Verbal Symbols and Demonstrations Across Modalities

1 Motivations for a New Ontology

In this short paper, I introduce a new theory of the ingredients of semantic composition in the verb phrase that I recently proposed in Ramchand (2018), and argue that although it was proposed in that work to deal with a central problem in the cartographic layering of meaning found in all of the world's languages, it turns out to be well suited to creating a common system for integrating symbolic content in human languages across modalities. Specifically, the design features of the solution proposed allow the simultaneous integration of both descriptive and demonstrative content (in the sense of Davidson 2015), and can be adapted to both onomatopæic and co-gestural effects in spoken languages. The system also allows one to formulate explicit hypotheses concerning the commonalities and differences between languages in the spoken and signed modalities respectively.

In this introduction, I first lay out the cartographic and psycholinguistics motivations for the proposal made in Ramchand (2018). The second section presents the technical aspects of the proposal in detail and shows how it solves the problems raised in the introduction. The third section applies the framework to the integration problem for parallel modalities, formulating a technical vocabulary for combining parallel demonstrative and descriptive content, and integrating it with the extensional semantic system. The final section discusses the implications that this kind of system has for the compositional semantics of signed languages and what similarities and differences it predicts with respect to spoken languages.

1.1 Cartographic Generalizations

Cartography tells us that there are robust crosslinguistic generalizations about the ordering of meaning elements in an extended functional projection. Specifically, in the verbal domain we have robust evidence for a universal functional sequence, or extended projection rooted in the lexical verb ((1)).

(1) $\text{Mood}_{speechact} > \text{Mood}_{evid} > \text{Mod}_{epist} > \text{T} > \text{Mod}_{circ} > \text{Asp} > \text{Voice} > \text{Cause} > \text{V}$
        Cinque (1999), Grimshaw (1991), Julien (2000)

Even those who are uncomfortable with the universalist claims coupled with fine grained 'cartography' , nevertheless subscribe to the C > T > V template of extended verbal projections and language specific rigid ordering (see also Ramchand and Svenonius 2014).

Moreover, at the bottom of every functional sequence, we find evidence for a kind of substantive, conceptual, rich, yet flexible kind of meaning, as denoted by open class items. This is often expressed in theories as a difference between lexical (open class) and functional (closed class) items, and extended projections in the sense of Grimshaw (1990) are all stated with the lexical formatives at the base. Evidence for this kind of *layered meaning* is pervasive and exceptionless crosslinguistically, and yet it currently looks 'accidental' or 'templatic' from the point of view of our formal ontologies. While syntacticians perhaps implicitly believe that a deeper semantic understanding will eventually release them from templatic formulations, it is clear to the formal semanticists that their own ontologies and primitives do not provide a ready made solution for meaning layering. In fact, morphosyntactic layering is seen by semanticists as a *syntactic* fact, something that their formalism can track or follow, but not something that will be explained by that formalism.

The semanticists are right. They use extensional formal ontologies where situations themselves, or referents, are fully specified particulars, and there is no semantic reason why some features of a situational particular should be linguistically privileged over any other. Situations are particulars that exist in worlds and at times. Events are minimal subparts of situations (the kinds expressed by vPs, according to Kratzer 2014), but are full real world particulars, objects in the extensional world. The fact that events are particulars with both causal and temporal properties for example, means that these two pieces of information are not distinguished in a principled way. So

the fact that causational information occurs hierarchically lower in the verbal extended projection than temporal information in actual languages is not a property of the situation per se, for those properties are not intrinsically ordered as far as the real world particular is concerned. This makes the internal structuring of propositions a templatic matter for the semanticists as well.

Looking more closely at the verbal extended projection we see that over and over in the literature on the interface between syntax and semantics, we find evidence for the existence of *event kinds* which are built up prior to the information about a particular situation in the world. For example, Gehrke (2013), Gehrke (2015) and subsequent work has argued from the interpretation of adjectival passive participles for the existence of event kinds (see also Gehrke and McNally 2015) as a conceptual primitive, parallelling the conceptual primacy of kinds in the nominal domain (Carlson (1977), Zamparelli (2000), Chierchia (1998), Hinzen 2016)

In fact, the formal semanticist does have the formal tools in place to talk about the abstract concepts of 'kinds' in both the nominal and verbal domains. In both these cases (nominal kinds and event kinds) they are built by generalizing over particulars, with the aid of the device of possible worlds (Lewis 1986). In fact, with the exception of Gehrke and collaborators, the work cited in these domains all use some kind of modal/possible worlds apparatus to formally describe kind level meanings, and systematically represent kinds as being built out of particulars.[1]

If semantics is to follow morphosyntax in such a way that what is simple and underived in the one system corresponds to what is simple and underived in the other, then *essence* must precede *existence* in the cumulative building up of a natural language proposition. On the other hand, the functional sequence tells us that there is some notion of kind/property that resides lower down, close to the root and is the basis for the build up of reference to partic-

---

[1]Event properties, in the work of McNally and Gehrke (Gehrke and McNally 2015, Gehrke 2015, Grimm and McNally 2015), is an attempt to introduce the kind or essence of the event (cf. Fine 2005) as a primitive of the ontology. These authors are very clear that they think of the notion of event essence as *preceding* information about instantiation in the verbal functional sequence. But there are severe problems with formalizing essences directly, in a semantics grounded in particulars. In particular, details of compositionality within the properties domain itself are difficult to make precise. In what follows, I will propose a different strategy, although I am sympathetic to the general research strand, and to the philosophical position of Kit Fine on these matters.

ulars. So In morphosyntax, particulars are built out of essences, exactly the inverse of the ontological commitments of the extensional semantic model. If we believe that semantic theories constitute a symbolic system for representing the meaning of syntactic structures, and that there is no necessity for the ingredients or ordering of the primes of that symbolic representation to track either the syntactic primes, or the primes of actual cognitive composition in human brains, then there is no need to be disturbed by this fact.

But if possible worlds can in principle be invoked this low down, within root meanings, and across syntactic categories, then what is the source for the layering pattern found in the morphosyntax of the verbal extended projection where modality is universally expressed high, quite far from the verbal root? It cannot come from the semantics per se, but from some other factor.

1.2   Internalism meets Externalist Semantics

I will argue that the other factor we need to consider is what might be thought of as a 'third factor' property related to the nature of language itself. To see how this plays out, we need to revisit another old debate concerning internalist vs. the externalist which is directly reponsible for the inverse ontological commitments of the semanticist and the syntactician.

The externalist argues that there is a crucial 'aboutness' to language, and that if we attempt to ground our theories in internalist notions then we are condemned to theories that make no sense of the intersubjectivity of language and which end up being at best unfalsifiable, and mystical at worst. The internalist on the other hand is trying to account for the fact that language is represented in the mind-brain of actual individual speakers. We would like to know what precisely it is that they have memorized, and how it is deployed in a particular linguistic context.

With respect to our internal representations, I think that Chomsky is entirely right when he points out that mental phenomena, and the meanings of our natural language symbols, do not seem to be explicable in purely extensional terms, not even with the help of Lewisian possible worlds. Citing Nagel 1993, Chomsky argues that:

> ' It is a hopeless task to "complete the materialist world picture" by translating accounts of "mental phenomena" in terms of a "description that is either explicitly physical or uses only terms that can apply to what is entirely physical" or perhaps give "as-

sertability conditions" on "externally observable grounds". (Nagel 1993. pg.37)
Chomsky 1995. pg 4 '

Instead, Chomsky advocates a more naturalistic enquiry into the internal products of a natural language system, which will be consistent with the position I will take in what follows.

> " Neurologist Rodolfo Llinás (1987) puts the matter well when describes perception as "a dream modulated by sensory input", the mind being a "computational state of the brain generated by the interaction between the external world and an internal set of reference frames" . But the internal frames that shape the dreams are far more intricate and intriguing than often assumed, even at the level of the lexicon, still more so when we turn to expressions formed by the computational processes."
> Chomsky (1995), pg 23

First of all, I will continue to assume a prominent role for situations as real world particulars in the system I envisage, and still maintain the externalist agenda for semantics in the sense of building a system that makes sense of language's relation to the world. Rather than make propositions denote truth values directly, I will adopt a kind of truthmaker semantics for situations (Barwise and Perry 1983, Kratzer 2014, Fine 2014) such that "Something on the side of the world— in this case, a state of affairs— verifies something on the side of language.". The verbal extended projection builds up a description of a full situational particular, which is then verified, or not, by the current state of affairs.

According to Fine "Truthmaking is a not a guide to metaphysics." . . . "But if our aim is to understand language, then our focus should be on immediate truth makers, not the ultimate truthmakers, and the question of *how* they make the statements of the language true will be of greatest concern." In fact, I will argue that in language, the verbal extended projection structures the situational description in a very particular way, which has implications for the immediate truth makers we have to assume in natural language and how they function.

In order to integrate a truthmaker situation semantics with the internalist agenda, we need to focus on the role of the symbolic primes of a natural language system and what they need to be able to do. Barwise and Perry

(1983) point out that, in addition to the 'aboutness' and potential 'intersubjectivity' of language, we need to account for the *reusability* and *perspectival relativity* of the symbolic prime.

2   Utterances as Generalized Demonstrations

In order to make sense of both internalist concerns and meaning layering, we can take our inspiration from an apparently extreme and exotic case in natural language semantics, that of the 'ideophone'. As Henderson (2015) puts it, work on the formal semantics of ideophones is scarce because of the 'difficulty in formalizing the distinction between descriptive meaning and depictive meaning, which ideophones seem to traffic in'. In giving his own account, Henderson explores a formal foundation for the notion of *demonstrations* from Davidson (2015) and extends it to account for the ideophonic data. Intuitively, demonstrations are a special type of communicative event that stand in a similarity relation with the event demonstrated.

The radical move I will make here is to take the lesson of ideophones as telling us something about the general case. In other words, I will propose that *all* linguistic symbol use should be implemented in terms of a demonstration event, not just obviously ideophonic elements.

The preview of the system I will propose goes like this. For human language to get off the ground, we need to have (i) common possession of symbols that are *abstractions over the different actual situations encountered in the learning phase*, and (ii) a speaker to deploy those symbols as a means of characterizing new situations in the world as she comes across them. Symbols do not denote a new essential 'sort' but denote artifacts of the language itself. These in turn are partial descriptions of events (worldly particulars), but because it is the symbol that is being combined and manipulated, no unwanted existential implications (in this, or any other possible world) are generated.

In order to do this we need to add to the usual model, a domain D_μ which is the domain of well-formed linguistic entities of type μ, after Potts (2007). These linguistic objects are triples, consisting of a < `phonological string`, **syntactic features**, SEMANTICS>.

Full expressions of type μ will be written in sans serif . So for example, the verb run might have the denotation:

$[[\ \mathsf{run}\ ]] = <\ \mathtt{run},\ <\ \mathbf{init,\ proc}\ >\ ,\ \lambda e[\mathrm{run}(e)]\ >$

For convenience, we adopt the convention in Henderson (2015) which uses the bottom corner notation to pick out the semantic part of the triple denoted by something of type μ. Thus,

$\llcorner\ \mathsf{run}\ \lrcorner = \lambda e[\mathrm{run}(e)\ ]$

Next, we need to see that the 'meanings' of the symbols themselves need to be devoid of temporal or worldly information. They form the hierarchically inner core which is then clothed with the contingent information of time, place and world, to link descriptions to actual particulars. The symbols of a person's language are acquired through actual experience of the world, both sensory and cognitive, but they are, importantly, abstractions over the particular instantiations exposed to. This abstraction process is probably partly mechanical but definitely automatic, and certainly filtered through our general cognitive predilections concerning what we consider worthy of attending to. The act of abstracting away from the inessential in a situation to form a representation of the properties that are to be encoded by the symbol being learned is a complicated question way beyond the remit of this paper. However, one thing is clear—- this process is a natural and immediate one and takes place (perhaps imperfectly) on the basis of one or two exposures. The result (which may be fine tuned over the life of the speaker) is a partial description of essential properties for the event that, on grounds of reusability, are *necessarily* devoid of information related to the particularity of the instantiation (time, place and the reference of the participants). The remarkable thing about possessing an inventory of symbols, which are then consciously deployed in multiple situations, is that this must go hand in hand with the instinct to abstract away from the particular to represent the essential, or reusable content of that sign. It is this third factor property of the symbolic system itself that leads to temporal and spatial, and referential information being represented morphosyntactically external to the memorized sign.

> **Lexical Symbols as Elements of the Ontology**
> (a) Symbols of the language constitute the domain D_μ, which are triples consisting of a $<$ `phonological string`,**syntactic features**, SEMANTICS$>$
> (b) The semantics of a verbal LI are partial descriptions based on sensory and cognitive abstractions over experience.
> (c) The syntactic part of the information in a triple that is a member of D_μ, is a subtree of the language. The merge of $u_1 \in D_\mu$ and $u_2 \in D_\mu$, creates a derived element of D_μ, u_3, which has the syntactic representation built by merging the syn-rep of u_1 with the syn-rep of u_2, and a semantics is composed by ordinary argument identification of ⌞ u_1 ⌟ and ⌞ u_1 ⌟.

After the symbolic layer (possibly complex) has been construced, it is necessary to convert these partial descriptions (properties of symbols) into a property of events. This is done by means of an explicit deployment operation. Let us take a look at how Henderson (2015) makes the conversion in his work on ideophones. The conversion makes use of a demonstration event d and a predicate TH$_\delta$ which combines with demonstration events to give the 'theme' of that event, namely the use of a particular linguistic object u. The DEMO predicate relates d and a general event variable e, saying that d 'demonstrates' or has certain structural properties in common with e.

(2)   QUOTE : $\lambda u \lambda d \lambda e [\text{TH}_\delta(d) = u \wedge \text{DEMO}(d,e)]$

In my own proposal, I will carry over this QUOTE function to the general case and propose that the reusable essential symbolic content of any perfectly ordinary sign is the equivalent of Henderson's ideophone. In other words, a symbol is a conventionalized ideophone used to invoke and describe an event. In order to make this fully general, we need to replace Henderson's DEMO predicate with what I will call CONVEY, and replace the exotic idea of a demonstration event with the ordinary garden variety UTTERANCE act itself. I replace the δ subscript on th with α since we are no longer talking specifically of demonstrations. Thus, Ramchandian 'quotation', built at the level of EvtP in this system, will have the following general form (3).

> **Deployment of Symbolic Content at EvtP**
> I. EvtP : $\lambda d\lambda e[\text{UTTERANCE}(d) \wedge \text{TH}_\alpha(d)=u \wedge \text{CONVEY}(d,e)]$
> Property of of an UTTERANCE event d and event e, which has u as its theme, and where d deploys u ($\in D_\mu$) to convey e.
> II. In the case of purely conventional (i.e. non-depictive) LIs, uttered with sincerity and without metaphor or hyperbole,
> '$\text{TH}_\alpha(d)=u \wedge \text{CONVEY}(d,e)$' $\longrightarrow$ '$\llcorner u \lrcorner (e)$'

The second part of the definition says that in the case of purely conventional (non-ideophonic or iconic signs), the CONVEY operation by an UTTERANCE simply means that the speaker intends that the semantic partial description connected to the symbol does indeed apply to the event in question.

This is a representational encoding of the intuition that reference involves a speaker and a context in addition to the symbol she is deploying. But it is not just a matter of a speaker X using the symbol Y to refer to the object Z, we need to leave room also for the contextual circumstances and mode of deployment of the symbol in question. Once again Chomsky (1995) puts it well,

> " More generally, person X uses expression E with its intrinsic semantic properties to talk about the world from certain intricate perspectives, focusing attention on specific aspects of its, under circumstances C, with the "locality of content" they induce (in Bilgrami's sense). "
> Chomsky (1995), p. 43

We will see further in section 3 that in order to accommodate non conventional and iconic aspects of meaning, we need to state a number of systematic conventions attaching to the deployment event itself. But before we turn to a more explicit treatment of these interesting cases, there is one further aspect of the system that needs to be put in place.

Recall that the one of the motivations for creating this lower symbolic zone in the VP was the fact of morphosyntactic layering across language, and the existence of symbolic/essential content at the core of all actual linguistic compositionality. Once the symbolic layer has been 'deployed' by the speaker, the complex representation so formed is now ready to be directly modified to received temporal and spatial information related to the actual

instantiation of the event/situation being described. In my own implementation, I follow Champollion (Champollion 2015) in going for a quantificational event semantics, where the event variable itself is existentially bound quite low (at the edge of the VP) and where what is carried up for modification is a variable over properties of events. This is independently necessary, as Champollion argues, because the event variable always scope underneath other quantificational elements. To illustrate, here is Champollion's denotation for the verb phrase *see Mary*.

(3) $[[\text{ see Mary }]] = \lambda f \, \exists e [\text{see}(e) \wedge f(e) \wedge \text{th}(e) = \text{Mary}]$

The verb phrase now denotes a property of event properties, a move that is required to allow further properties of the event to be added after existential closure. The locus of closure of the event variable converges in my own implementation with the introduction of a variable representing *spatiotemporal/worldly properties* of events which I propose is a complex property relating d and e. Put another way, f is a property of events e, *anchored in* d. Thus, the AspP built up by my quotational quantificational system, will therefore look as in (4).

(4) $[[\text{ AspP }]] = \lambda f_{<v,<v.t>>} \lambda d \exists e [\textbf{UTTERANCE}(d) \wedge \llcorner u \lrcorner (e) \wedge f(d)(e)]$

So at the level of AspP we have a property of Relations that link the UT-TERANCE context d with an existing event that is being demonstrated/described in d. That event has conceptual/perceptual properties as characterized by ⌞u⌟. At this point temporal information can be added to the event description that was impossible before. The quantificational aspect of the compositional system is important for the system defended and described in Ramchand (2018) and the data analysed there, but will be less important for our purposes in looking at gestural and iconic content in section 3. It is presented here for completeness.

Now, semanticists understand very well the need for incorporating contextual information to build meanings that have actual truth conditions, and the relevance of the UTTERANCE itself (the kaplanian context so to speak) is especially crucial for the conversion of indexical elements with a 'character' to what he calls 'content' which can then be verified, or assigned truth conditions. One way of thinking about the proposal here is that the symbolic zone represents kaplanian 'character' (most obviously different from 'content' in the case of so called indexical elements), but which is now the definition of

10

the whole lower symbolic zone.

The semantics that makes crucial reference to UTTERANCE context is not controversial, but reifying the speech event this low down in the structure certainly is. One might wonder why the speech event should be reified in this way at all. Why put it in the representation in this literal fashion, instead of simply invoking it in the model or in the process of interpretation? Relatedly, why represent deployment explicitly and treat language symbols as elements of the ontology in this linguistic representational way?

The full justification for this way of building up the compositional semantics can only be seen in the extent to which it proves a productive and fruitful way of looking at old and new language puzzles and in generating new predictions that can be verified. One such example of productive extension will in fact be explored in section 3. But in addition, as the closing salvo for this section, I offer a few more general and conceptual considerations in favour of the radical rethink I am proposing.

The first conceptual consideration is that by reifying the symbolic zone and making claims about what exactly is being represented there, we get a better understanding of the representation of lexical items, and systematic polysemies and reusabilities that are exploited for lexical symbols. By isolating this area of lexical symbology (in distinction to syntactic features or functional items), we create a more algorithmic theory of how meaning is built up psycholinguistically with claims and predictions that are more commensurable with the primes being investigated by psycholinguists. The underwriting of the natural language generalizations in terms of meaning layering were the primary motivation for the system in the first place and so the fact that it reduces one robust aspect of the V < T < C template to virtual conceptual necessity, given conditions of reusability, is a strong point in its favour.

Intuitively, in Quantificational Quotational Semantics (QQS) we build up a to a representation of a situation in four stages, as shown in the box below.

> **QQS: Layered Semantic Composition in the Clause**
> (i) The putting together of lexical items which encode certain event properties. This stage needs to be productive and compositional, but with no reference to temporal or world parameters. (The Interior)
> (ii) The assertion by the speaker of the existence of an event with those properties. (Deployment)
> (iii) Addition of temporal and world properties to the event. (Referential/Instantiational domain)
> (iv) Anchoring of the worldly and temporal properties via the Origo (the speaker and her contextual coordinates).
> (see also Hinzen and Sheehan 2015)

3 Applying QQS to Demonstrative Content

It is perhaps not surprising that this system is well suited to expressing demonstrative (as opposed to just descriptive) content because it was originally designed to integrate the iconic content of ideophones into the descriptive system. In generalizing the system as above, I am claiming that the ideophones are just a special case of what is always true, namely that a symbol is deployed by a speaker in a context to create a description of a real world event. In the limit, the deployment event is just the act of uttering. But what if the UTTERANCE event is fancier than that? What happens if the UTTERANCE event comes with a particular intonation, co/speech gestures like eye rolls or hand movements? What if the nature of the sign deployed carries presuppositions telling us that the deployment event is iconically similar to the event being evoked? We need to ensure that by generalizing the Henderson/Davidson/Potts treatment of demonstrations to conventional descriptive items, we still leave the flexibility to account for the more demonstrative kind of content.

Because of the reification of the deployment event d, we allow ourselves access to a variable on which to express differences of this kind. The CONVEY predicate was chosen to be maximally general, but we could state principles which choose different flavours of 'conveying' depending on the nature of the symbol involved and how it is being deployed. Let us take a look again at the general schema for denotations of EvtP in QQS, repeated here in (5).

(5)   **EvtP** : $\lambda d\lambda e$[**UTTERANCE**(d) $\wedge$ **TH**$_\alpha$(d)=u $\wedge$ **CONVEY**(d,e)]
Property of of an **UTTERANCE** event d and event e, which has u as its theme, and where d deploys u ($\in$ D$_\mu$) to convey e.

In the boring case (i) the **UTTERANCE** is straightforward and sincere and has no significant other properties and (ii) the symbol deployed carries straightforward conventionalized descriptive content. I claimed in Ramchand (2018) that this case, the formula can be reduced as follows.

(6)   **TH**$_\alpha$(d)=u $\wedge$ **CONVEY**(d,e) $\equiv$ '$\llcorner$ u $\lrcorner$ (e) (pure **DESCRIPTION**)

In this case, we could invent the predicate **DESCRIBE** to link d and e.

(7)   **EvtP** : $\lambda d\lambda e$[**UTTERANCE**(d) $\wedge$ **TH**$_\alpha$(d)=u $\wedge$ **DESCRIBE**(d,e)]
Property of of an **UTTERANCE** event d and event e, which has u as its theme, and where d deploys u ($\in$ D$_\mu$) to describe e.

However, in the case of demonstrations, the symbol in question *carries a crucial presupposition* that the **UTTERANCE** in addition bears a similarity relation to the event being conveyed (in Henderson 2015 's formulation). This similarity can be at the level of sound (onomatopeia) or at the level of spatial iconicity. In this case, the relation between d and e is more of an explicit demonstration of the properties of e. We might represent this more specifically as in (8)

(8)   **EvtP** : $\lambda d\lambda e$[**UTTERANCE**(d) $\wedge$ **TH**$_\alpha$(d)=u $\wedge$ **DEMO**(d,e)]
Property of of an **UTTERANCE** event d and event e, which has u as its theme, and where d deploys u ($\in$ D$_\mu$) to demonstrate e, because d bears some kind of similarity relation to e.

We could further propose more fine grained predicates relating d and e, depending on the nature/modality of the similarity evoked.

(9)   **DEMO** (iconic/auditory): (the symbol used in) d demonstrates or has certain auditory properties in common with e (onomatopeia).
DEMO(iconic/visual): (the symbol used in) d demonstrates or has certain structural properties in common with e (spatial iconicity).

However, I think this move would be a mistake. The problem is that these different cases are not really cleanly separable from each other in practice. Moreover, the degree of iconicity involved seems to be dependent more on the

lexical symbol itself than on the speaker or general contextual or real world knowledge. I propose therefore that the best formal representation uses the general predicate CONVEY which is neutral between the division of labour between demonstration and depiction.

In fact, many instances of CONVEY *combine* different ways of representing the event in question. Reifying the symbol as an element of the ontology and the properties of the deployment event, allows all of these ways of representing the event to be integrated deep inside the compositional semantics. For languages in the auditory modality, onomatopœia is a common phenomenon, but exists alongside descriptive content, like for example the case that here we have a verb that expresses an unaccusative motion event without telos.

(10)   The water whooshed into the room.
        EvtP : $\lambda d \lambda e[$**UTTERANCE**$(d) \wedge$ **TH**$_\alpha(d)=$ whoosh $\wedge$ '$\llcorner$ whoosh $\lrcorner$ (e)'
        $\wedge$ '$\ulcorner$ whoosh $\urcorner$ has properties in common with the sound of e']
        (Where I use top corner brackets by convention to pick out the
        phonological part of the triple corresponding to a symbol of D_μ.)

So, the claim here is that degree of iconicity is not built into the choice of predicate relating d and e, contra Henderson (2015), but to the presuppositional properties of the symbol deployed.

The second way in which the act of deployment can be enriched is when the UTTERANCE event d, is endowed with properties over and above the act of uttering the symbols. This encompasses a range of effects from intonation, to eye-rolling, to co-speech gestures.

I give examples of all three of these situations in turn, with an indication of how I think the information should be represented at EvtP. Sarcastic intonation on the symbol itself is used to express speaker doubt about the appropriateness of the symbol deployed. In the case of descriptive content this amounts to an expression of speaker disbelief in the applicability of the symbol's descriptive content to the event being described.

(11)   *Dialogue 1: Sarcasm*
        Speaker A: What did you have for dinner yesterday?
        Speaker B: I am trying to lose weight, so I had a huge bowl of kale with some lentils mixed in.
        Speaker A: Ooh. That sounds delicious. (where red text indicates sarcasm alert intonation).

EvtP : $\lambda d \lambda e[\textsc{utterance}(d) \wedge \textsc{th}_\alpha(d) = \mathsf{sound\ delicious} \wedge \textsc{speaker-of-d\ does\ not\ believe}(\ '\llcorner\ \mathsf{sound\ delicious}\ \lrcorner\ (e)')\ ]$

When it comes to co-speech gesturing, the demonstrative content of the gesture can be integrated low down in this system by introducing a modifier of d. For consistency, I have added this in (12) by means of another flavour of the **TH** predicate over d, which I will call $\textsc{th}_{gesture}$ and represent the content with a place holder shape symbol $\triangle$. We can state as a meaning postulate that all co-speech gestures explicitly timed by the speaker to coincide with the **UTTERANCE** and integrated with the event d, and hence are used to add 'thematic' properties of d. This means that a unified, but more complex d is now being used to **CONVEY** e, and since this gestural component is not part of the conventional meaning of the symbol, it is necessarily interpreted as demonstrative rather than descriptive. We do not need to change **CONVEY** in such cases because it remains neutral, and we should not change the presuppositions of the symbol being deployed, but because d has acquired extra properties, we get an implication/entailment of similarity between d and the event e being described.

(12) *Dialogue 2: Co-speech Gesture*
   Speaker A: You won't believe what happened when my pipes burst on me last night!
   Speaker B: OMG what was it like?!?
   Speaker A: The water whooshed into the room. (with wave-like hand gesture simultaneous with pronunciation of verb)
   EvtP : $\lambda d \lambda e[\textsc{utterance}(d) \wedge \textsc{th}_{gesture}(d) = \triangle \wedge \textsc{th}_\alpha(d) = \mathsf{whoosh}$
   $\wedge\ '\llcorner\ \mathsf{whoosh}\ \lrcorner\ (e)'\ \wedge\ '\ulcorner\ \mathsf{whoosh}\ \urcorner$ has properties in common with the sound of e' $\wedge$ 'The performance of d has properties in common with the spatial trajectory of e']

In the latter case, we can see that the deployment event d can in principle contain information in addition to the actual symbol being deployed.

4   The Visual Modality

How does this system of semantic composition, with its reification of the symbol and the deployment operator, bear on the issue of sign language

argument structure?

In some sense, the natural hope is that the formal semantics of signed languages is not fundamentally different from the formal semantics of spoken languages. This is broadly true in the present system as well. However, unlike the classical formal semantics for natural language, which has difficulty integrating demonstrative content with descriptive content, the present proposal is automatically set up to do this as a matter of course. The place where signed languages differ from spoken languages lies not in the descriptive content of the symbol (which I assume to have the same or similar properties across both language types), but in the degree and nature of iconicity that gets built in to that content.

Unlike other systems of semantic composition, we do not have to squeeze iconic content in right at the end of the compositional process. Rather, in this system, iconic content is integrated early, at the level of the lexical composition, or in the terms of Ramchand (2008), already within the first phase.

The visual modality for language is probably special for a number of independent cognitive reasons. It is clear even from looking at spoken languages that spatial gestures and representations have strong analogical force when it comes to the specification of descriptive content, even stronger perhaps than in the domain of sound symbolism. Co-speech gestures in spoken language are natural and ubiquitous.[2] How interesting it becomes therefore when the dominant language modality is also gestural. It means that the conventional association of the sign with descriptive content is rather more difficult to separate from demonstrative content in principle.

In case of co-speech gestures discussed above, we were able to invoke two distinct properties or 'thematic contents' for d. In the case of the speech symbol deployed, we said that $\text{TH}_\alpha = u$ indicated that particular symbol had been uttered by the speaker. But in the case of signed languages, how much of the symbol content of the hand shape is conventionalized, and how

---

[2]There is a considerable literature now the meaning contribution of co-speech gesture and its status as implicature/presupposition etc. Important in this regard is the work of Philippe Schlenker and colleagues on cosuppositions (Schlenker 2018, Lyn Tieu and Chemla 2017). The precise nature of the meaning contributed by co-speech gesture is beyond the scope of this particular paper, but in addition it is not clear how the distinctions proposed in the literature change their nature when embedded in the kind of compositional semantic system such as the one I am proposing. A detailed comparison between the present system and the cosuppositional literature will have to wait for future work.

much is iconic, i.e. dependent on the details of the demonstration ?

By way of illustration, consider the hypothetical deployment by a speaker of signed language L_s of the sign $\triangle$ in a language-signing event d. We assume that the conventional symbol $\triangle$ has a particular formal semantic descriptive contribution ($\llcorner \triangle \lrcorner$) and we assume that by using the sign in question, the signer signals that the descriptive content of the sign holds of the event being described. But what about the shape of the sign, or its speed or size? What effect do these have on the meaning of what is said? One natural thought is that the implementational, aspects of spatial signs to the extent that they are under the control of the signer, and known to be so, also convey demonstrative, or iconic content. One might speculate that in the case of signed languages, the demonstrative component might be more non-negotiable, always present, at least for certain symbolic items. In the case of spoken languages, onomatopœia is more pervasive than has been acknowledged in the literature, blended together with descriptive content. Even when a sign is completely conventionalized, the manner of uttering a word can add contentful nuance to its meaning. It seems to me that performance effects on content are pervasive and have only been left out of consideration because so far they have seemed to lie outside the system of formal semantics proper. However, there are enough examples of these sorts of meaning interacting with compositional semantic content (see for example the phenomenon of metalinguistic negation in Horn 1989 as a particularly obvious example), that it seems like it is probably correct to integrate these meanings, and integrate them early.

The difference between signed and spoken languages would therefore be a difference in degree and not a difference in kind. So far, we have been using the subscripts α and to distinguish different kinds of thematic content to the utterance event d. It seems however, that the distinction should not be modality specific, but more to do with whether the thematic content is a conventionalized symbolic contribution, or whether it is more related to the actual performance of d. We might therefore signal the difference between conventional and performance properties of the utterance event by use of the subscripts α and β, where α is convention-related content and β is performance-related content. In that case, the signed utterance would always have both α-content and β-content, the former coming from its conventional meaning and the latter from the properties of how the sign is performed.

(13) EvtP : $\lambda d \lambda e[\textsc{utterance}(d) \wedge \textsc{th}_\alpha(d) = \triangle \wedge \textsc{th}_\beta(d) = \textsc{manner}($

$\triangle$) $\wedge$ '$\llcorner \triangle \lrcorner$ (e)' $\wedge$ 'd has properties in common with the spatial trajectory of e']

Of course, this is only an illustration, and many details remain to be worked out. The point is that the existence of the deployment event in the representation language allows for space where these different contributions to meaning can be represented and integrated low down in the compositional system. The use of deployment means that the different themes of the utterance/deployment event do not incorporate a bias against gestural content, or parallel conventional and demonstrated content. In other words, this kind of system allows us in principle to ask certain questions that we were not able to ask before. Namely, is there a difference between conventional content and demonstrated content when it comes to implicational strength? Is there a difference between auditory and gestural demonstrative content in general? In other words is gestural content more iconic and integrated with the communication event across the board for human language? Or does the integration and strength of demonstrated content depend on the modality of the language in question? These questions must remain for further research.

5   Conclusion

To conclude, I have argued that symbolic self consciousness and the *reusability* of open class items is necessary for a generative meaning engine. Reusability and sensitive to speaker perspective are the crucial factor s that motivated the quotational system in the first place, and implementing those ideas allowed us to make a beginning on explaining certain core cartographic generalizations.

While the outer form of my proposal looks unusual, many of the effects are topological rather than substantive. Most results of formal semantic theory are not lost or disturbed by making this move. The change in ontology here is designed to reflect the primes actually employed in natural language composition to build meaning, and the topological transformation of the composition system is designed to reflect a more algorithmic view of how human minds compose meaning. To summarize the main features of QQS again, I have argued for:

(i) The reification of the symbol as object of the ontology.

(ii) The reification of the relationship between the deployment event (a com-

munication event situated in the here and now of the real world), early in the compositional history of the build up of the proposition.

(iii) The restriction of the symbol's lexical content to cognitive abstractions over time and locational instantiation.

When it comes to the last point in (iii), the lexical descriptive content of a conventional sign is argued to be modality independent, and hence applies to both signed and spoken languages. However, because of the reification of d, the deployment event, we immediately get a big difference between the signed and the spoken because d cannot avoid having spatial properties in the former case, while it cannot avoid having auditory properties in the latter case. Because d, the utterance event, is not exhausted by the deployment of the conventional content of the symbol, room is made for demonstrational content to be built in to d. (Indeed, the original system of Henderson 2015 reified d precisely in order to integrate demonstrational content).

Therefore, as an important side effect of this system, we saw that envisaging different presuppositions for symbolic entities could allow us to accommodate both demonstration and description within the same generalized CONVEY relation. At the same time, allowing more than one type of thematic TH predicate over communication events opened the way for both conventional and performance related content to be integrated. Neither of these possibilities is tied to a particular modality, so that in principle mixed effects in parallel across modalities (eg. co-speech gesture) are also accommodated.

Many issues remain. However, the nature of the system proposed here opens up many new fields of inquiry for the semantics of signed and spoken languages in the realm of verbal meaning.

References

Barwise, Jon, and John Perry. 1983. *Situations and attitudes*. Cambridge, Ma.: MIT press.

Carlson, Gregory Norman. 1977. Reference to kinds in english. PhD diss, University of Massachusetts at Amherst.

Champollion, Lucas. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38: 31–66.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6: 339–405.

Chomsky, Noam. 1995. Language and nature. *Mind* 104 (413): 1–61.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

Davidson, Kathryn. 2015. Quotation, demonstration and iconicity. *Linguistics and Philosophy* 38: 477–520.

Fine, Kit. 2005. *Modality and tense: Philosophical papers*. Oxford University Press.

Fine, Kit. 2014. Truthmaker semantics for intuitionist logic. *Journal of Philosophical Logic* 43 (2): 549–577.

Gehrke, Berit. 2013. Still puzzled by adjectival passives. In *Syntax and its limits*, eds. Cristina Sevdali Raffaella Folli and Robert Truswell, 175–191. Oxford University Press.

Gehrke, Berit. 2015. Adjectival participles, event kind modification: the case of frequency adjectives. *Natural Language and Linguistic Theory* 33: 897–938.

Gehrke, Berit, and Louise McNally. 2015. Distributional modification: the case of frequency adjectives. *Language* 91(4): 837–868.

Grimm, Scott, and Louise McNally. 2015. The *ing* dynasty: Rebuilding the semantics of nominalizations. In *Proceeding of SALT XXV*, eds. Mary Moroney Sarah D'Antonio and Carol Rose Little, 82–102.

Grimshaw, Jane. 1990. *Argument structure. Linguistic inquiry monographs; 18*. Cambridge: MIT Press.

Grimshaw, Jane. 1991. Extended projections Brandeis University.

Henderson, Robert. 2015. Pluractional demonstrations. paper presented at SALT 26.

Hinzen, Wolfram. 2016. On the grammar of referential dependence. *Studies in Logic, Grammar and Rhetoric* 46 (1): 11–33.

Hinzen, Wolfram, and Michelle Sheehan. 2015. *The philosophy of universal grammar*. Oxford University Press.

Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.

Julien, Marit. 2000. Syntactic heads and word formation: A study of verbal inflection. PhD diss, University of Tromsø.

Kratzer, Angelika. 2014. Situations in natural language semantics. In *Stanford encyclopedia of philosophy*, ed. Edward N. Zalta, Vol. 2014 Edition.

Lewis, David K. 1986. *On the plurality of worlds*. Oxford: Blackwell.

Lyn Tieu Robert Pasternak, Philippe Schlenker, and Emmanuel Chemla. 2017. Co-speech gesture projection: Evidence from truth value judgement and picture selection tasks. *Glossa* 2(1) (102): 1–27.

Potts, Christopher. 2007. The dimensions of quotation. In *Direct compositionality*, eds. Chris Barker and Pauline Jacobson, 405–431. Oxford University Press.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon*. Cambridge University Press.

Ramchand, Gillian. 2018. *Situations and syntactic structures: Rethinking auxiliaries and order in english. Linguistic inquiry monograph 77*. MIT Press.

Ramchand, Gillian, and Peter Svenonius. 2014. Deriving the functional hierarchy. *Journal of Language Sciences. Language Sciences*, http://dx.doi.org/10.1016/j.langsci.2014.06.013.

Schlenker, Philippe. 2018. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41 (3): 295–365.

Zamparelli, Roberto. 2000. *Layers in the determiner phrase*. New York: Garland.