## **Composition and projection across modalities**

#### January 2020

**Abstract** In this paper, I argue for a composition-driven approach to projection from under semantic operators that uniformly applies to compositionally integrated content in all modalities (lexicalized spoken expressions, gestures, facial expressions, suprasegmental morphemes). I focus on two composition strategies associated with two distinct projection behavior patterns, modification and supplementation. I show that this distinction manifests itself in the same way in all modalities and determines the projection behavior of a range of content types, including spoken adjectives and appositives, conventionalized and non-conventionalized gestures, degree modifiers exponed in various modalities, *phi*-features on spoken pronouns, and endpoints of directional gestures. This paper, thus, has two main implications. First, it offers a general explanatory approach to investigating projection of compositionally integrated content. Second, it establishes the need for treating secondary modality expressions as first class citizens at all levels of representation.

## Contents

1 Introduction			2		
2	Modifiers and supplements				
	2.1	Modifiers	4		
	2.2	Supplements	13		
	2.3	Section summary	17		
3	Com	position and projection in secondary modalities	17		
	3.1	Previous approaches to secondary modality content	17		
	3.2	Architectural assumptions	21		
	3.3	Non-conventionalized gestures	21		
	3.4	Conventionalized gestures	27		
	3.5	Degree modifiers cross-modally	29		
	3.6	Compositionally non-integrated secondary modality content	34		
4	Proj	ection of pronoun-internal modifiers	37		
	4.1	<i>Phi</i> -features on spoken pronouns	37		
	4.2	Endpoints of directional gestures	45		
5	Con	clusion	47		
References			48		

# **1** Introduction

Some types of semantic content have a potential or a requirement to project, i.e., to get interpreted outside the semantic scope of various operators, such as negation, conditional and question operators, modals, etc., despite appearing to be in their syntactic scope. This phenomenon is illustrated in (1a) for gender features on pronouns and in (1b) for appositives; in both cases, the expression of interest contributes an inference that projects from under *if* instead of making a local truth-conditional contribution.

- (1) a. If Skyler<sub>i</sub> brings her<sub>i</sub> dog, I'll give you \$10.
   → Skyler is female.
   ≠ If (Skyler is female and Skyler brings her dog)...
   b. If you invite Zoe, (who is) a stuntwoman, you should show her your muscle car.
  - b. If you invite Zoe, (who is) a stuntwoman, you should show her your must  $\rightarrow$  Zoe is a stuntwoman.
    - $\neq$  If (Zoe is a stuntwoman and you invite her)...

Recently, formal semantics/pragmatics has seen an upsurge in studying how secondary modality content (mostly hand gestures, but also facial expressions and suprasegmental morphemes) contributes to the meaning of spoken language utterances (Lascarides and Stone 2009; Ebert and Ebert 2014; Ebert 2017; Tieu et al. 2017, 2018; Hunter 2018; Schlenker 2018a,b, To appear; Zlogar and Davidson 2018, a.o.). Much of this work has focused on the projection behavior of non-conventionalized, purely iconic gestures, such as in (2).<sup>1</sup>



(2) a. Lea might bring  $\longrightarrow$  her dog<sub>LARGE</sub>.  $\rightarrow$  Lea's dog is large.



b. Zoe might shoot at the target<sub>LONGBOW</sub>.  $\rightarrow$  If Zoe shoots, she'll shoot a longbow.

Most of this literature has assumed that projection behavior patterns of secondary modality content are determined by modality-specific rules that furthermore rely on its surface properties, such as how it is linearized with respect to the primary modality content. This view is most prominent in Schlenker's (2018b) typology of "iconic enrichments", which aims to predict if/how a piece of secondary modality content X projects, based on: (i) whether X co-occurs with something in a more primary modality, and (ii) whether X is an "internal" or "external" enrichment, i.e., whether it is "syntactically eliminable". Thus, CO-SPEECH GESTURES, such as in (2), co-occur with speech, which a more primary modality, and are "external", as they can be omitted. Such content is predicted by Schlenker's typology to uniformly trigger assertion-dependent presuppositions of a special kind, "cosuppositions". This is in contrast to Ebert and Ebert's (2014) earlier claim that co-speech gestures are uniformly supplements, akin to appositives, and project as such.

<sup>&</sup>lt;sup>1</sup>I write gesture labels in ALL CAPS. Co-speech gestures are written as subscripts, with underlining approximately indicating their temporal alignment without making any syntactic claims. Illustrations are sometimes added at the approximate onset of the gesture.

I argue that such modality-specific typologies of projection behavior patterns relying on surface properties thereof are neither tenable nor necessary, as the same linguistic principles determine the choice of a projection behavior pattern for a given piece of content in all modalities. These principles are furthermore linked to the compositional properties of said piece of content. In this paper, I focus on two major composition strategies, MODIFICATION and SUPPLEMENTATION, associated with two distinct projection behavior patterns, as discussed in section 2.<sup>2</sup> In particular, MODIFIERS, i.e., roughly, expressions combining with  $\beta$  of type  $\tau$  and yielding  $\alpha$  of type  $\tau$ , such as attributive adjectives, only project when they are NON-RESTRICTING, i.e., truth-conditionally vacuous at the utterance level (a refined version of the definition from Leffel 2014), as in (3).

## (3) Context: The speaker believes that processed meat causes cancer. Maybe I shouldn't be eating so many deadly sausages. → Maybe I shouldn't be eating so many sausages. (truth-conditional vacuity)

 $\rightarrow$  All sausages are deadly.

(projecting inference)

Whether or not a given instance of a modifier is non-restricting is mostly guided by pragmatic considerations, and, unless a given subsective modifier is always composing with an expression that cannot be restricted any further, it always has the potential to be restricting, i.e., truth-conditionally non-vacuous, in its specific instances. In contrast, SUPPLEMENTS, i.e., roughly, expressions that pass the expression they combine with unchanged for compositional purposes, but assert the content of a proposition about said expression "on the side" (as in, e.g., Potts 2005, Koev 2013, or AnderBois et al. 2013), such as the appositive in (1b), project conventionally, i.e., much more systematically, and exceptional non-projecting instances of supplements are much more constrained than restricting modifiers.

In section 3, I show that the modifier vs. supplement distinction manifests itself in the same way across all modalities, which is expected under the view of the architecture of grammar that assumes late vocabulary insertion, whereby composition-based distinctions are blind to how exactly various parts of a given compositional structure get exponed. Modality-specific effects do, of course, emerge in the pragmatics and in the phonology and are usually gradient. For instance, co-occurrence with something in the primary modality might make a secondary modality modifier more likely to be interpreted as non-restricting. Furthermore, articulatory considerations affect multi-modal prosodic integration of expressions within an utterance, and the prosody/syntax interface imposes constraints on which syntactic-and, therefore, compositional-construals are compatible with the resulting surface structure. However, while these effects need to be studied in their own right, they do not directly determine the choice of a broad projection behavior pattern for a given expression. I show that this cross-modal, properly linguistic approach is efficient in explaining projection behavior of conventionalized and non-conventionalized gestures and degree modifiers exponed in various modalities. Previous approaches to projection of secondary modality content that rely on modality-specific, pre-linguistic distinctions such as "co-speech" vs. "postspeech" or "internal" vs. "external enrichment" fail to capture the rich empirical patterns observed for secondary modality content and, most importantly, the robust parallels across all modalities.

Having developed this cross-modal, composition-driven approach to projection, I also apply it to *phi*-features on pronouns in section 4. I propose that these are pronoun-internal modifiers

<sup>&</sup>lt;sup>2</sup>Naturally, this should not be taken to mean that I believe that these are the only composition strategies or projection behavior patterns that exist in natural language.

that always modify expressions whose extensions are singleton sets, which is why they are always non-restricting and, therefore, truth-conditionally vacuous. In this respect, *phi*-features behave like other pronoun-internal modifiers such as those in Khoekhoe, first described in Lee 2019, which, as I show, are also obligatorily non-restricting. This approach to *phi*-features is, thus, more explanatory and more morphosyntactically plausible than the standard treatment thereof as lexical presuppositions (Cooper 1983 et seq.). I also show that a similar treatment can be extended to endpoints of directional gestures, which have been previously argued to be akin to *phi*-features in Schlenker and Chemla 2018.

In section 5, I summarize the main points of this paper.

Videos or audios of some of the examples used in this paper can be downloaded at https://gofile. io/?c=TalDjp. When a file is available, its name is given in parentheses following the example. Some of the videos were excluded from this version of the paper for anonymization purposes.

## 2 Modifiers and supplements

#### 2.1 Modifiers

#### 2.1.1 Composition of modifiers

Modification is one of the basic and intuitive composition strategies used in natural language, even though it is not easy to define what counts as a modifier. I will define a MODIFIER as an expression  $\gamma$  that combines with an expression  $\beta$  of type  $\tau$  yielding  $\alpha$  of type  $\tau$ . I will also assume that there are further morphosyntax-based rules for determining which of the two sisters (if either) in a given configuration is the modifier and which one is the expression being modified. For example, in cases of intersective modification, the expression being modified is the one that projects its label or features to the parent node.

As is well-known, modifiers can be SUBSECTIVE, like *blond* and *skillful* in (4), and NON-SUBSECTIVE, like *alleged* in (5).

- (4) Zoe is a {blond, skillful} stuntwoman.  $\rightarrow$  Zoe is a stuntwoman.
- (5) Daisy is an alleged murderer.  $\not\rightarrow$  Daisy is a murderer.

In this paper I will for the most part focus on subsective modifiers, but I will come back to nonsubsective modifiers in section 3.4.

Intuitively, in a subsective modification configuration, the result of modification entails the expression being modified. What does it mean for two sub-propositional expressions to entail one another? A sub-propositional expression  $\alpha$  whose type ends in t can entail another expression of the same type  $\beta$  via GENERALIZED ENTAILMENT ( $\alpha \Rightarrow_{\forall} \beta$ ), which simply means that  $\alpha$  entails  $\beta$  via GENERALIZED MATERIAL IMPLICATION ( $\alpha \Rightarrow \beta$ ) for any values of the arguments  $\alpha$  and  $\beta$  take. The full definitions are given below. I will call the entailment relation between the result of subsective modification and the expression being modified the SUBSECTIVE ENTAILMENT.

#### (6) **GENERALIZED MATERIAL IMPLICATION<sup>3</sup>**

 $\llbracket \alpha_{\langle \tau_1 \dots \tau_n, t \rangle} \Rightarrow \beta_{\langle \tau_1 \dots \tau_n, t \rangle} \rrbracket = \lambda X_{\tau_1}^1 \dots \lambda X_{\tau_n}^n \cdot \llbracket \alpha \rrbracket (X^1) \dots (X^n) \to \llbracket \beta \rrbracket (X^1) \dots (X^n)$ 

(7) **GENERALIZED ENTAILMENT** 

 $[\![\alpha_{\langle \tau_1...\tau_n,t\rangle} \Rightarrow_\forall \beta_{\langle \tau_1...\tau_n,t\rangle}]\!] = 1 \text{ iff } \forall X^1_{\tau_1}...\forall X^n_{\tau_n}[[\![\alpha \Rightarrow \beta]\!](X^1)...(X^n)]$ 

#### (8) **SUBSECTIVE ENTAILMENT**

In a tree  $\alpha$  whose daughters are  $\gamma$  and  $\beta$  and  $\gamma$  modifies  $\beta$ , SUBSECTIVE ENTAILMENT obtains iff  $\alpha \Rightarrow_{\forall} \beta$ .

Subsective modifiers give rise to subsective entailment as their characteristic property, but this doesn't yet tell us anything about how they compose with the expressions they modify. I will maintain that all subsective modifiers use the composition strategy in (9), whose left conjunct assures that the subsective entailment goes through.

#### (9) **SUBSECTIVE MODIFICATION (composition strategy)**

In a tree  $\alpha_{\langle \tau_1...\tau_n,t\rangle}$  whose daughters are  $\gamma$  and  $\beta_{\langle \tau_1...\tau_n,t\rangle}$ ,  $\gamma$  is a SUBSECTIVE MODIFIER iff  $\gamma$  is a modifier and:

$$\llbracket \alpha \rrbracket = \lambda X_{\tau_1}^1 \dots \lambda X_{\tau_N}^n \cdot \llbracket \beta \rrbracket (X^1) \dots (X^n) \land (\dots)$$

The right conjunct can be as simple as  $[\![\gamma]\!](X_1)...(X_n)$  for intersective modifiers such as *blond* in (10). If we assume that intersective modifiers are of the same type as the expression they modify and compose via predicate modification (which I will, for the sake of simplicity, although nothing hinges on this assumption), the two conjuncts in the denotation of  $\alpha$  are just the result of that.

(10) a. Zoe is a blond stuntwoman. Zoe is a cellist.

 $\rightarrow$  Zoe is a blond cellist.

b. [[blond stuntwoman]] =  $\lambda x \lambda w$ .stuntwoman $(x)(w) \wedge blond(x)(w)$ 

Non-intersective subsective modifiers, such as *skillful* in (11), are usually assumed to take the expression they modify as an argument (i.e., they are of type  $\langle \tau, \tau \rangle$ ). I will assume that *skillful* and its kin still follow the schema in (9), though, with the left conjunct responsible for the subsective entailment instead of the latter being some lexical idiosyncratic property of the skillful relation. For example, for *skillful stuntwoman*, I will assume the denotation in (11b).

(11) a. Zoe is a skillful stuntwoman. Zoe is a cellist.

 $\not\rightarrow$  Zoe is a skillful cellist.

b. [[skillful stuntwoman]] =  $\lambda x \lambda w$ .stuntwoman $(x)(w) \wedge$  skillful(stuntwoman)(x)(w)

Under the standard Neo-Davidsonian event semantics (Carlson 1984; Parsons 1990, a.o.), modification in the verbal domain works very similarly to that in the nominal domain. Verbs start out as properties of events of type  $\langle v, st \rangle$  and are modified by other properties of events (arguments

<sup>&</sup>lt;sup>3</sup>Generalized material implication is a lifted version of material implication that can apply to sub-propositional expressions, just like generalized conjunction (Partee and Rooth 1983) is a lifted version of logical conjunction. The definition in (6) is adopted from Schlenker 2018a, fn. 25, with minor notational changes.

and adjuncts) via predicate modification, up until existential closure applies and binds the event variable. E.g., a VP *ran fast*, which entails *ran*, is composed as in (12b).

(12) a. Zoe ran fast.  

$$\rightarrow$$
 Zoe ran.  
b.  $[[ran fast]] = \lambda e \lambda w.ran(e)(w) \wedge fast(e)(w)$ 

Some non-intersective subsective modifiers can have even more involved denotations. For example, intuitively, degree modifiers like *very* and *completely* are subsective modifiers, which modify other modifiers and give rise to subsective entailment, as shown in (13a). To make them fit the schema in (9), I make an adjustment to the classical analysis of degree modifiers in Kennedy and McNally 2005 (where degree modifiers are not modifiers in the sense that they are not of type  $\langle \tau, \tau \rangle$ ) by separating the existential closure over degrees from the contribution of the modifier proper, as shown in (13b) for *completely full* (d ranges over degrees;  $d = \max(S_{\text{full}})$  reads roughly as 'd is the maximal degree on the scale of fullness').

(13) a. The glass is completely full.  $\rightarrow$  The glass is full.

b.

(i) *before existential closure* 

[completely full] =  $\lambda d\lambda x \lambda w.d = \max(S_{\text{full}})(w) \wedge \text{full}(x)(w) = d$ 

(ii) after existential closure  $[completely full] = \lambda x \lambda w. \exists d[d = \max(S_{\mathsf{full}})(w) \land \mathsf{full}(x)(w) = d]$ 

This adjustment is independently motivated by the fact that multiple degree modifiers are possible, as shown in (14), and makes composition of adjective phrases more similar to that of nominal and verbal projections, whereby we start out with an expression of a certain type ( $\langle e, st \rangle$ ,  $\langle v, st \rangle$ , or  $\langle d, \langle e, st \rangle \rangle$ ), and we can modify it by taking it and returning an expression of the same type however many times we need, until we merge a determiner-like element (determiner or existential closure) and turn it into a different type.

(14) Zoe is an [incredibly, [extremely, [amazingly talented]]] stuntwoman.

#### 2.1.2 **Projection of modifiers**

Now that we have established modification as a composition strategy, let us talk about when and how modifiers project.

All subsective modifiers are RESTRICTIVE, which means that they have the compositional potential to restrict the expressions they modify, i.e., roughly, to yield logically stronger expressions; this potential is the result of having the right conjunct in the schema in (9). However, not all specific instances of modifiers realize this potential. Leffel (2014) talks about NON-RESTRICTING MODIFIERS<sup>4</sup> (NRMs) as truth-conditionally vacuous modifiers, i.e., modifiers that don't affect the truth conditions of the utterance in which they appear. He illustrates this observation via examples like (15). Under the world-knowledge based assumption that not all chemicals are harmful, but all toxins are, (15a) doesn't entail a version of itself without the adjective, but (15b) does.

<sup>&</sup>lt;sup>4</sup>He uses the terms "non-restrictive" and "non-restricting" differently from me.

- (15) a. I will eliminate every harmful chemical.  $\Rightarrow$  I will eliminate every chemical. restricting modifier
  - b. I will eliminate every harmful toxin.  $\rightarrow$  I will eliminate every toxin.

More specifically, Leffel (2014) defines NRMs as (i) (subsective) modifiers intended by the speaker as truth-conditionally vacuous (the non-restrictingness condition proper), (ii) the inference contributed by the NRM is relevant in the discourse (the relevance constraint). Leffel's (simplified) definition of NRMs is given in (16).<sup>5</sup>

#### (16) Non-restricting modifiers in Leffel 2014, (3.58a) (simplified)

An occurrence  $\alpha_n$  of a modifier  $\alpha$  in the phrase  $[_{DP}...\alpha_n...N...]$  or  $[_{DP}...N...\alpha_n...]$  is non-restricting with respect to index *i* iff:

- a. the speaker believes that  $[DP]^i = [DP[\alpha_n/\varepsilon]]^i$ , where X[a/b] is the expression just like X except with a token of a replaced by one of b and  $\varepsilon$  is the empty string; and (non-restricting intention)
- b. there is a discourse relation between the non-restricting modifier inference and some discourse unit in the context. (relevance constraint)

I propose a refined version of Leffel's original definition in (17).

#### (17) NON-RESTRICTING MODIFIER (adopted here)

If an utterance u contains a subtree  $\alpha$  whose daughters are  $\gamma$  and  $\beta$  and  $\gamma$  is an occurrence of a subsective modifier,  $\gamma$  is NON-RESTRICTING at an index i iff:

- a. the speaker of *i* believes that  $\llbracket u \rrbracket^i = \llbracket u[\gamma/\varepsilon] \rrbracket^i$ , where X[a/b] is the expression just like X except with a token of *a* replaced by one of *b* and  $\varepsilon$  is the empty string; and (non-restricting intention)
- b. there is a discourse relation between the non-restricting modifier inference and some discourse unit in the context. (relevance constraint)

My definition is different from Leffel's in several respects. First, Leffel's definition is restricted to adnominal modifiers; I make it more general. Also, Leffel eventually excludes non-subsective adjectives like *former* and *alleged* from consideration, but this restriction is not part of his original definition; I make it explicit in the definition. Most importantly, I switch to global, utterance-level equivalence from local, constituent-level one. I introduce this last change because, as we will see shortly, Leffel's original definition is at odds with the projection behavior of inferences contributed by NRMs, and we need to separate the definitional utterance-level truth-conditional vacuity of NRMs, which holds globally, from the inference about the local equivalence of the expression being modified and the result of modification, which is sensitive to local contexts.

While NRMs are truth-conditionally vacuous, they still make a meaningful contribution in the form of NON-RESTRICTING MODIFIER INFERENCES (NRM inferences) that the expression being modified entails the result of modification ( $\beta \Rightarrow_{\forall} \alpha$ ).<sup>6</sup> For example, in (15b), the NRM inference

non-restricting modifier

<sup>&</sup>lt;sup>5</sup>He formalizes the relevance constraint; for the purposes of this paper, this is not necessary, so I am omitting his formalisms for simplicity.

<sup>&</sup>lt;sup>6</sup>Or, rather, that the two entail each other ( $\alpha \Leftrightarrow_{\forall} \beta$ ), however, the entailment from the result of modification to the expression being modified ( $\alpha \Rightarrow_{\forall} \beta$ ) is already assured by the modifier being subsective.

is that being a toxin entails being a harmful toxin, or, simply put, that all toxins are harmful.

However, unlike what is suggested by Leffel's definition in (16a), it is not the case that this entailment has to hold in all of the speaker's belief worlds. Rather, what the speaker has to believe is that this entailment holds relative to the LOCAL CONTEXT (Karttunen 1974; Stalnaker 1974; Heim 1983; Schlenker 2009, a.o.) of the modification configuration. For example, in both examples in (18), the speaker intends the modifier as truth-conditionally vacuous. However, in (18a), the (propositional) local context of the modification configuration is equivalent to the Stalnakerian global context set C (Stalnaker 1974), thus, we get a global inference that (the speaker believes that) being a sausage entails being a deadly sausage, i.e., that all sausages are deadly. In (18b), however, we get a conditionalized inference that (the speaker believes) that if processed meat causes cancer, all sausages are deadly.

# (18) a. *Context: The speaker believes that processed meat causes cancer.* I shouldn't be eating so many deadly sausages.

 $\rightarrow$  I shouldn't be eating so many sausages.

- $\rightarrow$  All sausages are deadly.
- b. Context: The speaker just read an article saying that processed meat might be causing cancer, but they are not ready to embrace it as a fact just yet.
  {Maybe processed meat causes cancer and, If processed meat causes cancer,} I shouldn't be eating so many deadly sausages.

 $\rightarrow$  {Maybe processed meat causes cancer and, If processed meat causes cancer,} I shouldn't be eating so many sausages.

- $\not\rightarrow$  All sausages are deadly.
- $\rightarrow$  If processed meat causes cancer, all sausages are deadly.

The same contrast holds in (19).

- (19) *Context: Pam only has one mother.* 
  - a. Pam is still taking care of her sick mother.
    - $\rightarrow$  Pam is still taking care of her mother.
    - $\rightarrow$  Pam's mother is sick.
  - b. Either Pam's mother has recovered, or Pam is still taking care of her sick mother.
    - $\rightarrow$  Either Pam's mother has recovered, or Pam is still taking care of her mother.
    - $\not\rightarrow$  Pam's mother is sick.
    - $\rightarrow$  If Pam's mother hasn't recovered, Pam's mother is sick.

For comparison, local context sensitivity of lexical presuppositions is illustrated for know in (20).

- (20) a. Zoe knows that she can't continue the race.
  - $\rightarrow$  Zoe can't continue the race.
  - b. (i) Maybe Zoe's car has been damaged and...
    - (ii) Either Zoe's car hasn't be damaged, or...
    - (iii) If Zoe's car has been damaged, ...

...she knows that she can't continue the race.

 $\not\rightarrow$  Zoe can't continue the race.

 $\rightarrow$  If Zoe's car has been damaged, she can't continue the race.

Now, with lexical presuppositions, we do often get stronger inferences than one would expect if local contexts are always taken into account. This issue is known as the proviso problem (Geurts 1996 et seq.), illustrated in (21).

- (21) a. If Bridget is a spy for the Allies, she won't expose her British contact.  $\rightarrow$  If Bridget is a spy for the Allies, she has a British contact.
  - b. If Bridget is smart, she won't expose her British contact.
    - - $\rightarrow$  Bridget has a British contact.

The proviso problem arises for NRM inferences as well:

- (22) a. If processed meat causes cancer, you shouldn't eat so many deadly sausages.  $\rightarrow$  If processed meat causes cancer, all sausages are deadly.
  - b. If you care about your health, you shouldn't eat so many deadly sausages.
    - If you care about your health, all sausages are deadly.
    - $\rightarrow$  All sausages are deadly.

To sum up, NRM inferences do indeed project similarly to lexical presuppositions. However, we don't have to maintain that they are triggered in the semantics, be it lexically or as part of a composition rule (as is proposed in Leffel 2014). I propose that NRM inferences are triggered pragmatically whenever the addressee concludes that the speaker intends for a given instance of a modifier to be non-restricting, i.e., truth-conditionally vacuous.

There are many reasons why one might think that a given instance of a modifier is intended as truth-conditionally vacuous. The most obvious case is when the modifier modifies an expression whose extension is known to be a singleton set, as in (19). There, if the addressee knows (or assumes) that Pam has one mother, they will interpret *sick* as non-restricting, i.e., truth-conditionally vacuous, and will generate the NRM inference that will assure said vacuity (i.e., that Pam's mother is sick in (19a), or that if Pam's mother hasn't recovered, she is sick in (19b)). If they didn't already believe the content of the NRM inference, they will then tacitly update the context set accordingly via global accommodation, or, if for some reason they are unwilling to do so, they can push back by saying something like *Hmm, I didn't know that Pam's mother was sick*. This global accommodation process is, thus, no different from global accommodation of lexical presuppositions, i.e., presuppositions arising as part of the lexical meaning of specific expressions such as *stop* or *know*, either, even though the triggering mechanism is different in the two cases.<sup>7</sup>

Note that there technically is a way to further restrict an expression whose extension is a singleton set, which is to yield the empty set. However, this doesn't seem to be empirically possible, as is illustrated in (23).<sup>8</sup>

(23) Context: Zoe, Lucy, and Pam are going on a camping trip. Pam is coming from Boston to New York to join the rest of the group. Zoe and Lucy are discussing how to get to the

<sup>&</sup>lt;sup>7</sup>This is not to say that none of the lexical presuppositions are triggered pragmatically. See, for example, Schlenker 2019 for a recent proposal that some presuppositions arising as part of the meaning of lexicalized and non-lexicalized expressions in various modalities are triggered pragmatically, as epistemic preconditions. This mechanism, however, cannot account for triggering of NRM inferences/"cosuppositions", as is acknowledged by Schlenker himself.

<sup>&</sup>lt;sup>8</sup>Bold indicates prosodic contrastive focus marking.

camping site from New York. ...

a. ... *They have agreed that they need a car, no matter how big. Lucy:* 

?Well, Pam is coming from Boston. I don't know if she **has** a car, but if she's coming in her **car**, we can use it to get to the camping site.

'...if (she has a car and she's coming in her car)...'

b. ...*They have agreed that they need a large car to fit all their supplies. Lucy knows that Pam has exactly one car. Lucy:* 

#Well, Pam is coming from Boston. I don't know if her car is **large**, but if she's coming in {her **large** car, her large **car**}, we can use it to get to the camping site.

Intended: '...if ({she has a large car, her car is large} and she's coming in her large car)...'

In (23a), the existence inference of the definite description *her car* cannot be satisfied globally, but it can be marginally interpreted locally under *if*, resulting in an imperfect but relatively acceptable utterance. In presupposition literature, local interpretations of normally projecting inferences are standardly derived via LOCAL ACCOMMODATION (e.g., Heim 1983; Schlenker 2009), which is a last-resort mechanism that is in principle available for presuppositions. Local accommodation can be operationalized differently, but at the end of the day yields the result where the presupposition is interpreted as a truth-conditionally non-vacuous conjunct at some local level, which is what seems to happen in (23a).

Such local interpretation should then be possible in (23b) as well, however, either it isn't or it's not enough to save the day; the utterance is crashingly bad. Why is that? I assume that every modifier instance is either restricting or non-restricting. Under the definition in (17), *large* in (23b) cannot be non-restricting, since it is not intended to be truth-conditionally vacuous. However, it doesn't seem to be able to be restricting either, as it looks like restricting modifiers have to pick out a (locally) non-empty part of the denotation of the expression they modify. I will not explore this property of restricting modifiers any further in this paper, however; I will just leave it as an empirical observation.

The discussion above highlights an important consequence of the view of NRM inferences developed here: even though they project as presuppositions, local accommodation isn't a meaningful notion for them. Local accommodation results in a truth-conditionally non-vacuous contribution of the affected piece of content, however, if a modifier isn't meant to be truth-conditionally nonvacuous, no NRM inference will be generated in the first place.

Another case when a modifier instance is likely to be non-restricting is when the lexical semantics of the modifier has an evaluative component. For instance, in (24), it would be reasonable for the addresee to construe of the adjective *obnoxious* as non-restricting, i.e., truth-conditionally vacuous, due to its evaluative nature. The addressee would then plausibly infer that the speaker considers all philosophers obnoxious (which would assure said truth-conditional vacuity), even if they didn't previously believe that, and will globally accommodate this inference. Or, alternatively, if they aren't willing to tacitly accommodate this inference, they might push back by saying something like *Hm*, *I didn't know you don't like philosophers*.

- (24) I don't want any obnoxious philosophers at my talk next week.
  - $\rightarrow$  I don't want any philosophers at my talk next week.
  - $\rightarrow$  All philosophers are obnoxious.

The reason why such modifiers have a penchant for being non-restricting seems to be two-fold. First, the evaluative component of their meaning makes adjectives like *obnoxious* always satisfy the relevance constraint on NRMs. Second, they don't seem to make great restricting modifiers, since the purpose of the latter is to delineate a (potentially) non-proper part of the denotation of the expression they modify, and the speaker's subjective attitudes might not be the best way to do so. That said, plain evaluative adjectives can be restricting, if needed, as shown in (25).

- (25) A: Which of her dogs is Lea bringing to the party?
  - B: The {**obnoxious**, **disgusting**, **lovely**} one.

Now, let us formalize NRM inferences. I will adopt Schlenker's 2009; 2010 general pragmatic approach to local contexts, which defines the local context c' of any expression d whose type ends in t as the strongest possible restriction one could make before interpreting d in a Stalnakerian context set C; a full definition is given in (26).

#### (26) LOCAL CONTEXT (adopted w/minor changes from Schlenker 2010, (7))

The LOCAL CONTEXT of an expression d that occurs in a syntactic environment  $a_b$  in a context set C is the expression with the strongest denotation c' which guarantees that for any expression d' of the same type as d, for all strings b' for which a d' b' is a well-formed sentence, the following holds in each world w in C:  $a (c' \text{ and } d') b' \leftrightarrow a d' b'$ 

We will need to keep track of two local contexts. One is the propositional local context c' computed via the mechanism in (26) for the clause that contains the modifier (e.g., by updating the global context C with the left conjunct under *maybe*, as in (18b), the negation of the left disjunct, as in (19b), or the antecedent of a conditional, as in (22)). The second local context is a sub-propositional local context c'', which is computed within the clause for the result of modification and is always relativized to c'.<sup>9</sup> In simple cases (which are the only ones I will consider in this paper) this means that the world argument of c'' is restricted to the worlds of c'.

We can now define NON-RESTRICTING MODIFIER INFERENCES as follows:

#### (27) NON-RESTRICTING MODIFIER INFERENCE

If a clause p contains a subtree  $\alpha$  whose daughters are  $\gamma$  and  $\beta$  and the addressee infers that the speaker intends  $\gamma$  as a non-restricting modifier (as defined in (17)), the addressee will generate the NON-RESTRICTING MODIFIER INFERENCE:

$$c'' \Rightarrow_\forall (\beta \Rightarrow \alpha),$$

where c'' is the local context of  $\alpha$  in the local context c' of p.

<sup>&</sup>lt;sup>9</sup>The need to keep track of two local contexts, a propositional and a sub-propositional one, is not specific to NRM inferences. The issue is not discussed explicitly in Schlenker's (or, to my knowledge, anyone else's) work on local contexts; he only discusses sub-propositional contexts for expressions in the scope of quantifiers and does not bring up cases in which both the propositional and the sub-propositional local contexts matter for projection of lexical presuppositions, such as (i), where to compute the projecting inference we need to first restrict our attention to the worlds in which the track is damaged and then to the individuals who are stuntwomen in those worlds.

 <sup>(</sup>i) If the track is damaged, each stuntwoman knows that she can't continue the race.
 → If the track is damaged, no stuntwoman can continue the race (i.e., each stuntwoman is such that she can't continue the race).

Now let us apply the definition above to (18a) and (18b). For simplicity, let us assume that the subpropositional local context c'' is maximally broad in both cases, i.e., that it denotes  $\lambda x \lambda w.D_e(x) \wedge [\![c']\!](w)$ , where  $D_e$  is the domain of individuals in the model. The propositional local context c' of the clause containing *deadly sausages* in (18a) denotes the global context set C, and we have agreed to treat c'' as maximally broad, so if the addressee recognizes that the speaker intends *deadly* as truth-conditionally vacuous, they will generate the inference in (28).

$$\begin{array}{ll} (28) & \forall x \forall w [\llbracket c'' \rrbracket (x)(w) \to (\llbracket sausage \rrbracket (x)(w) \to \llbracket deadly \ sausage \rrbracket (x)(w))] \\ & \forall x \forall w [(\mathbf{D}_e(x) \land \llbracket c' \rrbracket (w)) \to (\mathsf{sausage}(x)(w) \to (\mathsf{sausage}(x)(w) \land \mathsf{deadly}(x)(w)))] \\ & \forall x \forall w [(\mathbf{D}_e(x) \land \mathbf{C}(w)) \to (\mathsf{sausage}(x)(w) \to (\mathsf{sausage}(x)(w) \land \mathsf{deadly}(x)(w)))] \\ & \forall x \forall w [(\mathbf{D}_e(x) \land \mathbf{C}(w)) \to (\mathsf{sausage}(x)(w) \to \mathsf{deadly}(x)(w)))] \\ & \forall x \forall w [(\mathbf{D}_e(x) \land \mathbf{C}(w)) \to (\mathsf{sausage}(x)(w) \to \mathsf{deadly}(x)(w))] \end{array}$$

In (18b), however, the propositional context c' of the clause containing *deadly sausages* does not denote C, but rather the intersection of C with the set of worlds denoted by the left conjunct under *maybe*, i.e., the set of worlds in which processed meat causes cancer, which I will abbreviate as cancer(w). The NRM inference arising in (18b) is given in (29) (I omit the intermediate steps).

(29)  $\forall x \forall w [(\mathbf{D}_e(x) \land \mathbf{C}(w) \land \mathsf{cancer}(w)) \rightarrow (\mathsf{sausage}(x)(w) \rightarrow \mathsf{deadly}(x)(w))]$ 

#### 2.1.3 NRM inferences and cosuppositions

The definition of NRM inferences in (27) is reminiscent of COSUPPOSITIONS, i.e., assertiondependent inferences that project like presuppositions, originally proposed in Schlenker 2018a to account for inferences triggered by co-speech gestures across the board and extended to other types of content in Schlenker 2018b, To appear. In (30), I give the relevant part of Schlenker's generalized definition of cosuppositions from Schlenker To appear (with notational changes).<sup>10</sup>

#### (30) Schlenker's unified theory of cosupposition generation (Schlenker To appear, (67))

- a. A cosupposition is triggered when an elementary expression pp' has an entailment p' which is presented as being unimportant, and for this reason the global Context Set C should guarantee that, relative to its local context c', pp' should be equivalent to p, i.e.:
  - (i)  $c' \Rightarrow_{\forall} (pp' \Leftrightarrow p)$
- b. (i) is equivalent to the standard definition of cosuppositions in (ii):<sup>11</sup>
  - (ii)  $c' \Rightarrow_{\forall} (p \Rightarrow p')$
- c. An entailment p' might be presented as unimportant for different reasons:
  - (i) for reasons of manner, in case p' is contributed by a co-speech or co-sign gesture (which is parasitic and thus should not make an essential contribution);
  - (ii) for conceptual reasons, in case p' is understood not to matter given the context of the conversation.

Schlenker's algorithm in (30) is, on the one hand, unconstrained in that it isn't linked to any specific

<sup>&</sup>lt;sup>10</sup>While (30) only makes reference to the global context set C and a single local context c', Philippe Schlenker (p.c.) has confirmed that for a non-propositional pp' we need to compute two local contexts, a propositional one and a sub-propositional one, just like in (27).

<sup>&</sup>lt;sup>11</sup>Here Schlenker refers to the original formulation of cosuppositions for co-speech gestures in Schlenker 2018a.

compositional configuration and can even apply if p and p' don't combine compositionally and are instead two pieces of a single lexical entry pp'. This will lead to overgeneration in cases when the purported cosupposition trigger cannot be a modifier (as defined in this paper), as we will see in section 3. On the other hand, Schlenker assumes that p and p' are conjuncts within pp' (whether or not they combine compositionally) and that p' is thus an entailment of pp', which is too restrictive, as this prevents cosuppositions from arising in cases where p and p' are of different semantic types (pp' can't entail p' if they are of different types, nor can p and p' conjoin in this case). The consequence of that is that while Schlenker's cosupposition algorithm generates the same inferences for intersective NRMs as the algorithm in (27) (modulo the need for two local contexts rather than one), it can't apply to non-intersective subsective modifiers like *skillful* or *completely*, which are not of the same type as the expression they modify or the result of modification.

In fact, one of the cases mentioned in Schlenker To appear is a maximal-degree modifier 'completely' in Italian Sign Language (LIS) and spoken Italian discussed in Aristodemo 2017. In LIS, it is argued to be an obligatory part of signs such as *FULL* and *BALD*; in spoken Italian, it's a separate co-speech gesture combining with closed-scale adjectives such as 'full', 'straight', etc. Aristodemo argues that both give rise to projecting conditional inferences (e.g., 'full' entails 'completely full') and analyzes both as cosuppositions, positing that in LIS, 'completely' is a "co-sign gesture". Schlenker's algorithm in (30) cannot assure this result, though, since 'completely full' would need to entail 'completely' for it to apply. However, 'completely' isn't of the same type as 'completely full', so this is not possible. Under my view, 'completely' in both cases above would be a preferably non-restricting modifier, which would give rise to the inference in (31), assuring that (relative to the local context) if an individual is full to some degree, that degree is maximal.

(31) non-restricting modifier inference for *completely full*:  $\forall d \forall x \forall w[\llbracket c'' \rrbracket (d)(x)(w) \rightarrow ((\mathsf{full}(x)(w) = d) \rightarrow (d = \mathsf{max}(S_{\mathsf{full}})(w)))]$ 

The inability to handle non-intersective subsective modifiers is also a problem for Leffel's (2014) analysis of NRM inferences as semantic presuppositions hard-coded into the composition rule for NRMs (more generally, it is not clear how Leffel's analysis of NRMs would generalize beyond adnominal intersective modification).

Now, let me emphasize that simply switching to the 'daughter entails parent' formulation of cosuppositions from 'parent entails daughter' (or 'sister entails sister') as a technical tweak to allow generating cosuppositions in cases of non-intersective modification doesn't address the need for constraining Schlenker's cosuppositions, which I take to be the main issue with this notion. Recognizing that cosuppositions are but NRM inferences in disguise, however, constrains them in the right way (as we will see in section 3), allows generating them in cases of non-intersective modification, and correctly captures the uniform cross-modal properties of NRMs without inventing a new label for an existing phenomenon. In other words, fixing the technical bug doesn't address the bigger issue, but addressing the bigger issue fixes the technical bug along the way.

#### 2.2 Supplements

#### 2.2.1 Composition of supplements

SUPPLEMENTS compose with other expressions, sometimes called ANCHORS, and contribute a proposition of a special kind about those anchors. Supplements are most obviously exemplified by

appositives (nominal appositives and appositive relative clauses). Most analyses of appositives assume they adjoin to their anchors in narrow syntax (e.g., Potts 2005; AnderBois et al. 2013; Koev 2013, etc.). Thus, adnominal appositives, whose anchors are individuals, adjoin to DPs (Determiner Phrases), as opposed to adnominal modifiers, which adjoin to NPs (Noun Phrases). Further differences among specific analyses of appositives are immaterial for the purposes of this paper, but I will assume the majority view that adnominal appositives adjoin to DPs in narrow syntax.<sup>12</sup>

For example, in (32), the appositive (who is) a stuntwoman<sup>13</sup> composes with the DP Zoe denoting the individual Zoe and contributes the proposition of a special type that Zoe is a stuntwoman.

(32) I invited Zoe, (who is) a stuntwoman.  $\rightarrow$  Zoe is a stuntwoman.

Whichever specific account of appositives one assumes, the general outcome is the same: appositives are not restrictive. The way they compose with their anchors doesn't allow them to restrict anything. For example, Potts (2005) assumes bidimensional semantics whereby the anchor fills the argument slot of the appositive yielding a proposition of a special, conventional implicature type  $st^c$  (a simplified version is given in (33)). Koev (2013) and AnderBois et al. (2013) assume a dynamic unidimensional setup instead, whereby appositives introduce propositional discourse referents of a special kind, and the link between appositives and their anchors is anaphoric.

#### (33) **Composition of appositives in Potts 2005 (simplified)**



The consequence of this composition that appositives are non-restrictive is borne out empirically, as appositives cannot have restricting interpretations. For example, in (34a), the two contrasted nominals are prosodically integrated,<sup>14</sup> adjoin at the NP-level,<sup>15</sup> are, thus, modifiers and can, thus, be restricting. However, (34b) is infelicitous precisely because appositives, which are packaged into their own IPs and adjoin at the DP-level, can't ever be restricting.

- (34) a.  $(_{\rm IP} \text{ I invited Zoe the stuntwoman}), (_{\rm IP} \text{ not Zoe the politician}).$ 
  - b.  $#(_{IP} I invited Zoe), (_{PrP} the stuntwoman), (_{IP} not Zoe), (_{IP} the politician).$

Wang et al. (2005), and subsequently Nouwen (2014), give examples like (35), claiming that one-

<sup>&</sup>lt;sup>12</sup>See McCawley 1998; Schlenker 2013 for an alternative view on the syntax of appositives.

<sup>&</sup>lt;sup>13</sup>Here I assume that nominal appositives have some amount of silent syntactic structure that makes them similar to full-blown appositive relative clauses. This assumption isn't crucial, however, and whether there any syntactic differences between the two structures is immaterial for the purposes of this paper.

<sup>&</sup>lt;sup>14</sup><sub>IP</sub> stands for Intonational Phrase in MAE-ToBI (Beckman and Ayers 1997).

<sup>&</sup>lt;sup>15</sup>I follow Matushansky 2008 in that proper nouns have an NP layer denoting the property of being named in a certain way.

appositives are, in fact, restricting (and, thus, restrictive).

(35) If a professor, a famous one, publishes a book, they will make a lot of money.  $\not\rightarrow$  If a professor publishes a book, they are famous.  $\approx$  'If a famous professor publishes a book...'

First, let me point out that (35) doesn't really empirically distinguish between a restricting interpretation of the *one*-appositive and a non-projecting one, because the appositive there is adjoined to a low-scoping indefinite: 'If a famous professor publishes a book...' is equivalent to 'If (a professor publishes a book and they are famous)...'. That aside, I agree with the intuition in AnderBois et al. 2013 that *one*-appositives are more like corrections, or second thought elaborations, rather than bona fide restricting modifiers. AnderBois et al. (2013) adduce evidence that *one*-appositives are "appositive in prosody only"; in particular, they don't require discourse referents and can be easily targeted by direct responses in the discourse. However, these observations, albeit well-taken, prove that *one*-appositives are "at-issue", but don't prove that they are corrections or elaborations rather than bona fide restricting modifiers. This can be proven, however, with contrastive examples, which clearly show that *one*-appositives can't be used as planned restrictors, as in (36c)—as opposed, for example, to adjectives, as in (36a), or restrictive relative clauses, as in (36b).

- (36) a. If a **famous** professor publishes a book, they will make a lot of money, but if an **unknown** professor does so, they will make nothing.
  - b. If a professor that's **famous** publishes a book, they will make a lot of money, but if a professor that's **unknown** does so, they will make nothing.
  - c. #If a professor, a **famous** one, publishes a book, they will make a lot of money, but if a professor, an **unknown** one, does so, they will make nothing.

Appositives can also take verbal and propositional anchors, as shown in (37). In (37a), the appositive's anchor is the property of being a marathon-running event; in (37b), the anchor is the specific event of the roof collapsing; and in (37c), the anchor is the proposition that Zoe ran a marathon.

(37) a. Zoe has never run a marathon, which is a hard thing to do.

 $\approx$  'Zoe has never run a marathon, and running a marathon is a hard thing to do.'

- b. The roof collapsed, which injured several people.  $\approx$  'The roof collapsed, and that event injured several people.'
- c. Kim says that Zoe ran a marathon, which I find hard to believe.  $\approx$  'Kim says that Zoe ran a marathon, and I find the proposition that Zoe ran a marathon hard to believe.'

This flexibility is in contrast to appositives in the nominal domain, which seem to require their anchors to establish individual discourse referents, in a way reminiscent of cross-sentential anaphora:<sup>16</sup>

- (38) a. \*Lea didn't bring a dog, {who was large, a large animal}.
  - b. \*Lea brought no dog, {who was large, a large animal}.
  - c. Lea didn't bring a dog. \*It was large.
  - d. Lea brought no dog. \*It was large.

 $<sup>^{16}</sup>A \ dog \ in (38)$  is meant to be a low scope indefinite.

This constraint seems to be stronger for adnominal appositives than for cross-sentential anaphora, however. Thus, in (39a), the DP *any wombats* doesn't introduce an individual discourse referent. The appositive doesn't seem to be able to target the property of being a wombat (cf. the appositive in (37a)). Furthermore, if appositives contain anaphoric elements (as is assumed in Koev 2013; AnderBois et al. 2013), those anaphoric elements don't seem to be able to target implicitly introduced discourse referents, which is possible for ordinary pronouns, as shown in (39b), where *they* can refer to wombats in general.

a. Zoe doesn't have any wombats, {??who are nocturnal animals, \*nocturnal animals}.
b. Zoe doesn't have any wombats. They are nocturnal animals.<sup>17</sup>

It is not immediately clear what the reason for these contrasts are. It is possible that it has something to do with (i) there being a tight morphosyntactic link between an appositive and its anchor,<sup>18</sup> which distinguishes it from regular anaphora, and (ii) this link being more complex in the case of nominal anchors than in the case of verbal anchors. While I will not explore this intuition any further, these constraints on appositives will re-surface in the next section.

Apart from appositives, so-called "high" adverbs, as in (62), have been treated as supplements as well (Potts 2005), and I will adopt this categorization here.

(40) {Surprisingly, unfortunately, hopefully}, Lea is bringing her dog.

#### 2.2.2 **Projection of supplements**

Supplements project, and typically they project very strongly (i.e., they typically cannot be interpreted locally under semantic operators, not even under pressure):

- (41) a. If you invite Zoe, a stuntwoman, you should show her your muscle car.  $\rightarrow$  Zoe is a stuntwoman.
  - b. #I don't know if Zoe is a stuntwoman, but if you invite Zoe, a stuntwoman, you should show her your muscle car.
     Intended: '\_\_\_\_if (Zoe is a stuntwoman and you invite her) '.

Intended: '...if (Zoe is a stuntwoman and you invite her)...'

How exactly one accounts for supplement projection is immaterial for the purposes of this paper. For example, in both Potts 2005 and AnderBois et al. 2013, projection of supplements is assured by the propositions they contribute being special in a way that makes them impervious to semantic operators. In Potts 2005, these propositions have a special semantic type, and in AnderBois et al. 2013 they are a special type of discourse referents. Either option works for me.

Schlenker (2013) discusses apparent exceptions to the projection requirement on appositives, such as (42).<sup>19</sup>

(42) If tomorrow I call the Chair, who in turn calls the Dean, we will be in deep trouble.

 $\not$  If tomorrow I call the Chair, they will call the Dean.

 $\approx$  'If (tomorrow I call the Chair and they call the Dean)...'

<sup>&</sup>lt;sup>17</sup>Of course, a discourse link needs to be established between the two sentences for this utterance to be coherent. Here it is reasonable to interpret the second sentence as introducing the reason for Zoe not having any wombats.

<sup>&</sup>lt;sup>18</sup>Which is one of the reasons to maintain that appositives do adjoin to their anchors in the narrow syntax.

<sup>&</sup>lt;sup>19</sup>As far as I know, only appositive relative clauses, but not nominal appositives, exhibit such exceptional behavior.

Such interpretations are very constrained, however. Jasinskaja and Poschmann (2018) argue (convincingly) that appositive relative clauses only have such local interpretations when the discourse coherence relation between the matrix clause and the appositive one is a coordinating one. For example, in (42) we are dealing with an ordered sequence of events (Jasinskaja and Poschmann subsume this case under the discourse coherence relation of "Narration").

The crucial difference between modifiers and supplements with respect to projection is, thus, that for a modifier not to project means to be restricting, and this interpretation is in principle available to any modifier by default. Whether a given instance of a modifier is interpreted as restricting (and, therefore, non-projecting) or non-restricting (and, therefore, projecting) is determined by a variety of utterance-external factors. In supplements, however, projection is systematically triggered by a certain compositional configuration (e.g., that of an appositive or a sentence-level adverb), and local interpretations are only available under very limited circumstances.

This difference leads me to conclude that we do need two separate projection mechanisms for NRMs and supplements. A reductionist story like that in Morzycki 2008, whereby NRMs contribute Pottsian (2005) conventional implicatures, fails to capture this contrast in availability of local interpretations. Even if we were to use the same mechanics for assuring lack of interaction with semantic operators for the two types of projecting inferences (i.e., projection proper), we would still need two separate triggering mechanisms. In subsection 2.1.2 above, I proposed that triggering of NRM inferences is pragmatic in nature. I am not proposing any new mechanism for supplements, but it is clear that triggering of projecting inferences in supplements is much more systematic, much more conventional, and, thus, much more semantic, which is how it is implemented in most existing analyses of supplements.<sup>20</sup>

### 2.3 Section summary

The summary of the view on modifiers and supplements presented here is given in Table 1.

## **3** Composition and projection in secondary modalities

#### 3.1 Previous approaches to secondary modality content

As mentioned in the Introduction, most work on projection of secondary modality content has focused on non-conventionalized gestures, such as in (2), repeated below as (43).

- (43) a. Lea might bring  $her \ dog_{LARGE}$ .  $\rightarrow$  Lea's dog is large.
  - b. Zoe might shoot at the target<sub>LONGBOW</sub>.  $\rightarrow$  If Zoe shoots, she'll shoot a longbow.

<sup>&</sup>lt;sup>20</sup>Even Jasinskaja and Poschmann (2018), despite relying a lot on discourse relations, assume different syntactic construals for globally vs. locally interpreted appositives—which makes sense, given that the crucial distinction for them is between subordination and coordination.

Modifiers	Supplements
Compose with $\beta_{\tau}$ , yield $\alpha_{\tau}$	Compose with $\beta$ , pass it unchanged, and yield a proposition of a special type $st^c$ about $\beta$
If subsective $(\alpha \Rightarrow \beta)$ , can be restricting $(\beta \neq \alpha)$ or non-restricting $(\beta \Rightarrow \alpha)$	Can never be restricting
Trigger projecting inferences when non-restricting; triggering is pragmatic	Almost always trigger projecting inferences; triggering is conventional
<ul> <li>Examples (target content bolded):</li> <li>modifiers of NPs <ul> <li>adjectives (<i>her large dog</i>)</li> <li>restrictive relatives (<i>her dog that's large</i>)</li> </ul> </li> <li>modifiers of verbal projections (<i>shot a longbow</i>; <i>shot with a longbow</i>; <i>ran fast</i>)</li> <li>degree modifiers (<i>very drunk</i>)</li> </ul>	<ul> <li>Examples (target content bolded):</li> <li>supplements w/DP anchors (<i>her dog, (who is) a large animal</i>)</li> <li>supplements w/verbal anchors (<i>Zoe has never run a marathon, which is a hard thing to do</i>)</li> <li>supplements w/propositional anchors <ul> <li>appositives (<i>Kim says that Zoe ran a marathon, which I find hard to believe</i>)</li> <li>adverbs (<i>Surprisingly, Kim brought her husband</i>)</li> </ul> </li> </ul>

To my knowledge, the first specific proposal about the nature of the inferences in (43) is due to Ebert and Ebert (2014), who claimed that co-speech gestures are uniformly supplements and project as such. They propose to analyze such gestures within the dynamic system from Koev 2013 and AnderBois et al. 2013, although, once again, for our purposes the specific analysis of supplements doesn't matter. More recently, Ebert (2017) proposed a more refined analysis, whereby co-speech gestures are usually supplements, but there also exist NP-level co-speech gestures with "exemplification" semantics. Under this view, *her dog*<sup>LARGE</sup>, for example, typically has roughly the same semantics and projection behavior as *her dog, who is large*, in which case the gesture adjoins to the DP *her dog*. But the gesture *LARGE* can also sometimes be interpreted as adjoining to the NP *dog* and exemplifying a typical entity in its denotation.

Schlenker (2018a) argues against the supplemental analysis and proposes instead that cospeech gestures trigger "cosuppositions" (defined in (30)) across the board, however, he maintains that gestural supplements do exist, in the form of "post-speech" gestures, as in (44), whereby the gestures occupy their own time slot and follow the spoken expressions they associate with (and, which remains implicit in Schlenker's work, are packaged into their own IPs).

- (44) a. ( $_{IP}$  Lea might bring her dog) ( $_{IP}$  LARGE).
  - b.  $(_{IP} \text{ Zoe might shoot at the target}) (_{IP} \text{ LONGBOW}).$

One of Schlenker's original arguments for this analysis is that while both co-speech and "post-speech" gestures project, co-speech gestures don't seem to be subject to the same anaphoric constraints as adnominal appositives, illustrated previously in (38), but "post-speech" gestures are, as shown in (45). I'd like to note that in my experience the judgements on anchorless co-speech gestures under negation are gradient and variable, and most people do find them somewhat degraded (which is what I report in (45)), but there is still a contrast to be explained.

- (45) a. %?Lea didn't bring a dog<sub>LARGE</sub>.
  - b. %?Lea brought no  $\overline{\text{dog}_{\text{LARGE}}}$ .
  - c. #(<sub>IP</sub>Lea didn't bring a dog) (<sub>IP</sub> LARGE).
  - d.  $\#(_{IP}Lea brought no dog) (_{IP} LARGE).$

Tieu et al. (2017) provide further experimental evidence that co-speech gestures don't always project like appositives either. In particular, they look at examples like (46a) (among other things).



- (46) a. None of these ten guys helped<sub>LIFT</sub> his son.
  - b. None of these ten guys helped his son, which [= helping one's son] involved doing  $\underline{\text{this}}_{\text{LIFT}}$ .
  - c. \*None of these ten guys helped his son, which [= the event of helping x's son] happened like  $\underline{\text{this}}_{\text{LIFT}}$ .

The claim they put forward isn't that such examples should be ungrammatical under the supplement construal of the gesture, since spoken appositives in the verbal domain can have event predicates as anchors, as illustrated in (37a) and in (46b) (cf. (46c), where the anchor is meant to be an event, but since there is no available event discourse referent, the sentence is unacceptable).<sup>21</sup> Instead, they propose that, under the supplement construal of the gesture, sentences like (46a) should give rise to universal inferences (for all of these ten guys, if he was to help his son, that would have involved lifting) rather than existential ones (for some of these ten guys, if he was to help his son, that would have involved lifting), because, they claim, that's what sentences like (46b) do. They find, however, some evidence in favor of the existential projection pattern for sentences like (46a), which is what is expected under some analyses of presupposition projection (e.g., Beaver 2001).

Now, Tieu et al. (2017) don't provide comparable data either for spoken supplements or lexical presuppositions. Furthermore, the choice of the appositive for comparison in (46b) is not motivated in any obvious way. The fact that the appositive in (46b) bears the same tense as the matrix clause, to my mind, favors a reading in which the appositive describes homogeneous rules for all the ten guys within a specific situation. However, appositives with event predicate anchors can bear any tense or quantificational force, and it's unclear to me why gestures like *LIFT*, which contain no overt tense morphemes nor quantificational adverbs, should be compared to appositives like in (46b) rather than, say, generic appositives like in (47a), which, as far as I can tell, don't necessarily give rise to universal inferences regarding the ten guys in question (as some of them can be an exception to the generic rule), or even appositives with existential force like in (47b) or (47c).

- (47) a. None of these ten guys helped his son, which [= helping one's son] ({usually, typically}) involves doing this<sub>LIFT</sub>.
  - b. None of these ten guys helped his son, which [= helping one's son] sometimes involves doing this<sub>LIFT</sub>.

<sup>&</sup>lt;sup>21</sup>There is further discussion in Tieu et al. 2017 involving indicative vs. subjunctive appositives; however, since it's unclear to me what it would mean for a gesture to bear subjunctive mood, I will not rehash it here.

c. None of these ten guys helped his son, which [= helping one's son] could(, for example,) involve doing this<sub>LIFT</sub>.<sup>22</sup>

That aside, however, even if we assume a robust contrast in this respect (supplement construals always give rise to universal inferences; presupposition construals can give rise to existential inferences), the results in Tieu et al. 2017 cannot be used as an argument that gestures in examples like (46a) are never construed as supplements, only that they aren't always construed as such, since just because participants allow the construal that gives rise to an existential (weaker) inference, we cannot conclude that they don't also allow the construal that gives rise to a universal (stronger) inference. The same can be said about Schlenker's observations about anahoric constraints illustrated in (45). In other words, what Schlenker and Tieu et al.'s data show is that even if co-speech gestures are sometimes interpreted as supplements, the supplement strategy doesn't yield the entire range of possible interpretations for co-speech gestures.

Schlenker (2018b) extends this line of inquiry to other types of secondary modality content, aiming to develop a more general typology of projection patterns for "iconic enrichments". This typology aims to predict if/how a given piece of secondary modality content X projects, based on: (i) whether X has its own time slot or co-occurs with something in a more primary modality, and (ii) whether X is an "internal" or "external" enrichment, i.e., whether it is "syntactically eliminable". Thus, co-speech gestures like in (43) co-occur with speech, a more primary modality, and are "external", as they can be omitted. Schlenker claims that such content triggers "cosuppositions" across the board. "Post-speech" gestures are also "external", but they have their own time slot; such content is predicted to trigger supplements.

While Schlenker's (2018b) paper contains some extremely valuable empirical observations, the typology proposed there is not based on linguistically meaningful notions or principles and, thus, has little explanatory or predictive power. First, when discussing the various "enrichments", the paper rarely, if at all, makes explicit assumptions about their lexical semantics or how they compose, if at all, with the expressions they are "enriching". Second, the "internal" vs. "external" distinction is not defined in a theoretically meaningful way. The decisions about what's "internal" and what's "external" in speech and sign are made in a seemingly ad hoc fashion, and no coherent view of the architecture of grammar in which this distinction could be grounded is presented. The notion of "eliminability" is applied to at least two completely different distinctions: syntactic adjuncts vs. non-adjuncts (with the latter comprising verbs and their arguments) and segmental vs. suprasegmental morphemes (with the latter exemplified by vowel lengthening, discussed in subsection 3.5). Finally, while the distinction between content that has its own time slot and content that doesn't is potentially useful, it is not well-defined in Schlenker 2018b either, and prosodic phrasing is ignored, even though it can be indicative of syntactic structure (Price et al. 1991 et seq.) and is especially crucial in distinguishing between appositive and non-appositive content, which is central to Schlenker's notion of "post-speech" gestures.

All this makes Schlenker's typology hard to apply to new cases. However, when we do manage to do so, we start seeing this typology make the wrong predictions, as I show in the rest of this section. I also show that treating secondary modality expressions as linguistic objects at all levels of representation reveals that there is no need for a typology like Schlenker's and secondary modality expressions fit into the same typology of projection behavior patterns as primary modality ones.

<sup>&</sup>lt;sup>22</sup>A comparison with this appositive would be in line with the claim put forward in Ebert 2017 that co-speech gestures can have exemplification uses.

#### **3.2** Architectural assumptions

I will start by briefly laying out my assumptions about the architecture of grammar and the consequent view of how secondary modality content fits into the typology of projection behavior patterns. I assume the inverted Y model of grammar, whereby the syntax creates hierarchical, nonlinearized structures of abstract objects (labels, features, etc.), which get shipped off to the compositional semantics to be interpreted, on the one hand, and to the phonology, broadly construed, to create pronounceable sequences of (potentially overlapping) elements, on the other. Vocabulary insertion and linearization happen in the phonology branch after the split.

This view entails that at the level of syntax and compositional semantics no distinctions can be made between primary and secondary modality exponents, nor can there be distinctions made at this level based on the linearization properties of a given exponent. Any modality-specific specific effects emerge in the phonology, i.e., during linearization, prosodic phrasing, articulation, etc., or in the pragmatics, i.e., during post-compositional reasoning about the speaker's beliefs and intentions, which can be sensitive to the surface form of the utterance.

The modifier vs. supplement distinction, as set up in section 2, is first and foremost a compositional distinction, and it affects which projection mechanism is made available (in the case of modifiers) or enforced (in the case of supplements) for a given expression. Therefore, under the architectural assumptions above, it is impossible to maintain a view whereby secondary modality expressions uniformly project as supplements or as NRMs<sup>23</sup> based on their surface properties only, which insubstantiates the earlier debate about gesture projection in Ebert and Ebert 2014; Ebert 2017 vs. Schlenker 2018a and Schlenker's (2018b) typology.

#### 3.3 Non-conventionalized gestures

#### 3.3.1 Non-conventionlized co-speech gestures

Under the cross-modal, composition-driven approach to projection, gestures can in principle compose as modifiers or supplements, as long as their lexical meaning<sup>24</sup> is compatible with either construal. For instance, in <u>her dog<sup>LARGE</sup></u>, the gesture *LARGE* can be taken to represent the property of being of a certain size, in which case it will compose as a modifier adjoining to *dog*, similarly to the adjective *large* in *her large dog*. However, it can also be taken to represent a large object, in which case it will compose as a supplement adjoining to *her dog*, similarly to the nominal appositive in *her dog, a large animal*.

At this point let me add two quick notes on NP- vs. DP-adjoning gestures. I don't think that the temporal alignment of a co-nominal gesture is likely to be a good predictor of its level of adjunction in most cases. In particular, the preparatory phase of a co-speech gesture often starts way before the spoken expression it "associates with", and the stroke of the gesture will most likely align with the stressed syllable of the head noun in either case, especially if said noun is pitch accented (see Loehr 2004 on prosodic integration of co-speech gestures). Of course, one can try to construct examples with gestures co-occurring with different parts of DPs with very long determiners, but

<sup>&</sup>lt;sup>23</sup>Once again, Schlenker's "cosuppositions" are just NRM inferences in disguise.

<sup>&</sup>lt;sup>24</sup>To be clear, when I refer to a gesture's (or any expression's) lexical meaning, I make zero claims about whether said gesture is stored in the lexicon; it is whatever interpretation one posits for a given instance of a gesture, no matter how conventionalized said interpretation is.

I believe such investigations would need to be done experimentally, and I don't have any prior expectations about how sensitive non-signers are to precise gesture alignment. More importantly, the question of whether the exact temporal alignment of a gesture can predict its adjunction level is orthogonal to the issues that I am raising for analyses that implicitly or explicitly assume that co-speech gestures compose in the same way across the board. It is furthermore possible that some gestures (especially, conventionalized ones, as discussed in the next section) are more likely to be construed as modifiers or supplements based on their lexical meaning and/or morphosyntax. For size gestures, however, I don't see a reason to prefer one construal over the other (although, of course, there might be speakers who do have a relevant preference).

Similarly, both composition strategies should be in principle available for co-verbal gestures, as in *Zoe shot at the target*<sup>LONGBOW</sup>, however, this case is more ambiguous. The *LONGBOW* gesture itself can be plausibly interpreted as the property of longbow-shooting events, which can then compose as a modifier via predicate modification, similarly to the PP with a longbow in Zoe shot at the target with a longbow or the direct object a longbow (which, once again, is also a modifier under Neo-Davidsonian assumptions) in *Zoe shot a longbow at the target*. This gesture could also compose as a supplement with an event anchor ( $\approx$  Zoe shot at the target, which [= the event of Zoe shot at the target] was a longbow-shooting event) or with an event property anchor ( $\approx$  Zoe shot at the target, which [= shooting at the target] is usually a longbow-shooting event).

I would like to make it very clear at this point that the spoken language analogies above should be taken just as that, analogies, rather than 100% counterparts of the gestural examples. For example, the fact that *with a longbow* in one of my paraphrases above is a PP should not be taken to mean that I believe that *LONGBOW* in *shot at the target*<sup>LONGBOW</sup> is a PP. Similarly, whether or not there is a way to faithfully paraphrase an example with a gesture by replacing the gesture with a spoken expression does not directly determine whether a certain compositional construal is available for a given instance of a gesture. Assuming otherwise would be antithetical to my views on the architecture of grammar and how secondary modality content fits in. That said, this doesn't mean that speakers don't ever draw analogies between primary and secondary modality expressions when constructing representations at various levels.

Like other modifiers, modifier gestures can, thus, in principle be restricting or non-restricting. Esipova (2019b) reports experimental results showing that this is indeed the case, and, unlike appositives, co-nominal gestures can have restricting interpretations. However, unlike adjectives, restricting interpretations of co-nominal gestures are not as good as their non-restricting interpretations. The two relevant paradigms are given in (48).

- (48) *Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that Stephanie, ...* 
  - a. Non-restricting: ...who has two pets, a small cat and a large dog, is planning to bring along one of her pets. Anna, who has seen both Stephanie's pets before, says:
    Do you know which one of Stephanie's pets is coming with us? 'Cause if she's bringing...
    - (i) her small **cat**
    - (ii) her cat, a small animal
    - (iii) her cat<sub>SMALL</sub>

..., we'll be fine, but if she's bringing...

- (i) her large **dog**
- (ii) her **dog**, a **large** animal



- (iii) her **dog**<sub>LARGE</sub> ..., we should get a bigger van.
- B. Restricting: ...who has two dogs, a small Pug and a large Great Dane, is planning to bring along one of her dogs. Anna, who has seen both Stephanie's dogs before, says: Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing...
  - (i) her small dog
  - (ii) #her dog, a small animal
  - (iii) ?her **dog**<sub>SMALL</sub>
  - ..., we'll be fine, but if she's bringing...
  - (i) her large dog
  - (ii) #her dog, a **large** animal
  - (iii) ?her dog<sub>LARGE</sub>
    - ..., we should get a bigger van.

Neither the analysis of co-speech gestures as supplements across the board (Ebert and Ebert 2014), nor its version with "exemplification" uses added (Ebert 2017), can account for the availability of restricting interpretations of gestures.

The analysis of co-speech gestures as "cosupposition" triggers across the board (Schlenker 2018a,b) can derive restricting modifier interpretations thereof via local accommodation of cosuppositions contributed by gestures associated with spoken NPs (as suggested by Schlenker himself in the work cited) or via non-generation of cosuppositions in the first place. Either way, the gesture is interpreted as a conjoining to the spoken expression it associates with, which, in the case of NP-associating gestures, is equivalent to treating them as intersective modifiers (e.g.,  $\lambda x \lambda w. \text{dog}(x)(w) \wedge \text{LARGE}(x)(w)$ ).

However, under this view, DP-associating gestures should also be able to be interpreted as conjoining with the DPs they associate with. Depending on what semantics one posits for DP-associating gestures, this would result in different interpretations, none of which are attested. For instance, if one maintains that a DP-associating gesture has indefinite-like semantics, akin to *a large object*, the sentence *Lea brought her dog*<sub>LARGE</sub> should be able to be interpreted as 'Lea brought her dog and a large object', which is clearly not possible. Introducing an anaphoric link between the spoken DP and the gesture would allow for a more reasonable supplement-like interpretation of the gesture, however, it should still be able to be interpreted as a conjunct locally, i.e., *Lea brought her dog*<sub>LARGE</sub> should be able to have the interpretation 'Lea brought her dog and that dog is large'. Such interpretations are shown to be completely unavailable for co-nominal gestures (as well as for spoken adjectives and appositives) in Esipova 2019b, however:

(49) Context: We are going on a group tour. Anna and Maria are responsible for renting a van. Maria just told Anna that Stephanie is planning to bring along her dog. Anna knows that Stephanie only has one dog, but has never seen it. She says: Do you know how big Stephanie's dog is? 'Cause if she's bringing...

- a. #her small dog
- b. #her dog, a **small** animal
- c. #her  $\mathbf{dog}_{SMALL}$
- ..., we'll be fine, but if she's bringing...
- a. #her **large** dog
- b. #her dog, a **large** animal
- c. #her  $dog_{LARGE}$
- ..., we should get a bigger van.

Thus, no analysis that assumes a uniform projection behavior pattern for all co-speech gestures can correctly capture the actual, more nuanced empirical picture. Under the cross-modal, composition-driven analysis proposed here, however, it is expected that modifier gestures can in principle be non-restricting, in which case they trigger NRM inferences like other NRMs, or restricting, in which case they don't project. Supplement gestures project conventionally, like other supplements.

However, what still needs to be explained is the preference for non-restricting interpretations in co-speech gestures. I agree with the general intuition in the existing literature that it is the co-speech status of co-speech gestures that makes them more likely to be perceived as truth-conditionally vacuous (even though I disagree with the previous attempts to operationalize this intuition). One potential reason for that can be that secondary modality expressions that co-occur with primary modality material are not very salient and, thus, easy to overlook, and a cooperative speaker should probably avoid relying on such secondary modality expressions to convey truth-conditionally non-vacuous content. Another reason, specific to the data in Esipova 2019b, is that the restricting interpretations in those data all come along with contrastive focus on the gesture.<sup>25</sup> However, one can't accent a co-speech gesture vocally to mark this focus, so one has to accent the co-occurring spoken expression, which might be viewed as a suboptimal focus-marking strategy. Whatever the case is, this effect is gradient and variable (as shown also in Esipova 2019a), which makes it susceptible of pragmatic and/or prosody-related explanations.

#### 3.3.2 Non-conventionalized gestures with their own time slot

Let us now turn our attention to non-co-speech gestures. In his typology of "iconic enrichments", Schlenker (2018b) distinguishes between "pro-speech" and "post-speech" gestures. He defines the former as gestures that "replace" spoken expressions, and the latter as gestures that follow the spoken expressions they associate with. He furthermore clams that while both have their own time slot, the former are "internal enrichments", i.e., "ineliminable", and, therefore, at-issue (i.e., truth-conditionally non-vacuous), while the latter are "eliminable" and, therefore, supplements.

I am not sure how to use these non-linguistic definitions, but once we start looking at concrete examples of what Schlenker calls "pro-speech" gestures, we realize that those are gestures whose main stroke doesn't co-occur with any lexicalized spoken material (although it can co-occur with

<sup>&</sup>lt;sup>25</sup>This should not be taken to mean that restricting interpretations of modifiers always require focus, nor that focus in and of itself forces restricting interpretations of modifiers. The latter has been claimed to be the case in Umbach 2006; Leffel 2014, which I think is a simplification, but, in the interest of space, I will not discuss the relationship between focus and restrictingness in any detail here.

non-lexicalized vocalizations) and that are furthermore non-adjuncts syntactically, i.e., verbs and their arguments (this is why they are "ineliminable"), which indeed are truth-conditionally non-vacuous by default, regardless of the modality in which they are exponed. I give some examples of naturally occurring gestures of this kind in (50) below.

- (50) a. Well, you're a wonderful chef, you know, aside from being super-talented with the VIOLIN. ('Better Call Saul', S02E05, violin.mp4)
  - b. You can get a steak here daddy-o. Don't be a SQUARE.

('Pulp Fiction', square.mp4)

- c. Anyhow. I grab his knife and STAB-STAB-STAB.
- ('The Marvelous Mrs. Maisel', S01E03, stab-stab-stab.mp4)
  d. I need a summary judgement hearing with Michaelis for Frey V. O'Connell, and I'm hoping for Thursday at 10, before he uh... DRINK<sup>GULP</sup>.
- ('Better Call Saul', S02E05, drink.mp4) e. So... did you actually HIT-WITH-HAMMER<sup>CLICK</sup> White Girl?
  - ('iZombie', S04E12, hammer.mp4)
- f. Look, if this option is too expensive you can still always STAB<sup>GRUNT</sup> Badger in the chow line. ('Breaking Bad', S02E08, stab.mp4)
- g. And then the rest of the movie GO-DOWN<sup>PRLL</sup> just straight down from there.

('Cracked' YouTube channel, go-down.mp4)

h. Well, with Nancy and Gulliver CUT-THROAT<sup>GURGLE</sup>, who's most likely to win this thing, that's the pair that would benefit most. ('iZombie', S05E03, cut-throat.mp4)

In (50), all the gestures integrate with rest of the utterance syntactically, linearly, and prosodically in the same way spoken expressions would. Thus, in (50a) and (50b), the gestures are NPs within argument DPs, linearly appearing where you'd expect similar spoken NPs to appear and packaged prosodically into the same prosodic phrases as the rest of the VP, just like similar spoken NPs would. In the rest of the examples, the gestures are verbal constituents of some kind, which, too, linearize and integrate prosodically exactly like similar spoken expressions would. While in all the examples in (50) various strategies are employed to facilitate articulatory integration of speech and gesture within the same prosodic phrase, such as prolongation of the preceding syllable, pausing, vocalizations on the gestures,<sup>26</sup> etc., there are no instances of phrasal boundaries (marked by boundary tones and pitch reset) where you wouldn't expect them in a fully spoken utterance. See some further exploratory work on prosodic integration of such gestures in Harris 2020.

The same is true, however, for the gestures that Schlenker calls "post-speech" and that were previously illustrated in (44), repeated below in (51).

- (51) a.  $(_{IP} \text{ Lea might bring her dog}) (_{IP} \text{ LARGE}).$ 
  - b.  $(_{IP} \text{ Zoe might shoot at the target}) (_{IP} \text{ LONGBOW}).$

Configurationally, these gestures resemble English appositives (and nothing much else), which

<sup>&</sup>lt;sup>26</sup>Schlenker (2018b) argues that such vocalizatons are used to maximize iconicity. While I agree that iconicityrelated considerations do often play a role, they can't be the only relevant factor, since some vocalizations accompanying such gestures are not iconic. I thus maintain that one of the functions of these vocalizations is to facilitate prosodic integration of gestures with speech. Harris (2020) further surmises that yet another function of such vocalizations might be to draw attention to the gesture, making sure the addressee doesn't miss it.

also linearize as following their anchors and are packaged into their own IPs (see Selkirk 2005 and references therein for this latter point), and are, thus, construed as such. It is thus entirely unsurprising that such gestures cannot have restricting interpretations:

(52) #Do you know which one of Lea's dogs is coming with us? 'Cause if she's bringing her dog, (IP **SMALL**), we'll be fine, but if she's bringing her dog, (IP **LARGE**), we should get a bigger van.

We, thus, conclude that gestures that compositionally integrate into otherwise spoken utterances and don't linearize as co-speech follow the same rules of linear and prosodic integration as spoken expressions with the same syntactic status. The "pro-speech" vs. "post-speech" distinction is, thus, not a linguistically meaningful one; "post-speech" gestures are just "pro-speech" gestures whose linearization and prosodic grouping properties correspond to those of English appositives. I would like to preserve the co-speech vs. pro-speech terminological distinction, however, but define co-speech gestures as gestures whose main stroke co-occurs with lexicalized spoken material and pro-speech gestures as gestures whose main stroke doesn't co-occur with any lexicalized spoken material. This distinction is meaningful for discussing prosodic integration of gestures, and there furthermore seem to be constraints on which types of syntactic objects can linearize as co-speech, which are worth exploring in future research. While I don't make this part of the respective definitions, in this paper I only talk about content-bearing co-speech and pro-speech gestures (as opposed to beat gestures), and I furthermore mostly focus on gestures that compositionally integrate with speech (or those that can be plausibly construed as such). I come back to the issue of gestures that integrate with speech prosodically, but not compositionally in the last subsection of this section.

As argued in the previous subsection, the co-speech status of a gestural modifier makes it preferably non-restricting at least in some cases. One would then expect that pro-speech gestural modifiers don't exhibit such a preference by default. However, in English, pro-speech modifiers don't seem to be common; the reason seems to be that the articulatory cost of prosodic integration of such modifiers with speech is too high, and since gestural modifiers can be easily linearized as co-speech, normally, speakers opt for co-speech linearization. Even in English, however, we sometimes see exceptions to this tendency, in particular, when the vocalization part of the prospeech gesture, or, more generally, pro-speech demonstration,<sup>27</sup> is crucial for interpreting it, and, thus, co-speech linearization is not an option, as in the two naturally-occurring examples below:

(53) a. ...the SCRATCH<sup>KRRR</sup> car... (p.c.)<sup>28</sup>
 b. Was that a... like HEAD-TILT<sup>HMM</sup> experience? (The Cracked Podcast, hmm.mp4)

It is easier to imagine modifier pro-speech gestures in languages like Russian, whose more liberal syntactic displacement rules allow for configurations like (54b), where NPs can move out of DPs (the default word order is given in (54a)), leaving modifiers, spoken or gestural, behind. The modifiers in (54b) are, furthermore, the focus in a contrastive topic–focus configuration and are, thus, packaged into their own prosodic phrases,<sup>29</sup> which facilitates prosodic integration of the gestures.

<sup>&</sup>lt;sup>27</sup>Here I use the term *demonstration* along the lines of Davidson 2015, without necessarily subscribing to Davidson's theoretical analysis of demonstrations.

<sup>&</sup>lt;sup>28</sup>The speaker, a native speaker of English, was telling a story of how they scratched someone else's car, and at some point they referred to the car they scratched in this way, in order to distinguish it from the car they were driving.

<sup>&</sup>lt;sup>29</sup>The exact inventory of prosodic phrases in Russian is unclear, but the PrPs in (54) seem to roughly correspond to



Note that technically the gestures in (54b-ii) satisfy Schlenker's definition of "post-speech" gestures, as they are adjuncts that linearly follow the expressions they adjoin to, but they are restricting modifiers, not supplements. This once again highlights the importance of a properly linguistic approach to gestures and the inadequacy of pre-linguistic notions such as "post-speech" gestures.

#### 3.4 Conventionalized gestures

The previous section showed that even if we only focus on non-conventionalized gestures, we can already see that modality-specific typologies of projection behavior patterns that rely on surface properties of secondary modality expressions fail to capture their—much more linguistic—behavior. However, non-conventionalized gestures are, on the one hand, too unconstrained, because their semantic type is determined on the spot based on how their iconic content is interpreted. On the other hand, they are too constrained, because the types of meaning that can be encoded with purely iconic means are limited. Studying non-conventionalized gestures only is, thus, potentially misleading, since if most non-conventionalized gestures can be construed of as subsective modifiers, and independent post-syntactic considerations make it hard for modifier gestures to have their own time slot in English, it is easy to convince oneself that gestures project in a uniform way depending on their linearization. In this subsection, I show that once we expand our empirical scope to include conventionalized gestures, we start seeing even more diverse projection behavior, which cannot be captured by a typology like Schlenker's (2018b), but fits naturally into the same typology of projection behavior patterns as spoken expressions.

English intermediate phrases.

<sup>&</sup>lt;sup>30</sup>The English translation in (54b) might mislead one to believe that in such examples the contrastive topics are base-generated, and the focused constituents are nominals with silent heads; however, the case assignment facts show clearly that this is not the case, and we do in fact have fronted NPs with AdjPs left behind.

In particular, some conventionalized gestures can be subsective modifiers as well, and, thus, exhibit the projection behavior pattern proper to subsective modifiers. For instance, in both examples in (55), the contribution of the gesture does seem to project by default, even though a restricting interpretation of the gesture would be perfectly reasonable and can be obtained under pressure, as shown in (56). *DRINK* is a Russian conventionalized gesture that means 'drink (alcohol)' and involves flicking one's finger on one's neck (or tapping one's neck with the back of one's hand).



- (55) a. If you bring A = 1 a semanticist<sub>CRAZY</sub> to my talk, I'll likely fight with them.  $\rightarrow$  If you bring a semanticist to my talk, I'll likely fight with them.  $\rightarrow$  All semanticists are crazy.
  - b. If we wanna 2222 <u>celebrate<sub>DRINK</sub></u> my defense, we better go to a store now.  $\rightarrow$  If we wanna celebrate my defense, we better go to a store now.  $\rightarrow$  If we celebrate my defense, we'll do so by drinking alcohol.
  - $\rightarrow$  If Wim brings has brother I'll fight with him but if she brings h
- (56) a. ?If Kim brings her **brother**<sub>CRAZY</sub>, I'll fight with him, but if she brings her **normal** brother, that's  $OK.^{31}$ 
  - $\approx$  If Kim brings her **crazy** brother...
  - b. ?If we wanna <u>celebrate\_DRINK</u> my defense, we better go to a store now, but, of course, we can also celebrate **without** alcohol.
    - $\approx$  If we wanna celebrate my defense by drinking **alcohol**...

Schlenker's typology could treat these gestures in the same way as non-conventionalized gestures like *LARGE* and *LONGBOW* (and would then have to face the same objections as those raised in the previous subsection). However, other conventionalized gestures seem to be non-subsective modifiers, e.g., *AIR-QUOTES* in (57a) is akin to *so-called* or *quote-unquote*.



- (57) a. Kim is bringing her  $\frac{1}{1000}$  friend<sub>AIR-QUOTES</sub> to the party.
  - b. Kim is bringing her {so-called, quote-unquote} friend to the party.
  - (a,b):  $\not\rightarrow$  Kim is bringing her friend to the party.

However we go about describing and analyzing projection of non-subsective modifiers, Schlenker's "cosuppositions" (which are, once again, equivalent to NRM inferences in the case of modifier gestures) aren't the way to go, because the modifiers in (57) aren't truth-conditionally vacuous to begin with (this follows from the lack of subsective entailment in non-subsective modifiers).

Finally, some conventionalized gestures can only be supplements, e.g., *FINGERS-CROSSED*, which is akin to sentence-level optative adverbs like *hopefully* or a parenthetical *fingers crossed*:<sup>32</sup>

<sup>&</sup>lt;sup>31</sup>I apologize for the potentially sanist nature of this example.

<sup>&</sup>lt;sup>32</sup>Such optatives don't only adjoin to clauses, actually, at least not on the surface. Here's a naturally occurring example of a DP-adjoining *FINGERS-CROSSED* (cf. DP-adjoining epistemic adverbs, conveniently also illustrated in (i), discussed in Ernst 1984 et seq.); a spoken parenthetical *fingers crossed* would work here, too:

<sup>(</sup>i) That was Clive's friend in Vice, who I've enlisted in my search for that Beanpole Bob guy and potentially

(58)

If A strend of mine wins the race<sub>FINGERS-CROSSED</sub>, I'll buy them a drink. a. If a friend of mine, {hopefully, fingers crossed}, wins the race, I'll buy them a drink. b. (a,b):  $\checkmark \rightarrow$  I want a friend of mine to win. (a,b):  $X \approx$  If (a friend of mine wins and I wanted them to)...

Again, Schlenker's typology makes the wrong prediction for FINGERS-CROSSED in (58a): it's a co-speech gesture, but doesn't trigger a "cosupposition"/NRM inference; instead, it projects conventionally, like a supplement.

The typology of meanings that can be encoded by conventionalized gestures is even richer (see, for example, Esipova 2019c on gestures in Russian polar responses and Ippolito 2019 on Italian gestures in questions), but the examples above suffice to demonstrate that a typology like Schlenker's can't possibly account for all the diverse projection behavior patterns of conventionalized gestures. If we exclude all conventionalized gestures from a typology of projection behavior patterns that includes gestures, we would be missing the parallel between gestures like LARGE and gestures like CRAZY. Also, doing so would force us to take a stance on what's conventionalized and what isn't, treating this as a categorical distinction, which seems implausible. It would be impossible to exclude some, but not other conventionalized gestures without making reference to their syntactic and/or lexical properties, which would require assuming a properly linguistic approach to gestures. But assuming said approach reveals that a typology like Schlenker's isn't needed, as gestures fit into the same typology of projection behavior patterns as spoken content.

#### 3.5 **Degree modifiers cross-modally**

#### Degree modifiers in the primary modality 3.5.1

Open-scale<sup>33</sup> degree modifiers are persistently truth-conditionally non-vacuous (i.e., restricting) by default. This is true not only for run-of-the-mill degree modifiers like very, extremely, truly, etc., but also for degree modifiers with an attitudinal component like surprisingly and even for expressives used as degree modifiers, which are never restricting in their purely expressive uses:

(59) a. If the movie is {very, extremely, truly, surprisingly, fucking, bloody, damn} good, I'll stay till the end of the credits.

 $\not\rightarrow$  If the movie is good, I'll stay till the end of the credits.

- Which of her dogs is Lea bringing? b. A:
  - The {lovely, disgusting, #fucking, #bloody, #damn} one. **B**:

Two further relevant cases of degree modification in the primary modality are (i) modifier repetition, as in (60), which has a gradient iconic effect, and (ii) so-called "contrastive reduplication", whereby a string of the form x x gets interpreted along the lines of '{true, proper, prototypical} x', as in (61), which also has iconic roots. Both are truth-conditionally non-vacuous as well.

('iZombie', S05E08, fingers-crossed.mp4)

a zombie cure<sub>FINGERS-CROSSED</sub>.

<sup>&</sup>lt;sup>33</sup>I'm making this caveat here, because, as mentioned in section 2.1.3, the maximal-degree modifier 'completely' has been argued to project by default in spoken Italian and Italian Sign Language in Aristodemo 2017.

(60) a. You're really, really sick.

(Ghomeshi et al. 2004, fn. 3) (Ghomeshi et al. 2004, fn. 3)

- b. You are a sick, sick man.
- c. If the movie is very, very, very good, I'll stay till the end of the credits.  $\not\rightarrow$  If the movie is good, I'll stay till the end of the credits.
- (61) a. **Like'em**-like'em? Or, I'd-like-to-get-store-credit-for-that-**amount** like'em?
  - (corpus example from Ghomeshi et al. 2004, (1b))
  - b. I'm up, I'm just not **up**-up. (corpus example from Ghomeshi et al. 2004, (1d))
  - c. If Lea's bringing a dog-dog, not a chihuahua, we'll need a bigger van.

     → If Lea's bringing a dog, we'll need a bigger van.

#### 3.5.2 Degree modification via facial expressions

Like spoken mirative adverbs, the mirative facial expression whose most salient feature is eyes wide open, *OO*, can act as a proposition-contributing supplement or as a degree modifier:<sup>34</sup>

- (62) a. Yesterday there was a party, and, [ $\langle$  surprisingly, impressively $\rangle$ , [Mia got drunk]].  $\approx$  It is  $\langle$  surprising, impressive $\rangle$  that Mia got drunk.
  - b. Yesterday there was a party, and Mia got [ $\langle$ surprisingly, impressively $\rangle$  [drunk]].  $\approx$  Mia got drunk to a(n)  $\langle$ surprising, impressive $\rangle$  extent.
- (63) a. Yesterday there was a party,  $\overline{\text{and }[[\text{Mia got drunk}]^{OO}]}$ .  $\approx$  It is surprising that Mia got drunk.
  - b. Yesterday, there was a party, and Mia got  $[[\overline{drunk}]^{OO}]$ .  $\approx$  Mia got drunk to a {surprising, high} extent.

Now, more needs to be said eventually about the interaction of the temporal alignment of *OO* and its syntactic construals. For instance, one relevant observation is that in its supplement use, *OO* is focus-sensitive, just like *surprisingly* and its kin, which, I believe, does affect the docking of *OO*. For the purposes of this paper, however, what matters is that both interpretations are in principle available for a co-speech *OO*.

The degree modifier *OO* often comes with an intonational morpheme on the modified predicate, discussed in the next subsection. However, *OO* can make this contribution independently, as shown in (64), where *OO* co-occurs with the Russian gesture *DRINK* from the previous section.

(64)	a.	Yesterday, there was a party, and [[Mia got DRUNK] <sup>OO</sup> ].	
		pprox It is surprising that Mia got drunk.	
	b.	Yesterday there was a party, and Mia got $[\overline{DRUNK}]^{OO}]$ .	

 $\approx$  Mia got drunk to a {surprising, high} extent.

Aptly, *OO* projects like a supplement when it is interpreted as one, and is restricting by default when it is used as a degree modifier (just like spoken mirative adverbs like *surprisingly*):

- (65) a. When [[a friend of mine gets  $\overline{DRUNK}]^{OO}$ ], I always comment on that.
  - $\rightarrow$  When a friend of mine gets drunk, I always comment on that.
  - $\rightarrow$  When a friend of mine gets drunk, that is surprising.

<sup>&</sup>lt;sup>34</sup>Overlining approximately indicates temporal alignment for facial expressions. Photo and video illustrations are available for selected examples, but excluded from this version of the paper for anonymization purposes.

- b. When, a friend of mine gets [[DRUNK]<sup>OO</sup>], I always comment on that.
   → When a friend of mine gets drunk, I always comment on that.
   → When a friend of mine gets drunk, they do so to a {surprising, high} extent.
- c. #When a friend of mine gets DRUNK, I don't say anything, but when [[a friend of mine gets DRUNK]<sup>OO</sup>], I always comment on that. Intended: '...when (a friend of mine gets drunk and I am surprised by that)...'.

Again, Schlenker's typology can't predict this variable projection behavior for co-speech/co-gesture facial expressions.<sup>35</sup> But an approach that treats facial expressions as linguistic objects and starts by asking the question about their lexical semantics and compositional integration, once again, reveals that they fit neatly into a cross-modal typology of projection behavior patterns: a sentence-level *OO* patterns with sentence-level spoken adverbs, and a degree modifier *OO* patterns with other degree modifiers, whatever the reason for their persistently restricting nature.

#### 3.5.3 Suprasegmental degree modification

The aforementioned intonational degree modifier, *DEG-INT*, written as a subscript on the word whose stressed syllable it docks to, can make its contribution independently, too. Its precise phonetic and phonological properties remain to be established, as is whether they are distinct from those of the modified focus marking accompanying purely mirative, non-degree-modifier OO, and whether there is any internal compositionality to this morpheme.<sup>36</sup> One thing that seems to emerge relatively constantly is lower pitch on the target syllable, which can be achieved through various strategies (and there seems to be inter- and intra-speaker variation as to which strategy is used in a given instance of DEG-INT): a low target in a pitch accent (e.g., the leading tone in an L+H\* or an L\*+H), overall lowering or compression of the pitch range, or downstep on a high target. Another robust property of DEG-INT is lengthening of the accented syllable, in particular, of the onset and the nucleus. DEG-INT can also come with higher intensity and creaky phonation (the latter might be due to the lower pitch). There is also often increased juncture before the spoken item bearing *DEG-INT* (lengthening of the preceding syllable, pausing), not unlike what we see in prosodic integration of pro-speech gestures. It's possible that the increased juncture is used to draw attention to the secondary modality morpheme (as was speculated to be the case for vocalizations on pro-speech gestures in Harris 2020), but it is also possible that speakers need increased juncture for physiological reasons, to effectively exert the increased articulatory effort associated with DEG-INT. See also some potentially relevant observations about pauses after what the authors describe as mirative uses of *like* in Beltrama and Hanink 2019; in fact, I suspect many of the examples in Beltrama and Hanink 2019 also have DEG-INT on the expression following like.

Like contrastive reduplication, *DEG-INT* is morphosyntactically promiscuous and, like all other open scale degree modifiers, it is restricting by default:

(66)	a.	If the movie is good <sub>DEG-INT</sub> , I'll stay till the end of the credit	as. $\approx$ very good
		$\not\rightarrow$ If the movie is good, I'll stay till the end of the credits.	
	b.	Lea has a dog <sub>DEG-INT</sub> .	$\approx$ big dog or proper dog

<sup>&</sup>lt;sup>35</sup>Schlenker (2018b) claims that co-speech facial expressions trigger cosuppositions. He only looks at the "disgusted" facial expression, which I believe should be fit into a typology of attitudinal expressions along with expressives like *fucking* and plain evaluative adjectives like *disgusting*.

<sup>&</sup>lt;sup>36</sup>See the audio files with examples of different realizations of *DEG-INT* in the accompanying zip file.

c. When I saw that snake, I ran<sub>DEG-INT</sub>.

Now, Schlenker (2018b) discusses what he calls "iconic vowel lengthening" in (67).

(67) If the talk is looong, I'll leave before the end.  $\rightarrow$  If the talk is long, I'll leave before the end.

Schlenker claims that such vowel lengthening is an "internal enrichment" without its own time slot, and those can, but don't have to be at-issue. It's not clear to me in what sense such suprasegmental events are "internal" (nor how it relates to the sense in which non-adjuncts are "internal"), but that aside, this approach clearly misses the generalization about degree modifiers cross-modally.

More needs to be said about the role of iconicity, however. I do agree that examples like (67) can contain extra segment lengthening beyond that associated with *DEG-INT*, however, this extra lengthening in (67) isn't obligatory to convey the high degree reading. Schlenker also notes that having the same extra lengthening in *shooort* is weird. This seems correct, although I don't think the same holds about the regular *DEG-INT* on *short*. Similarly, *DEG-INT* seems to be equally good with *slow* and *fast*, but only the former allows for the extra lengthening.

(68) It was {slow<sub>DEG-INT</sub>, slooow<sub>DEG-INT</sub>, fast<sub>DEG-INT</sub>, #faaast<sub>DEG-INT</sub>}.

There might still be some remnant effects of iconicity in *DEG-INT*, however, due to the low pitch, associated with larger objects and possibly larger degrees (see, e.g., Ohala 1994; Clark et al. 2013 for discussions of iconic effects of pitch),<sup>37</sup> and the syllable lengthening, which might be constraining the distribution of this morpheme, but overall, *DEG-INT* is a more general phenomenon.

#### 3.5.4 Uniform cross-modal degree modifier semantics

In line with my architectural assumptions, I maintain that degree modifiers cross-modally expone a DegP head. The bare-bones semantics of any degree modifier morpheme DEG is given in (69).

(69) 
$$\llbracket \mathsf{DEG} \rrbracket (\llbracket \alpha_{\langle d, \langle \tau_1 \dots \tau_n, st \rangle \rangle} \rrbracket) = \lambda d\lambda X^1_{\tau_1} \dots X^n_{\tau_n} \lambda w. \llbracket \alpha \rrbracket (d) (X^1) \dots (X^n) (w) \land \mathsf{deg}(d) (w)$$

For example, compositionally, *sick sick* is the scalar predicate *sick* combined with the contrastive reduplication morpheme CRED, as shown in (70).

(70) a. 
$$\llbracket sick \rrbracket = \lambda d\lambda x \lambda w. \operatorname{sick}(x)(w) = d$$
  
b.  $\llbracket \operatorname{CRED} \rrbracket (\llbracket \alpha_{\langle d, \langle \tau_1 \dots \tau_n, st \rangle \rangle} \rrbracket) = \lambda d\lambda X_{\tau_1}^1 \dots X_{\tau_n}^n \lambda w. \llbracket \alpha \rrbracket (d)(X^1) \dots (X^n)(w) \wedge \operatorname{high}_{\llbracket \alpha \rrbracket}(d)(w)$   
c.  $\llbracket \operatorname{CRED} \rrbracket (\llbracket sick \rrbracket) = \lambda d\lambda x \lambda w. \operatorname{sick}(x)(w) = d \wedge \operatorname{high}_{\operatorname{sick}}(d)(w)^{38}$ 

Predicates born non-scalar first type-shift into scalar versions of themselves, e.g., *dog* type-shifts from a property of individuals into a function that maps individuals onto the scale of dogness:

(71) a. 
$$[[SCALAR]]([[\alpha_{\langle \tau_1...\tau_n, st \rangle}]]) = \lambda d\lambda X^1_{\tau_1}...X^n_{\tau_n} \lambda w.\mathsf{scale}_{[[\alpha]]}(X^1)...(X^n)(w) = d$$
  
b. 
$$[[dog]] = \lambda x \lambda w.\mathsf{dog}(x)(w)$$

 $<sup>^{37}</sup>$ A relevant question here is how these iconic effects of pitch connect to the claim in Pierrehumbert and Hirschberg 1990 that bitonal pitch accents L\*+H and L+H\* both evoke scales.

 $<sup>^{38}</sup>$ One might want to argue that CRED requires that *d* is within the prototypical range for the input expression rather than simply high on the scale; such lexical particulars are orthogonal to the goal of this paper, however.

c. 
$$[[SCALAR]]([[dog]]) = \lambda d\lambda x \lambda w.scale_{dog}(x)(w) = d$$

As shown before in subsection 2.1.1, the d variable is existentially closed off after all degree modifiers have composed.

Degree modifiers like *surprisingly* and *OO* have an extra attitudinal component that is easy to add to their lexical semantics. One might ask if the attitudinal component of such items can project independently of the degree modifier component proper. I am not sure what exactly the empirical picture is like, but the two components can definitely be treated separately in our semantics. We can have the degree modifier component make a regular, truth-conditional contribution, which is furthermore truth-conditionally non-vacuous (i.e., restricting) by default, and the attitudinal component can either be truth-conditional, but truth-conditionally vacuous (i.e., non-restricting) or non-truth-conditional to begin with. In the latter case, it can operate in the expressive dimension, along the lines of Potts 2007. Such a split is clearly needed for items like *fucking*, which can simultaneously make a truth-conditional, restricting contribution as degree modifiers and perform their expressive function as emotional outlets. Consider this naturally occurring example, where the speaker clearly uses *fucking* as both a restricting degree modifier and a non-truth conditional (and at least in the second occurrence, compositionally vacuous) outlet for his anger:

(72) Context: Daniel Craig, in an interview, when asked if Phoebe Waller-Bridge was a "diversity hire" for 'Bond':

Look, we're having a conversation about Phoebe's gender here, which is fucking ridiculous. She's a great writer. Why shouldn't we get Phoebe onto Bond? (...) I know where you're going, but I don't actually want to have that conversation. I know what you're trying to do, but it's wrong. It's absolutely wrong. She's a fucking great writer. One of the best English writers around.<sup>39</sup>

One thing that remains uncaptured is the iconic effects of modifier repetition in the primary modality, as in (60), and of extra vowel lengthening, as in (67). In the case of modifier repetition, it is possible that each occurrence of the modifier is interpreted independently, as a further restriction on the input (compare the effect of modifier repetition obtained in (60) to that of multiple degree modifiers in (14)). An extra component would need to be introduced, assuring that this step-wise restriction always restricts towards the higher degree of the scale; this can be done in many ways, but I will leave it at that. How we should operationalize gradient iconicity effects such as extra vowel lengthening is an extremely interesting question from the architectural perspective, which I leave for future research. What is clear is that this problem needs a properly linguistic approach.

Another question that I will leave for future research is how to treat simultaneous exponents of degree modifiers in different modalities in a principled way. For instance, (73) is a naturally occurring example in which the image of a larger, fuller chest is simultaneously conveyed by *OO*, *DEG-INT*, and a co-speech gesture. The compositional possibilities here are numerous, and it is not clear if and how we should distinguish among them.

(73) *Context: The speaker is giving advice on how to build pec muscles.* You won't have a chest that looks like this<sub>TOUCHES-ANOTHER'S-CHEST</sub>, you'll actually have

<sup>&</sup>lt;sup>39</sup>https://www.esquire.com/entertainment/movies/a29696991/daniel-craig-phoebe-waller-bridge-james-bond-diversity-hire/

 $a \overline{\text{chest}_{\text{DEG-INT}}}_{\text{LARGE-ROUND}}$ .

The final direction for future research that I will mention relates to the claims in the literature that some languages either don't have degree variables at all or can't bind them, because they lack spoken lexical items that would operate on degree variables, such as degree modifiers or comparatives (e.g., Motu in Beck et al. 2009, Washo in Bochnak 2015). It would be worthwhile investigating whether speakers of such languages can use and perceive secondary modality degree modification. This could shed light on how deep this cross-linguistic asymmetry lies.

### 3.6 Compositionally non-integrated secondary modality content

Now, the view I have been arguing for in this paper does not imply that all secondary modality content that integrates prosodically with some primary modality content also necessarily integrates with that content compositionally. In fact, I maintain that the cross-modal typology of expressions that "parasitize" on host utterances at some level of representation, e.g., in the prosody or even in the morphosyntax, but don't integrate with them compositionally or do so vacuously (i.e., without actually interacting with the content of the expression they compose with) is quite vast and is worth studying on its own. Notably, such "parasites" exist in all modalities we have looked at so far.

An example of spoken expressions that exhibit such behavior are expressives, which can be used to express an emotion that doesn't seem to be compositionally linked to anything in the utterance—a feat that regular evaluative items like *lovely* aren't capable of. Some of such expressives don't even seem to integrate with the host utterance syntactically, as in (74).<sup>40</sup>

(74) Will you please (never) {fucking, bloody, goddamn, \*lovely, \*disgusting} stop?

Facial expressions often act in a way similar to compositionally non-integrated expressives; for example, the previously discussed mirative facial expression *OO* can certainly express immediate surprisal of the speaker at something external to the utterance, as in (75b), contrasted with a docked, focus-sensitive *OO* compositionally conveying non-immediate surprisal in (75a).

- (75) a. (i)  $\overline{\mathbf{Kim}}^{OO}$  brought her husband to the party.  $\approx \{\text{Even, surprisingly,}\}$  **Kim** brought her husband to the party.
  - (ii) Kim brought her **husband**<sup>OO</sup> to the party.  $\approx$  {Surprisingly, Kim brought her **husband** to the party., Kim even brought her **husband** to the party.}
  - b. *Context: Kim's husband comes in, which the speaker neither expected nor wanted.* Great, what do I do now?!<sup>OO</sup>

Similarly, utterances can be produced with a choppy meter and accompanied with punctuated gestures aligned with prominent vocal prosodic events to convey an expressive meaning. Thus, adding choppy meter and backhand claps to (74) would signal an even more heightened emotional state:<sup>41</sup>

(76) Will<sup>T\*</sup><sub>CLAP</sub> you please<sup>T\*</sup><sub>CLAP</sub> fucking<sup>T\*</sup><sub>CLAP</sub> stop<sup>T\*</sup><sub>CLAP</sub>?!

<sup>&</sup>lt;sup>40</sup>See also the previously discussed example in (72), where at least the second occurrence of *fucking* doesn't have its expressive component compositionally linked to anything in the rest of the utterance.

<sup>&</sup>lt;sup>41</sup>The superscript T\* indicates a pitch accent on the stressed syllable of the word without specifying its type.

Snapping fingers to the meter of the utterance can be used to convey a sense of urgency or impatience and, in particular, to tell the addressee to hurry up:

(77)  $We_{SNAP}^{T*} have_{SNAP}^{T*} five_{SNAP}^{T*} minutes_{SNAP}^{T*}!$ 

The choppy metter can be indicated in written communication with periods:

(78) Will. You. Please. Fucking. Stop.

In written web communication, clapping emoji placed between words throughout the string, as in (79), are now broadly used to add extra emphasis.<sup>42</sup>

(79)	mcc @mcclure111	Follow ~
	"Biological sex" 👏 is 👏 anyone 👏 using 👏 that 👏 is 👏 probably ≶ do 👏 oppressive	fake 👏 and 👏 : 👏 terminology ing 👏 something
	9:37 AM - 28 Mar 2016	
	261 Retweets 454 Likes 🧑 🌑 🗐 😎 🧟	0 🚱 🏖 🚯 📚

There even exist instances of other emoji used for a similar purpose, but arguably with an additional meaning component:<sup>43</sup>

(80) THIS ■ IS ■ NOT ■ A ■ FUCKING ■ MONARCHY ■ https://t.co/Vg85LGcaLd

— Sarah Lerner (@SarahLerner) July 8, 2017

Intonational contours can be "parasitic", too. For instance, "uptalk" used to build rapport or signal a persona does not compose semantically with the host utterance, but rather hijacks its boundary contour. This is in contrast to rising declaratives that are actually used to question some component of the utterance they inhabit (see, e.g., Jeong 2018 for a recent overview of rising declaratives).

Given the rich nature of such meanings and the fact that they emerge in various modalities, including the primary one, I believe that we need to resist the urge to throw them all into a "paralinguistic" basket<sup>44</sup> and should study them in their own right. That said, we also need to distinguish such cases from cases when secondary modality content does integrate compositionally with primary modality content. This is especially important when discussing the problem of projection. Naturally, we don't expect compositionally non-integrated content to interact with any semantic operators in the host utterance. That should not obscure the fact that compositionally integrated secondary modality content can project from under such operators in a non-trivial way.

<sup>&</sup>lt;sup>42</sup>These have been claimed to imitate a gesture that is widely used in the speech of Black people (especially women): https://www.vice.com/en\_us/article/jpyajg/the-clap-and-the-clap-back-how-twitter-erased-black-culture-from-an-emoji (the example in (79) comes from there).

<sup>&</sup>lt;sup>43</sup>The example in (80) comes from this blog post, which also discusses where such emoji go in the written string: https://makingnoiseandhearingthings.com/2017/07/13/.

<sup>&</sup>lt;sup>44</sup>As is often the case, for instance, for meanings expressed via intonation (e.g., Ladd 2008).

In most, if not all, of the examples above, the expressions of interest are making non-truthconditional contributions (signalling the speaker's emotions or performing a social function).<sup>45</sup> It is also possible that we can have two truth-conditional utterances produced simultaneously in two different channels and establish some meaning correspondences between the two noncompositionally. Some examples of what could be analyzed as compositionally non-integrated co-speech gestures that are linked to the spoken content they co-occur with via some discourse relation are discussed by Hunter (2018). For example, she claims that examples like (81a) can be analyzed similarly to (81b), i.e., as two independent utterances linked by a discourse relation only.

- (81) a. Do we have any more paper<sub>SCRIBBLE</sub>?
  - b. I want to write/draw something. Do we have any paper?

While I agree that a priori there is nothing that excludes an analysis along these lines for examples like (81a),<sup>46</sup> there is also nothing that enforces it. The *SCRIBBLE* gesture in (81a) can also be construed of as a property, say, the property of being something one could draw or write on, in which case it would be a compositionally integrated modifier and will (by default) give rise to a projecting inference that paper is used for drawing or writing.

In fact, as mentioned before in subsection 2.1.2, NRMs need to be licensed by a discourse relation (Leffel 2014), so, under the modifier analysis, in the right extralinguistic context, the addressee might infer that the speaker intends to write or draw on the paper, which is why they are using the NRM. Thus, the two analyses, in a way, end up being mirror images of one another. On one analysis, the gesture is not compositionally integrated (at least not as a modifier) and directly communicates the speaker's intentions to use the paper for drawing or writing; the inference akin to an NRM one emerges via a salient discourse relation. On the other analysis, the gesture is integrated as a modifier and (by default) gives rise to an NRM inference; it is the inference about the speaker intending to use the paper for drawing or writing that arises via a discourse relation. Most importantly, the two analyses correspond to two different construals of the example in (81a), both of which I believe to be in principle possible (it's a separate question if both are equally likely), rather than two different theories of co-speech gestures.

A less ambiguous example was given in Hunter 2019, cited in (82), where the example with the gesture in (82a) is supposed to mean something like (82b).

#### (82) *Context: The speaker is observing someone trying to fit a sheet of paper into an envelope.*

- a. It won't fit<sub>FOLD</sub>.
- b. It won't fit. (You have to) fold it.

To the extent that (82a) is acceptable, it does indeed seem to consist of two compositionally independent utterances. One piece of evidence in favor of this construal as the preferred one is that if we try to embed (82a) in a way that blocks this construal, the result is hard to make sense of:

(83) a. ??If <u>it won't fit<sub>FOLD</sub></u>, let me know.
b. ??I doubt that it won't fit<sub>FOLD</sub>.

<sup>&</sup>lt;sup>45</sup>With the possible exception of the snapping gesture in (77), which could arguably be interpreted as a regular, truth-conditional imperative telling the speaker to hurry up.

<sup>&</sup>lt;sup>46</sup>Although I'm a bit skeptical about how far the parallels between the two specific examples in (81) go, given that there is nothing obvious about the speaker's desires encoded in the *SCRIBBLE* gesture itself.

Even if we posit a conditional inference in (82a) (along the lines of 'If it won't fit, you have to fold it'), it doesn't arise as an NRM inference, because no restricting construal is possible for it (for instance, (83b), if it means anything, clearly can't mean 'I doubt that (it won't fit and you have to fold it)'). Under the unconstrained cosuppositional theory of co-speech gestures from Schlenker 2018a, it would again be puzzling why local interpretations of this conditional inference are unavailable when they are available in other cases. If one were to maintain any compositional link between the two utterances in (82a), one natural analogy would be a sentence-level appositive along the lines of (84), which makes the discourse relation between the two utterances explicit.<sup>47</sup>

(84) It won't fit, which [= the fact that it won't fit] is why you have to fold it.

We can also produce two utterances in the two modalities simultaneously without having any link between them whatsoever, even directing them at two different addressees. For example, I can be talking to someone while simultaneously beckoning someone else via a gesture. Or I can nod or show a thumbs-up gesture to the chair of a conference session, acknowledging that I've seen them showing the '10 minutes left' sign, without interrupting my talk. I would imagine, however, that there are cognitive limits on how semantically rich such simultaneous utterances can be.

## **4 Projection of pronoun-internal modifiers**

## 4.1 *Phi*-features on spoken pronouns

The composition-driven approach to projection can be efficiently applied to obtain more explanatory accounts of some other instances of projection. In this subsection, I propose to treat *phi*features on pronouns as modifiers that are always non-restricting and, thus, give rise to NRM inferences, as they always modify expressions whose extension is a singleton set. This makes for a more explanatory and morphosyntactically plausible analysis of *phi*-features than the standard treatment thereof as lexical presuppositions.

#### 4.1.1 Projection of *phi*-features

(1a), repeated here as (85), showed that *phi*-features on pronouns contribute projecting inferences.

(85) If Skyler<sub>i</sub> brings her<sub>i</sub> dog, I'll give you \$10.  $\rightarrow$  Skyler is female.

The inferences contributed by *phi*-features on pronouns project very strongly, i.e., they cannot be interpreted locally, not even under severe pressure, as shown in (86) below. In (86a), the inference contributed by the gender feature on *her* is allowed to project, so the sentence is felicitous. In (86b), the gender inference can't project because of the context, so it has to be interpreted locally, but this option is not available for *phi*-features, so the sentence is infelicitous.

(86) *Context: In a magical universe, the speaker is inspecting a victim of a curse that is especially powerful when a woman casts it. ...* 

<sup>&</sup>lt;sup>47</sup>The mere existence of such appositives pushes the boundaries of what we call compositional integration to the point where we might say that any two utterances that are linked by a discourse relation are integrated compositionally.

- a. ...*The speaker thinks that either Lucius or Bellatrix cast the curse*. I don't know if it was Lucius or Bellatrix<sub>i</sub> who cast the curse, but if she<sub>i</sub> cast it, the victim is unlikely to recover.
  → The referent of *she* is female.
  b. ...*The speaker has no suspects*.
- #I don't know the gender of [the person who cast the curse]<sub>i</sub>, but if she<sub>i</sub> cast it, the victim is unlikely to recover.
  → The referent of *she* is female.
  Intended: '...if (x<sub>i</sub> is female and she<sub>i</sub> cast it)...'

The contrast in (86) is intuitively similar to the contrast we have seen for NRMs. In (86a), there are two potential referents, and the question under discussion (QUD) is about which of the referents is the culprit. The two referents differ along a certain dimension, namely, gender, but the gender on the pronoun only helps you identify the referent of *she* (in the absence of overt indices in spoken language); the content of the gender feature does not address the QUD and is thus free to project. This is similar, in particular, to the scenario in (48a), where we also have two potential referents (Stephanie's only cat and her only dog) that differ along some dimension (size), which is in some way relevant in the context, but does not address the QUD and is thus free to project.

In (86b), however, we have a single referent with an unknown property along the gender dimension, and we are trying to use the gender feature on *her* to locally establish this property, but fail. Something similar is happening in (49), where we also have one referent (Stephanie's only dog) with an unknown property (size), which cannot be established locally via a modifier.

I propose that this similarity obtains because *phi*-features on pronouns are in fact modifiers, but they are always interpreted as non-restricting due to how denotations of pronouns are built. But first, let's take a look at the standard analysis of *phi*-features on pronouns as lexical presuppositions.

#### 4.1.2 Lexical presupposition analyses of *phi*-features

Projecting inferences of *phi*-features on pronouns have been traditionally analyzed as lexical presuppositions. Below I review two representative analyses within this approach.

In Heim and Kratzer 1998 (also, e.g., Cooper 1983; Sudo 2012), the pronoun starts out as an index variable of type e whose value is supplied by the assignment function g. *Phi*-features then merge one by one and check their respective presuppositions about the value of this variable. If its presupposition is satisfied, the feature passes this value on; if it isn't, the feature returns a failure. In (87), I provide Heim & Kratzer's structure for *she* (for simplicity, I only include gender).

#### (87) **Pronouns in Heim and Kratzer 1998**



Elbourne (2005) treats pronouns as definite descriptions, which have an et-type NP layer. For

him, not all pronouns bear indices, but in those that do, this NP denotes the property of being the value of the index.<sup>48</sup> This NP combines with a determiner that is similar to the definite article, but comes with extra presuppositions contributed by the *phi*-features. Elbourne further assumes that in English pronouns, the NP doesn't have an overt exponent, and morphemes like *she* spell out the determiner. Elbourne's structure for an indexed *she* is given in (88) (again, I only include gender).

#### (88) Indexed pronouns in Elbourne 2005



There are several issues with the two analyses of *phi*-features on pronouns above.

First, if *phi*-features contribute lexical presuppositions, those have to be "strong", in the sense that they cannot be locally accommodated, even under pressure, as shown in (86b). The "weak–strong" distinction posited for presuppositions is a descriptive, not an explanatory one. Much of the presupposition literature explicitly or implicitly assumes that why some expressions trigger presuppositions, as well as how strong those presuppositions are, is determined lexically. The literature that does try to come up with a principled triggering algorithm (see Abrusán 2011 for an overview and one of the proposals) typically targets weak triggers (usually by design). Strong triggers thus remain a mystery. Either one needs to explain why a given triggering algorithm cannot be undone for such triggers, or we should just give up on non-lexical triggering algorithms for strong triggers and assume that their presuppositions are purely lexical. The latter option is particularly vexing in the case of *phi*-features, since it is puzzling then why languages systematically divide the lexical content of pronouns into at-issue and presupposed. An analysis that offers a principled explanation as to why local accommodation is not an option for *phi*-features is thus more appealing.<sup>49</sup>

Next, the composition strategy used by *phi*-features in (87) is unusual. Do we have other natural language expressions that only check presuppositions (and are not susceptible of the analysis I am proposing for *phi*-features)?<sup>50</sup> If not, then, all other things being equal, an analysis that doesn't posit a new composition strategy for *phi*-features will have an edge over an analysis that does.

Elbourne's analysis doesn't run into this issue, but only because he doesn't put *phi*-features into separate nodes in the syntax in the first place, instead lumping them all in D. This is incompatible with any framework that builds words in the syntax, such as Distributed Morphology (Halle and Marantz 1993 et seq.) or Nanosyntax (see Starke 2010 for a brief overview), given that in many

<sup>&</sup>lt;sup>48</sup>I am simplifying Elbourne's treatment of indices in a way that is not crucial for anything here.

<sup>&</sup>lt;sup>49</sup>I want to point out that the mere existence of other inexplicably "strong" projecting inferences (e.g., those contributed by *too* or *again*) doesn't undermine the need for explaining the behavior of *phi*-features. There are no reasons to believe that these other "strong" inferences are "strong" for the same reason as those contributed for *phi*-features, and having less unexplained cases is always better than having more.

<sup>&</sup>lt;sup>50</sup>Now, as shown in subsection 3.6, we do have expressions that only perform non-truth-conditional functions, like expressing emotions, but in (87), *phi*-features contribute truth-conditional content, except it is purely presuppositional.

languages pronouns are morphologically complex.

Heim & Kratzer's analysis does posit separate nodes for *phi*-features, but is morphosyntactically implausible for a different reason.<sup>51</sup> There have been arguments in the literature that pronouns do need a property-denoting layer going back as early as Postal 1966. Postal adduces examples like *we linguists* to argue that English pronouns can be modified. However, one could argue that in examples like this *linguists* is actually a prosodically integrated appositive, as it doesn't interrupt the pronoun. Some recent striking evidence that pronouns can be modified internally comes from Lee's (2019) work on Khoekhoe pronouns, which can be modified by  $\langle e, (s)t \rangle$ -type properties, including compositionally complex ones, wedging between the morpheme exponing person and those exponing gender and number, as shown in (89).

(89) sáã-‡<sup>h</sup>űpíχầ-nàmã-t-ó
 2-loud-Nama-MIX-PL
 'you loud Namas<sup>52</sup>(men and women, more than 2)' (Khoekhoe)

In recent joint fieldwork, my collaborator and I confirmed that such pronoun-internal modifiers behave like NRMs. First, these modifiers can never be used to restrict a set of potential referents. For example, (90) is OK if uttered while pointing at ten women all of whom are doctors, but bad if only some of these women are doctors and the speaker is trying to use *doctor* to restrict the set.

(90) ||<sup>2</sup>í-||ắé-tí-?áo-tì ké tíĩ-tắ ní sắó
3-heal-?-person-F.PL DECL I FUT follow
'They doctors (female, more than 2) will follow me.'
✓ 'They are all doctors, and they will all follow me.'
✗ 'Those of them who are doctors will follow me.' (Khoekhoe)

Furthermore, the inferences contributed by such pronoun-internal modifiers must project, just like in the case of *phi*-features, as shown in (91), which is OK in a context when the speaker knows that the woman they are talking about is a doctor, but bad in a context when they don't know if she is.

(91) ||<sup>2</sup>í-||űé-tí-?áo-s ká sãó té ?ós(ä) ké ||<sup>2</sup>í-s-äà ní hùí
3-heal-?-person-F.SG follow if/SEMELFACT me then DECL 3-F.SG-OBL FUT help
'If she doctor follows me, she will help.'
✓ 'She is a doctor, and if she doctor follows me, she'll help.'
✗ 'If (she is a doctor and she follows me), she'll help.'

To sum up, a morphosyntactically plausible structure for pronouns is one where (i) pronouns have an  $\langle e, (s)t \rangle$ -type layer, and (ii) *phi*-features have their own nodes within this layer.

#### 4.1.3 *Phi*-features as non-restricting modifiers

I propose that we can avoid all the issues above and have an explanatory and morphosyntactically plausible analysis of *phi*-features on pronouns by treating them as modifiers that are, further-

<sup>&</sup>lt;sup>51</sup>It's also unlikely that *phi*-features are DP-level adjuncts, which is also why I am not proposing to treat them as supplements. But the labels in (87) are not a crucial part of Heim & Kratzer's analysis.

<sup>&</sup>lt;sup>52</sup>Namas are the largest Khoekhoe-speaking ethnic group.

more, obligatorily non-restricting. The reason for the latter is that *phi*-features always modify the Elbourne-style property of being the value of the pronoun's index, whose extension is always a singleton set (containing one individual, atomic or plural).<sup>53</sup> As established before in subsection 2.1.2, if a modifier instance modifies an expression whose extension is a singleton set, it can't be restricting and, thus, has to be non-restricting. After all the *phi*-features have merged, a determiner applies to the resulting set and returns the individual in it, which is the value of the index. Thus, the internal composition of pronouns resembles that of other nominals much more closely: they start out as  $\langle e, (s)t \rangle$ -type properties, can be modified as such,<sup>54</sup> and finally combine with a determiner yielding an individual. This composition is exemplified in (92) for the English pronoun *she* (once again, for now I am only showing how things work for gender).

#### (92) **Pronouns in the proposed system (incomplete)**



In (92), I am assuming intersective composition of the feature and the property it merges with for simplicity. I'm also agnostic about the labels and the morphosemantic mapping within a pronoun, i.e., which morpheme spells out which part of the compositional structure. Finally, the lexical semantics I give to the [fem] feature in (92) is a simplification and needs to be refined to account, for example, for uses of masculine plural pronouns to refer to a mixed-gender group (e.g., French *ils* 'they.MASC') and for various gender-neutral uses of gendered pronouns. Those adjustments, however, are needed under any analysis of gender features, and they are relatively easy to make without affecting the main point of my proposal, so I won't explore this issue further.

As any other NRMs, *phi*-features contribute NRM inferences; as things stand, in (92), the inference would be  $\forall x \forall w[\llbracket c'' \rrbracket (x)(w) \rightarrow (\lambda x.x = g(i) \rightarrow \text{female}(x)(w))]$ , where c'' is the sub-propositional local context of the pronoun in the propositional local context c' of the clause containing it. This, too, is a simplification, to which I will briefly come back at the end of this subsection.

Now, Elbourne explicitly adds presuppositions of existence and uniqueness to the pronoun determiner, because for him not all pronouns bear indices, and the property the determiner combines with can be some contextually salient property, in which case he has to rely on the existence presupposition of the determiner to assure projection of the *phi*-feature inferences. For me, all pronouns bear indices and, thus, all *phi*-features are NRMs and project as such.<sup>55</sup> Making projection of

<sup>&</sup>lt;sup>53</sup>Once again, for Elbourne (2005), not all pronouns bear indices. His motivation for having pronouns that are indexless definite descriptions is to handle donkey anaphora. My motivation for treating pronouns as definite descriptions is completely independent of donkey anaphora-related considerations, and I assume that all pronouns bear indices.

<sup>&</sup>lt;sup>54</sup>Language-specific restrictions on how pronouns can be modified could come from which lexical items a given language has to spell out which chunks of structure. For example, if English *she* can only spell out an uninterrupted structure consisting of the property of being the index value and all the relevant *phi*-features, no additional modifiers will be allowed, since there will be no way to spell out the resulting structure.

<sup>&</sup>lt;sup>55</sup>Naturally, I don't need a separate uniqueness requirement either, since the uniqueness of the output individual is

*phi*-feature inferences an epiphenomenon of projection of a lexical existence presupposition would be empirically problematic, since existence inferences of definites are not as strong as *phi*-feature inferences—or other NRM inferences, for that matter, as shown before in (23). If *phi*-feature inferences project as part of lexical existence presuppositions, it's unclear why local accommodation of said existence presuppositions couldn't save the day in (86b). The existence and/or uniqueness presuppositions can still be part of the  $\iota$  operator; they just aren't doing any work for me.

Now, what about number and person? The nature of the content contributed by person features is such that the empirical parallels with NRMs are not as easy to make as for gender (i.e., it is hard to imagine in context in which the QUD would be about who is the speaker or addressee of the utterance being produced). The contrast in (86), however, can be replicated for number:<sup>56</sup>

- (93) Context: In a magical universe, the speaker is inspecting a victim of a curse that is powerful when cast by a single person, but weak if two or more people jointly cast it. ...
  - a. ... The speaker thinks that either Bellatrix (standing on the left of the speaker) cast the curse alone or the curse was cast jointly by the siblings Alecto and Amycus (standing on the right of the speaker).

I don't know who cast the curse, but if  $\underline{she}_{POINT-LEFT}$  cast it, the victim is unlikely to recover.

 $\rightarrow$  The referent of *she* is singular.

b. ...*The speaker has no suspects, but they know that only women can cast this curse.*#I don't know how many women cast the curse, but if **she** cast it, the victim is unlikely to recover.

 $\not\rightarrow$  The referent of *she* is singular.

Intended interpretation: '...if ( $x_i$  is an atomic individual and she<sub>i</sub> cast it)...<sup>57</sup>

In the proposed system, person and number are meant to compose in the same way as gender, as shown in (94) ( $\leq$  stands for 'is a mereological part of'; I omit worlds for simplicity); the NRM inferences are computed for each modifier ( $X_1 \Rightarrow X_2, X_2 \Rightarrow X_3$ , and  $X_3 \Rightarrow X_4$ ; each of these expressions would need to be entailed by the local context c''). Once again, the details might need to be refined (for example, one might want to split person into [ $\pm$ speaker] and [ $\pm$ addressee] or alter the lexical entries for some features), but I believe (94) to be a viable rough sketch. The flexible nature of the proposed compositional structure is compatible with a wide variety of specific morphosyntactic analyses of pronouns, as long as the semantic types work out.

assured by it being identical to the value of the index. I, furthermore, assume that the (local) existence of the individual identical to the value of the index is independently assured by the mechanics of the assignment function.

<sup>&</sup>lt;sup>56</sup>The point is a bit harder to make for the plural feature, since if the plural is inclusive (i.e., it is compatible with the referent being singular), the inference about the referent being plural arises, when it does, via competition with the singular. While scalar implicatures can be at-issue, especially under contrastive focus, as shown below, it is unclear how systematic this is.

<sup>(</sup>ii) If you read **some** of the assigned books, this won't be enough to pass.  $\approx$  'If you read some but not all of the assigned books...'

<sup>&</sup>lt;sup>57</sup>Cf. I don't know how many people cast the curse, but if there was only one of them...

#### (94) **Pronouns in the proposed system (complete)**



Now, as promised let me add a brief note on the local context sensitivity of *phi*-features. If *phi*-feature inferences are NRM inferences, they should be sensitive to local contexts. However, things are not as clear-cut for all *phi*-features, in a way that goes beyond the proviso problem. In (95), the presupposition of *too* is not satisfied globally, but because it is entailed by the local context created by the antecedent of the conditional, the use of *too* is felicitous. This is true both for counterfactual conditionals (the "counterfactual scenario") and for indicative ones (the "ignorance scenario").

(95)	a.	If Mia was in the library, Lea would be there, too.	counterfactual scenario
	b.	If Mia is in the library, Lea is there, too.	ignorance scenario

Yanovich (2010) claimed that gender on pronouns is not sensitive to counterfactual local contexts, i.e., the form corresponding to the actual gender of the referent should be used in examples like (96a), but in ignorance scenarios, as in (96b), the form justified by the local context can—and, in fact, has to—be used.

- a. Context: Sasha is a girl. I will buy Sasha a toy train. If Sasha were a boy, I would buy {#him, her} a doll. (Yanovich 2010, (13))
  b. Context: Sasha's gender is unknown.
  - I am at the end of my wits. If Sasha is a boy, I should buy him a doll. But if Sasha is a girl, I'd rather buy her a toy car. (Yanovich 2010, (11))

Yanovich proposed to treat gender on pronouns as indexical presuppositions, refining the original proposal in Cooper 1983 and outlining the directions for technical implementation of this idea.

I believe that Yanovich's insight is ultimately correct, but its implementation could benefit from a more careful empirical work that takes into account variation in the use of pronoun forms

in ignorance scenarios. In particular, in my experience, some speakers prefer to use gender-neutral *they* in examples like (96b). This variation is not unlike what we see in languages with the T–V distinction such as Russian, where, as shown in (97), most speakers I consulted do not allow the more marked T form in counterfactual scenarios despite it being "licensed" in the local context, but in ignorance scenarios there is more variation.<sup>58</sup> I believe a controlled experimental study is needed to establish which judgement patterns are robust and what the magnitude of variation is, which should inform the more specific analysis of gender and T–V features as indexicals.

#### (97) *Context: The speaker is talking to a single person.*

- a. Esli by my s vami byli na ty, ja by {#tebja, vas} nazyvala Anja. if SUBJ we with you.PL were on you.SG I SUBJ you.SG you.PL called Anya 'If you and I were on the T form basis, I would be calling you Anya.'
- b. Ja ne pomnju, na ty my ili na vy, no esli na ty, %{ty možeš', I not remember on you.SG we or on you.PL but if on you.SG you.SG may.2SG vy možete} nazyvat' menja Anja.
  you.PL may.2PL call me Anya
  'I don't remember if we are on the T or on the V form basis, but if we are on the T form basis, you may call me Anya.' (Russian)

Now, why would treating an expression as indexical allow us to affect its local context sensitivity? Indexicals are evaluated with respect to some part of the context parameter  $c_p^{59}$  on the evaluation function that stores the speaker, the addressee, the time and place of the utterance, etc. (see, e.g., Schlenker 2018c for an overview). For example, gender on pronouns can be linked to the index  $c_g$  within  $c_p$  that stores preferred third person pronoun forms for individuals, and T–V features can be linked to  $c_{TV}$  that stores appropriate second person pronoun forms for (ordered) pairs of individuals. Certain environments can introduce a context parameter  $c'_p$  different from the actual one, and some indexicals can shift and be evaluated with respect to  $c'_p$ ; we can also allow parts of  $c_p$  to shift independently of one another. This process can be kept independent of local context computation for presupposition projection, i.e., local contexts don't have to introduce  $c'_p$ , and even if they do, indexicals don't have to shift.

Crucially, treating gender or T–V on pronouns as indexical is independent from treating them as obligatorily non-restricting modifiers giving rise to an NRM inferences. Except instead of something like 'for all individuals in all the worlds w of the local context c', being g(i) entails being female in w', we will be getting something like 'for all individuals in all the worlds w of the local context c', being g(i) entails having *she* as the preferred pronoun form in  $c_g$  of  $c_p$ '. This inference can still project like a presupposition, but as before, it doesn't have to be hardcoded lexically into any feature entries. This independence of the indexical nature of a given expression and the nature of the inference it contributes (i.e., whether it is presuppositional, and if it is, how this presupposition is triggered) also holds for person features on pronouns, which are also linked to  $c_p$ . I leave a more in-depth investigation of indexical *phi*-features for future research.

<sup>&</sup>lt;sup>58</sup>Schlenker 2007 proposes to analyze T–V features as indexical presuppositions, but he doesn't discuss their local context sensitivity, nor does he draw parallels between T–V and gender on pronouns.

<sup>&</sup>lt;sup>59</sup>I use the p subscript to distinguish between c used throughout the paper to refer to the local context computed for presupposition projection and  $c_p$  as the context parameter for indexical interpretation.

#### 4.2 Endpoints of directional gestures

Schlenker and Chemla (2018) draw empirical parallels between *phi*-features on pronouns and endpoints of directional gestures (and endpoints of directional signs in sign language), similar to these in (98),<sup>60</sup> which they call "height specifications"—a term which I will adopt here as well.

(98) Context: Zoe is a stuntwoman. The crew just filmed a scene in which she was fighting an extra, while the director of the movie Uma was away. Uma originally wanted Zoe to stab the extra in that scene, but she just learned that Zoe might have punched the extra in the face instead. Uma says:

If Zoe punched the extra<sub>PUNCH-HIGH</sub>, we'll have to reshoot the scene.  $\rightarrow$  The extra that Zoe was fighting is taller than Zoe.

Similar constraints apply to projection of height specifications on gestures, as in the case of ordinary modifiers, spoken or gestural, and *phi*-features on pronouns. This point is illustrated in (99).

- (99) *Context: Zoe is a stuntwoman. The crew just filmed a scene in which Zoe was punching an extra while the director of the movie Uma was away. ...* 
  - a. ... Uma knows who the extra was and that he is much taller than Zoe. Uma originally wanted Zoe to punch him in the face, not in the sternum. Uma says:

If Zoe **punched** him<sub>PUNCH-HIGH</sub>, that's OK,

but if she **Carlow** punched him<sub>PUNCH-LOW</sub>, we'll have to reshoot the scene.

b. ... Uma doesn't know who the extra was, but she knows that he was punched in the face, and she originally wanted the extra to be much taller than Zoe. Uma says:

%??If Zoe **<u>punched</u>** him<sub>PUNCH-HIGH</sub>, that's OK, but if she <u>**punched**</u> him<sub>PUNCH-LOW</sub>, we'll have to reshoot the scene.

Intended: '...if (his face is  $\langle$  higher than Zoe's face, at Zoe's face level $\rangle$  and Zoe punched him)...'

In (99a), we have two potential areas of contact for the punching movement, the extra's face or his sternum, whose height with respect to Zoe's face is known and fixed. In (99b), we only have one potential point of contact, the extra's face; its height, however, is not known, and we are trying to locally establish it via a height specification on the co-verbal gesture, but fail. The second scenario should be reminiscent of the scenarios in (49) (a dog of unknown size) and (86b) (a suspect of unknown gender). The first one could be compared to either the scenarios supporting NRM inter-

<sup>&</sup>lt;sup>60</sup>Except Schlenker & Chemla looked at pro-speech gestures, and I am looking at co-speech gestures whose content is, furthermore, mostly redundant given the spoken verbs they co-occur with (except for the height specifications). I believe that the same observations can be made for co-speech gestures whose content is not redundant or for pro-speech gestures. The reason I am avoiding the former is to make the inferences of the examples simpler, and the reason I am avoiding the latter is to avoid any articulatory integration issues typical of pro-speech gestures. However, I believe any analysis of cases like in (98) should generalize to non-redundant co-speech gestures and to pro-speech gestures.

pretations in (48a) (a dog and a cat of unknown sizes) and (86a) (two suspects of different genders referred to by two pronouns with different indices) or those supporting restricting interpretations as in (48b) (two dogs of different sizes).

Schlenker and Chemla (2018) assume a presuppositional analysis for both height specifications on gestures and *phi*-features on pronouns. Naturally, treating the inferences of height specifications on gestures as lexical presuppositions raises the same questions of how these presuppositions are triggered and why they resist local accommodation (although some people find (99b) very marginally acceptable, which I come back to at the end of this section). Moreover, under this view, the nature of the empirical parallels between the two types of content remains a mystery.

I propose analyzing height specifications on gestures as modifiers and further assimilating them to *phi*-features on pronouns or to ordinary modifiers, such as adjectives. The most straightforward way to do so is to say that the gestures in (99) contain incorporated nominal arguments specifying the area of contact for the punching movement (akin to *in the face* and *in the sternum*, but without the assumption that there is a PP layer in the syntax). The height specifications on the gestures then modify these arguments. The rest depends on what exactly the nature of these arguments is.

We could say that the incorporated arguments in (99) are essentially pronouns, standing for the extra's face  $(x_i)$  and sternum  $(x_j)$ . Then the analysis of height specifications would be the same as for *phi*-features on pronouns. For example, the denotation of the "face" argument in (99a) would be  $\lambda w. \iota x. \text{high}(x)(w) \wedge x = g(i)$ , and we would get a projecting NRM inference that the extra's face is located high(er than Zoe's face). Nothing blocks the projection of these inferences in (99a), so the utterance is felicitous. In (99b), however, the NRM inferences of the height specifications on the gestures can't project, given the context, so a failure obtains.

Alternatively, we could argue that the incorporated arguments in (99) are ordinary definite descriptions, with a contextually determined descriptor (e.g.,  $\lambda w.\iota x.high(x)(w) \wedge body-part(x)(w)$ ). The height specifications in (99a) would then be restricting modifiers, just like the adjective *large* in *her large dog*, when there are two contextually relevant dogs of different size. Height specifications in (99b) would then be NRMs (as there is only one relevant body part), but since they can't project in the context, (99b) would fail in the same way as similar examples with adjectives.

Can we distinguish between these two options? A possible direction would be to compare the behavior of height specifications on co-verbal gestures under ellipsis and *only* to that of *phi*-features on pronouns and ordinary modifiers, spoken and gestural (Schlenker and Chemla (2018) claim that height specifications on gestures behave just like *phi*-features on pronouns, but they don't compare this behavior to that of regular NRMs). This would be a natural direction for future research.

Another direction for future research is extending the modifier-based analysis proposed here for *phi*-features on pronouns and endpoints of directional gestures to various geometric properties of signs in sign language and pointing in spoken language (see Ahn 2019; Grosz 2019; Ebert 2019 for some potentially relevant recent discussions).

Finally, why is (99b) marginally acceptable for some? The construal whereby the gestures in (99) incorporate nominal arguments might not be the only one available. It might be possible, at least for some people, to construe the gestures in (99) as containing no nominal arguments whatsoever, with the height specifications being adverbials specifying the direction of motion (akin to *upwards*). Under this construal, the height specifications in both examples in (99) would be restricting modifiers and should be (marginally) acceptable. Because the contrast in (99) still obtains, either this construal is dispreferred for some reason, or there is an independent reason why (99b) is still degraded under this construal. This is yet another issue I leave for future research.

## **5** Conclusion

In this paper, I have put forward a uniform, composition-driven approach to projection of compositionally integrated content across modalities. I have focused on two composition strategies, which come with associated projection behavior patterns: modification and supplementation. I have argued that this distinction exists in both the primary modality (lexicalized spoken expressions in spoken language) and various secondary modalities (gestures, facial expressions, suprsegmental morphemes), with uniform consequences for projection. I have classified several types of projective or potentially projective content as modifiers or supplements. I summarize the modifier vs. supplement distinction and my claims about the different types of content in Table 2, which is an updated version of Table 1.

Modifiers	Supplements
Compose with $\beta_{\tau}$ , yield $\alpha_{\tau}$	Compose with $\beta$ , pass it unchanged, and yield a proposition of a special type $st^c$ about $\beta$
If subsective $(\alpha \Rightarrow \beta)$ , can be restricting $(\beta \Rightarrow \alpha)$ or non-restricting $(\beta \Rightarrow \alpha)$	Can never be restricting
Trigger projecting inferences when non-restricting; triggering is pragmatic	Almost always trigger projecting inferences; triggering is conventional
<ul> <li>Examples (target content bolded):</li> <li>modifiers of NPs: <ul> <li>spoken adjectives (<i>her large dog</i>; <i>her so-called friend</i>)</li> <li>restrictive relatives (<i>her dog that's large</i>)</li> <li>gestures/demonstrations <ul> <li>(<i>her [[dog]LARGE]</i>; <i>a [[semanticist]CRAZY]</i>;</li> <li><i>her [[friend]AIR-QUOTES]</i>; <i>a HEAD-TILT<sup>HMM</sup> experience</i>)</li> </ul> </li> <li>modifiers of verbal projections: <ul> <li>spoken arguments and adjuncts (<i>shot a longbow</i>; <i>shot with a longbow</i>; <i>ran fast</i>)</li> <li>gestures (<i>shot at the targetLONGBOW</i>(<i>v,st</i>))</li> </ul> </li> <li>degree modifiers: <ul> <li>spoken (<i>very drunk</i>; <i>fucking drunk</i>; [DOG [CRED]] → <i>dog<sup>T*</sup> dog</i>)</li> <li>intonational (<i>drunkDEG-INT</i>)</li> <li>facial expressions (<i>DRUNK<sup>OO</sup></i>)</li> </ul> </li> <li><i>phi</i>-features on spoken pronouns (<i>her<sup>[3,sg,fem]</sup></i>)</li> </ul></li></ul>	<ul> <li>Examples (target content bolded):</li> <li>supplements w/DP anchors: <ul> <li>spoken appositives (her dog, (who is) a large animal)</li> <li>gestures ([[her dog]LARGE]; (IP her dog) (IP LARGE))</li> </ul> </li> <li>supplements w/verbal anchors <ul> <li>spoken appositives (Zoe has never run a marathon, which is a hard thing to do)</li> <li>gestures (Zoe shot at the targetLONGBOW((v,st),stc); (IP Zoe shot at the target) (IP LONGBOW))</li> </ul> </li> <li>supplements w/propositional anchors <ul> <li>spoken appositives (Kim says that Zoe ran a marathon, which I find hard to believe)</li> <li>spoken adverbs (Surprisingly, Kim brought her [husband]<sub>F</sub> to the party)</li> <li>gestures (If a friend of mine wins<sub>FINGERS-CROSSED</sub>)</li> <li>facial expressions (Kim brought her</li> </ul> </li> </ul>

For spoken content, one can usually tell whether a given expression is a modifier or a supplement based on its morphosyntactic makeup (although in some cases the difference is only signalled by prosody). For secondary modality content, which is often morphosyntactically poor, this is not always immediately clear, and one and the same string can sometimes be ambiguous between the modifier and the supplement construal (or the surface differences might be very subtle).

The two construals do, however, come apart in many cases. For example, they do so when one construal cannot yield the right interpretation, as is the case with the target expressions interpreted as restricting modifiers, since supplements can't be restricting. They also come apart in the case of co-nominal gestures without an individual discourse referent to serve as an anchor for them, in which case the supplement construal is not possible either, as adnominal supplements require individual anchors. Next, the two construals come apart in many cases of pro-speech gestures (i.e., those whose main stroke doesn't co-occur with any lexicalized spoken material) integrated into otherwise spoken utterances. Those follow the rules of linearization, syntax/prosody mapping, and syntactic displacement of the spoken language they are embedded into. Combined with articulatory considerations of speech-gesture integration, these rules often block the modifier construal for pro-speech gestures in English, however, exceptions exist even in English, and since many of the syntactic and phonological rules are language-specific, it is easier to get modifier construals for pro-speech gestures in other languages. Finally, conventionalized secondary modality expressions can come with various lexicalized constraints that affect their compositional possibilities, which, in turn, affects their projection behavior. For example, the gesture CRAZY seems to always be an adjective morphosyntactically and behaves as such when it comes to composition and projection, and gestures like FINGERS-CROSSED are always sentence-level supplements and project as such. An especially curious case of the modifier and supplement construals for secondary modality content coming apart is the distinction between the degree modifier and the sentence-level supplement uses of the facial expression encoding surprisal, OO. The two interpretations associated with the two construals are very distinct, even if the surface alignment of the facial expression is not always a very reliable predictor for which construal is intended.

Importantly, the compositional construal—and, consequently, the projection behavior—of a given piece of secondary modality content cannot be directly predicted based on whether it cooccurs with something in a more primary modality or whether it is syntactically optional (even though there might be some interface-mediated correlations with either factor). The "co-something" status might make a given piece of content preferably truth-conditionally vacuous (as seems to be the case for many co-speech gestures), but it doesn't directly determine how this truth-conditional vacuity is assured; furthermore, this preference can be overridden by other considerations. By demonstrating the diversity and complexity of linguistic behaviors exhibited by secondary modality content, this paper has thus challenged both approaches that sweep such content under a "paralinguistic" rug or posit modality-specific rules for such content based on non-linguistic notions.

## References

Abrusán, M. (2011). Predicting the presuppositions of soft triggers. *Linguistics and Philosophy*, 34(6):491–535.

Ahn, D. (2019). ASL IX to locus as a modifier. Talk given at the 50th Meeting of the North East Linguistic Society (NELS), Massachusetts Institute of Technology.

- AnderBois, S., Brasoveanu, A., and Henderson, R. (2013). At-issue proposals and appositive impositions in discourse. *Journal of Semantics*, 32(1):93–138.
- Aristodemo, V. M. V. (2017). *Gradable constructions in Italian Sign Language*. PhD thesis, Institut Jean-Nicod.
- Beaver, D. I. (2001). *Presupposition and assertion in dynamic semantics*. CSLI publications, Stanford, CA.
- Beck, S., Krasikova, S., Fleischer, D., Gergel, R., Hofstetter, S., Savelsberg, C., Vanderelst, J., and Villalta, E. (2009). Crosslinguistic variation in comparison constructions. *Linguistic variation yearbook*, 9(1):1–66.
- Beckman, M. E. and Ayers, G. (1997). Guidelines for ToBI labelling. Version 3.0. The Ohio State University Research Foundation.
- Beltrama, A. and Hanink, E. A. (2019). Marking imprecision, conveying surprise: *Like* between hedging and mirativity. *Journal of Linguistics*, 55(1):1–34.
- Bochnak, M. R. (2015). The degree semantics parameter and cross-linguistic variation. *Semantics* and *Pragmatics*, 8(6):1–48.
- Carlson, G. N. (1984). Thematic roles and their role in semantic interpretation. *Linguistics*, 22(3):259–280.
- Clark, N., Perlman, M., and Johansson Falck, M. (2013). Iconic pitch expresses vertical space. In Dancygier, B., Borkent, M., and Hinnell, J., editors, *Language and the creative mind*, pages 393–410. CSLI Publications, Stanford.
- Cooper, R. (1983). Quantification and syntactic theory. Reidel, Dordrecht.
- Davidson, K. (2015). Quotation, demonstration, and iconicity. *Linguistics and philosophy*, 38(6):477–520.
- Ebert, C. (2017). Co-speech vs. post-speech gestures. Talk given at *Language and cognition* workshop in memory of Peter Bosch, Osnabrück, February.
- Ebert, C. (2019). Pointing—and its meaning. Invited talk given at the 22nd Amsterdam Colloquium.
- Ebert, C. and Ebert, C. (2014). Gestures, demonstratives, and the attributive/referential distinction. Talk given at *Semantics and Philosophy in Europe (SPE 7)*, Berlin.
- Elbourne, P. D. (2005). Situations and individuals. MIT Press, Cambridge, MA.
- Ernst, T. B. (1984). *Towards an Integrated Theory of Adverb Position in English*. PhD thesis, Indiana University.
- Esipova, M. (2019a). Acceptability of at-issue co-speech gestures under contrastive focus. *Glossa: A Journal of General Linguistics*, 4(1):11.
- Esipova, M. (2019b). Composition and projection in co-speech gestures. In Blake, K., Davis, F., Lamp, K., and Rhyne, J., editors, *Proceedings of Semantics and Linguistic Theory (SALT)* 29, pages 111–137.
- Esipova, M. (2019c). Polar responses in Russian across modalities and across interfaces. Talk given at *Formal Approaches to Slavic Linguistics (FASL)* 28, Stony Brook University.
- Geurts, B. (1996). Local satisfaction guaranteed: A presupposition theory and its problems. *Linguistics and Philosophy*, 19(3):259–294.
- Ghomeshi, J., Jackendoff, R., Rosen, N., and Russell, K. (2004). Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory*, 22(2):307–357.
- Grosz, P. G. (2019). Pronominal typology and reference to the external world. In Schlöder, J. J., McHugh, D., and Roelofsen, F., editors, *Proceedings of the 22nd Amsterdam Colloquium*, pages

563-573.

Halle, M. and Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Hale, K. and Keyser, S. J., editors, *The View from Building 20*, pages 111–176. MIT Press, Cambridge, MA.

Harris, A. (2020). Prosodic integration of pro-speech gestures. Junior Paper, Princeton University.

- Heim, I. (1983). On the projection problem for presuppositions. In Barlow, M., Flickinger, D., and Wescoat, M., editors, *Proceedings of the Second West Coast Conference on Formal Linguistics*, pages 114–125.
- Heim, I. and Kratzer, A. (1998). Semantics in generative grammar. Blackwell, Oxford.
- Hunter, J. (2018). Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics and Philosophy*, pages 1–16.
- Hunter, J. (2019). Relating gesture to speech. Talk given at the workshop *Linguistic Investigations Beyond Language: Gestures, Body Movement and Primate Linguistics,* ZAS Berlin.
- Ippolito, M. (2019). Gestures as markers on non-canonical questions. In Schlöder, J. J., McHugh, D., and Roelofsen, F., editors, *Proceedings of the 22nd Amsterdam Colloquium*, pages 575–583.
- Jasinskaja, K. and Poschmann, C. (2018). Attachment in syntax and discourse: towards an explanation for the variable scope of non-restrictive relative clauses. In Maspong, S., Stefánsdóttir, B., Blake, K., and Davis, F., editors, *Proceedings of Semantics and Linguistic Theory (SALT)* 28, pages 433–453.
- Jeong, S. (2018). Intonation and sentence type conventions: Two types of rising declaratives. *Journal of Semantics*, 35(2):305–356.
- Karttunen, L. (1974). Presuppositions and linguistic context. *Theoretical Linguistics*, 1:181–194.
- Kennedy, C. and McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81(2):345–381.
- Koev, T. (2013). Apposition and the structure of discourse. PhD thesis, Rutgers University.
- Ladd, D. R. (2008). Intonational Phonology. Cambridge University Press, Cambridge.
- Lascarides, A. and Stone, M. (2009). A formal semantic analysis of gesture. *Journal of Semantics*, 26(4):393–449.
- Lee, N. (2019). Khoekhoe pronominal morphosyntax: Gender on Root-attached little *n*. Talk given at *The 93rd Annual Meeting of the LSA*, New York City.
- Leffel, T. (2014). *The semantics of modification: Adjectives, nouns, and order*. PhD thesis, New York University.
- Loehr, D. P. (2004). Gesture and intonation. PhD thesis, Georgetown University.
- Matushansky, O. (2008). On the linguistic complexity of proper names. *Linguistics and Philoso-phy*, 31(5):573–627.
- McCawley, J. D. (1998). *The syntactic phenomena of English*, volume 2. The University of Chicago Press, Chicago IL.
- Morzycki, M. (2008). Nonrestrictive modifiers in nonparenthetical positions. In Kennedy, C. and McNally, L., editors, *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, Studies in Theoretical Linguistics, pages 101–122. Oxford University Press, Oxford.
- Nouwen, R. (2014). A note on the projection of appositives. In McCready, E., Yabushita, K., and Yoshimoto, K., editors, *Formal Approaches to Semantics and Pragmatics*, pages 205–222. Springer, Dordrecht.
- Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In Hinton, L., Nichols, J., and Ohala, J. J., editors, *Sound symbolism*, pages 325–347. Cambridge

University Press, Cambridge.

Parsons, T. (1990). Events in the semantics of English. MIT Press, Cambridge, MA.

- Partee, B. and Rooth, M. (1983). Generalized conjunction and type ambiguity. In Bäuerle, R., Schwarze, C., and von Stechow, A., editors, *Meaning, use and interpretation of language*, pages 361–383. de Gruyter, Berlin.
- Pierrehumbert, J. and Hirschberg, J. B. (1990). The meaning of intonational contours in the interpretation of discourse. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, *Intentions in communication*, pages 271–311. MIT Press, Cambridge.
- Postal, P. (1966). On so-called pronouns in English. In Dinneen, F. P., editor, *Report on the 17th Annual Round Table Meeting of Language and Linguistics Studies*, pages 177–206. Georgetown University Press, Washington, D.C.
- Potts, C. (2005). The logic of conventional implicatures. Oxford University Press, Oxford.
- Potts, C. (2007). The expressive dimension. *Theoretical linguistics*, 33(2):165–198.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., and Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, 90(6):2956–2970.
- Schlenker, P. (2007). Expressive presuppositions. *Theoretical Linguistics*, 33(2):237–245.
- Schlenker, P. (2009). Local contexts. Semantics and Pragmatics, 2(3):1–78.
- Schlenker, P. (2010). Presuppositions and local contexts. Mind, 119(474):377-391.
- Schlenker, P. (2013). Supplements without bidimensionalism. Ms., Institut Jean-Nicod and New York University.
- Schlenker, P. (2018a). Gesture projection and cosuppositions. *Linguistics and Philosophy*, 41(3):295–365.
- Schlenker, P. (2018b). Iconic pragmatics. Natural Language & Linguistic Theory, 36(3):877–936.
- Schlenker, P. (2018c). Indexicals. In Hansson, S. O. and Hendricks, V. F., editors, *Introduction to Formal Philosophy*, pages 297–321. Springer.
- Schlenker, P. (2019). Triggering presuppositions. Ms., Institut Jean-Nicod and New York University.
- Schlenker, P. (To appear). Iconic presuppositions. Natural Language & Linguistic Theory.
- Schlenker, P. and Chemla, E. (2018). Gestural agreement. *Natural Language & Linguistic Theory*, 36(2):587–625.
- Selkirk, E. (2005). Comments on intonational phrasing in English. In Frota, S., Vigário, M., and Freitas, M. J. a., editors, *Prosodies: with special reference to Iberian languages*, pages 11–58. Mouton de Gruyter.
- Stalnaker, R. (1974). Pragmatic presuppositions. In Kunitz, M. and Unger, P., editors, *Semantics and philosophy*, pages 197–213. New York University Press.
- Starke, M. (2010). Nanosyntax: A short primer to a new approach to language. Nordlyd, 36(1):1-6.
- Sudo, Y. (2012). On the semantics of phi features on pronouns. PhD thesis, Massachusetts Institute of Technology.
- Tieu, L., Pasternak, R., Schlenker, P., and Chemla, E. (2017). Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: A Journal of General Linguistics*, 2(1):102.
- Tieu, L., Pasternak, R., Schlenker, P., and Chemla, E. (2018). Co-speech gesture projection: Evidence from inferential judgments. *Glossa: A Journal of General Linguistics*, 3(1):109.
- Umbach, C. (2006). Non-restrictive modification and backgrounding. In Gyuris, B., László, K., Piñón, C., and Varasdi, K., editors, *Proceedings of the Ninth Symposium on Logic and Language*,

pages 152–159, Budapest. Hungarian Academy of Sciences.

- Wang, L., Reese, B., and McCready, E. (2005). The projection problem of nominal appositives. *Snippets*, 10(1):13–14.
- Yanovich, I. (2010). On the nature and formal analysis of indexical presuppositions. In Nakakoji, K., Murakami, Y., and McCready, E., editors, *New Frontiers in Artificial Intelligence. JSAI-isAI* 2009. Lecture Notes in Computer Science, pages 272–291. Springer, Berlin, Heidelberg.
- Zlogar, C. and Davidson, K. (2018). Effects of linguistic context on the acceptability of co-speech gestures. *Glossa: A Journal of General Linguistics*, 3(1):73.