallest *v*P that dominates DP.[34]

(184) Local Co-valuation:

Let S be a root LF evaluated against (global) context C, and let DP1 and DP2 be two DPs dominated by some constituent X in S. Then,

DP1 is *co-valued with DP2 in X* iff $[\![\text{DP1}]\!]^{w',g'} = [\![\text{DP2}]\!]^{w',g'}$ in every $< w',g' > \in LOCO_{c,s}(X)$.

(DP1 and DP2 are co-valued in the local context of X if they co-refer throughout the local context of X.)

With this, the definition of condition A is in (185):

(185) **Condition A**:

Let $\text{pro}_{i_{[\text{REFL:}\_\_]}}$ be a pronominal variable in a derivation S in context C, and let X be the syntactic local domain of $\text{pro}_{i_{[\text{REFL:}\_\_]}}$:

a. $\text{pro}_{i_{[\text{REFL:}\pm]}}$, if $\text{pro}_{i_{[\text{REFL:}\_\_]}}$ is co-valued with a c-commanding DP in X.

b. $\text{pro}_{i_{[\text{REFL:}\_\_]}}$, otherwise.

(Value the reflexively feature with +, if the pronominal variable has a local c-commanding DP co-valued with it locally. Otherwise, leave unvalued.)

How does this work? given the strict-reading LF in (177b), we can determine that pro_i will be valued + for reflexivity, iff the context of utterance assigns *Mary* as its semantic value. This is because, in that case the two DPs will be locally (and globally, as it happens) co-valued in the *v*P, which is the local domain of the variable. This accounts for strict reflexives. The story about strict reflexives in ellipsis is analogous.

(186) a. Base generate: only [Mary_F defended pro_i]

[34]The details of the local domain depends on one's favourite syntactic theory, and should not matter for our purposes.

b. Information from Context: $g_c = [i \rightarrow Mary]$.

c. After checking for condition A: $\text{pro}_{i_{[\text{REFL:}+]}}$

d. PF: *herself* (with assignment of the gender feature-value too.)

With the same LF representation the variable could be left unchecked, and therefore be spelled out with a non-reflexive, if the information from context was different, namely if pro_i did not refer to Mary across the local context.

Turning to sloppy readings, the LF now has variable binding; but the valuation procedure is the same.

(187) only $[_{vP}\ \text{Mary}_F\ \lambda_i\ t_i$ defended $\text{pro}_{i_{[\text{REFL:}\_]}}]$

Here too, the pronominal index in the object position has a local c-commanding DP—two, in fact—which are co-valued with it locally: on every assignment in the LOCO of *v*P, the object and the subject co-refer. This crucially obtains, remember, because focus values don't matter for the construction of local contexts.

(188) In (187), $[\![\text{subject}]\!]^{w',g'} = [\![\text{object}]\!]^{w',g'} = Mary$ in every $< w',g' > \in LOCO_{c,s}(vP)$.

Thus rule (185) assigns a + to pro_i in (187).

ECM in (181) is not an issue anymore. Condition A refers to the syntactic local domain of the relevant variable (cf. 183), and in ECM constructions, according to syntactic theory, the local domain includes the pronoun's potential antecedent. That part of the condition for the insertion of *self* is syntactic, not semantic.

I have constructed the system in such a way that reflexives **compete** with non-reflexive pronouns for insertion at PF. If the condition for insertion of the reflexive feature is met, the rule says that one must insert *self*. This predicts that there can never be cases where a non-reflexive pronoun is co-valued with its antecedent in the (local) context. There is a famous class of examples due to Heim (2007), following Evans (1980) and Reinhart (1983), that aims to show that this prediction is not borne out. These are so-called obviation of condition B examples:

(189) You know what Mary, Sue and John have in common?

Mary admires John$_1$, Sue admires him$_1$, and **John** admires **him**$_1$ too!

If the competition story here is correct, this must mean that examples like (189) involve a sophisticated representation of context on which the two DPs are not co-valued; that is the only way condition A in (185) allows a non-reflexive to surface. That there is no co-valuation here has been suggested also by Schlenker 2005; Grodzinsky & Sharvit 2007, but it remains to be seen whether this is ultimately right.

5

Fake features in relative clauses

The rules of logical syntax must follow of themselves, if we only know how every single sign signifies.

– Ludwig Wittgenstein, *Tractatus Logico-Philosophicus, 3.334*

Abstract

This chapter zooms in on fake features in relative clauses such as *I am the only one who likes* ***my*** *parents* (Kratzer 2009; Wurmbrand 2017a), a construction that is known to raise various challenges to a unified theory of the phenomenon of fake features. The main puzzle concerns the syntactic relationship between the element that hosts the fake feature and its antecedent, which is more distant than a theory of the syntax-semantics interface allows. I develop and argue for a new analysis of this construction. The main novelty towards overcoming the challenges is the claim that the underlying LF-syntax is not the way it appears at first glance. The Valuation-from-Context framework developed in previous chapters also plays an important role.

5.1 Introduction and Overview

5.1.1 Fake Indexicals

The final chapter discusses fake ϕ-features in relative clauses, a configuration that raises puzzles above and beyond the ones posed by the structures I so far dealt with. The construction in question goes back to Partee (1989) and has been investigated further by Kratzer (2009), Wurmbrand (2015, 2017a), and Bassi (2019) (among others).

(190) I am the only one who is willing to admit that **I** could be wrong.

$\forall x \neq me$: *x isn't willing to admit that* $\boldsymbol{x}$ *could be wrong*. (based on Partee 1989, fn.3)

On one of the readings of this sentence, as the paraphrase makes clear, 'I' in the scope of adjectival *only* is read as a bound variable. On this reading, the relative clause characterizes the set of individuals {x | x is willing to admit that **x** is wrong}, and (190) says that this set contains nothing but the speaker. Just like the ellipsis and focus-association cases discussed thus far, then, the person feature on the embedded pronoun must somehow be discharged of its normal semantic contribution. The facts are the same if the occurrences of 1st person pronouns are replaced with 2nd person. One can also illustrate the puzzle with other adjectival quantifiers besides *only*, such as superlative (e.g. *tallest*) and ordinal (*first*) adjectives.

(191) a. I was the first one to reveal **my** cards.

$\forall x \neq me$: *x's didn't reveal* $\boldsymbol{x}$*'s cards before I did*.

b. *(Context: at the gym. Looking around, I say:"I'm so flexible!...*

I'm the tallest one here who can reach **my** toes with **my** fingers."

$\forall x \neq me$: *if x can reach* $\boldsymbol{x}$*'s toes with* $\boldsymbol{x}$*'s fingers, x is shorter than me*.

In what follows, all my examples will be illustrated with adjectival *only* and with 1st person as the representative of fake indexicals.

I will call instances of this type of structure **Partee sentences**. The rest of this section is devoted to establishing key empirical generalizations about Partee sentences, some of them known

and some new, and to introducing the main theoretical challenges that they raise.

Like before, the basic explanandum is the very fact that a person feature that surfaces is ignored semantically: what is the mechanism allowing that? But Partee sentences also exhibit interesting properties that set them apart from the previous cases that involved focus association with a focus-sensitive operator, e.g. the truth-conditionally equivalent to (190) *Only I am willing to admit that I could be wrong*. The different properties presumably stem in one way or another from the particular syntax of this copular construction. Partee sentences are thus interesting to study because they might reveal to us non-trivial interactions between variable binding and independent principles of grammar.

The first difference concerns *optionality*: a bound reading here is possible also with a 3rd person pronoun, with no perceivable change in meaning, whereas in association with focus cases this is out, see (192).

(192) *Optionality of person form in Partee sentences*

a. I am the only one who didn't get **their/his** paycheck yet. (✓ bound)

b. Only I didn't get **their/his** paycheck yet. (✗ bound)

A second difference has to do with *variation*. Whether fake indexicals are licensed in such a construction is subject to curious cross-linguistic, and often cross-speaker, differences (Kratzer 2009, Wurmbrand 2017a, Wurmbrand 2017b, Ivan & Mirrazi 2019, Bassi 2019). Many language—including Dutch, Hebrew, Greek, French—are like English in allowing a bound reading in a configuration like (190). Other languages are more restrictive. German and Hungarian do not allow (singular) fake indexicals in Partee sentences, see (193)-(194), although they allow them in association with focus (*only I...)*, and they allow the bound-variable reading to surface when **1sg** in (193)-(194) is replaced with **3sg** pronoun/affix like in (192a). Here and throughout, '✗'/'✓' mark (un)acceptability on the bound-variable reading.

(193) ***German*** *Partee sentences do not allow fake indexicals (in the singular)*[1] (Kratzer, 2009)

[1]This is the reported judgment in the literature. One speaker I asked didn't find the bound reading of (193) so

Ich bin die einzige die **meine** Kinder versorg-t (✗)
I am the.sg.fem only who.sg.fem **my** children takes.care.of-3sg
*'only the speaker is an x such that x takes care of {the speaker's/*x's} children.'*

(194) ***Hungarian*** *Partee sentences do not allow fake indexicals*[2]

Én vagyok az egyetlen a-ki le ad-ta-∅ a beadandó-**m**-at (✗)
I am the only.one the-who.sg PART give-past-3.sg the paper-**1sg**.poss-acc
*'only the speaker is an x such that x submitted {the speaker's/*x's} paper. '*

A theory of fake indexicals in Partee sentences must both explain what makes them possible in some languages, as well as address the variation. This chapter focuses primarily on the first question, and I will reason mostly (though not exclusively) based on English data. I touch on the cross-linguistic question in section 5.6.

5.1.2 Fake Gender

Fake Indexicality is but one puzzle of fakeness brought about by Partee sentences, and one that the literature since Kratzer's (2009) pioneering work has focused on (Maier & de Schepper 2010; Wurmbrand 2017a; Ivan & Mirrazi 2019, a.o.). But the puzzle is not only about person ('indexical') features. *Gender* features on pronouns are also ignored in a similar fashion, a fact which has gone somewhat under the radar. If Barbara Partee utters (195), which is minimally different from (190) by having *she* instead of *I*, this is not infelicitous in a situation where some of the people she is asserting to be incapable of acknowledging mistakes are men.

(195) I am the only one who is willing to admit that **she** could be wrong. (✓)

$\forall x \neq me$, ***male or female****: x isn't willing to admit that x could be wrong*.

In other words, the feminine gender on the bound pronoun only tells us something about the gender of the matrix subject; it does not restrict the domain that the adjective quantifies over to female individuals only. The point can also be made with a 3rd person antecedent as in (196).

bad, and assigned it a question-mark. Two other speakers rejected it. I'll proceed on the assumption that (193) is unacceptable. Kratzer (2009:ex.7-8) reports that for pronouns in the plural the facts are different, an observation I will account for at the end of the chapter. For now I focus on singular antecedents.

[2] I thank Dóra Takács and Lilla Magyar for providing judgments on (194).

(196) Alex is the only one here who is willing to admit that **she** could be wrong. (✓)

$\forall x \neq$ *Alex,* ***male or female****: x isn't willing to admit that x could be wrong.*

This is in contrast to 'standard' cases of binding like *no student admitted* ***she*** *was wrong*, in which the feminine pronominal form does constrain the whole domain of quantification. In this respect Partee sentences bear similarity to the association-with-focus cases discussed in previous chapters, where only the overt antecedent controls features on a bound variable.

Recall from (192a) that English can also realize a bound variable in Partee structures with a gender-neutral, semantically unmarked *they* pronoun. English, then, has three possibilities to express what appears to be the same reading: one with the gender-neutral (semantically unmarked) form *their* (192a), one where the variable matches with the antecedent in gender but not necessarily person (195), and one with matching person with the antecedent (190). How to account for this three-fold optionality?[3]

5.1.3 Compositional issues: how is the variable linked to the matrix subject?

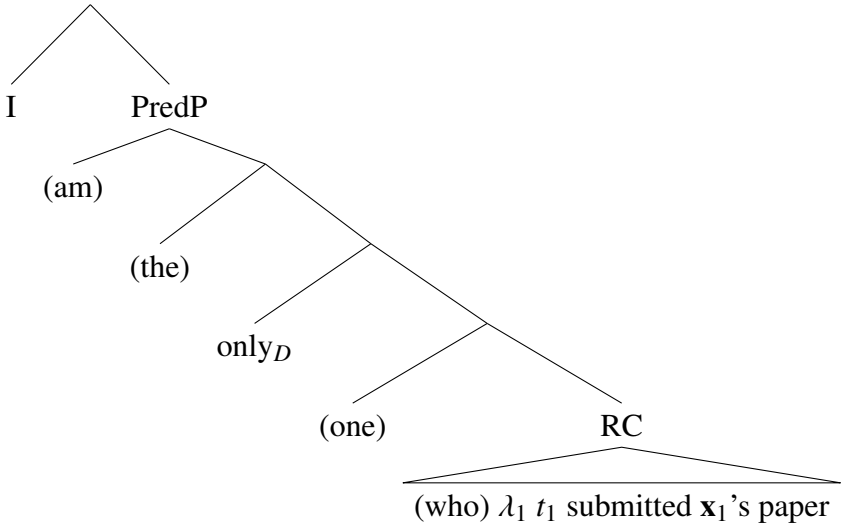
Given normal assumptions about the syntax and compositional semantics, Fake Indexicality and fake gender in Partee sentences pose an analytical problem for the theory I have outlined so far. In a nutshell, the problem is that the matrix subject, the intuitive source of fake features in Partee sentences, is too high in the structure to be considered the binder of the variable and therefore to have the chance of determining its morphological shape. This problem also affects Feature Transmission theories, if they rely only on binding relationships to explain the surface form of variables.

To appreciate this, consider the schematic LF in (197b) and assume that it is the syntactic input for all three realization options in (197a) .

(197) a. I am the only one who submitted my/her/their paper.

[3]I should point out that a minority of English speakers have difficulty accepting fake indexicals across the board, and some find a *they* pronoun in certain Partee sentences marginal. Most of my claims here are thus only relevant for those English speakers who accept all three realization options. I am not sure what accounts for the inter-speaker variation. Perhaps generational differences are a factor.

b. LF:



The possessive variable $\mathbf{x}_1$ is λ-bound at the edge of the relative clause by *who*—or by whatever operator is responsible for creating the relative clause predicate. Indeed, in order to get the relevant readings it is necessary that the variable be λ-closed within the scope of the quantification introduced by adjectival *only*. I assume a run-of-the-mill lexical entry for adjectival *only* that takes a predicate and returns a predicate, in (198).

(198) $[\![_{adj}\text{only}_D\ \text{P}_{\langle et\rangle}]\!]^g =$

$$\lambda x : \underbrace{x \in D \wedge \forall y \in D, [\![\text{P}]\!]^g(y) \text{ is defined} \wedge\ [\![\text{P}]\!]^g(x)=1}_{\text{definedness conditions (presupposition)}} .\ \forall y_{\neq x} \in D, [\![\text{P}]\!]^g(y)=0.$$

Adjectival *only* comes with an implicit (contextually-supplied) domain of individuals D, takes a predicate P and an individual $x \in$ D as arguments, and says that the predicate is not satisfied by anyone in D other than x. It also appropriately presupposes that the presuppositions projecting from P are satisfied by every member of D (this is the second conjunct in the definedness conditions). The predicate outputted by *only*$_D$ *P* is eventually saturated by the matrix subject.[4]

[4]I am ignoring any contribution of the determiner *the* in this position and assume that the meaning of (198) is inherited by the meaning of the whole post-copular phrase PredP, so that PredP *(am) the only NP* means the same as *only NP* here. As is well-known (see Coppock & Beaver 2015 and references therein), in post-copular positions a phrase of the form *the only/adj-est NP* behaves as if it denotes a predicate rather than an individual whose existence is presupposed, unlike occurrences of this phrase in argument positions. Coppock & Beaver (2015) assign *the* a meaning that maps a predicate to another predicate. The details are complicated and do not have a bearing on the present discussion (they will not help solve our issues), so I will assume for simplicity throughout the chapter that *the* in semantically inert in a configuration like (197b). Hence it is marked in parentheses.

Now, the theory of fake features advocated in this thesis says that the morphological features of a bare variable are determined by its local semantic value. Evidently, the local semantic value of the variable here is the set of individuals that make up the domain of quantification D, and this is what the current system predicts (to be more precise, the local context of the variable is a set of assignments that differ from the original assignment g only in that they assign some member of D as a value of the variable 1[5]). 1st person feature-assignment in (197a) is then never expected, because D in felicitous uses of quantification always contains more individuals than just the matrix subject; and feminine feature-assignment is not expected when D contains a mixed-gender group. Only the default *they* should surface, contrary to fact.[6]

I will propose that the problem lies not in the present system for constructing local contexts, nor in the system for valuating a pronoun's features based on local contexts, but rather in the assumed LF syntax. In the following sections I will develop and defend on analysis in which a Partee construction admits an LF in which the variable is, despite initial appearance, directly bound by the matrix subject, so that its local semantic value is the same as its intuitive antecedent. I will show that the proposal makes further correct predictions based on independent diagnostics for structure.

Before I turn to my solution in detail, it would be good to briefly see how others have tried to deal with the present issue. In the next subsection I review a proposal by Wurmbrand (2017a), who

[5]To see concretely that this is the predicted local context, consider (i) below, which is the interpretation assigned by the system to the RC enriched with a LC variable adjoined to the bound pronoun. When this is composed with the meaning of adjectival *only*, which was given in (198), the result is in (ii).

(i) $[\![(\text{who})\ \lambda_1\ [_S\ t_1\ \text{submitted}\ ^{C'}x_1\text{'s paper}]]\!]^{g,\ [C'\to G]} = \lambda y :\ \underline{[\![S]\!]^{g[1\to y],\ [C'\to G]}\ \text{is defined}}.\ [\![S]\!]^{g[1\to y],\ [C'\to G]} = 1.$
$= \lambda y : \underline{g^{[1\to y]} \in G}...$ (by LC interpretation rule)

(ii) $[\![\text{only}_D]\!]^{g,\ [C\to G]}\left([\![\text{i}]\!]^{g,\ [C\to G]}\right) = \lambda x : x \in D \wedge \underline{\forall y \in D, g^{[1\to y]} \in G} \wedge\ ...$

Due to the underlined part (which comes from the universal projection encoded in *only*'s meaning), the smallest set of assignments G that will make the whole LF compose smoothly is a set that has one assignment per individual in D.

[6]In Bassi 2019 I followed a strategy to explain fake features in Partee constructions (*I am the only one who...*) by reducing these cases to whatever explains fake feature in focus-association configurations (*only* I_F...). To this end, I worked up a syntax wherein adjectival *only* is a focus-sensitive clausal operator in disguise that associates with silent focus on the trace of the head of the RC. For several reasons, in this chapter I will abandon that line of approach and explore a different compositional route to explaining Partee sentences, one where adjectival operators contribute ordinary quantification as in (198) (no quantification over focus alternatives involved). One reason for this is that an analysis that relies on the adjectival operator being a focus-sensitive operator does not seem to easily extend to cases of fake features in partitive structures discussed in section 5.4.

embeds her account within a Feature Transmission under Binding approach (Heim 2008; Kratzer 1998). As mentioned above, the compositionality challenge poses itself for that line of approach too, and it is easy to see why: since in (197b) the matrix subject is not the binder of the variable,[7] a mechanism of transmission that reads only binding relationships is not enough to explain fake features in Partee sentences.

After reviewing Wurmbrand's solution, I will present in section 5.1.5 new data that I take to be problematic for any Feature Transmission approach, and which will constitute the first empirical motivator for my own proposal.

5.1.4 Wurmbrand (2017a)

Wurmbrand (2015, 2017a) follows Kratzer 2009 in presenting a Feature Transmission account to fake indexicals in Partee sentences (neither Wurmbrand nor Kratzer tackle the issue of fake gender). Wurmbrand's main concern is the cross-linguistic question, but she does also provide a solution to the (more basic) linking problem. She enriches feature transmission under binding with two additional local dependencies: *predication* and *relativization*. The three dependencies—binding, relativization and predication—transitively link the matrix subject to the minimal pronoun:

(199) I am the only one who (only one) who λ_1 t_1 didn't get $\emptyset_1$ paycheck
predication *relativization* *binding*

Wurmbrand (2017a) proposes that this 'transitive linking' establishes the required connection for transmission and allows the matrix subject to share its features with the variable, if certain morpho-syntactic conditions—some general, some language-specific—are met. Those conditions, whose details do not need to concern us, are met in English, so the matrix subject is allowed to transfer a person feature to the pronoun, (200). The PF output after the feature sharing/valuation is with *my*.

(200) I am the only one who (only one) who λ_1 t_1 didn't get $\emptyset_1$ paycheck
predication *relativization* *binding*
sharing of 1ST *person*

[7] By 'binder' here I mean the (semi-)technical sense of the word: the sister of the closest λ-operator co-indexed with the variable (Heim & Kratzer 1998).

Again ignoring the details, in Wurmbrand's system sharing of 1ST person under transitive linking is not obligatory (as opposed to merely under binding), and if it does not occur the spell-out is with 3rd person, as in (192a). While Wurmbrand does not discuss fake gender, she could perhaps assume that the matrix subject has another option of transmitting just a gender feature to the variable.[8,9]

5.1.5 Fake features not on bound pronouns

Wurmbrand's theory (as well those of other researchers: Kratzer 2009, Bassi 2019 and Ivan & Mirrazi 2019) is designed to explain why features on *bound variables* get to be (sometimes) semantically inactive in Partee sentences. This, however, misses a broader generalization. The last and main empirical point I wish to make in this introduction section is that any adequate analysis of Partee sentences must acknowledge that the puzzle affects not just features that surface *on* bound pronouns, but also those on full NPs/DPs that *dominate* a bound variable. That is, similarly to the ellipsis and association-with-focus cases we have covered in previous chapters, number features and gender features on full NPs/DPs that host a pronoun bound by the edge of the RC are allowed to be ignored too:

(201) *Number features on full DPs can be ignored in Partee structures*

a. Alex was the only one in the swimming pool who paid close attention to [$_{DP}$ her child].

Can entail: $\forall x \neq$ *Alex: x didn't pay close attention to x's child or children*.

[8] Although, this idea would likely require some revisions to other parts of Wurmbrand's system, which I have not presented, concerning the crucial role of gender specification in licensing/blocking Fake Indexicals cross-linguistically.

[9] Kratzer (2009) presents a different version of the Feature Transmission approach for Partee sentences. In Kratzer's (2009) theory, as opposed to Wurmbrand's, the matrix subject plays no role in the mechanics of transmission; the features that initiate transmission originate RC-internally—specifically, on the embedded little *v* which introduces the relative pronoun—and they are transferred onto the bound minimal pronoun at PF. On Kratzer's proposal, in (197b) the embedded *v*/relative pronoun could enter the derivation with [1ST-SG] and then transfer down to the minimal pronoun at PF. There is a conundrum, however: if the features originate RC-internally, in the scope of adjectival *only*, something needs to ensure that at the end of the day they are devoid of any semantic contribution (so as to not catastrophically constrain *only*'s domain of quantification). Kratzer (2009:pp.197-8) acknowledges this gap in her proposal and leaves it unresolved, and I do not see a non-stipulative way to resolve it within her framework for feature interpretation. It seems better to keep to the intuition that the ultimate source of the fake features is the matrix subject.

b. Alex is the only one here who made up [$_{DP}$ the story she told].

Can entail: ∀*x* ≠ *Alex: x didn't make up the story or stories x told.*

Take for example (201b), and consider this situation: we are in a group discussion and everyone told a story about themselves, some more than one story. Alex told just one. In such a situation (201b) does not exclude from the domain of quantification people who told more than one story; it entails that everyone who told any number of stories did not make them up.

Like before, it seems that the culprit is the number feature. A lexical modifier cannot be analogously ignored, see (202). The sentences in (202) do have a bound-variable reading for the pronoun pronoun *her*/*she*, but only if all the individuals in the domain of quantification have just one child/told just one story.

(202) *A lexical modifier cannot be ignored*

a. Alex was the only one in the swimming pool who paid close attention to her single child.

??/# ∀*x* ≠ *Alex: x didn't pay close attention to x's (one or more) children.*

b. Alex is the only one here who made up the single story she told.

??/# ∀*x* ≠ *me: x didn't make up the (one or more) stories x told.*

That gender features on full DPs can be similarly inactive is illustrated below from Hebrew.

(203) *Gender features on full DPs can be ignored (Hebrew)*

a. ani lo ha-yaxid Se- sone et [$_{DP}$ ha-mxanex-**et** Seli]
I not the-only that- hates ACC [$_{DP}$ the-head.teacher-**FEM.SG** of-me]
'I am not the only one who hates my head teacher'

b. ani ha-yaxid po Se- [$_{DP}$ ha-mitmoded-**et** Se- hicbati avur-a] alta Salav?
I the-only here that- [$_{DP}$ the-contestant-**FEM.SG** that- voted.1SG for-her] went.up level?
'Am I the only x such that the contestant that x voted for qualified to the next round?'

I contend then that an analysis is called for that unifies the cases of features on bound pronouns with cases like (201) and (203). But there is no formal link between a full DP with fake features and an 'antecedent', so it is not possible to use a framework like Wurmbrand's (2017a) which operates purely on formal syntactic relationships to account for (201) and (203).

5.1.6 Overview of the proposal

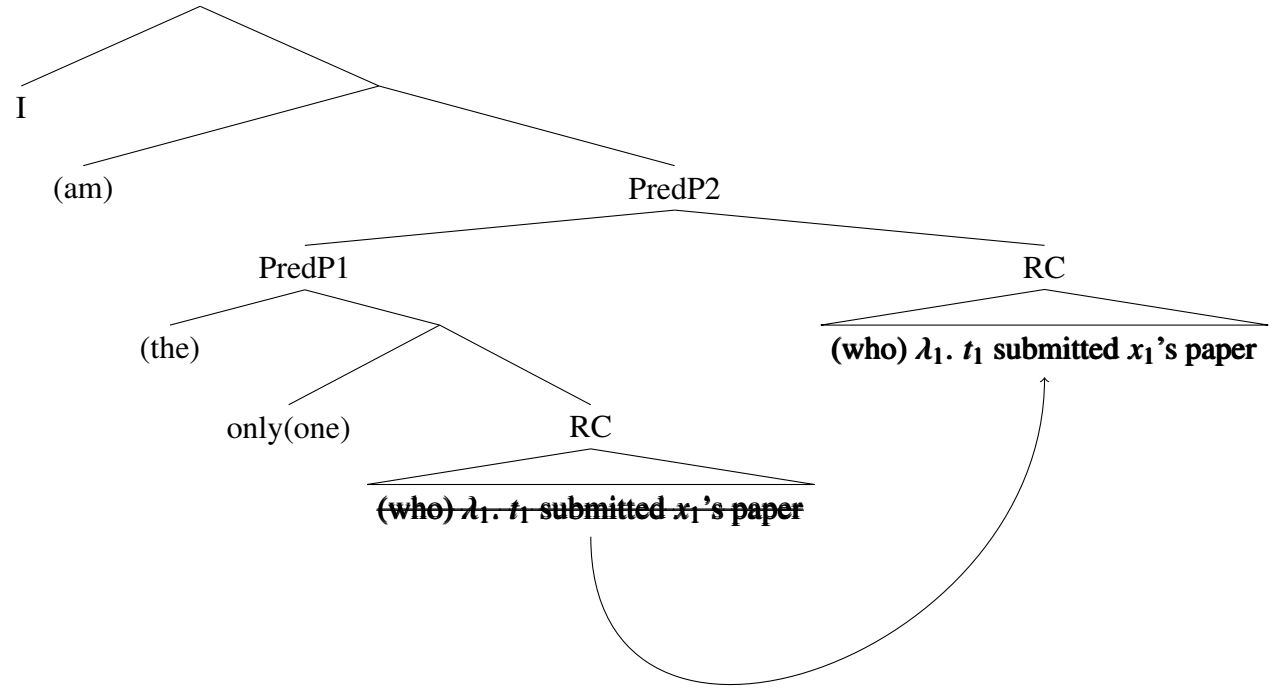
What we learn from the above observations is that feature transmission to bound variables misses a broader generalization, for the same reasons as I argued for in the earlier parts of this dissertation. And, ideally, the facts should be amenable to the same general solution advocated in this thesis, namely feature-valuation from local contexts.

In this chapter, then, I propose and defend an account of Partee sentences along the same line as in the previous chapters: late feature insertion based on local semantic value of an inherently feature-less DP (pronominal or full). I suggest that Partee sentences can host fake features because there is an optional process that creates a structure in which the local context for the relative clause in Partee sentences includes only the subject as relevant binder of the pronoun.

I will lay out my proposal about the optional process that achieves this desideratum in section 5.2, but let me give a summary of the proposal now. The central claim is that Partee sentences admit an LF in which the surface position of the relative clause (RC) is, despite appearance, not in the scope of the adjectival modifier. More specifically, when the features are fake the RC occupies **two** positions at LF: there is a silent copy in the scope of the adjectival modifier, and another copy—the overt one—in a higher position where it is predicated directly on the matrix subject. To illustrate with fake indexicality, the syntax of *I am the only one who submitted my paper* looks schematically like (204b).

(204) a. I am the only one who submitted my paper.

b. LF:



Given a semantics that intersects PredP1 and RC, the structure is interpreted in a way roughly equivalent to *I submitted my paper and I am the only one (who did).*[10] Both copies of the RC are interpreted, although, as indicated by the strike-through, only the higher position is pronounced. The variable x is inherently feature-less, which is what underlies fakeness. But once (204b) is shipped to the realization module and according to the rules for assigning ϕ-features, the spell out of DPs (pronominal and nominal alike) will be according to the semantic value in the local context. Crucially, the local context of any DP in the higher position is determined by the matrix subject,

[10] Another possible paraphrase is with a non-restrictive relative clause: *I, who submitted my paper, am the only one (who did).* This paraphrase might be closer to the intuitive meaning of (204a) since it conveys the information that the speaker submitted the speaker's paper at some backgrounded level of meaning rather than at the level of the assertive content (see Potts 2005 for the view that non-restrictive modifiers belong to the backgrounded dimension of 'conventional implicatures'). My perspective on the difference between English and German in section 5.5 will make use of the idea that the (uttered) relative clause should be thought of as some kind of a non-restrictive relative clause. Making this idea precise in the semantics, , requires thinking seriously about how the semantic composition of (204a) works in a way that integrates the non-restrictive RC with PredP1. For now we can abstract away from this detail, and I continue to assume standard intersection.

which directly binds into the RC; in (204b) the binder denotes the speaker, so the bound pronouns will be realized as 1ST-SG.

In section 5.2 I go over the details of the proposal. In section 5.3 I show that the hypothesized syntax, according to which the overt RC is in a high position, makes correct predictions about the interaction with NPI licensing and certain restrictions on combining fake features. This will lend further support to the (rather counter-intuitive) claim that the RC can be attached as high as shown in (204b). In section 5 I touch on the cross-linguistic question and suggest a way to understand the difference between English in German in the licensing of fake features (cf. 193), although some of the suggestions there will remain tentative.

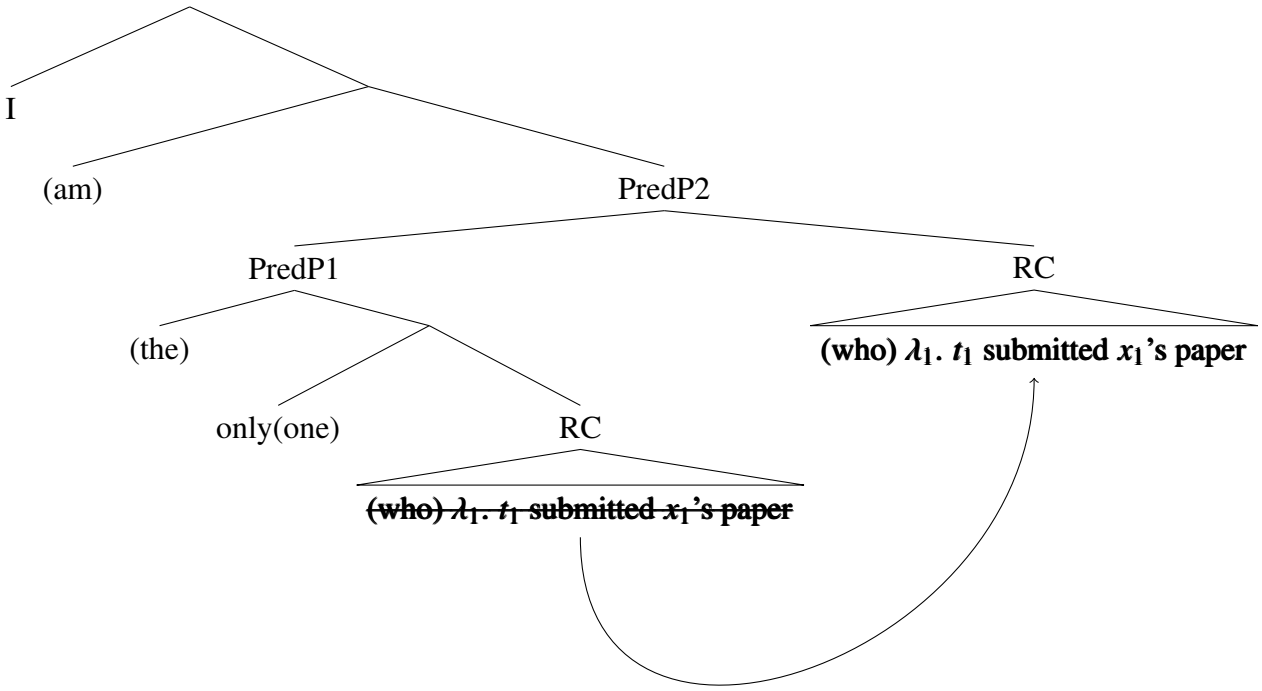
5.2 Core proposal

5.2.1 Syntax and interpretation

I assume that a structure like (204b), repeated below in (205) is generated by rightward movement of the RC.[11] I want to show now that it represents the correct reading, given defensible assumptions about semantic composition. Essentially, the high copy of the RC does not disrupt the overall interpretation of the structure, nor contribute additional information, due to the existence of the low copy.

(205) LF:

[11] Throughout, I represent the gap position inside the RC as an impoverished trace, but nothing in my account changes if it is rather a full copy that gets trace converted (e.g. Fox 2002) or undergoes a similar process.


I
(am)
PredP2
PredP1
RC
(the)
(who) λ1. t1 submitted x1's paper
only(one)
RC
(who) λ1. t1 submitted x1's paper

First let's compute PredP1, by applying the entry for adjectival *only* in (206), repeated from (198), to the interpretation of the (silent) relative clause. The result is in (207) (see footnote 4 about my ignoring *the* in this position):

(206) $[\![_{adj}\text{only}_D\ \text{P}_{\langle et\rangle}]\!]^g =$

$\lambda x : \underbrace{x{\in}D \wedge \forall y \in D, [\![\text{P}]\!]^g(y) \text{ is defined} \wedge\ [\![\text{P}]\!]^g(x){=}1}_{\text{definedness conditions (presupposition)}} .\ \forall y_{\neq x} \in D, [\![\text{P}]\!]^g(y){=}0.$

(207) $[\![\text{PredP1}]\!]^g =$

$\lambda x : \forall y \in D, [\![\lambda_1\ \text{t}_1 \text{ submitted } \text{x}_1\text{'s paper}]\!]^g(y) \text{ is defined } \wedge [\![\lambda_1\ \text{t}_1 \text{ submitted } \text{x}_1\text{'s paper}]\!]^g(x){=}1.$

$\forall y_{\neq x} \in D, [\![\lambda_1\ \text{t}_1 \text{ submitted } \text{x}_1\text{'s paper}]\!]^g(y) = 0.$

=

$\lambda x : \forall y \in D, y \text{ has a paper } \wedge x \text{ submitted x's paper.}$

$\forall y_{\neq x} \in D, y \text{ didn't submit y's paper.}$

In (207), for reasons of space, I dropped the restriction '$x \in D$' from the definedness conditions of the predicate and I will continue to do so. The reader should remember that the argument of the predicate is always taken to be a member of the contextually supplied domain of quantification.

Next, I assume that PredP1 composes with the higher copy of RC by predicate intersection. I adopt the rule of Predicate Modification from Heim & Kratzer 1998, in (208).

(208) **Predicate Modification (PM)**: if a node has two daughtors {A, B} such that both $[\![\text{A}]\!]^g$ and $[\![\text{B}]\!]^g$ are of type $\langle e, t\rangle$, then:

$[\![\text{A B}]\!]^g = \lambda z : [\![\text{A}]\!]^g(z) \text{ is defined} \wedge [\![\text{B}]\!]^g(z) \text{ is defined. } [\![\text{A}]\!]^g(z) = 1 \wedge\ [\![\text{B}]\!]^g(z) = 1$

Note that I am being explicit about how PM projects definedness conditions (presuppositions). This, in particular the part that projects the definedness conditions of $[\![\text{A}]\!]$, is important for ensuring that the extra copy of the RC does not matter much to the overall interpretation.[12]

The computation of PredP2, then, is given in (209). For perspicuity, I am representing in the definedness conditions of the end result only the information that projects from PredP1 (the left

[12]The part that projects the definedness conditions of $[\![\text{B}]\!]$ will play a role later when we compute local contexts for elements in the high copy of the RC.

predicate), because the information projecting from the definedness conditions of RC is anyway redundant.[13]

(209) $[\![\text{PredP2}]\!]^g =$ (by PM)

$\lambda x : [\![\text{PredP1}]\!]^g(x)$ is defined ($\wedge$ $[\![\text{RC}]\!]^g(x)$ is defined).

$[\![\text{PredP1}]\!]^g(x)=1 \ \wedge \ [\![\text{RC}]\!]^g(x)=1$

=

$\lambda x : \forall y \in D, [\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(y)$ is defined $\wedge [\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(x)=1$.

$\forall y_{\neq x} \in D, [\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(y) = 0 \wedge \underline{[\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(x)=1}$.

The underlined part, which is contributed by the overt copy of RC, is semantically redundant; it conveys information encoded already in the definedness conditions. Whenever the function in (209) is defined, it is already entailed that the matrix subject submitted their paper, so repeating it in the output of the function (the at-issue content) doesn't add anything. (209) can then be shortened to (210), where the underlined part has been removed, and which is exactly what was gotten for PredP1 in (207) before the extra copy of RC was integrated.

(210) $[\![\text{PredP2}]\!]^g =$

$\lambda x : \forall y \in D, [\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(y)$ is defined $\wedge [\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(x)=1$.

$\forall y_{\neq x} \in D, [\![\lambda_1 \text{ t}_1 \text{ submitted x}_1\text{'s paper}]\!]^g(y) = 0$.

After composing (210) with the matrix subject *I*, this gets the intuitively correct meaning: the speaker submitted the speaker's paper, and every contextually-relevant y besides the speaker that has a paper did not submit y's paper.

5.2.2 Local valuation

The structure in (205) is submitted to PF realization, during which feature valuation takes place. In particular the bound pronoun, x_1, which is feature-less at the point (205) is interpreted, needs to be supplied with ϕ-features based on its semantic value in its local context. What is its local context?

[13] This follows from the fact that the quantificational domain D is restricted to contain the argument x.

Crucially, there are two copies of x_1: one where it is λ-closed in the scope of $only_D$ and a higher one outside of it. I assume that the system can distinguish between copies for the purpose of determining local contexts. That is, we can talk about x_1's local context in the position in which it is realized, the higher position, which may not necessarily be the same as its local context in the lower position. I will now show that the local context for x_1 in the higher position in (205) is such that the value of x_1 with respect to it is the same as the value of the matrix subject.

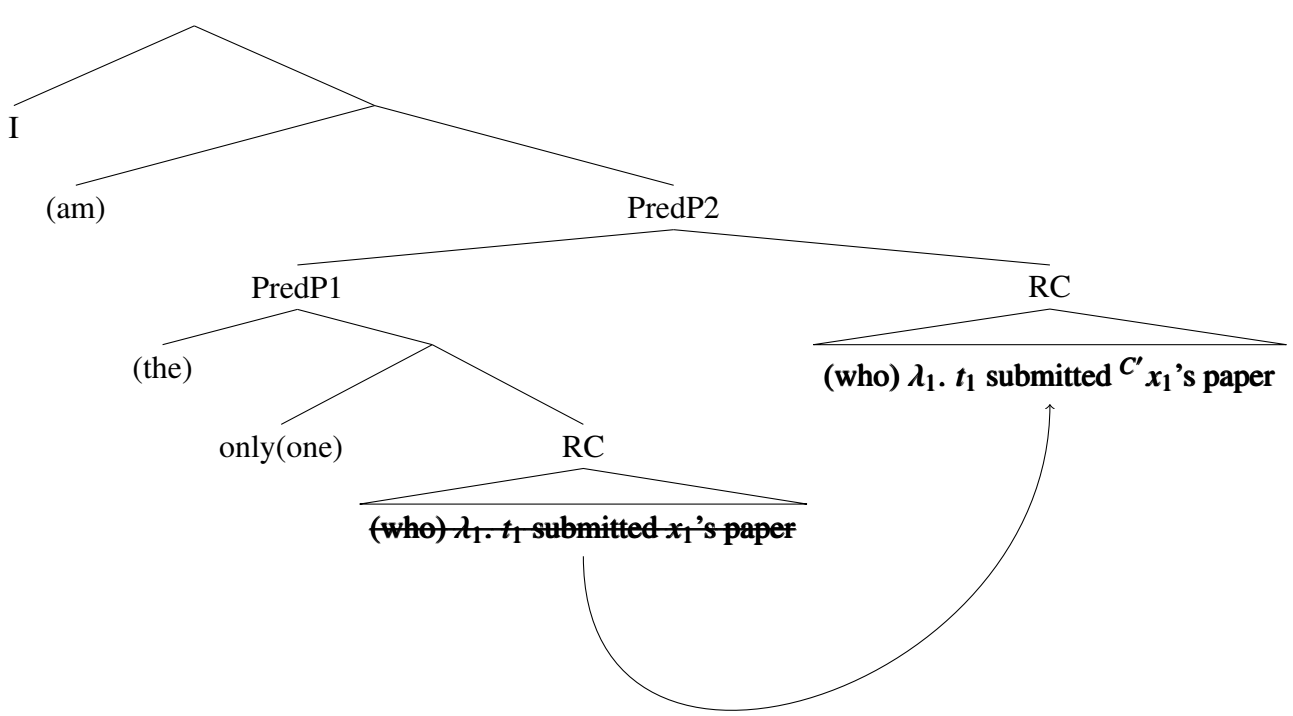
Before the details, I give the intuition: in a structure of the form $[A\ \lambda_1[...x_1...]]$, recall, the local context for x_1 maps x_1 to $[\![\text{A}]\!]^g$ (for any g). If we add a sister P to the λ_1-predicate, such that the structure now looks like $\Big[A\ \Big[[_P...][\lambda_1[...x_1...]]\Big]\Big]$ and the two predicates are intersected to create a more complex predicate, this should not make any difference; it is still the case that the local value of x_1 is $[\![\text{A}]\!]^g$. Because $[\![\text{A}]\!]^g$ is still the only argument that saturates the λ_1-predicate in its high occurrence.[14]

Technically, we derive this as we have done before by attaching the VC variable, C', to the overt occurrence of the bound pronoun, see (211) below. The goal is to find the smallest value of C' that will make the whole LF in (211) defined.[15]

[14]I would still call A the 'binder' of x_1 in this configuration, although this evidently would necessitate giving up on the derived definition of a '(semantic) binder' often used in static (non-dynamic) frameworks for binding. On that definition (see footnote 7 and Heim & Kratzer 1998), a binder of a variable is the sister of the closest c-commanding λ co-indexed with the variable. But in (205) the matrix subject is not the sister of the λ-predicate. This notational convenience, however, is also abandoned in non-static frameworks for binding.

[15]There is a question about the assumptions that allow one to adjoin a VC variable just in one syntactic copy. I think of VC variables as useful notational devices to allow us to define local contexts for embedded constituents based on static truth-conditional semantics. As such, I do not consider them part of the lexicon of the object language, on the same level as e.g. *apple*, *the* and λ. To be maximally concrete, I assume that they are special elements that can be inserted onto constituents in an LF at a stage of derivation after the LF has been shipped to semantic interpretation (but before it has been submitted to phonological realization, of course).

(211)



The rest of this subsection is devoted to proving that, given an initial assignment g, the Local Context of $\boldsymbol{x_1}$ in the high position is $\{g^{1 \to \textit{the speaker}}\}$. As a reminder, in (212) I repeat the rule of VC Variable Interpretation: it records the active assignment function in a set of assignments, and otherwise does nothing. As before, it doesn't matter if C' attaches to x_1 or anywhere else inside the scope of (the higher) λ_1.

(212) **LC Variable Interpretation** (=(116))

For any assignment g' and a context G (set of assignments) that C' denotes:

$[\![^{C'}\psi]\!]^{g', [C' \to G]}$ is defined iff $\boldsymbol{g' \in G}$ and $[\![\psi]\!]^{g', [C' \to G]}$ is defined. If defined, $= [\![\psi]\!]^{g', [C' \to G]}$

Let's compute bottom up. I start in (213) with the higher copy of RC. I concentrate only on the definedness conditions of the functional denotations, because definedness is the only part that matters for local contexts in this system. Hence I place '...' instead of writing the output of functions.

(213) For any assignment g and any set of assignments G:

$[\![[_{\text{RC}}$ (who) λ_1 [S t_1 submitted $\mathbf{C'}$ x_1's paper]]$]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}$ = (by PA rule)

λy : $[\![\text{S}]\!]^{g[1\rightarrow y],\,[\mathbf{C'}\rightarrow\mathbf{G}]}$ is defined. ... = (by VC interpretation rule)

λy : $\underline{g^{[1\rightarrow y]} \in G}$ $\wedge$ $[\![\text{S}]\!]^{g[1\rightarrow y]}$ is defined. ...

The function we got records in G, for any argument y, a modification of g mapping 1 to y. Now we compute the next node up, PredP2.

(214) $[\![\text{PredP2}]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}$ = (by PM rule)

λx : $[\![\text{PredP1}]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}(x)$ is defined $\wedge$ $[\![\text{RC}]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}(x)$ is defined. ...

Using the result gotten in (213), we can rewrite (214) as follows:

(215) $[\![\text{PredP2}]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}$ =

λx : $[\![\text{PredP1}]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}(x)$ is defined $\wedge$ $\underline{g^{[1\rightarrow x]} \in G}$ $\wedge$ $[\![\text{S}]\!]^{g[1\rightarrow x]}$ is defined. ...

Just like in (213), the function in (215) records in G, for any argument x, a modification of g mapping 1 to x. And because the VC variable C' doesn't occur in PredP1, it is easy to see that expanding $[\![\text{PredP1}]\!]^{g,\,[\mathbf{C'}\rightarrow\mathbf{G}]}$ in (215) is unnecessary; it would not change the basic result. The function still only records $g^{[1\rightarrow x]}$ in G.

Finally, the function in (215) is applied to the matrix subject. This means that the whole LF in (211) is defined w.r.t to g and G only if $g^{[1\rightarrow \textit{the speaker}]} \in G$. For any g, the smallest value of C' that makes (211) defined is $\{g^{1\rightarrow \textit{the speaker}}\}$. This is the local context.

(216) Derived: The Local Context for x_1 (in the higher position) in (211) given g is $\{g^{1\rightarrow \textit{speaker}}\}$.

5.2.3 Predicting ('fake') features on bound pronouns

The rules for assigning ϕ-features based on local semantic value predict that the bound-variable pronoun in (211) is to be realized with the bundle of features **1-sg** features: the variable's semantic value is the speaker throughout its local context (see 216). This concludes my basic explanation of Fake Indexicals in Partee sentences.

(217) Predicted PF of (211):

I am the only one who did my homework

Fake gender in a sentence like (218) (a version of (196)) is also explained, and the structure of the explanation is the same. The syntax that feeds realization is in (219).

(218) Alex is the only one who submitted **her** paper. (✓)

$\forall x \neq$ *Alex,* ***male or female****: x did not submit x's paper.*

(219) Alex is [[the only one ~~[who$_1$ t_1 submitted x_1 paper]~~] [who$_1$ t_1 submitted x_1's paper]]

The local context of the pronoun x_1 in the position in which it is pronounced is $\{g^{1\rightarrow Alex}\}$. We predict the spell-out to be *her* if Alex is identified as female. The domain of quantification of *only* does not have any influence on the choice of gender on the pronoun, essentially because the overt pronoun is not 'bound' by the quantificational domain of *only*; the local context of x_1 is not $\{g^{1\rightarrow x} \mid x \in D\}$.

The theory still doesn't predict cases of partial agreement such as (262) (a similar example to (195)), where the antecedent is 1st person and the bound pronoun agrees with it in gender but not in person.

(220) I am the only one who submitted **her** paper. (✓)

$\forall x \neq$ *me,* ***male or female****: x did not submit x's paper.*

(221) I am [[the only one ~~[who$_1$ t_1 submitted x_1 paper]~~] [who$_1$ t_1 submitted x_1's paper]]

This is because the rules of feature assignment do not allow for optionality. If the structure for the relevant reading is in (221), repeated from (211), 1st person assignment to x_1 is obligatory. And, as mentioned in section 5.1.3 (see especially footnote 5), a simpler syntax without the extra extraposition step, where the RC that gets spelled out is in the scope of *only*, is predicted to allow *her* as the spell-out of x_1 only if the whole domain of quantification consists of female-identified individuals, which doesn't seem to be right. I will get back to this issue in section 5.6.

5.2.4 Predicting ('fake') features on full NPs/DPs

The theory correctly generalizes to the observations made in section 5.1.5 that number and gender features on full DPs can also be 'fake'. I repeat the data in (222)-(223).

(222) *Number features on full DPs can be ignored in Partee structures* (=201)

a. Alex was the only one in the swimming pool who paid close attention to [$_{DP}$ her child].

Can entail: $\forall x \neq$ *Alex: x didn't pay close attention to x's child or children*.

b. Alex is the only one here who made up [$_{DP}$ the story she told].

Can entail: $\forall x \neq$ *Alex: x didn't make up the story or stories x told*.

(223) *Gender features on full DPs can be ignored (Hebrew)* (=203)

a. ani lo ha-yaxid Se- sone et [$_{DP}$ ha-mxanex-**et** Seli]
I not the-only that- hates ACC [$_{DP}$ the-head.teacher-**FEM.SG** of-me]
'I am not the only one who hates my head teacher'

b. ani ha-yaxid po Se- [$_{DP}$ ha-mitmoded-**et** Se- hicbati avur-a] alta Salav?
I the-only here that- [$_{DP}$ the-contestant-**FEM.SG** that- voted.1SG for-her] went.up level?
'Am I the only x such that the contestant that x voted for qualified to the next round?'

Remember the claim from previous chapters which was motivated based on ellipsis and focus association data: full DPs, just like pronouns, come into the derivation without features, and they are supplied with features according to their semantic value in their local context. This assumption carries over to our Partee construction, importantly. For e.g. (222b), number on the full DP is under-specified at LF, and the reason that feature valuation from context outputs singular on DP is that the syntax of Partee constructions allows an LF with an extraposed RC like before, where the matrix subject binds into the extraposed RC in the same way as before and thus affects embedded contexts in the same way. Any RC-internal DP in need of features, no matter if it is a pronoun bound at the edge of the RC (as in the previous examples) or a full DP that contains one (as in these examples), will be evaluated against the same local context.

To illustrate, extraposition of the RC creates the structure in (224) for (222b):

(224) Alex is ⌈ [the only one ~~[who$_1$ t_1 made up [$_{DP}$ the STORY x_1 told]]~~]
[who$_1$ t_1 made up [$_{DP}$ the STORY x_1 told]] ⌉

The DP is generated feature-less, so its encoded meaning does not prejudge number; this makes the sentence felicitous in a mixed-number-of-stories scenario. The local context of anything inside the extraposed RC is $\{g^{1\rightarrow Alex}\}$. If the (global) context entails that Alex told just one story, the semantic value of DP in the higher position throughout its local context is the (single) story Alex told. Therefore, according to the valuation rules, the D head of DP is supplied with **sing** feature, and the PF end result is 'story' in the singular.

5.2.5 Taking stock

As an intermediate summary, I have proposed that fake features in Partee sentences are possible because Partee structures support a syntax in which the relative clause surfaces higher, not in the position one would initially think it is, and is predicated directly over the matrix subject. That makes the matrix subject the sole determinant of what the local context is of elements in the (realized copy of) the RC. This, together with an independent theory of feature valuation from local contexts, allows the matrix subject to be the ultimate source of the fake features.

In the next section I investigate two predictions stemming from this account. One has to do with restrictions on possible combinations of fake features ('Fake Together', subsection 5.3.1), and the other concerns interaction with NPI licensing (subsection 5.3.2). The results will provide further evidence for the hypothesis that it is the high syntactic position of the RC that licenses fake features based on (i) restricted possibilities for fakeness ('Fake Together'),

5.3 Further predictions

5.3.1 Prediction: 'Fake Together'

I claimed in subsection 5.2.4, examples (222), that a full DP that is 'bound into' in a Partee structure is felicitous in a mixed-number scenario because the RC that hosts it is attached high. This leads to a prediction: a singular DP should *not* be felicitous in a mixed-number scenario, if we can fix the position of the overt RC to be in the lower position (in the scope of *only*). We can do so by changing the feminine *she* in (222) to the gender-neutral *they*. The reasoning goes as follows: if a bound pronoun in a Partee construction is the semantically-unmarked *they*-pronoun, it cannot correspond to a structure like (224) with an extraposed RC, at least not when the gender of the matrix subject is known (i.e. when a more specific option than the elsewhere element is available for spell-out). We hypothesized that a *they*-pronoun indicates low attachment for the (overt) RC, with no extraposition; this is schematized in (225).

(225) *A* they*-pronoun fixes a low scope for the RC*

a. Martha is [the only one [who$_1$ t_1 ... *they*$_1$ ...]]

b. * Martha is $\Big[$ [the only one ~~[who$_1$ t_1 ... *they*$_1$ ...]~~] [who$_1$ t_1 ... *they*$_1$ ...] $\Big]$

If so, an RC-internal DP containing this *they* should also be trapped in the low position. But then, its local semantic value must range over the whole domain of quantification (i.e., the local context for the DP is $\{g^{1\to x} \mid x \in D\}$, where D is the contextually-supplied domain). This should not allow a mixed-number scenario.

The prediction is borne out (although judgments might be subtle). In the scenario in (226), there is a contrast between (226a) and (226b). The same goes for (227).

(226) *Context: Some participants told one or more stories. Martha told just one.*

a. Martha is the only participant here who made up [$_{DP}$ **her** story]. (✓)

b. ?? Martha is the only participant here who made up [$_{DP}$ **their** story].

(227) *Context: In the big raid, some officers arrested a number of criminals. Martha arrested just one.*

a. Martha is the only officer who treated [$_{DP}$ the criminal **she** arrested] nicely. (✓)

b. ?? Martha is the only person who treated [$_{DP}$ the criminal **they** arrested] nicely.

Furthermore, the [b] sentences dramatically improve if we focus on a non-mixed number scenario, i.e. one that entails that all the people told just one story/had just one marble next to them. This is again predicted.

This constitutes suggestive evidence in favor of the hypothesis that the possibilities for feature realization are correlated with the attachment site of the RC.

In fact, it is an instance of a more general prediction which I will call FAKE TOGETHER:[16] if one feature-option surfaces in a way that controls for the attachment site of the RC, all other DPs in the RC must align. This prediction can be tested with fake indexicals, and the facts in (228) conform to the prediction.

(228) a. ✓ I am the only one who submitted **my** paper after you told **me** to.

b. ✓ I am the only one who submitted **their** paper after you told **them** to.

c. ✗ I am the only one who submitted **their** paper after you told **me** to.

d. ??/✗ I am the only one who submitted **my** paper after you told **them** to.

If the overt RC is attached high, we get all-**1st** person in (228a); if it attaches low we get all-**they** in (228b). But mixing is not allowed, since once the position of the RC is fixed it is not possible to violate the feature valuation rules.

5.3.2 Interaction with NPI licensing

Negative Polarity Items (NPIs) like *any*, *ever* and *give a damn* are licensed in Strawson Downward Entailing (SDE) environments (Fauconnier 1975; Ladusaw 1979; von Fintel 1999). The scope of

[16] To echo the term 'Shift Together' used in the literature on indexical shift (Anand & Nevins 2004 and others).

adjectival *only* (and of superlative adjectives generally) is a SDE environment, as exemplified with the intuitively valid argument in (229). Therefore it is unsurprising that NPIs are licensed in this environment, (230).

(229) *The scope of adjectival* only *is a SDE environment*

(i) John is the only one who owns a bike.

(ii) John owns a red bike.

∴ John is the only one who owns a red bike.

(230) *NPIs are licensed in the scope of adjectival* only

a. John is the only professor who ever reads anything.

b. Mary is the only one who gives a damn.

This derives a prediction. If fake features are the result of an LF in which the RC is not in the scope of *only*, we expect an interaction with NPI licensing. RCs that host an NPI can attach low as in (231a), but they cannot extrapose lest the NPI would lose its SDE environment, as illustrated in (231b).

(231) *NPIs fix a low scope for the RC*

a. X is the [only one [who$_1$ t_1 ... NPI ...]]

b. * X is [[the only one ~~[who$_1$ t_1 ... NPI ...]~~] [who$_1$ t_1 ... NPI ...]]

Fake features are therefore predicted impossible with an NPI. For instance, we expect that an NPI will not be able to be combined with a fake indexical. Speakers I have consulted with indeed perceive a reliable contrast in the following examples:

(232) *Fake Indexicals and NPI licensing*

a. I am the only one who bought {✓**their**/✗**my**} spouse anything at the airport.

(cf. *I am the only one who bought* ✓*my spouse something at the airport.*)

b. I am the only one here who ever gave a perfect score to one of {✓**their**/✗**my**} students.
(cf. *I am the only one here who gave a perfect score to one of* ✓*my students.*)

c. I am the only one here who gives a damn about {✓**their**/✗**my**} team.
(cf. *I am the only leader who cares about* ✓*my team.*)

These facts constitute a strong argument that the possibility for fake features correlate with whether the fake indexical is bound in the scope of the adjectival operator or not.

5.4 Agreement Mismatches in Partitive Relatives

There are a number of loose ends about Partee sentences that need to be tied, among them the cross-linguistic question regarding fake indexicals (section 5.1.1) and the question about partial agreement (section 5.2.3). I will address them in sections 5.5-5.6. Before we get there, I want now to draw the reader's attention to a puzzle closely related to Partee constructions, dubbed Agreement Mismatch in Partitive Relatives (**AMPR**) by Longenbaugh (2019), who was the first to study it as far as I know. In this section I argue that AMPR should be subject to the same mechanism underlying Partee constructions. I will then use AMPR to expand the range of construction on which we can test the theory.

5.4.1 Data

Some examples of AMPR are in (233).

(233) *Agreement Mismatch in Partitive Relatives* (*AMPR*) (Longenbaugh 2019)

a. Paula is one of the few linguists who understand**s** Chomsky.

b. John is one of those/the many mathematicians who respect**s** set theory.

c. I met one of the many linguists who like**s** Aspects yesterday.

d. *one of* [$_{\text{DP}}$ {*the/those/the few/the many/the only/...*} *NP.**pl*** [$_{\text{RC}}$ ... *V.**sg*** ...]]

In AMPR structures, there is a partitive construction (*one of..*) that hosts a restrictive RC in the semantic scope of the head NP (*linguists*, *mathematicians*). The head NP is plural, but can interestingly license singular agreement on the verb in the RC. The mismatch is acceptable only when a partitive structure is involved:

(234) * The few linguists who understand**s** Chomsky smiled.

The question is what allows number mismatch in partitive constructions like (233) but not outside of them.

5.4.2 Longenbaugh's (2019) proposal

Longenbaugh (2019) builds an account that starts from the sensible assumption that the RC in AMPR attaches to the lower of the two NPs; in (233a), for example, the RC attaches to *linguists* and not the higher, singular, *one of the linguists*. This is virtually necessary in order to respect the fact that the RC content forms part of the semantic restriction of the lower NP: (233a) entails that there are few linguists who understand Chomsky, not that there are few linguists simpliciter.

(235) Longenbaugh's (2019) LF:

John is one of the few [linguists [~~linguist~~ who$_1$ t_1 understands Chomsky]]

Why is the number mismatch, then? according to Longenbaugh, it is because: (i) the internal head of the RC in AMPR, marked with strike-through in (235), is a full NP that must be elided under identity with a head-external NP,[17] and (ii) the antecedent for this ellipsis can exceptionally be the farther, singular NP *one (linguist)*, rather than the closer (head-external) plural NP *mathematicians*. Since the antecedent is singular, the internal (elided) head is singular as well, which licenses singular agreement on the verb. This syntax is coupled with a semantics that allows a plural NP (*linguists*) to compose with a RC headed by a singular NP.

[17] This in turn follows from the assumption that the RC in AMPR is an instance of the so-called 'Matching' sort, which requires ellipsis under identity with an external NP.

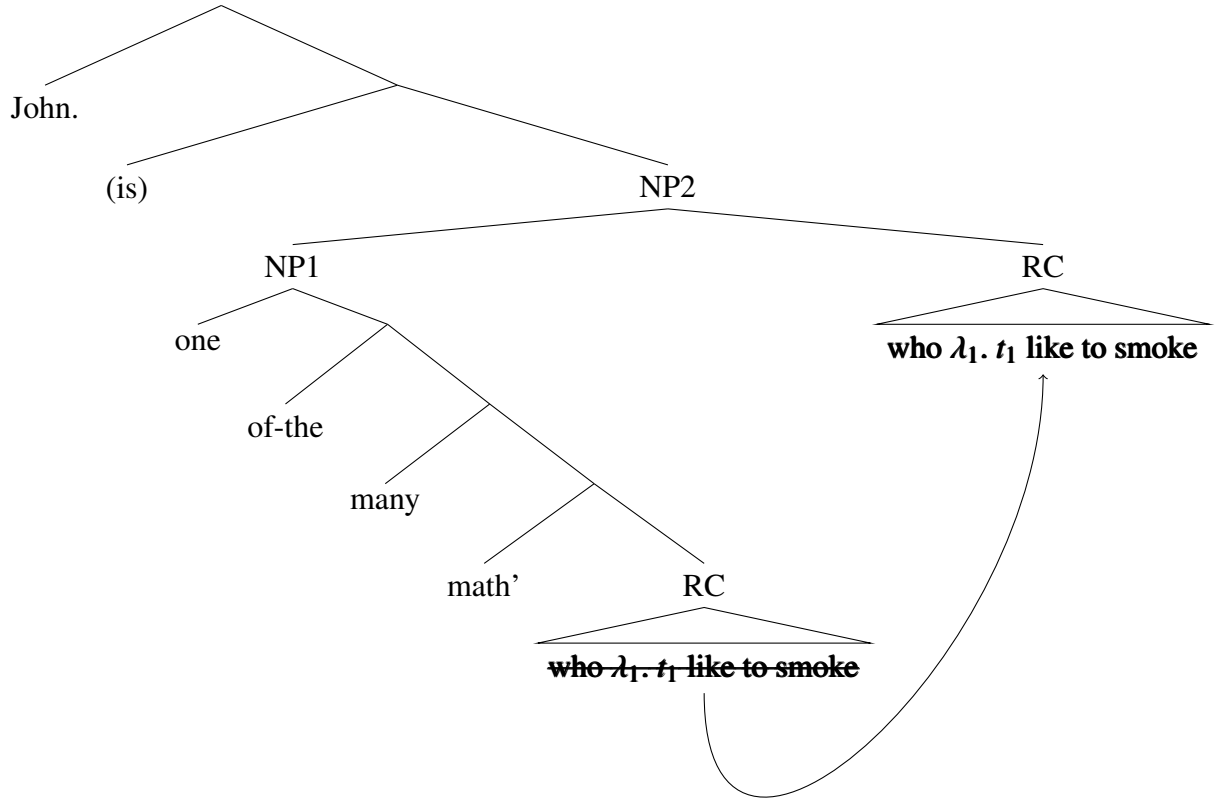
Longenbaugh argues for parts of his analysis based on intricate predictions which I will not review or evaluate here (see his paper for details). As he acknowledges, however, relaxing the identity condition between an RC-head and its antecedent in this way opens up the door to some overgeneration problems. This should motivate us to look for an alternative explanation which keeps the tight syntactic relationship usually assumed to be involved in constructing RCs.

5.4.3 High-attachment theory of AMPR

It is reasonable to suppose, and I will, that the surprising appearance of the mismatching feature in AMPR is due to the same underlying mechanism that is behind fake features in Partee sentences. As the reader perhaps might have guessed, my alternative analysis denies the assumption that the RC is attached **only** low. I will argue for the structure in (236b).

(236) a. John is one of the many mathematicians who likes to smoke.

b. LF:



As in Partee sentences, we have two copies of the RC. The low copy is responsible for the fact that the mathematicians in question are restricted by the content of the RC.[18] The high RC copy modifies the partitive, and I assume that the two compose by predicate intersection like before. But this (like before) does not affect the overall truth conditions thanks to the existence of the low copy. The resulting meaning of (236b) can be paraphrased as *John is one who likes to smoke of the many mathematicians who like to smoke*. It is routine to carefully show how these truth conditions

[18] A couple of remarks: in (236b) the low noun *mathematician* is not part of the constituent that gets copied, but I can be agnostic about that. As far as I can see it would not make a difference if the extraposed constituent included *mathematician*. The trace could also be represented as a full noun. I am also keeping neutral about what kind of RC derivation is in hand - it could be a Matching/Head-External analysis or a Head-Internal analysis. These remarks apply to Partee constructions too.

come about by providing suitable lexical entries for the pieces, but in order not to jade the readers I leave it to them as an exercise.

Crucially, there is no number information in the RC at the input to interpretation. The point is that the verb gets to be in the singular because of the mechanics of feature valuation. Specifically, the local context of the trace of RC in the high position is $\{g^{1\rightarrow John}\}$. The rule for number assignment sees that the local value of the trace is an atomic individual and therefore assigns **sing** to the trace, after which verb agreement takes place and the result is (236a).

5.4.4 AMPR and NPIs

The account immediately makes a prediction about interaction with NPI-licensing. NPIs are in general licensed in the scope of some adjectives that host a partitive structure:

(237) a. John is one of the few students who ever read anything.

b. Mary is one of only four students who ever read anything.

If AMPR is a result of the hypothesized syntax which takes the RC outside the scope of these NPI-licensing adjective, we predict NPIs to be unacceptable with AMPR. This is borne out:

(238) *AMPR does not allow NPI licensing*

a. *John is one of the few students who ever **reads** anything.
(cf. *John is one of the few students who reads class material.*)

b. Mary is one of only 3 employees who {**dare/*dares**} to say anything to the boss.
(cf. *Mary is one of only 3 employees who dares to talk to the boss.*)

c. He is one of only a dozen-or-so people that {**have/*has**} ever stepped on the moon.
(cf. *He is one of only a dozen-or-so people that has visited the moon.*)

d. She is one of the very few students who {**give/*gives**} a damn about my class.

I take this data to furnish a strong argument that AMPR-licensing RCs attach high.

5.4.5 Fake Together and other predictions

Longenbaugh (2019) observed that when the verb shows mismatch in number with the head (AMPR), a pronoun bound by the head of the RC must do so too. This is an instance of Fake Together.

(239) *Bound pronouns agree with the verb in AMPR*

a. Sally is one of the many students who **is/#are** submitting **her** paper on time.

i. many students ∈ {x: x submits x's paper} (bound)

ii. # many students ∈ {x: x submits Sally's paper} (referential (ok with *are*))

b. Sally is one of the many students who **#is/are** submitting **their** paper on time.

(239a) is predicted by the current analysis. Mismatching verbal agreement (singular agreement) indicates an extraposed structure, illustrated in (240). The internal trace of the RC is assigned singular number – and so must a co-indexed pronoun, due to the obligatory nature of feature assignment.

(240) Sally is [[one of the many students ~~[who λ_1 t_1 BE submitting x_1's paper]~~]
[who λ_1 t_1 BE submitting x_1's paper]]

By contrast to (240), in a structure with no extraposition as in (241) the RC surfaces in the position where it makes its central semantic contribution, the restrictor position of *the many*.

(241) Sally is [one of the many students [who λ_1 t_1 BE submitting x_1's paper]]

The trace position, I assume, can get valued for plural in (241) by syntactic agreement with the external plural head (which in turn gets plural based on the semantic value-in-context of the containing DP *the many...*). That the pronoun gets plural number (cf. (239b)) also follows from the theory of local contexts under certain assumptions, but I will not show that here.[19]

[19]Without elaborating, I will hint that deriving a LOCO for a variable bound from inside the restrictor position of *the* runs into an interesting difficulty given: (i) the general system of deriving LOCOs from chapter 4, (ii) the maximality-based entry for *the* from chapter 3, and (iii) standard assumptions about plurality. One possible fix to the challenge involves making the switch in the entry for *the* that I described in fn.10 in chapter 3. Discussing the challenge in more detail will take us too far afield.

5.4.6 Gender and Person in AMPR

In (239a) we saw an AMPR example where a bound pronoun is in the singular. It is also transparent, from the perspective of the current theory, why it is assigned feminine gender: if Sally (the denotation of the matrix subject) is identified as female, the local context forces feminine assignment.

There is also fake indexicals (fake person) in AMPR. In (242) I collected three real-life examples from a number of hits I found in a Google search.[20]

(242) *Fake Indexicals in AMPR*

a. I'm one of those who has **my** own favorite deviled egg recipes.
(www.thedomesticfront.com/deviled-eggs-with-capers-and-parsley).

b. I'm one of those who loves to make things **my**self from scratch.
(https://forum.kde.org/viewtopic.php?t=83569).

c. I'm one of those who likes to get minutes into **my** legs in pre-season.
(https://www.manchestereveningnews.co.uk/sport/football/football-news/bolton-doyle-wimbledon-league-one-21318238).

One further mystery is why in cases of fake personal pronouns, the *verb* doesn't indicates fake person; cf. *loves*, not *love* in (242b)). This is true also for Partee RCs in English: *I am the only one who is/*am.....* I will address this question in the next section, where I make a suggestion to reduce it to independent observations about the behavior of non-restrictive relative clauses. I will use the suggestion to explain also the difference between English and German about AMPR and Partee constructions.

[20] Longenbaugh's (2019) theory of AMPR is designed to account specifically for the number mismatch between the apparent head of the RC and the verb, and requires at least significant modifications to extend to deal with fake gender or person.

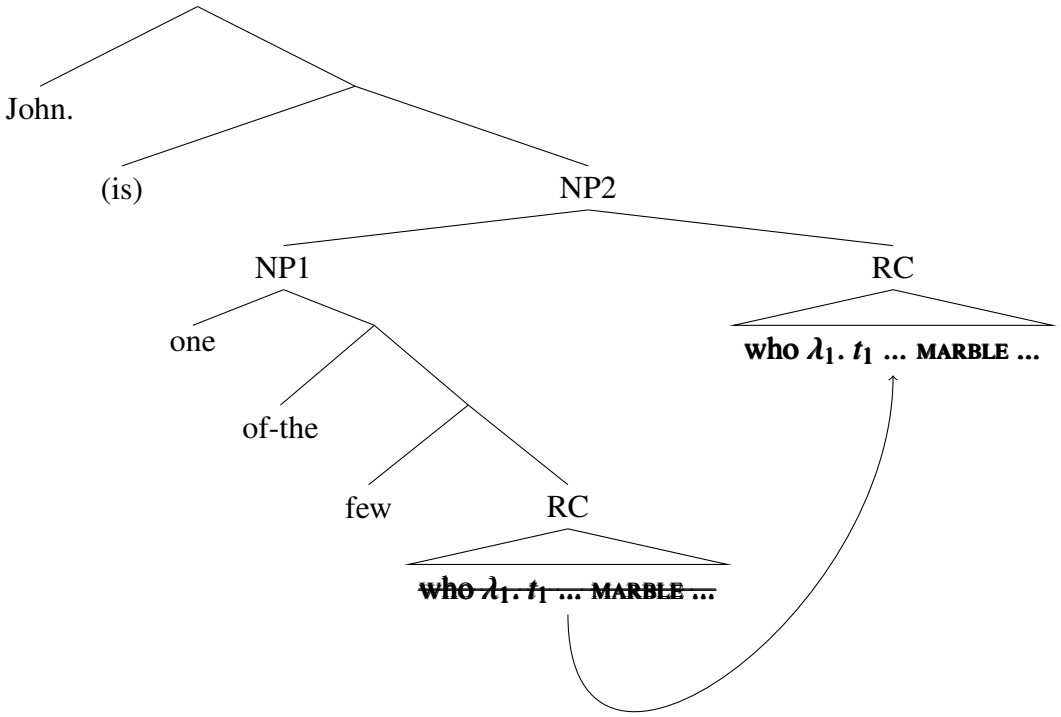
5.4.7 Number on full NPs/DPs in AMPR

Finally, there are also fake features on full NPs in AMPR, see (243a); the analysis I give is in (243b), the same way Partee sentences were handled.

(243) a. John is one of the few who picked up the marble next to him.

Can entail: *There are few people who picked up the marble or marbles next to them.*

b. LF:



5.5 Remarks on cross-linguistic variation

In this section I address the difference between German and English in licensing Fake Indexicals. I repeat the German example in (244).[21]

[21] While I focus just on German, I hope that my main conclusions can carry over to a wider range of languages that show restrictions on fake indexicals.

(244) (=193) ***German*** *does not allow Partee fake indexicals (in the singular)* (Kratzer, 2009)

Ich bin die einzige die **meine** Kinder versorg-t (✗)
I am the.sg.fem only who.sg.fem **my** children takes.care.of-3sg
'only the speaker is an x such that x takes care of {*the speaker's/**x's*} *children.'*

My perspective will link this fact to a known, though relatively underappreciated, fact due to Ito & Mester (2000) about how non-restrictive relative clauses behave in German.

Fake Indexical Traces

I will approach the German puzzle indirectly, by first addressing a seemingly remote problem, the one we encountered towards the end of the previous section. It was observed there that English never allows person agreement on RC-internal verbs in Partee constructions or AMPR:

(245) a. * I am the only one who **am** happy.

b. * I am one of the few students who **am** reading the article now.

Why is that? the question becomes more interesting when we learn that this fact is not universal. At least Hebrew, French, Greek and Brazilian Portuguese allow RC-internal person agreement, or what I will call 'Fake Indexical Trace' in Partee structures. The example below illustrate (notice that the sentences host a bound indexical pronoun too, but fake indexical traces are also ok without the pronoun):[22]

(246) a. ***French*** *person agreement in Partee structures*[23]

Je suis la seule qui **suis** partie de chez **moi** (✓)
I be.1sg the.fem only who be.**1sg** left from house of.**me**
'I am the only one who left from (my) home'

b. ***Greek*** *person agreement in Partee structures*

[22] In some languages, I found variation for some of the examples reported in this chapter. The judgments I report throughout reflect the intuitions of most consulted speakers.

[23] I thank Keny Chatain, Vincent Reuillard and Paul Marty for French judgments; Filipe Hisao Kobayashi and Suzana Fong for Brzilian Portuguese; and Sabine Iatridou for Greek. Analogous facts to (246) are reported to hold in Icelandic (Wurmbrand 2017b) Farsi (Ivan & Mirrazi 2019) as well.

Ime i moni pu ipeval**a** [tin dulia **mu**] egeros (✓)
am the only.fem.sg that submitted.**1sg** [the work **my**] on-time
'I am the only one who submitted my work on time'

c. ***Brazilian Portuguese*** *person agreement in Partee structures*

Eu sou o único que submeti um artigo que **eu** (mesmo) escrevi (✓)
I am the unique who submitted.**1sg** an article that **I** (SELF.1) wrote
'I am the only one who submitted an article I wrote'

d. ***Hebrew*** *person agreement in Partee structures*

ani ha-yexida Se- rai-**ti** ma kara la-horim Sel-**i** (✓)
I the-only.sg.fem that- saw-**1.sg** what happened to.the-parents of-**me**
'I am the only one who saw what happened to her parents'

Fake Indexical Traces appear also in AMPR. Here is an example from Hebrew:

(247) ***Hebrew*** *person agreement in AMPR*[24]

ani exad (me-)ha-anaSim ha-meatim Se- hiclaxti-**ti** lirot ma kara Sam. (✓)
I one (of-)the-people the-few that- managed-**1.sg** to.see what happened there
'I am one of the few people who (I-)managed to see what happened there'

It requires an explanation, then, why English prohibits fake indexical traces, even though it allows fake indexical pronouns. I suggest to reduce this fact to the observation, corroborated by a number of speakers I consulted, that non-restrictive relative clauses in English also disallow or disfavor this person agreement. Consider:

(248) *English non-restrictive agreement in 3rd even if the head is 1st(/2nd)*

a. ??I, who take it easy usually, didn't care about that.

b. I, who take**s** it easy usually, didn't care about that.

a. ??You, who take it easy usually, didn't care about that.

b. You, who take**s** it easy usually, didn't care about that.

[24] I have found numerous examples of this online and in real life.

These striking facts also hold in clefts constructions: many speakers strongly prefer *It is I/me who **is**...* over *It is I/me who **am**....*[25] And, importantly, these facts do not seem to hold in other languages. At least Hebrew and Brazilian Portuguese speakers require or strongly prefer to agree in person between the verb in a non-restrictive RC and its antecedent.

(249) *Hebrew person agreement in non-restrictive RCs*[26]

ani, Se- lakax-**ti**/*∅ et ze bekalut, ... (Hebrew)
I, that- took-**1.sg**/***3.sg** ACC that easily, ...
'I, who took it easily, ...'

While I do not have a deep explanation to the English peculiarity in (248), I propose that it underlies why English does not allow Fake Indexical Traces in Partee sentences and AMPR. To this end I state the condition in (250) and treat it as a hard-coded principle of English.

(250) Condition: English does not allow 1st(/2nd) person specification on verbs that agree with heads/traces of RCs.

This condition, I assume, trumps what the rules for feature valuation prescribe. To illustrate, a Partee sentence like (251a) is still assigned the LF in (251b) with extraposition, because of the fake indexical pronoun. As such, the verb should in principle show 1st person too (cf. the languages in (246)). But (250) kicks in and prevents 1st person agreement – either because English traces of RCs resist being assigned person in the first place, or because it is the verb that doesn't tolerate person agreement with a trace (we can remain neutral about that). The result is to recourse to 3rd person.

(251) a. I am the only one here who like**s my** parents. (✓)

[25]There are a number of debates online among speakers (with varying degrees of linguistic awareness) who are disturbed by this phenomenon. Here is a quote from one of the discussions:

> "The grammar of the pronoun 'who' has changed for many people so that it inherits number but not person from its antecedent: it now has (for them) a fixed third-person number (thus 'you who has', 'I who is')." (https://forum.fiverr.com/t/a-buyers-rights-to-use/27645)

[26]I thank Suzana Fong for confirming to me that Brazilian Portuguese behaves similarly. I haven't investigated this in other languages, but I expect the same to be true in any language that has fake indexical traces. .

b. I am [[the only one ~~[who$_1$ t_1 LIKE x_1 parents]~~] [who$_1$ t_1 LIKE x_1's parents]]

And sentences like (245) are bad in English (though not in many other languages) because of the condition in (250).

From the perspective of the current theory, the idea that (245) reduces to the behavior of non-restrictive RCs in (248) is not too surprising. Non-restrictive RCs have the semantic status of backgrounded information, not part of the assertive/at-issue content (Potts 2005 and others). And recall that the semantics proposed in this chapter for extraposed RCs in Partee sentences and AMPR make it function very similarly to a non-restrictive RC in this sense: while the RC is technically integrated as part of the standard assertive/truth-conditional level, its contribution is locally redundant (given the existence of the low copy). It is therefore useful to think of RCs in Partee sentences and AMPR as non-restrictive in some sense.

Back to German

With this fresh perspective, it is time to get back to German. German non-restrictive relative clauses show yet more peculiarity. Ito & Mester (2000) notice that German non-restrictive RCs headed by 1st/2nd (singular) pronouns do not allow a gap at all inside the RC—regardless of the person specification on the verb (252a,b). Rather, German must double the head of the RC, and then agreement proceeds as standard (252c).

(252) *German 1st person doubling in non-restrictive RCs (Ito & Mester 2000)*

a. * Ich, der sechzig bin, ...
I, who.MASC sixty am, ...
'I, who is/am sixty years old, ...'

b. * Ich, der sechzig ist, ...
I, who.MASC sixty is, ...

c. Ich, der **ich** sechzig bin/*ist, ...
I, who.MASC **I** sixty am/*is, ...

The facts are exactly the same for 2nd person:

(253) *German 2d person doubling in non-restrictive RCs (Ito & Mester 2000)*

a. * Du, der sechzig bist, ...
you.SG, who.MASC.SG sixty be.2SG, ...
'You, who is/are sixty years old, ...'

b. * Du, der sechzig ist, ...
you.SG, who.MASC.SG sixty is, ...

c. Du, der **du** sechzig bist/*ist, ...
you.SG, who.MASC.SG **you.SG** sixty be.2SG/*is, ...

With third person antecedents, however, the facts are reversed. Here doubling is disallowed, and the gap is ok:

(254) *German 3rd person non-doubling in non-restrictive RCs (Ito & Mester 2000)*

a. Er, der sechzig ist, ...
He, who.MASC sixty is, ...
'He, who is sixty years old, ...'

b. * Er, der **er** sechzig ist, ...
He, who.MASC **he** sixty is, ...

Based on these facts, I propose the condition in (255):

(255) Condition: If the trace (base position) of the internal head of a relative clause in German is specified for 1st(/2nd) person singular, that trace cannot be left silent, and must be realized overtly (with the relevant features); with 3rd person heads, the trace must be left silent.

Again, I am not proposing a deep explanation of (255), but it does justice to the emerging generalization. I want to show now that taking (255) under consideration helps account for the restricted distribution of Fake Indexicals in German. Recall:

(256) (German, =244)

Ich bin die einzige die **meine** Kinder versorg-t (✗)
I am the.sg.fem only who.sg.fem **my** children takes.care.of-3sg
*'only the speaker is an x such that x takes care of {the speaker's/*x's} children.'*

Importantly, fake indexicals are also impossible in the minimally different (257) (Kratzer 2009; Wurmbrand 2017b) where the verb shows 1st person agreement but without realizing the RC-internal gap position:

(257) *German 1sg verbal agreement still doesn't allow fake person* (Wurmbrand, 2017b)

Ich bin die einzige die **meine** Kinder versorg-**e** (✗)
I am the.sg.fem only who.sg.fem **my** children takes.care.of-**1sg**

But also importantly, Wurmbrand (2017b) reports that some German speakers allow doubling in this configuration, and for them the fake indexical reading then surfaces:

(258) *Pronoun doubling allows fake person for some speakers* (Wurmbrand, 2017b)

Ich bin die einzige die **ich meine** Kinder versorg-**e** (%)
I am the.sg.fem only who.sg.fem **I** **my** children takes.care.of-**1sg**

Based on my general account and (255), we account for this complex array of facts: both (256) and (257) are unacceptable because the trace inside the relative clause failed to be realized overtly ('pronoun doubling'). More specifically, recall that the only syntactic representation that gives fake indexicality a chance is one where the RC is extraposed and attached high, (259a). In that position, the trace of the RC gets assigned 1st person by feature valuation. The condition in (255) prescribes that the trace must be realized overtly. (256)-(257) fail to do that, whereas (258) doesn't.[27]

(259) a. I am [[the only one ~~[who$_1$ t_1 TAKE care of x_1 son]~~] [who$_1$ t_1 TAKE care of x_1's son]]

b. I am [the only one [who$_1$ t_1 TAKE care of x_1 son]]

The only syntactic representation that allows the trace not to be realized according to (255) is if the RC attaches low as in (259b). But a low attachment doesn't allow a fake-indexical reading of the pronoun.

Further predictions

As opposed to the case of singular, German plural antecedents do allow Fake Indexicality ((260), from Kratzer 2009):

[27] I am not sure why doubling the pronoun is apparently possible only for a subset of German speakers. Perhaps for the other speakers the condition in (255) is even stricter and disallows pronoun-doubling to rescue the structure when the RC is linearly remote from its antecedent.

(260) *German **does** allow fake indexicals in the plural* (Kratzer, 2009)

Wir sind die einzigen die **unseren** Sohn versorg-en (✓)
we are the.pl only.pl who **our** son take.care.of-1/3pl
'No one$_i$ other than us take care of our/their$_i$ son'

Given the perspective to link FI facts to non-restrictive relative clauses, (260) should correlate with plurals' behavior in non-restrictive RCs. The prediction is correct: as Ito & Mester (2000) show, no doubling need happen with plural 1st person antecedents.

(261) *Ito & Mester (2000)*

Wir, die sechzig sind, ...
We, who.PL sixty are, ...
'We, who are sixty years old, ...'

A high attachment of the RC is then possible for (260) even without doubling, which explains why fake indexicals are allowed.

5.6 A loose end

I end this chapter with a loose end: partial agreement in gender. The data is repeated:

(262) I am the only one who submitted **her** paper. (✓)

$\forall x \neq$ *me,* ***male or female****: x did not submit x's paper.*

If the structure for the relevant reading is in (263), repeated from (221), the system so far predicts that 1st person assignment to x_1 is obligatory.

(263) I am [[the only one ~~[who$_1$ t_1 submitted x_1 paper]~~] [who$_1$ t_1 submitted x_1's paper]]

I will not solve the issue here, but I would like to make some observations which I believe are relevant.

Given the NPI diagnostic, the following fact suggests that the RC that hosts a partially-agreeing pronoun ideed attaches high as represented in (263).

(264) I am the only one in this house who ever puts {*__her__/__their__} cloths in the hamper!
(*said by an angry mom to her husband and two boys.*)

If this is correct, the remedy seems to require a relaxation of the feature valuation rules; after all, in the high position the semantic value of the bound pronoun in its local context should be the speaker.

One possible way to formulate a generalization from this is as follows:

(265) Possible Gemeralization: If a pronoun is bound by a relative operator (*who*), then even if its ultimate antecedent (and therefore its local semantic value) is the speaker, its person features can be neutralized/deleted.

The following data suggest that the above generalization is independently correct. Namely, the behavior of relative clauses when positioned (more transparently) in predicative positions point to similar facts.

(266) a. I am not someone who thinks highly of **my/her**/??their mother.
b. I am not a person who thinks highly of **my/her**/??their mother.
c. I am not someone who likes to make a fool out of **my**self(/**him**self/?**them**selves).

How to formally encode (265), whether it can reduce to something else, and what are its consequences for the theory of local valuation, I leave for future research.

THIS PAGE INTENTIONALLY LEFT BLANK

Bibliography

Ackema, Peter, & Neeleman, Ad. 2018. *Features of person: From the inventory of persons to their morphological realization*. Linguistic Inquiry Monographs, vol. 78. MIT Press.

Adger, David, & Svenonius, Peter. 2011. Features in minimalist syntax. *The Oxford handbook of linguistic minimalism*, **1**, 27–51.

Alexiadou, Artemis. 2011. Plural mass nouns and the morpho-syntax of number. *Pages 33–41 of: Proceedings of the 28th west coast conference on formal linguistics*. Cascadilla Proceedings Project Somerville, MA.

Alexiadou, Artemis. 2019. Morphological and semantic markedness revisited: The realization of plurality across languages. *Zeitschrift für Sprachwissenschaft*, **38**(1), 123–154.

Anand, Pranav, & Nevins, Andrew. 2004. Shifty operators in changing contexts. *Pages 20–37 of: Semantics and linguistic theory*, vol. 14.

Bacovcin, Hezekiah Akiva, Zehr, Jérémy, & Schwarz, Florian. 2018. To accommodate or to ignore?: The presuppositions of again and continue across contexts. *Glossa: a journal of general linguistics*, **3**(1).

Bar-Lev, Moshe E. 2020. An Implicature account of Homogeneity and Non-maximality. *Linguistics and Philosophy*, 1–53.

Bassi, Itai. 2019. Fake Indexicals and their Sensitivity to Focus. *In: Proceedings of the North East Linguistic Society (NELS) 49, (Preprint: https://ling.auf.net/lingbuzz/004648)*.

Bassi, Itai, & Longenbaugh, Nick. 2018. Features on Bound Pronouns: an Argument Against Syntactic Agreement Approaches. *In: Proceedings of NELS 48*.

Beaver, David, & Velleman, Dan. 2011. The communicative significance of primary and secondary accents. *Lingua*, **121**(11), 1671–1692.

Beaver, David I, Clark, Brady, Flemming, Edward Stanton, Jaeger, T Florian, & Wolters, Maria. 2007. When semantics meets phonetics: Acoustical studies of second-occurrence focus. *Language*, **83**(2), 245–276.

Beaver, David Ian, & Clark, Brady. 2008. *Sense and Sensitivity: How Focus Determines Meaning*. Wiley-Blackwell.

Beck, Sigrid. 2016. Focus Sensitive Operators. *In:* Fery, Caroline, & Ishihara, Shinichiro (eds), *The Oxford Handbook of Information Structure*. Oxford University Press.

Bjorkman, Bronwyn M. 2017. Singular they and the syntactic representation of gender in English. *Glossa: a journal of general linguistics*, **2**(1), 80.

Bobaljik, Jonathan David, & Wurmbrand, Susi. 2012. Word order and scope: Transparent interfaces and the 3/4 signature. *Linguistic inquiry*, **43**(3), 371–421.

Bobaljik, Jonathan David, & Zocca, Cynthia Levart. 2011. Gender markedness: the anatomy of a counterexample. *Morphology*, **21**(2), 141–166.

Bonet, Eulàlia. 1995. Feature structure of Romance clitics. *Natural language & linguistic theory*, **13**(4), 607–647.

Bruening, Benjamin. 2019. *Generalizing the Presuppositional Approach to the Binding Conditions*. Unpublished Ms. Draft: https://udel.edu/ bruening/Downloads/BindingPresupp5.pdf.

Büring, Daniel. 2016. *Intonation and Meaning*. Oxford Surveys in Semantics and Pragmatics.

Büring, Daniel. 2019. Topless and Salient—Convertibles in the Theory of Focus. *Pages 137–155 of: The Semantics of Plurals, Focus, Degrees, and Times*. Springer.

Charlow, Simon. 2009. Strong predicative presuppositions. *In: Slides, workshop on 'New Directions in the Theory of Presupposition', ESSLLI*.

Charlow, Simon. 2019. *GIVENness and Local Contexts*. Talk at SALT 29, UCLA.

Charnavel, Isabelle. 2017. *Presupposition failure and intended pronominal reference: Person is not so different from gender after all*. LingBuzz MS.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Vol. 11. MIT press.

Chomsky, Noam, & Lasnik, Howard. 1977. Filters and control. *Linguistic inquiry*, **8**(3), 425–504.

Cooper, Robin. 1979. The interpretation of pronouns. *Pages 61–92 of:* Heny, F., & Schnelle, H. (eds), *Selections from the Third Groningen Round Table, Syntax and Semantics,*, vol. 10. New York: Academic Press.

Coppock, Elizabeth, & Beaver, David. 2015. Definiteness and determinacy. *Linguistics and Philosophy*, **38**(5), 377–435.

Corbett, Greville G. 1991. *Gender*. Cambridge University Press.

Crain, Stephen, Ni, Weijia, & Conway, Laura. 1994. Learning, parsing and modularity. *Perspectives on sentence processing*, 443–467.

Dahl, Östen. 1973. On so-called'sloppy identity'. *Synthese*, 81–112.

Danon, Gabi. 2011. Agreement and DP-Internal Feature Distribution. *Syntax*, **14**(4), 297–317.

de Swart, Henriëtte, & Farkas, Donka. 2010. The semantics and pragmatics of plurals. *Semantics and pragmatics*, **3**, 1–54.

Evans, Gareth. 1980. Pronouns. *Linguistic inquiry*, **11**(2), 337–362.

Fauconnier, Gilles. 1975. Polarity and the scale principle. *Chicago Linguistics Society*, **11**, 188–199.

Féry, Caroline, & Samek-Lodovici, Vieri. 2006. Focus projection and prosodic prominence in nested foci. *Language*, **82**(1), 131–150.

Fiengo, Robert, & May, Robert. 1994. *Indices and identity*. MIT press.

von Fintel, Kai. 1994. *Restrictions on quantifier domains*. Ph.D. Dissertation, University of Massachusetts, Amherst.

von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, **16**(2), 97–148.

Fox, Danny. 1999. Focus, Parallelism, and Accommodation. *Pages 70–90 of:* Matthews, T., & Strolovitch, Devon (eds), *Proceedings of the Ninth Conference on Semantics and Linguistic Theory*. Cornell University, Ithaca, NY: CLC Publications.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry*, **33**(1), 63–96.

Fox, Danny, & Johnson, Kyle. 2016. QR is restrictor sharing. *Pages 1–16 of: Proceedings of the 33rd West Coast Conference on Formal Linguistics. Somerville, Mass.: Cascadilla Proceedings Project*.

Fox, Danny, & Katzir, Roni. 2011. On the characterization of alternatives. *Natural Language Semantics*, **19**(1), 87–107.

Grimm, Scott. 2013. Plurality is distinct from number-neutrality. *Pages 247–258 of: Proceedings of the North East Linguistic Society*, vol. 41. University of Massachussetts at Amherst Amherst, MA.

Grinder, John, & Postal, Paul M. 1971. Missing antecedents. *Linguistic inquiry*, 269–312.

Grodzinsky, Yosef, & Sharvit, Yael. 2007. Coreference and self-ascription. *Manuscript, McGill University*.

Halle, Morris. 2000. Distributed morphology: Impoverishment and fission. *Amsterdam Studies in the Theory and History of Linguistic Science*, 125–150.

Hankamer, Jorge, & Sag, Ivan. 1976. Deep and surface anaphora. *Linguistic inquiry*, **7**(3), 391–428.

Harley, Heidi, & Noyer, Rolf. 2014. Distributed morphology. *Pages 463–496 of: The second Glot International state-of-the-article book*. De Gruyter Mouton.

Harley, Heidi, & Ritter, Elizabeth. 2002. Person and number in pronouns: A feature-geometric analysis. *Language*, 482–526.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts, Amherst.

Heim, Irene. 1983. On the projection problem for presuppositions. *Formal semantics–the essential readings*, 249–260.

Heim, Irene. 1997. Predicates or formulas? Evidence from ellipsis. *Pages 197–221 of: Semantics and linguistic theory*, vol. 7.

Heim, Irene. 1999. *Notes on Superlatives*. MIT lecture notes.

Heim, Irene. 2007. Forks in the Road to Rule I. *In: Proceedings of the Thirty-Eigth Annual meeting of the North East Linguistic Society (NELS 38)*.

Heim, Irene. 2008. Features on bound pronouns. *Phi theory: Phi-features across modules and interfaces*, 35–56.

Heim, Irene, & Kratzer, Angelika. 1998. *Semantics in generative grammar*. Vol. 13. Blackwell, Oxford.

Hestvik, Arild. 1995. Reflexives and ellipsis. *Natural Language Semantics*, **3**(2), 211–237.

von Heusinger, Klaus. 2007. Alternative semantics for definite NPs. *On Information Structure, Meaning and Form–Generalizations across languages*, **100**, 485–508.

Hirsch, Aron. 2017. *An inflexible semantics for cross-categorial operators*. Ph.D. thesis, Massachusetts Institute of Technology.

Hirsch, Aron, & Wagner, Michael. 2019. Only reconstruction and backwards association. *In: In Proceedings of the Amsterdam Colloquium*.

Ippolito, Michela. 2007. On the meaning of some focus-sensitive particles. *Natural Language Semantics*, **15**(1), 1–34.

Ito, Junko, & Mester, Armin. 2000. Ich, der ich sechzig bin: An Agreement Puzzle. *Jorge Hankamer WebFest*, **132**, 134.

Ivan, Rodica, & Mirrazi, Zahra. 2019. Farsi Fake Indexicals and embedded T Agreement: Predication Matters! (In Prep.). *In:* Farrell, Patrick (ed), *Proceedings of LSA 4*.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. MIT Press.

Jackendoff, Ray. 1997. *The architecture of the language faculty*. MIT Press.

Jacobson, Pauline. 1992. Antecedent contained deletion in a variable-free semantics. *Pages 193–214 of: Semantics and Linguistic Theory*, vol. 2.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and philosophy*, **22**(2), 117–185.

Jacobson, Pauline. 2012. Direct compositionality and 'uninterpretability': The case of (sometimes) 'uninterpretable' features on pronouns. *Journal of Semantics*, **29**, 305–343.

Kanazawa, Makoto. 2001. Singular donkey pronouns are semantically singular. *Linguistics and Philosophy*, 383–403.

Katz, Jonah, & Selkirk, Elisabeth. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, 771–816.

Katzir, Roni. 2013. A note on contrast. *Natural language semantics*, **21**(4), 333–343.

Keenan, Edward L. 1971. Names, quantifiers, and the sloppy identity problem. *Research on Language & Social Interaction*, **4**(2), 211–232.

Kitagawa, Yoshihisa. 1991. Copying identity. *Natural Language & Linguistic Theory*, **9**(3), 497–536.

Kramer, Ruth. 2016a. The location of gender features in the syntax. *Language and linguistics Compass*, **10**(11), 661–677.

Kramer, Ruth. 2016b. A split analysis of plurality: Number in Amharic. *Linguistic Inquiry*, **47**(3), 527–559.

Kratzer, Angelika. 1991. The representation of focus. *Pages 825–834 of:* von Stechow, Arnim, & Wunderlich, Dieter (eds), *Semantik: Ein internationales Handbuch der zeit-genössischen Forschung*. Walter de Gruyter.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. *Pages 92–110 of: Proceedings of SALT*, vol. 8.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, **40**(2), 187–237.

Kratzer, Angelika, & Selkirk, Elisabeth. 2020. Deconstructing information structure. *Glossa: a journal of general linguistics*, **5**(1).

Krifka, Manfred. 1991. A compositional semantics for multiple focus constructions. *Pages 127–158 of: In: Proceedings of Semantics and Linguistics Theory (SALT) I. Published under Cornell Working Papers in Linguistics 10*.

Krifka, Manfred. 2004. Focus and/or context: A second look at second occurrence expressions. *Context-dependence in the analysis of linguistic meaning*, 187–207.

Krifka, Manfred. 2006. Association with focus phrases. *Pages 105–136 of: The architecture of focus*, vol. 82. Walter de Gruyter.

Kripke, Saul A. 2009. Presupposition and anaphora: Remarks on the formulation of the projection problem. *Linguistic Inquiry*, **40**(3), 367–386.

Kučerová, Ivona. 2018. ϕ-Features at the Syntax-Semantics Interface: Evidence from Nominal Inflection. *Linguistic inquiry*, **49**(4), 813–845.

Ladusaw, William. 1979. Polarity Sensitivity as Inherent Scope Relations. Ph.D. dissertation, University of Texas at Austin.

Lakoff, George. 1971. On generative semantics. *1971*, **232**, 296.

Lakoff, George. 1973. Some thoughts on transderivational constraints.

Landau, Idan. 2016a. Agreement at PF: An argument from partial control. *Syntax*, **19**(1), 79–109.

Landau, Idan. 2016b. DP-internal semantic agreement: A configurational analysis. *Natural language & linguistic theory*, **34**(3), 975–1020.

Link, Godehard. 1983. The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach. *In:* Baeuerle, R., Schwarze, C., & Stechow, Arnim von (eds), *Meaning, Use and Interpretation of Language*. DeGruyter.

Löbner, Sebastian. 1989. Germanschon-erst-noch: An integrated analysis. *Linguistics and philosophy*, **12**(2), 167–212.

Löbner, Sebastian. 2000. Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences. *Linguistics and Philosophy*, **23**(3), 213–308.

Longenbaugh, Nicholas. 2019. Agreement mismatch in partitive relatives. *Linguistic Inquiry*, **50**(4), 847–861.

Maier, Emar, & de Schepper, Kees. 2010. Fake indexicals in Dutch: A challenge for Kratzer (2009). *Manuscript: Radboud University Nijmegen*.

Martí, Luisa. 2020. Inclusive plurals and the theory of number. *Linguistic Inquiry*, **51**(1), 37–74.

Matushansky, Ora. 2013. *Gender confusion*.

Mayr, Clemens. 2015. Plural definite NPs presuppose multiplicity via embedded exhaustification. *Pages 204–224 of: Semantics and Linguistic Theory*, vol. 25.

McKillen, Alanah. 2016. *On the interpretation of reflexive pronouns*. Ph.D. thesis, McGill University.

Merchant, Jason. 2014. Gender mismatches under nominal ellipsis. *Lingua*, **151**, 9–32.

Merchant, Jason. 2019. Ellipsis: A survey of analytical approaches. *The Oxford handbook of ellipsis*, 18–46.

Ouwayda, Sarah. 2014. *Where Number Lies: Plural marking, numerals, and the collective-distributive distinction*. University of Southern California dissertation.

Panagiotidis, Phoevos. 2021. *Towards a (minimalist) Theory of Features: preliminary notes*. Manuscript, retrieved from https://ling.auf.net/lingbuzz/005615/current.pdf?_s=O4sLizOrg3azbWC8.

Partee, Barbara. 1989. Binding implicit variables in quantified contexts. *Pages 342–365 of:* Wiltshire, Caroline, Graczyk, Randolph, & Music, Bradley (eds), *Papers from the 25th Regional Meeting of the Chicago Linguistic Society. Part One, The General Session*. Chicago Linguistic Society.

Pesetsky, David. 2013. *Russian case morphology and the syntactic categories*. Vol. 66. MIT Press.

Pesetsky, David, & Torrego, Esther. 2007. The syntax of valuation and the interpretability of features. *Phrasal and clausal architecture: Syntactic derivation and interpretation*.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford University Press.

Rappaport, Gilbert C. 2013. Determiner Phrases and Mixed Agreement in Slavic[1]. *The nominal structure in Slavic and beyond*, **116**, 343.

Reinhart, Tanya. 1983. Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy*, 47–88.

Ritter, Elizabeth. 1995. On the syntactic category of pronouns and agreement. *Natural Language & Linguistic Theory*, **13**(3), 405–443.

Romero, Maribel, & Novel, Marc. 2013. Variable binding and sets of alternatives. *Pages 174–208 of: Alternatives in semantics*. Springer.

Rooth, Mats. 1985. *Association with focus*. Ph.D. thesis, University of Massachusetts, Amherst.

Rooth, Mats. 1992a. Ellipsis redundancy and reduction redundancy. *In: Proceedings of the Stuttgart ellipsis workshop*, vol. 29. Citeseer.

Rooth, Mats. 1992b. A theory of focus interpretation. *Natural language semantics*, **1**(1), 75–116.

Ross, John. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, Massachusetts Institute of Technology.

Rullmann, Hotze. 2004. First and second person pronouns as bound variables. *Linguistic Inquiry*, **35**, 159–168.

Sag, Ivan A. 1976. *Deletion and logical form.* Ph.D. thesis, Massachusetts Institute of Technology.

Sauerland, Uli. 2003. A new semantics for number. *Pages 258–275 of: Semantics and Linguistic Theory*, vol. 13.

Sauerland, Uli. 2005 (January). *On the Semantic Markedness of* Φ*-Features*.

Sauerland, Uli. 2008. Implicated presuppositions. *The discourse potential of underspecified structures*, **8**, 581–600.

Sauerland, Uli. 2013. Presuppositions and the alternative tier. *Pages 156–173 of: Semantics and Linguistic Theory*, vol. 23.

Sauerland, Uli, & Alexiadou, Artemis. 2020. Generative Grammar: A Meaning First Approach. *Frontiers in Psychology*, **11**, 3104.

Sauerland, Uli, Anderssen, Jan, & Yatsushiro, Kazuko. 2008. The plural is semantically unmarked. *Pages 413–434 of: Linguistic evidence*. De Gruyter Mouton.

Schlenker, Philippe. 2005. Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics*, **13**(1), 1–92.

Schlenker, Philippe. 2008. Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics*, **34**(3), 157–212.

Schlenker, Philippe. 2009. Local Contexts. *Semantics and Pragmatics*, **2**(3), 1–78.

Schwarzschild, Roger. 1994. Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics*, **2**(3), 201–248.

Schwarzschild, Roger. 1999. GIVENness, AvoidF and other constraints on the placement of accent. *Natural language semantics*, **7**(2), 141–177.

Schwarzschild, Roger. 2019. The Representation of Focus, Givenness and Exhaustivity. *In:* Bhatt, Rajesh, Frana, Ilaria, & Menéndez-Benito, Paula (eds), *Making Worlds Accessible, A Festschrift for Angelika Kratzer*. UMass ScholarWorks.

Seuren, Pieter AM. 1984. Operator lowering. *Linguistics*, **22**(5), 573–627.

Sigurðsson, Halldór Ármann. 2019. Gender at the edge. *Linguistic Inquiry*, **50**(4), 723–750.

Singh, Raj. 2011. Maximize Presupposition! and local contexts. *Natural Language Semantics*, **19**(2), 149–168.

Spathas, Giorgos. 2010. *Focus on anaphora*. Ph.D. thesis, Universiteit Utrecht.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. *Presuppositions and implicatures in compositional semantics*, 243–281.

von Stechow, Arnim. 2003. Feature deletion under semantic binding. *Pages 377–403 of:* Kadowaki, Makoto, & Kawahara, Shigeto (eds), *Proceedings of North East Linguistic Society (NELS) 33*. Amherst: University of Massachusetts.

Steriopolo, Olga, & Wiltschko, Martina. 2010. *Distributed Gender Hypothesis*. na.

Stokke, Andreas. 2020. Features of referential pronouns and indexical presuppositions. *Inquiry*, 1–33.

Sudo, Yasutada. 2012. *On the Semantics of Phi Features on Pronouns*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Sudo, Yasutada. 2014. Dependent plural pronouns with Skolemized choice functions. *Natural Language Semantics*, **22**(3), 265–297.

Sudo, Yasutada, & Spathas, Giorgos. 2016. Nominal ellipsis and the interpretation of gender in Greek. *Pages 712–729 of: Proceedings of Sinn und Bedeutung*, vol. 20.

Sudo, Yasutada, & Spathas, Giorgos. 2020. Gender and interpretation in Greek: Comments on Merchant (2014). To appear in *Glossa*.

Sugawara, Ayaka. 2016. *The role of question-answer congruence (QAC) in child language and adult sentence processing*. Ph.D. thesis, Massachusetts Institute of Technology.

Takahashi, Shoichi, & Fox, Danny. 2005. MaxElide and the re-binding problem. *Pages 223–240 of: Proceedings of SALT*, vol. 15.

Tancredi, Christopher. 1992. *Deletion, Deaccenting and Presupposition: Massachusetts Institute of Technology Ph. D.* Ph.D. thesis, dissertation.

Tomaszewicz, Barbara Maria. 2015. *Superlative ambiguities: A comparative perspective*. Ph.D. thesis, University of Southern California.

Tomioka, Satoshi. 1999. Sloppy identity puzzle. *Natural Language Semantics*, **7**(2), 217–241.

Wagner, Michael. 2006. Association by movement: evidence from NPI-licensing. *Natural Language Semantics*, **14**(4), 297–324.

Wagner, Michael. 2012. Focus and givenness: A unified approach. *Contrasts and positions in information structure*, 102–47.

Williams, Edwin S. 1977. Discourse and logical form. *Linguistic inquiry*, 101–139.

Wold, Dag E. 1996. Long distance selective binding: The case of focus. *Pages 311–328 of: Proceedings of Semantics and Linguistic Theory*, vol. 6.

Wurmbrand, Susi. 2015 (November). *Fake indexicals, feature sharing, and the importance of gendered relatives*. MIT Colloquium.

Wurmbrand, Susi. 2017a. Feature sharing or how I value my son. *Claire Halpert, Hadas Kotek & van Urk (eds.), The Pesky Set: Papers for David Pesetsky*, 173–182.

Wurmbrand, Susi. 2017b. Icelandic as a partial null subject language: Evidence from fake indexicals. *Pages 339–345 of:* Baile, Michelle Sheehan & Laura R. (ed), *Order and structure in syntax: Subjecthood and argument structure*, vol. II. Language Science Press.

Yanovich, Igor. 2010. On the nature and formal analysis of indexical presuppositions. *Pages 272–291 of: JSAI International Symposium on Artificial Intelligence, 2009*. Springer.

Yatsushiro, Kazuko, & Sauerland, Uli. 2006. [feminine] in a high position. *Snippets*, **13**, 11–12.

Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy*, **32**(4), 353–407.