Towards a uniform super-linguistic theory of projection*

Maria Esipova

Princeton University, Princeton, NJ, USA
`mesipova@princeton.edu`

Abstract

Formal semantics/pragmatics has recently seen an upsurge in studying secondary modality content like gestures and facial expressions. Much of this work has focused on how such content projects from under semantic operators, but most of it has assumed that projection of secondary modality content is governed by independent rules. I argue that projection behavior of compositionally integrated content is guided by the same linguistic principles in all modalities, but to see that we need to treat secondary modality morphemes as linguistic objects across the board. I apply this approach to two novel case studies: conventionalized co-speech gestures and degree modifiers in various modalities.

1   Introduction

The question of how gestures accompanying speech contribute meaning has lately been gaining traction in formal semantics/pragmatics (Lascarides & Stone 2009; Ebert & Ebert 2014; Hunter 2018; Schlenker 2018a, a.o.). This line of inquiry has been extended to other types of secondary modality content, e.g., facial expressions and suprasegmental morphemes, in Schlenker 2018b. Much of this work has focused on how contributions of non-conventionalized, iconic gestures, such as in (1), project from under various semantic operators, such as negation, *if*, modals, etc.[1]

(1)   a.   Lea might bring her dog$_{\text{LARGE}}$.
            $\rightarrow$ Lea's dog is large.

       b.   Zoe might shoot at the target$_{\text{LONGBOW}}$.
            $\rightarrow$ If Zoe shoots, she'll shoot a longbow.

Most of this literature has assumed that projection patterns of secondary modality content are determined by modality-specific rules. 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 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's "syntactically eliminable". Thus, *co-speech gestures* like in (1) co-occur with speech, a more primary modality, and are "external", as they can be omitted. Such content is predicted by Schlenker's typology to trigger "cosuppositions" (assertion-dependent presuppositions).

[1] I write gesture labels in ALL CAPS. Underlining approximately indicates temporal alignment of the gesture without making any syntactic claims. Illustrations are sometimes added at the approximate onset of the gesture.

While Schlenker's (2018b) paper contains some extremely valuable empirical observations, the typology proposed there is not based on linguistically meaningful notions or principles. For one thing, 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".

Next, 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 both verbs and their arguments) and segmental vs. suprasegmental morphemes (with the latter exemplified by vowel lengthening, discussed in subsection 4.3).

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 empirical phenomena. However, when we do manage to do so, we start seeing this typology make the wrong predictions, as I will show in this paper. I will also show that approaching secondary modality content as bona fide linguistic objects at all levels of representation reveals that there is no need for a typology like Schlenker's, since secondary modality expressions fit into the same typology of projection patterns as primary modality expressions. I will ground this super-*linguistic* approach in the theoretical assumptions from Esipova 2019a,b, summarized in section 2. I will then discuss two new case studies supporting this approach over Schlenker's: conventionalized gestures (section 3) and degree modifiers in various modalities (section 4). Section 5 concludes. Audios and videos of selected examples from this paper are available at: `https://osf.io/5gu4d/`.

2   Background

One important distinction I'll be making is between MODIFIERS and SUPPLEMENTS. Modifiers, exemplified by attributive adjectives in (2), are, roughly, expressions of type $\langle \tau, \tau \rangle$. SUBSECTIVE MODIFIERS exhibit SUBSECTIVE ENTAILMENT, i.e., they combine with an expression α yielding β that entails α via generalized entailment. For example, in (2a), *blond* and *skillful* are subsective modifiers. NON-SUBSECTIVE MODIFIERS, like *alleged* in (2b), don't exhibit subsective entailment.

(2)   a.   Zoe is a {blond, skillful} stuntwoman.
            → Zoe is a stuntwoman.
      b.   Daisy is an alleged criminal.
            ↛ Daisy is a criminal.

Supplements, exemplified in (3) by appositives (an appositive relative clause and a nominal appositive), combine with an expression, pass it unchanged for subsequent composition, but also yield a proposition of a special kind about it. I'll remain agnostic about how exactly the special status of this proposition is operationalized (see, e.g., Potts 2005; Koev 2013).

(3)   I will invite Zoe, (who is) a stuntwoman.

Specific instances of subsective modifiers can be RESTRICTING, when they are used to pick out

a smaller part of the denotation of the expression they are combining with, as in (4a), or NON-RESTRICTING. The latter are truth-conditionally vacuous modifiers that are used to add some extra information about the denotation of the expression they are combining with, as in (4b).[2]

(4) a. I should be eating less **saturated** fats and more **healthy** fats.
         $\not\to$ I should be eating less fats and more fats.
         $\not\to$ All fats are {saturated, healthy}.
     b. *Context: The speaker believes that processed meat causes cancer.*
         I shouldn't be eating so many deadly hot dogs.
         $\to$ I shouldn't be eating so many hot dogs.
         $\to$ All hot dogs are deadly.

Non-restricting modifiers give rise to inferences about the truth-conditional equivalence of the expression they modify and the result of modification; in (4b), it is the inference that being a hot dog is equivalent to being a deadly hot dog. Crucially, as noted in Esipova 2019a,b as a refinement of the original definition of non-restricting modifiers in Leffel 2014, this equivalence is assessed relative to the local context of the modified expression, as shown in (5).

(5) *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.*
     If processed meat causes cancer, I shouldn't be eating so many deadly hot dogs.
     $\not\to$ All hot dogs are deadly.
     $\to$ If processed meat causes cancer, all hot dogs are deadly.

In Esipova 2019a,b, I argue that these non-restricting modifier inferences arise pragmatically when the addressee recognizes the speaker's intention to use a modifier as a truth-conditionally vacuous one. Various pragmatic factors affect whether a given modifier instance is restricting (e.g., adjectives with an attitudinal meaning component like *lovely* are often non-restricting).

Whichever specific analysis one assumes for supplements, it's an empirical fact that, in contrast to modifiers, supplements can never be restricting, as shown in (6), and project very strongly, as shown in (7), in a way that suggests conventional rather than pragmatic triggering.

(6) #($_{IP}$ I invited Zoe), ($_{IP}$ (who is) a **stuntwoman**), ($_{IP}$ not Zoe), ($_{IP}$ (who is) a **politician**).[3]

(7) #I don't know if Zoe is a stuntwoman, but if you invite Zoe, (who is) a stuntwoman, you should show her your muscle car.
     Intended: '...if (you invite Zoe and she is a stuntwoman)...'

The exceptional cases in which appositives do interact with semantic operators, noted first in Schlenker 2013, are subject to strict discourse constraints (Jasinskaja & Poschmann 2018), which is distinct from the default availability of restricting uses for subsective modifiers.

For a more detailed and technical discussion of modifiers and supplements, the viewer is referred to Esipova 2019a,b. For the purposes of this paper, the discussion above will suffice.

Next, I assume an inverted Y model of grammar whereby the syntax creates hierarchical, non-linearized structures of abstract objects (labels, features, etc.), which get shipped off to the compositional semantics to be interpreted and to the phonology to eventually create pronounceable sequences of (possibly overlapping) elements. Vocabulary insertion and linearization

---

[2]I use **bold** to indicate prosodic (contrastive) focus marking; for English spoken material, this is usually an (L+)H* pitch accent on the stressed syllable of the bolded word in MAE-ToBI terms (Beckman & Ayers 1997).

[3]IP stands for 'Intonational Phrase' in MAE-ToBI terms; the appositives in (6) are not to be confused with prosodically integrated nominal modifiers as in *($_{IP}$ I invited Zoe the stuntwoman)*.

happen 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- or linearization-specific specific effects emerge in the phonology broadly construed (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.

The modifier vs. supplement distinction is first and foremost a compositional one, 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 gestures that are linearized in a certain way uniformly project as supplements or as non-restricting modifiers, which unsubstantiates the earlier debate about gesture projection in Ebert & Ebert 2014 vs. Schlenker 2018a.[4]

In Esipova 2019a,b, I showed that non-conventionalized gestures like those in (1) should be analyzed under the assumptions above. Thus, $LARGE$ can be construed of as iconically representing the property of being large and compose as a modifier (like the adjective *large*), giving rise to a projecting inference only if non-restricting. Or it can be construed of as iconically representing a large object, composing and projecting as a supplement (like the nominal appositive *a large animal*). I furthermore maintain that co-speech modifier gestures prefer to be non-restricting for pragmatic and possibly prosodic reasons, although restricting interpretations, as in (8), are in principle available for them, which is supported by experimental results.

(8)   *Context: We are going on a group tour and need to rent a van. Stephanie, who has a Pug and a Great Dane, is planning to bring one of her dogs with her.*
        ?%Do you know which one of Stephanie's dogs is coming with us? 'Cause if she's bringing her her **dog**$_{SMALL}$, we'll be fine, but if she's bringing her her **dog**$_{LARGE}$, we should get a bigger van.   (video available)

However, non-conventionalized gestures like in (1) 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 modifiers, and independent post-syntactic considerations make it hard for modifier gestures to have their own time slot in English (as I argue is the case in Esipova 2019b), it is easy to convince oneself that gestures project in a uniform way depending on their linearization. Once, however, we start expanding our empirical scope to include more conventionalized secondary modality content, we start seeing even more clearly that the projection behavior of a given piece of content can't possibly be determined by the parameters assumed in Schlenker's typology.

3   Conventionalized gestures

Some conventionalized gestures can be subsective modifiers, too, and, thus, exhibit the projection behavior pattern proper to modifiers. For instance, in both examples in (9), 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 (10). $DRINK$ 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).

---

[4]As argued in Esipova 2019a,b, Schlenker's cosuppositions are non-restricting modifier inferences in disguise.

(9) a. If you bring a semanticist$_{\text{CRAZY}}$ to my talk, I'll likely fight with them.
         → If you bring a semanticist to my talk, I'll likely fight with them.
         → All semanticists are crazy.

     b. If we wanna celebrate$_{\text{DRINK}}$ my defense, we better go to a store now.
         → If we wanna celebrate my defense, we better go to a store now.
         → If we celebrate my defense, we'll do so by drinking alcohol.

(10) a. ?If Kim brings her **brother**$_{\text{CRAZY}}$, I'll fight with him, but if she brings her **normal** brother, that's OK.[5]
          ≈ If Kim brings her **crazy** brother...

      b. ?If we wanna **celebrate**$_{\text{DRINK}}$ my defense, we better go to a store now, but, of course, we can also celebrate **without** alcohol.
          ≈ If we wanna celebrate my defense by drinking **alcohol**...

Schlenker's typology could treat these gestures like those in (1) (with the same objections as those raised in Esipova 2019a,b). However, other conventionalized gestures seem to be non-subsective modifiers, e.g., *AIR-QUOTES* in (11a) is akin to *so-called* or *quote-unquote*.[6]

(11) a. Kim is bringing her friend$_{\text{AIR-QUOTES}}$ to the party.
      b. Kim is bringing her {so-called, quote-unquote} friend to the party.

However we go about describing and analyzing projection behavior of non-subsective modifiers, Schlenker's cosuppositions (which are, once again, just non-restricting modifier inferences) aren't the way to go, because if *friend$_{AIR\text{-}QUOTES}$* doesn't even entail *friend*, the two certainly can't be truth-conditionally equivalent.

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*.[7]

(12) a. If a friend of mine wins the race$_{\text{FINGERS-CROSSED}}$, I'll buy them a drink.
      b. If a friend of mine, {hopefully, fingers crossed}, wins the race, I'll buy them a drink.
          ✓ → I want a friend of mine to win.
          ✗ ≈ If (a friend of mine wins and I wanted them to)...

Again, Schlenker's typology makes the wrong predictions for *FINGERS-CROSSED* in (12): it's a co-speech gesture, but doesn't trigger a cosupposition; instead, it projects conventionally, like a supplement. But if we exclude all conventionalized gestures from a typology of projection

---

[5] I apologize for the potentially sanist nature of this example.

[6] I thank Patrick Grosz (p.c.) for drawing my attention to both *AIR-QUOTES* and *FINGERS-CROSSED*.

[7] 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:

(i) That was Clive's friend in Vice, who I've enlisted in my search for that Beanpole Bob guy and potentially a zombie cure$_{\text{FINGERS-CROSSED}}$.   ('iZombie', S05E08, video available)

patterns that includes gestures, we would be missing the parallel between gestures like *LARGE* and gestures like *CRAZY*. Also, this would require one to take a stance on what's conventionalized and what isn't, treating it 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 necessary, as gestures fit into the same typology of projection patterns as spoken content.

4   Degree modifiers cross-modally

4.1   Degree modifiers in the primary modality

Open-scale[8] 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:

(13)   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.
       b.   A: Which of her dogs is Lea bringing?
             B: The {lovely, disgusting, #fucking, #bloody, #damn} one.

Two further relevant cases of degree modification in the primary modality are (i) modifier repetition, as in (14), which has a gradient iconic effect, and (ii) so-called "contrastive reduplication", whereby a string of the form $\boldsymbol{x}\,x$ gets interpreted along the lines of '{true, proper, prototypical} x', as in (15), which also has iconic roots. Both are truth-conditionally non-vacuous as well.

(14)   a.   You're really, really sick.   (Ghomeshi et al. 2004, fn. 3)
       b.   You are a sick, sick man.   (Ghomeshi et al. 2004, fn. 3)
       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.

(15)   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.   Lea doesn't have a chihuahua, she has a **dog**-dog.

It is not obvious to me how exactly the gradient mapping between the form and meaning should be operationalized in cases of modifier repetition like in (14). It is possible, however, that each occurrence of the modifier is interpreted independently, as a further restriction on the input. Compare the effect obtained in (14) to that of multiple degree modifiers:

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

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 and will focus on a perhaps more straight-forward phenomenon of contrastive reduplication. In line with my architectural assumptions, I maintain that contrastive

---

[8]I'm making this caveat here, because the maximal-degree modifier 'completely' has been argued to project by default in spoken Italian and Italian Sign Language in Aristodemo 2017.

reduplication expones a DegP head, CRED, whose (coarse-grained) semantics is given in (17). I mostly assume the degree semantics from Kennedy & McNally 2005, except, following Esipova 2019b, I separate out the existential closure that they build into their lexical entries for degree modifiers so as to make sure that degree modifiers are in fact modifiers (i.e., expressions of type $\langle \tau, \tau \rangle$) and to allow for multiple degree modification like in (16). The prototype relation in (17) requires that the degree d is within the prototypical range for the input expression.

(17) $[\![\text{CRED}]\!]([\![\alpha_{\langle d, \langle \tau_1 \ldots \tau_n, st \rangle \rangle}]\!]) = \lambda d \lambda X_{\tau_1}^1 \ldots X_{\tau_n}^n \lambda w. [\![\alpha]\!](d)(X^1) \ldots (X^n)(w) \wedge \text{prototype}_{[\![\alpha]\!]}(d)(w)$

Inherently scalar predicates like *sick* combine with CRED directly, as in (18).

(18)   a. $[\![sick]\!] = \lambda d \lambda x \lambda w. \text{sick}(x)(w) = d$
         b. $[\![\text{CRED}]\!]([\![sick]\!]) = \lambda d \lambda x \lambda w. \text{sick}(x)(w) = d \wedge \text{prototype}_{\text{sick}}(d)(w)$

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:

(19)   a. $[\![\text{SCALAR}]\!]([\![\alpha_{\langle \tau_1 \ldots \tau_n, st \rangle}]\!]) = \lambda d \lambda X_{\tau_1}^1 \ldots X_{\tau_n}^n \lambda w. \text{scale}_{[\![\alpha]\!]}(X^1) \ldots (X^n)(w) = d$
         b. $[\![dog]\!] = \lambda x \lambda w. \text{dog}(x)(w)$
         c. $[\![\text{SCALAR}]\!]([\![dog]\!]) = \lambda d \lambda x \lambda w. \text{scale}_{\text{dog}}(x)(w) = d$

The d variable is existentially closed off after all degree modifiers have composed.
    With these tools at our disposal, let us look at secondary modality degree modifiers.

4.2   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:[9]

(20)   a.   Yesterday there was a party, and, [⟨surprisingly, impressively⟩, [Mia got drunk]].
              ≈ It is ⟨surprising, impressive⟩ that Mia got drunk.
         b.   Yesterday there was a party, and Mia got [⟨surprisingly, impressively⟩ [drunk]].
              ≈ Mia got drunk to a(n) ⟨surprising, impressive⟩ extent.

(21)   a.   Yesterday there was a party, and [[Mia got drunk]^OO].
              ≈ It is surprising that Mia got drunk.
         b.   Yesterday, there was a party, and Mia got [[drunk]^OO].
              ≈ Mia got drunk to a {surprising, high} extent.

Now, more needs to be said eventually about the interaction of temporal alignment of *OO* and its syntactic construals. As pointed out by Patrick Grosz (p.c.), in its supplement use, *OO* is focus-sensitive, just like *surprisingly* and its kin, which, I believe, does affect the docking of *OO* (see some preliminary observations in Esipova 2019c). For the purposes of this paper, however, what matters is that both interpretations are in principle available for a co-speech *OO*.
    The degree modifying *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 (22), where *OO* co-occurs with the Russian conventionalized gesture *DRINK* from the previous section (which here is a predicative adjective morphosyntactically, showing that the gesture itself is a root).

---

[9]Overlining approximately indicates temporal alignment for facial expressions.

(22) a. Yesterday, there was a party, $\overline{\text{and [[Mia got \overline{\text{DRUNK}}]}^{\text{O-O}}]}$.   (video available)
         ≈ It is surprising that Mia got drunk.
     b. Yesterday there was a party, and Mia got [[$\overline{\text{DRUNK}}$]$^{\text{O-O}}$].   (video available)
         ≈ 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 *surprisingly* and the like):[10]

(23) a. $\overline{\text{When [[a friend of mine gets \overline{\text{DRUNK}}]}^{\text{O-O}}]}$, I sometimes comment on that.
         → When a friend of mine gets drunk, that is surprising.
     b. When, a friend of mine gets [[$\overline{\text{DRUNK}}$]$^{\text{O-O}}$], I sometimes 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,
         $\overline{\text{but when [[a friend of mine gets DRUNK]}^{\text{O-O}}]}$, I sometimes 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/gesture facial expressions.[11] 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 patterns: a propositional *OO* patterns with propositional spoken adverbs, and a degree modifier *OO* patterns with other degree modifiers, whatever the reason for their persistently restricting nature.

4.3   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,[12] but it seems that it can involve an L*+H pitch accent, creaky phonation, lengthening of the accented syllable (not just the vowel, the onset lengthens, too, if it can), and higher intensity (audio available).

Like CRED, *DEG-INT* is morphosyntactically promiscuous and restricting by default:

(24) a. If the movie is good$_{\text{DEG-INT}}$, I'll stay till the end of the credits.   ≈ *very good*
         ↛ If the movie is good, I'll stay till the end of the credits.
     b. Lea has a dog$_{\text{DEG-INT}}$.   ≈ *big dog* or *proper dog*
     c. When I saw that snake, I ran$_{\text{DEG-INT}}$.   ≈ *ran fast* or *ran properly*

DEG-INT can be given the same semantics as CRED or a version thereof. I think there are some discrepancies between the two, but leave a proper exploration of these for future research.

Schlenker (2018b) discusses what he calls "iconic vowel lengthening" in (25).

(25)   If the talk is loooong, I'll leave before the end.
         ↛ If the talk is long, I'll leave before the end.

---

[10]A reviewer asks if the attitudinal component of a degree modifier *OO* can project. The judgements are subtle, but, if need be, we can separate the 'high degree' and 'that degree is surprising' components for *OO* or its spoken kin and have the latter project (cf. items like *fucking*, which can simultaneously make a truth-conditional contribution as degree modifiers and perform their expressive function as emotional outlets; Esipova 2019c).

[11]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 evaluative adjectives like *disgusting*, whose beginnings can be seen in Esipova 2019c.

[12]I am currently working with Zoe Kahana and Reis White on establishing these properties and whether they are distinct from those of the modified focus marking accompanying purely mirative, non-degree-modifier *OO*.

Schlenker maintains that 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 entirely clear to me in what sense such suprasegmental events 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 (25) can contain extra segment lengthening beyond that associated with *DEG-INT*. However, this extra lengthening in (25) isn't obligatory to convey the high degree reading, and a more general phenomenon is at play, as shown in (24). 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.

(26)   It was {slow$_{\text{DEG-INT}}$, slooow$_{\text{DEG-INT}}$, fast$_{\text{DEG-INT}}$, #faaast$_{\text{DEG-INT}}$}.

There might still be some remnant effects of iconicity in *DEG-INT*, however, due to the low pitch (coming from the L* target, which tends to be extra-low in *DEG-INT*) and the overall syllable lengthening, constraining the distribution of this morpheme. How we should operationalize all these iconicity effects architecturally is an extremely interesting question, which I leave for future research. What is clear, however, is that this problem needs a properly linguistic approach.

Another question that I will leave for future research is how to treat simultaneous combinations of various degree modifiers in a principled way. For instance, (27) 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.

(27)   *Context: The speaker is giving advice on how to build pec muscles.*
        You won't have a chest that looks like this$_{\text{TOUCHES-ANOTHER'S-CHEST}}$, you'll actually have a chest$_{\text{DEG-INT}}$$_{\text{LARGE-ROUND}}$.   ('Athlean-X' YouTube channel, video available)

One 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).[13] 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.

5   Conclusion

In this paper, I hoped to show that by putting *linguistics* into *super linguistics*, i.e., by treating secondary modality content such as gestures, facial expressions, and intonational morphemes as linguistic objects at all levels of representation, we gain predictive and explanatory power. In particular, this approach reveals parallels in projection behavior between primary and secondary modality content that void the need for modality-specific typologies of projection patterns. I believe other classifications of meaning-bearing expressions should be cross-modal, too. For example, I develop the beginnings of a cross-modal typology of attitudinal and expressive content in Esipova 2019c and establish the need for a cross-modal typology of expressions that integrate with the host utterance at some level(s) of representation, but not compositionally (e.g., pure expressives or some morphemes with purely social functions) in Esipova To appear.

---

[13]I thank Yining Nie (p.c.) for pointing out this literature to me.

References

Aristodemo, Vita Maria Valentina. 2017. *Gradable constructions in Italian Sign Language*: Institut Jean-Nicod dissertation.

Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Stefan Hofstetter, Christiane Savelsberg, John Vanderelst & Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9(1). 1–66. doi:10.1075/livy.9.01bec.

Beckman, Mary E. & Gayle Ayers. 1997. Guidelines for ToBI labelling. Version 3.0. The Ohio State University Research Foundation.

Bochnak, M. Ryan. 2015. The degree semantics parameter and cross-linguistic variation. *Semantics and Pragmatics* 8(6). 1–48. doi:10.3765/sp.8.6.

Ebert, Cornelia & Christian Ebert. 2014. Gestures, demonstratives, and the attributive/referential distinction. Talk given at *Semantics and Philosophy in Europe (SPE 7)*, Berlin.

Ernst, Thomas Boyden. 1984. *Towards an integrated theory of adverb position in English*: Indiana University dissertation.

Esipova, Maria. 2019a. Composition and projection in co-speech gestures. To appear in *Proceedings of Semantics and Linguistic Theory (SALT) 29*.

Esipova, Maria. 2019b. *Composition and projection in speech and gesture*: New York University dissertation.

Esipova, Maria. 2019c. Emotion and composition. Talk given at *SURGE*, Rutgers University.

Esipova, Maria. To appear. Polar responses in Russian across modalities and across interfaces. *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 28*.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory* 22(2). 307–357. doi:10.1023/B:NALA.0000015789.98638.f9.

Hunter, Julie. 2018. Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics and Philosophy* 1–16. doi:10.1007/s10988-018-9244-0.

Jasinskaja, Katja & Claudia Poschmann. 2018. Attachment in syntax and discourse: towards an explanation for the variable scope of non-restrictive relative clauses. In Sireemas Maspong, Brynhildur Stefánsdóttir, Katherine Blake & Forrest Davis (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 28*, 433–453. doi:10.3765/salt.v28i0.4438.

Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381. doi:10.1353/lan.2005.0071.

Koev, Todor. 2013. *Apposition and the structure of discourse*: Rutgers University dissertation.

Lascarides, Alex & Matthew Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26(4). 393–449. doi:10.1093/jos/ffp004.

Leffel, Timothy. 2014. *The semantics of modification: Adjectives, nouns, and order*: New York University dissertation.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199273829.001.0001.

Price, Patti J, Mari Ostendorf, Stefanie Shattuck-Hufnagel & Cynthia Fong. 1991. The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America* 90(6). 2956–2970. doi:10.1121/1.401770.

Schlenker, Philippe. 2013. Supplements without bidimensionalism. Ms., Institut Jean-Nicod and New York University.

Schlenker, Philippe. 2018a. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. doi:10.1007/s10988-017-9225-8.

Schlenker, Philippe. 2018b. Iconic pragmatics. *Natural Language & Linguistic Theory* 36(3). 877–936. doi:10.1007/s11049-017-9392-x.