

Focus on what's not at issue: Gestures, presuppositions, appositives under contrastive focus¹

Maria ESIPOVA — New York University

Abstract. This paper is an attempt to systematically investigate how contrastive focus interacts with various types of not-at-issue content (co-speech and post-speech gestures, lexical presuppositions, and appositives). I look, in particular, at when focus forces at-issue interpretations of typically not-at-issue content, when it doesn't, and when such at-issue interpretations are impossible even to satisfy focus-related requirements. I conclude that the main factors affecting how a given type of content aligns along these dimensions are its prosodic (in)dependence and level of attachment in the syntax. The two factors also interact in a non-trivial way, in particular, for gestures, which I use as a basis for an analysis of gestures that doesn't assume that their temporal alignment directly determines their semantics (contra Ebert and Ebert 2014, Ebert 2017, Schlenker 2017), but instead relies on syntax/semantics and syntax/prosody interaction.

Keywords: focus, not-at-issue content, gestures, presuppositions, appositives

1. Introduction

Contrastive focus has been observed to affect presupposition projection (Abusch 2002, Simons et al. 2017, a.o.). For example, (1) typically gives rise to a global inference, triggered by *stop*, that Umbridge used to drink firewhisky, but (2) doesn't give rise to any global inferences about her previous drinking habits, despite the presence of the presupposition triggers *stop* and *start*.

- (1) Umbridge might have stopped drinking firewhisky.
→ Umbridge used to drink firewhisky.
- (2) *Context: The faculty at Hogwarts have to report to Madam Pomfrey whenever they significantly change their drinking habits. Ron knows that Umbridge has filed such a report, but he doesn't know how exactly her habits have changed; he says:*
Umbridge might have **stopped** drinking firewhisky, but she also might have **started** doing so.
↗ Umbridge used {to drink, to not drink} firewhisky.

Standard theories of presupposition projection (e.g., Heim 1983, Schlenker 2009) can handle examples like (2) by resorting to *local accommodation*, which is a process of making a presupposition part of the at-issue content by treating it as a conjunct at the level at which it is triggered. For example, applying local accommodation in (2) makes it roughly equivalent to:

¹For empirical and/or theoretical discussions at various points of this project I'd like to thank Amir Anvari, Lucas Champollion, Ailís Courbane, Kathryn Davidson, Cornelia Ebert, Edward Flemming, Irene Heim, Sabine Iatridou, Paloma Jeretič, Rob Pasternak, Jon Rawski, Roger Schwarzschild, Ildikó Emese Szabó, Anna Szabolcsi, Adina Williams, and Kata Wohlmuth, as well as audiences at MACSIM VI, Philippe Schlenker's seminar at NYU in Fall 2016, LFRG reading group at MIT and Meaning & Modality lab at Harvard in Spring 2017, NYU Semantics Group in Summer 2017, and Sinn und Bedeutung 22. Special thanks to Philippe Schlenker for helpful discussions and feedback at all stages of this project.

- (3) Umbridge might have used to drink firewhisky and stopped, but she also might have used to not drink firewhisky and started.

A similar observation can be made for inferences triggered by *co-speech gestures*, i.e., content-bearing, non-conventionalized gestural adjuncts co-occurring with verbal expressions they adjoin to. Such inferences typically project from embedded environments (Ebert and Ebert 2014, Ebert 2017, Schlenker 2017; see also Tieu et al. 2017a, b for some experimental evidence):

- (4) *Context: The Yule ball at Hogwarts is tomorrow.*

If Hagrid brings a dog^{LARGE} [], it's gonna be a mess.
 → If Hagrid brings a dog, it will be a large one.²

The status of the inference in (4) is a matter of debate. Ebert (and Ebert) (2014, 2017) claim it is a Pottsonian (2005) supplement, and Schlenker (2017) argues it is a special kind of presupposition. The two analyses make very similar predictions for (4). However, contrastive focus can sometimes force co-speech gestures to be interpreted as at-issue restrictive modifiers³:

- (5) If Hagrid brings a dog^{SMALL} [], it's gonna be OK,

[]
 but if he brings a dog^{LARGE} [], it's gonna be a mess.
 ↗ If Hagrid brings a dog, it will be a {small, large} one.
 ≈ If Hagrid brings a **small** dog... but if he brings a **large** dog...

At first glance, the parallel between (2) and (5) could be used as an argument in favor of the presuppositional analysis of co-speech gestures, but it turns out that the empirical picture is more complicated. This paper is an attempt (to my knowledge, the first one) to systematically investigate the interaction of contrastive focus with different types of not-at-issue content and explain the differences that arise in a principled way. I look at when focus-related considerations force at-issue interpretations of typically not-at-issue content and when they don't. I also show that in some cases at-issue interpretations are unavailable even when that would be the only way to make contrast felicitous. The data suggest that the two major factors affecting the

²For those unfamiliar with ‘Harry Potter’ lore, Hagrid is a half-giant with a fondness for large animals, so this inference is quite natural.

³The first discussion of examples like (i) I’m aware of is due to Rob Pasternak (p.c.), but such examples can be in principle analyzed via metalinguistic negation, which targets the form and not the content. In this paper I try to look at examples that are unlikely to involve metalinguistic uses of expressions containing projective content.

- (i) I don’t want a beer^{SMALL} [], I want a beer^{LARGE} []!

focus-related behavior of a given type of not-at-issue content are whether it's prosodically independent and where it attaches in the syntax. With this in mind, I sketch an analysis of gestures that relies on both of those factors, thus, diverging from both Ebert's and Schlenker's accounts.

The rest of the paper is organized as follows. In section 2 I clarify my notation, terminology, and data elicitation practice. In section 3 I explore when contrastive focus forces at-issue interpretations of typically not-at-issue content, when it doesn't, and when even contrast requirements can't force such interpretations. In section 4 I discuss the issues those data pose for Ebert's and Schlenker's analyses of gestures and propose an alternative. Section 5 concludes.

2. Notation, terminology, and data elicitation

I adopt the following notational conventions in this paper:

- In 'verbal expression' GESTURE, the gesture co-occurs with the verbal expression (the underlining loosely indicates temporal alignment of the gesture, without any syntactic claims).
- In 'verbal expression — GESTURE' the gesture follows the verbal expression.
- New gestures are illustrated by pictures after an underscore.
- A **word** written in bold indicates prosodic (contrastive) focus marking: primarily (L+)H* pitch accent in ToBI terms (Beckman and Ayers 1997) on the stressed syllable of the bolded word for verbal content; kinetic emphasis for standalone gestures.
- A *word* written in italics in examples indicates prosodic contrastive topic marking: (L+)H* or L*+H pitch accent on the stressed syllable of the italicized word. Note that it is very hard, if not impossible, to distinguish between contrastive topic and contrastive focus prosodic marking in isolation. The choice of marking is thus informed by the semantic considerations.
- (IP...) marks intonational phrases (IPs) in ToBI terms whenever it's relevant.

Throughout the paper I will use the term *not-at-issue* to talk about content that projects from embedded environments all the way to the global context, such as the presupposition in (1) or the gestural inference in (4). I will use the term *at-issue* to talk about content that doesn't project at all, i.e., is interpreted under the lowest operator under which it is embedded, as in (2) or (5). Thus, for the purposes of this paper I don't care, in particular, whether a given piece of content can be directly negated in the discourse. For example, it has been shown in Syrett and Koev 2014 that utterance-final appositive relative clauses (ARCs), as in (6) (adopted from Syrett and Koev 2014; crucially, the original example doesn't contain negation), can be relatively easily targeted by direct negation in the discourse. For me the content of the ARC in A's utterance in (6) is still not-at-issue, since it projects from under negation.

- (6) A: The symphony didn't hire my friend Sophie, who is a classical violinist.
 → Sophie is a classical violinist.
 B: That's not true! Sophie isn't a classical violinist.

As for the term *local accommodation*, while it was originally coined for presuppositions, I will use it to talk about treating any type of typically not-at-issue content as a conjunct at the local syntactic level. For presuppositions that's the level at which the trigger is merged; for appositives and gestures that would be the level at which they adjoin in the narrow syntax.

Finally, the data reported in this paper are based on introspective judgements of native speakers of English (all linguists); for each example judgements were elicited from at least three speakers (for most examples the number of speakers is more than five). Any variation or uncertainty in judgements is reported.

3. When contrastive forces local accommodation, when it doesn't, and when it can't

3.1. Contradictory inferences

We have already seen one case of at-issue interpretations of typically not-at-issue content under contrastive focus in (2) for lexical presuppositions and in (5) for adnominal co-speech gestures; (7) illustrates the same case for adverbial co-speech gestures.

- (7) If Hermione stirs her potion^{CLOCKWISE} , it will turn blue,



but if she stirs her potion^{COUNTERCLOCKWISE} , it will turn green.⁴

↗ If Hermione stirs her potion, she'll do so {clockwise, counterclockwise}.

Such examples can be accounted for without any sophisticated theory of interaction of not-at-issue content with focus. Projecting an inference *p* globally means imposing a requirement on the context set that it entail *p*. In each of the examples above, the two inferences *p* and *q* that would normally project contradict each other. Since no context set can entail a contradiction, it can't be the case that both *p* and *q* project. Technically, locally accommodating only one of the inferences in each case should suffice to avoid the contradiction, but this should be ruled out via further pragmatic reasoning on what the speaker believes to be possible (à la Gazdar's (1979) clausal implicatures or any alternative). For example, in (5) and (7), projecting one of the contradictory inferences and accommodating the other would render one of the antecedents false, but it's odd to utter a conditional whose antecedent the speaker believes to be false.

The reasoning above applies regardless of the type of focus involved or what the focus alternatives are. However, when the inferences contributed by the not-at-issue expressions within an utterance are not contradictory, the type of focus and the nature of the alternatives do matter for whether a given content ends up projecting. The next two subsections review such cases.

3.2. Contrastive topic + focus configuration

Oftentimes the content semantically in focus has to address the question under discussion (QUD; in the sense of Roberts 2012, a.o.) and thus is inevitably at-issue, as opposed to back-

⁴I believe, when they are not contrasted with each other, the CLOCKWISE and COUNTERCLOCKWISE gestures don't necessarily give rise to any inferences about the directionality of movement, but in a context when the directionality is the only locus of contrast, it naturally has to be interpreted maximally iconically.

grounded. This is the case, in particular, within a contrastive topic + focus (CT+F) configuration (in the sense of Büring 2003, a.o.). Consider, for instance, the following pair of examples:

- (8) a. *Hermione petted Buckbeak, and Luna petted Fang.*⁵
 QUD: Who petted whom?
 sub-QUDs: Who did Hermione pet? Who did Luna pet?
 b. *Hermione petted Buckbeak, and Luna fed Buckbeak.*
 QUD: Who did what to Buckbeak?
 sub-QUDs: What did Hermione do to Buckbeak? What did Luna do to Buckbeak?

In (8a), the part of content that specifies that it was Buckbeak who Hermione petted is new information addressing the immediate sub-QUD and thus cannot possibly be presupposed; that Hermione petted someone might very well be (perhaps weakly) presupposed (the same is true, mutatis mutandis, about Fang and Luna). In (8b), what is at-issue is what the girls did to Buckbeak; that they did something to Buckbeak might, once again, be presupposed.

Now, let us turn to perhaps less obvious CT+F examples with co-speech gestures:

- (9) *Context (applies to all animal fighting examples): At Hogwarts, when a small animal and a large animal find themselves in the same room, they usually fight.*
 If Flitwick brings a dog^[SMALL]_F, and Hagrid brings a dog^[LARGE]_F, they'll fight.⁶
 ↗ If {Flitwick, Hagrid} brings a dog, it will be a {small, large} one.
 ≈ If Flitwick brings a **small** dog, and Hagrid brings a **large** dog...

- (10) If *Hermione stirs her potion*^[CLOCKWISE]_F,
 and *Luna stirs her potion*^[COUNTERCLOCKWISE]_F, there will be an explosion.
 ↗ If {Hermione, Luna} stirs her potion, she'll do so {clockwise, counterclockwise}.
 ≈ If *Hermione* stirs her potion **clockwise**, and *Luna* stirs her potion **counterclockwise**...

The gestural inferences in the two examples above don't project, even though the two gestural inferences in either example don't contradict each other, so the simple reasoning described in the previous subsection is insufficient to account for their non-projection. Instead, the gestures in the examples above behave like ordinary modifiers under focus in a CT+F configuration, i.e., the prosodic prominence co-occurring with each word-gesture pair signals semantic focus on the gesture rather than on the verbal expression. Intuitively, the reason why prominence has to associate with the gestures in both examples is because the verbal content of the two items under prominence in each example is non-contrastive, and there is a general requirement for all F-items within a CT+F coordinated structure to contrast with each other:

- (11) #*Hermione petted Buckbeak, and Luna petted Buckbeak.*

At this point I will not offer a comprehensive story about what it means for two expressions to

⁵Buckbeak is a hippogriff, and Fang is a dog.

⁶Flitwick is a part-goblin and thus very short, so the inference that if he brings a dog to the Yule ball, it will be a small one would be a natural one.

be contrastive (but see, e.g., Katzir 2014 for a discussion). For the purposes of this paper, let us assume the — inevitably simplified — view that for any two expressions α and β of the same semantic type ending in t , α and β are properly contrastive if neither entails the other (assuming generalized entailment), and α and β of type e are contrastive if $[\alpha] \neq [\beta]$. Similarly, I won't engage in a discussion about the nature and status of this contrast requirement on F-items in CT+F coordinated structures. I'll just assume that it holds, at least in the examples above (but see van Rooij 2010, Büring 2016 for a potentially relevant discussion).

One could try to rebut the contrast-based intuition above by arguing that the gestures in (9) and (10) give rise to more general inferences; e.g., a dog^{SMALL} in *Hagrid brings a dog*^{SMALL} doesn't give rise to an inference about the size of a dog Hagrid would bring but rather about the size of dogs in general. If that's the case, then the inferences of a dog^{SMALL} and a dog^{LARGE} would be contradictory, and thus the reasoning from the previous subsection could derive non-projection in (9) without appealing to any focus-related considerations. To use the same argument for (10), one would have to assume that the gestures in (10) necessarily attach to the verb *stirs* rather than the whole VP *stirs her potion*, since the two VPs don't have the same semantics due to the different indices on the pronouns. This reasoning would predict that mere co-occurrence of mutually exclusive gestures with the same predicate within one utterance should force non-projection. However, this prediction isn't borne out; once we remove prominence from the word-gesture pairs, either by getting rid of the CT+F configuration altogether or by making something else the F-items, the gestural inferences can project:

- (12) If **Flitwick** brings a dog^{SMALL}, and **Hagrid**(, too,) brings a dog^{LARGE}, they'll fight.
→ If ⟨Flitwick, Hagrid⟩ brings a dog, it will be a ⟨small, large⟩ one.

- (13) If *Hermione* stirs **her** potion^{CLOCKWISE},
and *Luna* stirs **her** potion^{COUNTERCLOCKWISE}, there will be an explosion.
→ If ⟨Hermione, Luna⟩ stirs her potion, she'll do so ⟨clockwise, counterclockwise⟩.⁷
(A sample context: Hermione and Luna are brewing potions next to each other; Hermione's potion requires stirring clockwise, and Luna's potion requires stirring counterclockwise; stirring two nearby potions in different directions causes an explosion.)

This observation suggests that even if gestures can give rise to generic inferences about predicates, more narrow inferences are certainly also possible.

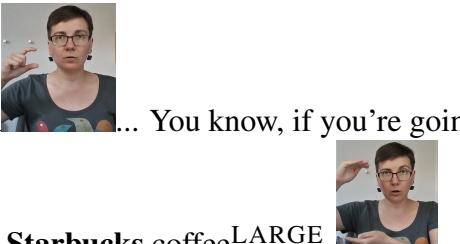
Furthermore, one could speculate that perhaps focus-marking prosodic prominence co-occurring with a gesture always makes it at-issue, without any additional contrast considerations. This, however, is not the case either. Once we make the verbal content of the word-gesture pairs under focus-marking prominence within a CT+F coordinated structure contrastive, as in (14), the prominence can associate with the verbal content. Co-speech gestures don't like to be at-issue unless under pressure, so the prominence will in fact preferably associate with the verbal content, thus, no longer forcing the gestures to be at-issue.⁸

⁷With the caveat that the contrast between the two gestures might be interpreted as an inconsistency in the default stirring gesture (see fn. 4), in which case the direction of movement wouldn't be interpreted iconically.

⁸It is as of now unclear to me what is the cause and what is the effect here. It might very well be that co-speech

- (14) If *Hagrid* brings a a [dog]_F^{LARGE}, and *Filch* brings a a [cat]_F^{SMALL}, they'll fight.
 → If ⟨Hagrid, Filch⟩ brings a ⟨dog, cat⟩, it will be ⟨large, small⟩.⁹

Here is also a naturally occurring example of this configuration produced by a Parisian guide:

- 
- (15) If you're going for a coffee^{SMALL} ... You know, if you're going for
 a real coffee^{SMALL} ..., not a Starbucks coffee^{LARGE} ...
 → ⟨Real, Starbucks⟩ coffees are ⟨small, large⟩.

So, as an intermediate summary: when a prosodically F-marked element within a CT+F coordinated structure is a word-gesture pair, the at-issue interpretation of the gesture is only forced when the verbal content in that pair doesn't properly contrast with the other F-items.

Now, does the same generalization apply to lexical presuppositions? One problem is that we can usually only speculate what the at-issue/not-at-issue content of a given trigger is. That said, let's look at *start* and *stop*. One option is that they contrast in both at-issue and presuppositional content: for *start P* the two components are, roughly, ‘P now’ and ‘not P before’, and for *stop P* they are ‘not P now’ and ‘P before’. If the generalization above applies to lexical presuppositions, having *stop* and *start* as prosodically F-marked elements within a CT+F coordinated structure should not force local accommodation, which indeed seems to be the case¹⁰:

- (16) If *Umbridge stopped* drinking firewhisky, and *McGonagall started* doing so, we're in trouble.
 → ⟨Umbridge, McGonagall⟩ used to ⟨drink, not drink⟩ firewhisky.

Now let's look at another pair of lexical items, *know* and *think*, which can be construed as having the same at-issue content, but differing in their not-at-issue content, with *know* but not *think* triggering a factive presupposition. If that's correct, the generalization above predicts local accommodation when the two are prosodically F-marked elements in a CT+F coordinated structure. It seems indeed that the inference typically triggered by *know* in non-contrastive contexts doesn't obtain when *know* is contrasted with *think* in a CT+F coordinated structure:

- (17) a. If *Hermione knows* that her parents are in danger, she'll talk to Dumbledore.
 → *Hermione's parents are in danger*.
 b. If *Ron thinks* that his parents are in danger, and *Hermione knows* that *hers*¹¹ are, they'll talk to Dumbledore.
 ↗ *Hermione's parents are in danger*.

gestures are preferably not-at-issue precisely because it is harder for focus-marking prominence to associate with them, which, given certain assumptions about givenness, can very well result in default projection.

⁹Here it is quite natural to get a secondary generic inference that dogs are in general larger than cats.

¹⁰Local accommodation, of course, is still possible, e.g., in a context similar to the one in (2).

There is an apparent problem with a local accommodation approach to (17b), though, since (17b) doesn't really have the reading whereby the complement of *know* is a conjunct under *if*:

- (18) If Ron thinks that his parents are in danger, and Hermione's parents are in danger and she thinks that they are, Ron and Hermione will talk to Dumbledore.

Instead, (17b), intuitively, seems to suggest a contrast in how much evidence about their parents being in danger it will take Ron and Hermione respectively to go talk to Dumbledore, regardless of whether their parents actually are in danger. There has been some attempts to revisit the lexical semantics of *know* in terms of the level of certainty of the attitude holder or amount of evidence available to them (see, e.g., Wiegand 2018 for a recent discussion). However, as things stand, those accounts are insufficient to predict the nature of inferences triggered by *know* and their projection behavior in all environments; in particular, they don't account for (17a). One (perhaps, unsatisfying) possibility would be to say that *know* is lexically ambiguous between a factive, presupposition-triggering predicate, whose at-issue content is roughly equivalent to that of *think*, and a non-factive, non-presuppositional predicate meaning something like 'believe with a great amount of certainty/evidence' — as opposed to *think*, which would mean something like 'believe with a moderate amount of certainty/evidence'. As things stand, both readings would be predicted to be possible in non-contrastive environments (so, if there is a preference for the factive reading, something additional needs to be said), but the second reading would become much more salient when *know* is contrasted with *think*.

Setting the peculiarities of attitude predicates aside, it would seem that the data for some lexical presupposition triggers, in particular, *start/stop* and *know*, are compatible with the generalization made for co-speech gestures. The content that would normally project doesn't do so and is instead treated as part of the at-issue content (via local accommodation or lexical adjustments) when it's necessary to make the necessarily at-issue F-items within a CT+F coordinated structure contrastive. When the contrast requirement can be satisfied by some other content, the typically not-at-issue content, both gestural and presuppositional, can remain not-at-issue.

3.3. Not-at-issue focus

Let us now turn to focus that doesn't require at-issueness of the content it targets. Such focus marks novelty of certain content without having it address the QUD, for example:

- (19) (IP If **Flitwick** brings a dog), (IP which will be **small**), (IP and **Hagrid**(, too,) brings a dog), (IP which will be **large**), (IP they'll fight).
→ If ⟨Flitwick, Hagrid⟩ brings a dog, it will be ⟨small, large⟩.
- (20) (IP If *Hermione* stirs **her** potion), (IP which she'll do **clockwise**), (IP and *Luna* stirs **her** potion), (IP which she'll do **countrerclockwise**), (IP there will be an explosion).
→ If ⟨Hermione, Luna⟩ stirs her potion, she'll do so ⟨clockwise, countrerclockwise⟩.

¹¹The focus-marking prominence on *hers* is not-at-issue focus, discussed in the next subsection.

Note that the ARCs in the two examples above aren't F-items in CT+F coordinated structures; (19) is an instance of additive coordination (which can be highlighted by adding an overt additive particle *too*), and in (20) the F-items are the two *her* pronouns with distinct antecedents. In fact, true appositives can't be F-items in a CT+F configuration at all. ARCs, and appositives more generally, are packaged into their own IPs (see Selkirk 2005 and references therein). Trying to impose CT+F prosody on the strings in (19) would change prosodic grouping, resulting in a restrictive RC interpretation (which in this case will also require tense changes in the RCs):

- (21) (IP If *Flitwick* brings a dog which is **small**), (IP and *Hagrid* brings a dog which is **large**), (IP they'll fight).¹²
 → If ⟨*Flitwick*, *Hagrid*⟩ brings a dog, it will be a ⟨small, large⟩ one.

Note that something similar happens with post-posed nominal adjuncts; (22a) has two bona fide non-restrictive nominal appositives, while in (22b) the adjuncts are interpreted as restrictive modifiers and require existence of two salient Fangs:

- (22) a. (IP If *Filch* brings **Mrs Norris**), (IP a **small** beast), (IP and *Hagrid* brings **Fang**),
 (IP a **large** beast), (IP they'll fight).¹³
 b. (IP If *Flitwick* brings Fang the **small** beast), (IP and *Hagrid* brings Fang the **large** beast), (IP they'll fight).

While it might be tempting to say that the restrictive interpretations of the adjuncts in (21) and (22b) are instances of local accommodation of appositives, the prosodic, morphosyntactic, and lexical changes alone suggest different structures to begin with. In fact, as I will show in the next section, local accommodation of adnominal appositives would yield different readings from those of restrictive modifiers, but is for the most part impossible.

Interestingly enough, similar not-at-issue focus can be argued to be possible for what Ebert and Schlenker call *post-speech gestures*, i.e., gestural adjuncts that follow the verbal expressions they adjoin to (even though the utterance below is quite awkward to pronounce, presumably due to the fact that post-speech gestures like to be utterance-final):

- (23) (IP If **Flitwick** brings a dog) — (IP **SIMALL**), (IP and **Hagrid** brings a dog) — (IP **LARGE**), (IP they'll fight).
 → If ⟨*Flitwick*'s, *Hagrid*'s⟩ brings a dog, it will be ⟨small, large⟩.

- (24) (IP If *Hermione* stirs **her** potion) — (IP **CLOCKWISE**), (IP and *Luna* stirs **her** potion)
 — (IP **COUNTER-CLOCKWISE**), (IP there will be an explosion).
 → If ⟨*Hermione*, *Luna*⟩ stirs her potion, she'll do so ⟨clockwise, counterclockwise⟩.

Ebert has at different points argued that post-speech gestures can have at-issue (Ebert and Ebert 2014) or “parenthetical” (Ebert 2017) semantics.¹⁴ Schlenker (2017) claims that post-speech

¹²Some people disprefer *which* as a relativizer in restrictive RCs and need to also replace *which* with *that* here.

¹³Mrs Norris is a cat.

¹⁴It's not entirely clear to me how “parenthetical” post-speech gestures are different from “supplemental” co-

gestures are “supplements”, which is compatible with them being more like appositives or more like parentheticals. For the purposes of this paper these differences likely don’t matter. What matters is that with regard to the possibility of bearing not-at-issue focus, not-at-issue post-speech gestures¹⁵ pattern with other prosodically independent not-at-issue content.

A natural question at this point is if there are counterparts of (21) and (22b) for post-speech gestures, i.e., if we can have an example with gestures linearly following the verbal expressions they adjoin to but still in the same IP as those verbal expressions. The answer seems to be ‘no’. For example, the following string is very hard to pronounce:

- (25) ??(IP If *Flitwick* brings a dog — **SMALL**), (IP and *Hagrid* brings a dog — **LARGE**), (IP they’ll fight).

There seems to be an articulatory constraint on gestures, which requires that once they share an IP with some verbal content they should be anchored to some vocal prosodic event, such as pitch accents in languages like English¹⁶, i.e., they have to be linearized as what we have been so far calling *co-speech gestures*. In other words, there are two major articulatorily non-taxing possibilities for alignment of gestures: as prosodically independent items in their own IPs or as prosodically dependent items within verbal IPs.

Now, going back to not-at-issue focus, it is easy to see that the interpretations of the sentences in (19)/(23) and (20) are essentially the same as in (12) and (13) (repeated below), respectively:

- (12) If **Flitwick** brings a dog^{SMALL}, and **Hagrid**(, too,) brings a dog^{LARGE}, they’ll fight.
- (13) If *Hermione* stirs **her** potion^{CLOCKWISE},
and *Luna* stirs **her** potion^{COUNTERCLOCKWISE}, there will be an explosion.

The difference, however, is in how prominence is marked. Intuitively, it’s OK to mark prominence on contrastive ARCs or post-speech gestures, since they are in their own IPs. Trying to mark co-speech gestures as prominent in (12) or (13) is awkward, possibly because it creates a confusion/garden-pathing environment regarding the structure, since the prosodic differences between CT+F and additive coordination in this case are very subtle. However, some speakers accept not-at-issue focus on the second word-gesture pair only in (12) and (13).

Now, what happens to presupposition triggers? One observation is that two different lexical items with the same (purported) at-issue content and different (purported) presuppositional content can’t participate in additive coordination, regardless of how prominence is marked:

speech gestures, considering that Potts calls both appositives and parentheticals supplements (Potts 2005, p. 6).

¹⁵Certainly not all gestures that follow some verbal material are not-at-issue. Some of them can be independent standalone utterances that have ordinary at-issue semantics. Some can have specification at-issue uses similar to those of *one* appositives (discussed, for example, in Nouwen 2014):

- (i) Bring me a beer, a small one.
(ii) Bring me a beer — SMALL.

Yet, the gestures in (23) and (24) seem to be making a contribution similar to that of ordinary appositives.

¹⁶E.g., Loehr (2004) shows that apexes of gesture strokes tend to align with pitch accents in English.

- (26) *If Ron thinks that his parents are in danger, and Hermione, too, knows that hers are, they'll talk to Dumbledore.

This is presumably because the additive presupposition of *too* can't ignore sublexical not-at-issue content¹⁷, i.e., *too* in (26) triggers the presupposition that Ron knows that his parents are in danger, not that he thinks that they are, which is not satisfied in the context. The presupposition of *too* can ignore lexically independent not-at-issue content such as gestures (whether co- or post-speech) and appositives, which allows them to participate in additive coordination. A similar generalization applies to how different types of not-at-issue content behave under ellipsis. Note that these facts require an additional explanation under Schlenker's presuppositional analysis of co-speech gestures, but this discussion is well beyond the scope of this paper.

That said, some people can to some extent accept the following utterance, without an additive particle and with not-at-issue focus on *knows* only (although the judgements are hard):

- (27) *Context: McGonagall doesn't know if Ron's parents are in danger, but she knows that Hermione's are; she doesn't know what Ron and Hermione think; she says:*
If Ron thinks that his parents are in danger, and Hermione **knows** that **hers** are, they'll talk to Dumbledore.

The take-home message of this and the previous subsection is that a major factor in how focus interacts with a given type of not-at-issue content is whether that content is prosodically independent. Co-speech gestures are prosodically anchored to the verbal content they share an IP with, and lexical presuppositions are a sublexical component of a lexical item that also contains some at-issue content¹⁸; as a result, it's hard for them to bear not-at-issue focus, but they can serve as F-items in CT+F coordinated structures (and sometimes they have to). Appositives and post-speech gestures necessarily occupy their own IPs; as a result, they can bear not-at-issue focus but can't be F-items in CT+F coordinated structures. Additionally, there might be subtler differences between lexically dependent (presuppositions) and lexically independent (gestures) not-at-issue content regarding not-at-issue focus, but the data are somewhat messy at this point. In the next subsection I will show that the empirical picture is even more complicated, since some not-at-issue content can't be locally accommodated even as a last resort.

3.4. When local accommodation is impossible

Let's start with an observation that appositives adjoining to nominals don't have the semantics of restrictive modifiers but instead contribute a proposition about the DP they associate with:

- (28) a. (IP Hagrid brought his dog), (IP who is large).
≈ Hagrid brought his dog, and, by the way, his dog is large.

¹⁷A notorious exception are gender features, which have often been given a presuppositional analysis but can be famously ignored under ellipsis and in additive coordination; I won't have much to say about this.

¹⁸Except triggers like *too* and *again*, which arguably only contribute presuppositional content. Interestingly enough, these triggers are typically considered "strong" in the sense that they don't easily allow for local accommodation in the first place, regardless of focus placement. Why this is so is beyond the scope of this paper.

- b. (IP Hagrid brought his dog), (IP a large beast).
 \approx Hagrid brought his dog, and, by the way, his dog is a large beast.

To know if this content can be at-issue, we want to see if those propositions can be treated as maximally local conjuncts. As it happens, they typically can't¹⁹, not even as a last resort:

- (29) *Context: Hermione knows that Hagrid has a single dog, but she doesn't know how big that dog is. Hagrid is planning to bring his dog to the Yule ball. Hermione says:*
- a. #(IP If Hagrid brings his dog), (IP who is **small**), (IP it's gonna be OK), (IP but if he brings his dog), (IP who is **large**), (IP it's gonna be a mess).
 - b. #(IP If Hagrid brings his dog), (IP a **small** beast), (IP it's gonna be OK), (IP but if he brings his dog), (IP a **large** beast), (IP it's gonna be a mess).

No antecedent in (29) can have a reading along the lines of ‘If Hagrid brings his dog and his dog is (a) {small, large} (beast)...’, even though that would have made the sentences meaningful. Surprisingly enough, such propositional readings cannot be accommodated for gestures either, regardless of whether they are linearized as co-speech or post-speech:

- (30) *Same context as in (29).*
- a. #(IP If Hagrid brings his dog^{SMALL}), (IP it's gonna be OK), (IP but if he brings his dog^{LARGE}), (IP it's gonna be a mess).
 - b. #(IP If Hagrid brings his dog) — (IP **SMALL**), (IP it's gonna be OK), (IP but if he brings his dog) — (IP **LARGE**), (IP it's gonna be a mess).

Note that the co-speech gestures in (30a) could get a restrictive modifier interpretation in a different context (one in which Hagrid has at least two dogs, one small and one large), but the post-speech gestures in (30b) can't (for reasons discussed in the previous subsection).

So, to sum up the data on when local accommodation is possible and when it isn't: adnominal appositives and post-speech gestures always have propositional semantics and cannot be locally accommodated; adnominal co-speech gestures can be accommodated when they have predicative semantics but not when they have propositional semantics.

A natural question is if similar restrictions apply to lexical presuppositions, i.e., if only predicative, but not propositional presuppositional content can be locally accommodated. Since we can typically only speculate about the exact form of the presuppositional content in any given case, it is hard to talk about its semantic type, and I won't attempt to do so here. That said, the analysis I sketch in section 4.2 suggests that it is not the propositional type of a given piece of not-at-issue content that makes it unaccommodatable, but rather the mismatch between the type of that piece of content and the type of the phrase it adjoins to in the narrow syntax.

¹⁹Schlenker (2013) discusses some apparent exceptions for ARCs:
(iv) If tomorrow I call the Chair, who in turn calls the Dean, then we will be in deep trouble.
 $\not\rightarrow$ If tomorrow I call the Chair, they will call the Dean.
 \approx If tomorrow I call the Chair, and they call the Dean...

Such examples routinely involve a description of a sequence of events; since they would be hard to, if not impossible, to replicate with gestures, I do not discuss them in this paper.

3.5. Summary of the data

Table 1 summarizes the data discussed in this section (with some simplifications).

Table 1: Different types of not-at-issue content: summary

content type	structural properties	semantic type	at-issue interpretations	not-at-issue focus
appositives	prosodically independent; lexically independent	propositional	impossible	possible
post-speech gestures	prosodically independent; lexically independent	propositional	impossible	possible
co-speech gestures	prosodically dependent; lexically independent	predicative or propositional	can be forced for predicative gestures as a last resort; impossible for propositional gestures	hard
lexical pre-suppositions	prosodically dependent; lexically dependent	??	can be forced as a last resort (for some triggers)	hard

In the next section I will discuss what issues these data raise for Ebert’s and Schlenker’s analyses of gestures and sketch an alternative analysis that avoids these issues.

4. Analyses of gestures

4.1. Issues for Ebert’s and Schlenker’s analyses

When it comes to post-speech gestures, as I have said before, the differences between Ebert’s parenthetical and Schlenker’s supplemental analyses (if any) likely don’t matter for the data at hand. As far as both analyses predict that post-speech gestures (at least those that we have looked at in this paper) cannot be locally accommodated, these data pose no problems for them.

Ebert’s (2017) analysis of co-speech gestures is two-fold. She claims that co-speech gestures in general have supplemental semantics akin to that of appositives, but she also allows for NP-level gestures with “exemplification” semantics. In other words, under her view, for example, *a dog*^{LARGE} typically has the same semantics as *a dog*, (IP which is/will be large), but — if my interpretation of her claims is correct — the gesture LARGE can also sometimes be interpreted as indicating the size of a typical entity in the denotation of the predicate *dog*.

As things stand, it is unclear if this view can predict the restrictive modifier interpretations of predicative co-speech gestures under pressure (as in (5), (7), (9), or (10)). As we have seen in the previous section, appositives only have propositional semantics, so even if they were capable of being locally accommodated (which they don't seem to be), the predicted readings wouldn't be those of restrictive modifiers. As for "exemplification" interpretations, it is unclear to me if they can be at-issue and what the predicted result would be, if they can.

Schlenker (2017) argues that co-speech gestures trigger assertion-dependent presuppositions he calls *cosuppositions*. A gestural cosupposition has the form $V \Rightarrow G$, where V is the verbal expression the gesture adjoins to, G is the gesture's content, and \Rightarrow is generalized entailment. When this cosupposition projects, the local context c' of $[[V]G]$ has to entail it: $c' \Rightarrow (V \Rightarrow G)$. When it is locally accommodated, it is conjoined to V : $V \& (V \Rightarrow G)$, which is equivalent to $V \& G$, where $\&$ is generalized conjunction. Note that for Schlenker's cosuppositional mechanism to apply, it is crucial that the denotations of V , G , and c' are all of the same type.

Without going into technical details, given certain assumptions about how local contexts are computed, Schlenker's cosuppositional analysis yields correct predictions for adnominal *et*-type co-speech gestures when they adjoin to *et* expressions, i.e., NPs, both for projection and local accommodation. For example, the following results obtain for *Hagrid brings his dog^{LARGE}*, if the gesture adjoins to the NP *dog*: if the cosupposition projects, we get a conditional inference, roughly, 'If Hagrid brings his dog, his dog is large' (from which it is easy to generalize to 'Hagrid's dog is large' tout court); if the cosupposition gets accommodated we get the at-issue content of the sentence to be 'Hagrid brings his large dog', which is exactly what we want.

Things become more complicated when gestures adjoin to expressions of type $\langle et, t \rangle$ or *e*, i.e., DPs. If all co-speech gestures trigger presuppositions, which seems to be Schlenker's claim, we have to assume that DP-level co-speech gestures are of the same type as the DP they adjoin to. Depending on our further assumptions about the denotation of the gesture and the local context, the cosuppositional mechanism can yield similar results for DP-level attachment of the gesture as for NP-level attachment when it comes to projection. However, as things stand, we will also predict local accommodation to be possible in this case. The exact result for local accommodation will depend on what denotation we assume for the DP-level gesture. For example, if we assume that a DP-level gesture *LARGE* denotes an existential quantifier 'a large object', we get the at-issue content of the sentence to be 'Hagrid brings his dog and a large object'. If we introduce an anaphoric link across the verbal expression *his dog* and the gesture *LARGE*, we can get the at-issue content of the sentence to be 'Hagrid brings [his dog]_i and it_i is large'. Neither is attested.

More generally, since the only attested at-issue interpretation of adnominal gestures is that of restrictive *et* modifiers, we should find a way to block any other at-issue interpretations. If one wants to maintain Schlenker's claim that all co-speech gestures trigger cosuppositions, they would have to stipulate either that adnominal co-speech gestures can only attach to *et* expressions, or that only *et* adnominal co-speech gestures can be locally accommodated. Neither option seems to be well-motivated. Furthermore, under such an approach, the fact that both post-speech gestures and non-predicative co-speech gestures can't be locally accommodated

seems entirely accidental. In the next subsection I will sketch an analysis of gestures, focusing specifically on adnominal gestures, that doesn't assume that the linearization of a gesture directly determines its semantics. Instead, I will propose that the semantics and projection properties of a gesture are determined by its level of attachment in the syntax, which will also restrict its linearization possibilities.

4.2. Proposal: syntax/semantics and syntax/prosody of gestures

I'm going to assume that all NP-level adjuncts, gestural or not, denote predicates of *et* type (which is not particularly controversial), and all DP-level adjuncts, gestural or not, denote propositions containing a pronoun anaphoric to the DP the adjunct merges with in the narrow syntax. This latter assumption is quite natural to make for ARCs and nominal appositives (which for our purposes can be just reduced ARCs), and I am generalizing it to DP-level gestures. For example, an NP-level gesture *LARGE* denotes $\lambda x.\text{large}(x)$, but a DP-level gesture *LARGE* denotes $\text{large}(x)$, where x is anaphoric to the DP the gesture adjoins to.

Now, Schlenker's cosuppositional mechanism can apply to NP-level gestures, because they have the same type as the verbal expression they modify. When the cosupposition projects, we get a conditional, assertion-dependent inference; when it is accommodated — in particular, to satisfy the contrast requirements under focus — the gesture behaves as a restrictive modifier.

However, the cosuppositional mechanism can't apply to DP-level gestures, since they are not of the same type as the DP they adjoin to. More generally, DP-level adjuncts can't be interpreted where they merge in the narrow syntax. Instead, they have to raise at LF and adjoin at some sentential level. As a first approximation, they adjoin at the highest possible level at which the discourse referent introduced by the DP they originally merged with is still available for them. Further assumptions about the status of not-at-issue DP-level adjuncts might be needed to derive their projection behavior when they don't end up having matrix scope even after raising. For example, Schlenker's (2013) semantic *translucency* of appositives will do the job (Schlenker's treatment of appositives is in general very much in line with the story I've been developing here, modulo some differences in syntactic assumptions). Regardless of those further assumptions, however, this general approach makes the mechanism of local accommodation, i.e., conjunction at the level of "triggering", inapplicable to DP-level adjuncts.

This story gives us a principled reason why different mechanisms apply to NP- vs. DP-level gestures, with potentially different results when it comes to at-issue interpretations. But how do we explain why post-speech adnominal gestures seem to be incapable of having predicative semantics and, subsequently, at-issue interpretations? I propose that that's because NP-level gestures can't be linearized as post-speech due to articulatory and prosody/syntax constraints. There are two conflicting requirements that ensure this result. On the one hand, NP-level adjuncts want to be in the same IP as the NPs they adjoin to (we have seen it, for example, for restrictive RCs and postposed restrictive nominal modifiers). However, as we have seen in section 3.3 (example (25) and discussion thereof), once a gesture is in an IP with some at-issue content, it can't be prosodically independent due to articulatory reasons, i.e., it has to be

co-speech. These two requirements can be formulated as OT-style constraints:

- (31) ANCHORG: Assign * for each gesture that is inside an IP containing verbal content but is not anchored to any vocal prosodic event.
- (32) WRAPNP (a narrow version of Truckenbrodt 1999's WRAPXP constraint): Assign * for each IP boundary inside an NP.

The constraints above are in principle violable, but since there is always a better candidate, with the NP-level gesture linearized as co-speech, post-speech NP-level gestures shouldn't emerge:

	[D [NP GESTURE]]	ANCHORG	WRAPNP
(33)	a. (IP...D NP ^{GESTURE} ...)		
	b. (IP...D NP — GESTURE...)	*!	
	c. (IP...D NP) — (IP GESTURE)		*!

Since DP-level gestures packaged into their own IPs don't violate WRAPNP, DP-level gestures can be linearized either as co-speech or post-speech:

	[[D NP] GESTURE]	ANCHORG	WRAPNP
(34)	a. (IP...D NP ^{GESTURE} ...)		
	b. (IP...D NP — GESTURE...)	*!	
	c. (IP...D NP) — (IP GESTURE)		

At this point I'm not committing to any specific constraint-based theory (even though the tableaux above are done in the style of the classical OT for simplicity). However, if we want to capture both variation and gradience in judgements, which is especially pertinent when dealing with gestures, theories that place constraints on a numerical scale, such as stochastic OT (Boersma 1997 et seq.), or have weighted constraints, such as Harmonic Grammar and variations thereof (Legendre et al. 1990 et seq.), might be better suited than the classical OT (Prince and Smolensky 1993/2004).

5. Conclusion

In this paper I have looked at how different types of not-at-issue content interact with contrastive focus. In particular, I have tried to address the following questions:

1. When do focus-related considerations force at-issue interpretations of typically not-at-issue content?
2. Can a given type of content bear the so-called not-at-issue contrastive focus (i.e., focus that marks contrast without addressing the QUD)?
3. When are at-issue interpretations of a given type of content impossible, even when that would be the only way to satisfy contrast requirements?

Regarding question 1, I have looked at cases when the semantically focused element has to be

addressing the immediate QUD and thus has to be at-issue, as is the case for F-items in a CT+F configuration. I have shown that at-issue interpretations of typically not-at-issue content are forced in this case only if it's necessary to make the F-items properly contrastive across the conjuncts. These considerations apply to prosodically dependent not-at-issue content only, namely, co-speech gestures and lexical presuppositions. Prosodically independent content, such as appositives and post-speech gestures, can't be an F-item in a CT+F configuration.

Regarding question 2, I have observed that it is much easier for prosodically independent not-at-issue content to bear not-at-issue contrastive focus. There also seem to be further subtle differences between lexically independent (co-speech gestures) and lexically dependent (presuppositions) content with respect to not-at-issue focus, which need to be investigated further.

The answers to the first two questions group together lexical presuppositions and co-speech gestures on the one hand, and appositives and post-speech gestures on the other, which goes in line with Schlenker's (2017) analysis of gestures (contra Ebert and Ebert 2014, Ebert 2017), even though it's unclear whether this patterning reveals anything about the semantics of the types of content at hand rather than the role of their structural properties.

The answer to the third question, however, emphasizes the role of the said structural properties for at least appositives and gestures, since only those adnominal adjuncts that match the phrase they adjoin to in semantic type (in particular, predicative co-speech gestures) can have at-issue interpretations under pressure. Adnominal appositives, post-speech gestures, and propositional co-speech gestures can't have at-issue interpretations even under pressure. Taking this observation as a pivotal point, I have sketched an analysis of adnominal gestures whereby NP-level gestures are predicative and thus can be locally accommodated by conjoining with the NP they adjoin to, but DP-level gestures are propositional, like appositives, and thus can't be locally accommodated. I have further proposed that DP-level gestures can be linearized as either co- or post-speech, but NP-level gestures can only be linearized as co-speech, due to articulatory and prosody/syntax constraints. A natural next step is to extend this approach to adverbial gestures.

References

- Abusch, D. (2002). Lexical alternatives as a source of pragmatic presuppositions. In B. Jackson (Ed.), *Proceedings of Semantics and Linguistic Theory 12*, pp. 1–19. Ithaca: Cornell University.
- Beckman, M. E. and G. Ayers (1997). Guidelines for ToBI labelling. version 3.0. The Ohio State University Research Foundation.
- Boersma, P. (1997). How we learn variation, optionality, and probability. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam 21*, pp. 43–58.
- Büring, D. (2003). On D-trees, beans, and B-accents. *Linguistics and philosophy* 26(5), 511–545.
- Büring, D. (2016). (Contrastive) topic. In C. Féry and I. Shin (Eds.), *Handbook of Information Structure*, pp. 64–85. Oxford University Press.
- Ebert, C. (2017). Co-speech vs. post-speech gestures. Talk given at *Language and cognition workshop in memory of Peter Bosch*, Cognitive Science Department, University of Os-

nabrück, February 2017.

- Ebert, C. and C. Ebert (2014). Gestures, demonstratives, and the attributive/referential distinction. Talk given at *SPE 7*, Berlin, June 28.
- Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.
- Heim, I. (1983). On the projection problem for presuppositions. In M. Barlow, D. Flickinger, and M. Wescoat (Eds.), *Second West Coast Conference on Formal Linguistics (WCFFL)*, pp. 114–125.
- Katzir, R. (2014). On the roles of markedness and contradiction in the use of alternatives. In S. Pistoia-Reda (Ed.), *Pragmatics, semantics and the case of scalar implicatures*, pp. 40–71. Palgrave Macmillan.
- Legendre, G., Y. Miyata, and P. Smolensky (1990). Can connectionism contribute to syntax? Harmonic grammar, with an application. In M. Ziolkowski, M. Noske, and K. Deaton (Eds.), *Proceedings of the 26th regional meeting of the Chicago Linguistic Society*, pp. 237–252. Chicago: Chicago Linguistic Society.
- Loehr, D. P. (2004). *Gesture and intonation*. Ph. D. thesis, Georgetown University.
- Nouwen, R. (2014). A note on the projection of appositives. In E. McCready, K. Yabushita, and K. Yoshimoto (Eds.), *Formal Approaches to Semantics and Pragmatics*, pp. 205–222. Springer.
- Potts, C. (2005). *The logic of conventional implicatures*. Oxford University Press.
- Prince, A. and P. Smolensky (1993/2004). *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA and Oxford, UK: Blackwell.
- Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6), 1–69.
- Schlenker, P. (2009). Local contexts. *Semantics and Pragmatics* 2(3), 178.
- Schlenker, P. (2013). Supplements without bidimensionalism. Ms., Institut Jean-Nicod and New York University.
- Schlenker, P. (2017). Gesture projection and cosuppositions. *Linguistics and Philosophy*. To appear.
- Selkirk, E. (2005). Comments on intonational phrasing in English. In S. Frota, M. Vigrio, and M. J. Freitas (Eds.), *Prosodies: With special reference to Iberian languages*, pp. 11–58.
- Simons, M., D. Beaver, C. Roberts, and J. Tonhauser (2017). The best question: Explaining the projection behavior of factives. *Discourse processes* 54(3), 187–206.
- Syrett, K. and T. Koev (2014). Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics* 32(3), 525–577.
- Tieu, L., R. Pasternak, P. Schlenker, and E. Chemla (2017a). Co-speech gesture projection: Evidence from inferential judgments. Ms.
- Tieu, L., R. Pasternak, P. Schlenker, and E. Chemla (2017b). Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa* 2(1), 102.
- Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic inquiry* 30(2), 219–255.
- van Rooij, R. (2010). Topic, focus, and exhaustive interpretation. In C. Lee (Ed.), *Proceedings of the 18th International Congress of Linguists*. Seoul: Korea University.
- Wiegand, M. (2018). ‘I just know it’: Intensification as evidence for non-presuppositional factivity. Talk given at the 92nd Annual Meeting of the LSA, Salt Lake City.