Compositionality and iconicity *

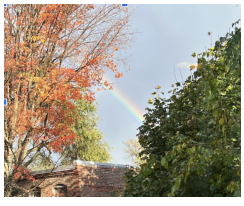
Kathryn Davidson

(Draft: January 8, 2024)

Abstract.

Although language is often taken to be a paradigmatic case of the use of arbitrary symbols to communicate ideas, it is also clear that linguistic communication across all modalities frequently incorporate elements of iconic depiction and non-arbitrariness. How exactly symbolic and iconic aspects of language interact is an area of active research on spoken and signed languages and gesture studies across the cognitive sciences. Here we overview approaches to formally modeling the contribution of iconic and symbolic meaning within natural language. The case is made that while both symbolic and iconic content are pervasive in language, they contribute primarily to two very different kinds of meaning. Propositional meaning is built entirely from symbolic abstractions (which may make reference to external particulars, such as objects, events, and even iconic depictions of them), and can be the input for negation or question formation, which involve reasoning over alternatives, a key feature contributed by abstract symbols. In contrast to symbolic descriptions, iconic depictions are represented as particulars, and must be reanalyzed as symbolic in order to evoke alternatives. In addition, we discuss non-depictive/descriptive iconicity within conventionalized symbolic lexicons, which seems to play important roles in acquisition and/or language processing.

Keywords. Depiction, description, iconicity, gesture, sign languages alternatives, compositionality, semiotics, psycholinguistics

1. Introduction. Language features dominantly in studies of human cognition because it is remarkably precise and creative, able to convey infinitely many new ideas with finite means. However, in this it is not unique: depiction, too, allows for remarkably precise and infinitely creative communications with finite means. So what exactly is the difference in meaning between, say, the sentence *I saw a rainbow yesterday* and the picture  which shows you the

* Thanks to Mark Dingemanse, Masha Esipova, Annemarie Kocab, Josh Martin, and Carla Umbach for comments on earlier drafts, Nozomi Tomita, James Waller, and Jeremy Kuhn for lengthy discussion of depiction and iconicity that influenced this writing, Jessica Tanner for filming ASL sentences, audience members at National University of Singapore, the European Society for Philosophy and Psychology, the joint SPP/ESPP conference in 2023, Susan Carey's "Origins of abstract thought" seminar at Harvard for sources of inspiration, and the editor and three reviewers who provided especially helpful suggestions. This work was supported by an NSF CAREER grant BCS-1844186.

rainbow that I saw? How do we interpret the relation between a picture and its caption, or an illustration in a book of text? And how does this relate to iconic content that seems to be more tightly woven into linguistic expression, such as onomatopeic words like the English *knock*, non-arbitrary signs in sign languages like American Sign Language sign 

, co-speech and co-sign gestures illustrating the shape of a rainbow while talking about it, and the many other places where primarily arbitrary language and iconic forms interact in human language?

This question is particularly timely as several separate strands of research have been increasingly engaged with the (non?)uniqueness of the logical structure of meaning in human language as compared to other aspects of human and non-human cognition. For example, a trend of work in (i) "super-semantics" seeks to extend the methodology and insights of formal semantic models of language to other domains of human activity, such as dance, music, comics, etc. as well as non-human communication (Patel-Grosz et al. 2022). This endeavour is primarily methodological in arguing that formal analyses capture important distinctions in these domains, in this sense sharing insights with structuralist approaches to language from the early 20th century (Saussure 1916), but in many cases it is secondarily arguing that the extention of formal methods to domains beyond language will make more sense to the extent that there are more similarities between the specific structures that we see within natural language meaning and outside of it. At the same time, cognitive scientists have been reviving the question of the (ii) viability of a Language of Thought as a model for cognition outside of natural language semantics, asking which aspects of abstract, compositional language are present in other areas of cognition that seem quite far from language, such as map-based navigation (Quilty-Dunn et al. 2022). Finally, (iii) advances in artificial intelligence via so-called Large Language Models have renewed interest in the ways in which language as a particular activity encodes information about the world and how this contrasts with other means of learning about the world through experience and/or reasoning outside of language, raising important questions about the domain of language as distinct from other ways of thinking and learning (Mahowald et al. 2023). All of these relate directly to the question of similarities and perhaps also differences in the way that we model the kind of meaning expressed in pictures and various forms of depictive communication versus the kind of meaning expressed in abstract symbolic language.

This paper will take the stance that even within human language taken broadly, meaning is not all of the same sort, and that it is necessary to contrast the ways that meaning can be conveyed in symbolic language from the ways that meaning can be conveyed in pictures, even many "pictures" that are very tightly woven into linguistic structure, as in co-speech and co-sign gestures, depictions in speech and sign, and other iconic language. We'll begin by separating, for clarity of argumentation, two different phenomena that have been called "iconic" within language:

iconic depictions, and iconicity in the sense of non-arbitrary mappings between linguistic forms and meanings. This is a distinction that has been made in different ways in different literature, and is perhaps most clearly aligned with the *descriptive/depictive* distinction (Clark 2016) especially as it relates to the way that both kinds of iconicity can be seen (and dissociated) in spoken language ideophones (Dingemanse 2015) and in sign language (Ferrara & Hodge 2018) (and relates to a distinction between object and concept-related interpretation of co-speech gestures by Fricke 2022). We'll argue in this paper that this distinction is a critical one, which those interested in natural language semantics and iconicity should be concerned with when it comes to understanding meaning. The outcome is that, perhaps surprisingly, language itself provides an excellent test case for understanding the similarities and differences between meaning as it is encoded in abstract symbolic structures and meaning as it is found in iconic depictions.

Section 2 will focus on clarifying the difference between depiction and description within language, providing examples from both spoken and sign languages. Section 3 will sketch an approach to formally modeling this distinction, categorizing the content of symbolic language as containing propositions and the content of depictions as identifying particulars. Section 4 will argue that this generates correct predictions when it comes to reasoning over propositional alternatives. Section 5 returns to consider, via contrast with iconic depiction, the notion of "iconicity" within the symbolic lexicon, as a possible faciliator for language acquisition and processing. Section 6 compares the approach taken here to existing approaches to modeling iconicity in both the cognitive, computational, and formal semantic traditions, and Section 7 concludes.

2. Iconicity in Depiction versus Description. In language, the arbitrariness of the sign (Saussure 1916) highlights the independence of some linguistic forms from their meaning, such as the fact that the form of the English word *rainbow* has no particular relation to its meaning. Arbitariness can be contrasted with iconicity, in which forms bear some relation to their meanings. Examples of iconic linguistic forms come in all modalities (including both spoken languages and sign languages), as well as a very wide variety of other modifications (e.g. repetitions, lengthenings) of speech, sign, gesture, etc. in which a form has some non-arbitrary relationship to its meaning.

Although there is recent interest in the extent of iconicity/non-arbitrariness among expressions across the world's spoken languages (Blasi et al. 2016; Winter & Perlman 2021) and a surge of research on both iconic and noniconic aspects of sign languages and gesture (Goldin-Meadow & Brentari 2017), there is perhaps surprisingly much less consideration to the role that iconicity plays in interpreting sentence meaning and how it intersects with the compositional nature of language. For example, if a form A is iconic, and composes with form B, is the iconicity in the resemblance of a form to its meaning interpreted as part of the whole ($A + B$), or is it ignored in the interpretation of the composite? What semantic contributions might an iconic form share, or not, with the majority of the language in both spoken and sign languages that are not so iconic? Within many

cognitive linguistic approaches to meaning, it is often assumed that iconicity is a mechanism for adding content, paired with observations on how iconicity can both support (Taub 2001; Emmorey 2014) and constrain (Meir 2010) non-compositional meaning like metaphors. Within formal semantic approaches to meaning, research on iconic co-speech gestures finds that they differ in their compositional properties from non-iconic words, such as their interaction with negation (Ebert & Ebert 2016; Tieu et al. 2017), but solutions to this puzzle have focused on the issue of gesture and speech existing in separate modalities (Schlenker et al. 2013; Schlenker 2018a; Esipova 2019b; Tieu et al. 2019), working also primarily under the assumption that the iconicity of the content has the same potential for compositionality as non-iconic content. Within formal approaches to sign language linguistics, the question of iconic compositionality arises and the approach is often to analyze it as truth-conditional, essentially using mapping to arrive at the same kind of content available in non-iconic language (Strickland et al. 2015; Schlenker et al. 2013; Kuhn & Aristodemo 2017; Aristodemo & Geraci 2018).

Here, we gain new insights into the question of iconic compositionality from a semiotic perspective by dividing iconic content by what it requires to be interpreted, directly following work on depiction in semiotics and psychology (Clark 2016; Dingemanse 2015; Dingemanse & Akita 2017; Ferrara & Hodge 2018; Kita 1997), consistent with a wide spectrum of philosophical approaches to depiction that separate the meaning conveyed by pictures from the kind found in compositional symbolic language (Peirce 1974; Fodor 2007; Burge 2018; Camp 2018; Block 2023). On the one hand, there are iconic form-meaning pairings which are also symbols that derive meaning via convention: these include onomatopeia like *chirp* in English, expressive sensory ideophones like *gelegele* 'shiny' in Siwu (Dingemanse 2019) or *gorogoro* 'heavy rollling object' in Japanese (Kita 1997), emblems like the thumbs up gesture 👍, visually iconic signs like BIRD in ASL which resembles a bird's beak opening and closing (Emmorey 2014), etc. We can categorize these forms as having **descriptive/symbolic iconicity**: these linguistic forms are not entirely arbitrary (in this sense, many call them "iconic"), but the meaning they convey in their most conventionalized form is a stable concept which has the *potential* to be represented by an arbitrary symbol, and is accessed via convention. The idea is that this contrasts with **iconic depiction**, which need not express stable concepts, and notably **could not be expressed equally as well by an arbitrary symbol**, and instead must be interpreted via a mapping of the same sort used to interpret pictures (Peirce 1974; Clark 2016; Dingemanse 2015; Fodor 2007; Camp 2018). These include modifications of onomoatopeia and ideophones to mimic a particular sound or perceptual experience, gestures of a particular size, depictive verbs in sign languages which illustrate a particular arrangement, and even quotation-like demonstrations that depict another's actions, attitudes, or speech (Clark & Gerrig 1990; Davidson 2015; Maier 2018). Descriptive iconicity involves conventionalized forms mapping to discrete meanings (or an area/vector of meaning, when we consider polysemy); depic-

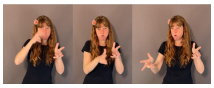
Descriptive iconocity	**Depictive iconicity**
(Some) lexical signs:  in ASL Classifier handshapes in ASL:  'vehicle' Ideophones in Siwu: *gelegele* 'shiny' Onomatopeia: *Margola chirped in her cage.* Categorical gestures: 👍	(Some) co-speech gestures: *This huge*  *tail* Classifier movements, locations in ASL:  'vehicles like this' Siwu ideophone w/depiction: *gelegele-gelegelegelegelegelegele* "Expressive" onomatopeia: *Margola went "chiiirrrrp!"* Quotations: *Alexis was like, "No, David!"*

Table 1. Examples of descriptive iconicity vs. depictive iconicity

tive iconicity involves neither but meaning derives from mapping a form directly to a particular pictoral representation. In many areas of lingustics these have both been categorized as "iconicity", including especially in disparate recent formal work (Schlenker 2021; Ramchand 2019), and one way this approach differs from previous approaches is in taking this difference as crucial to understanding the role of iconicity in compositionality. Table 1 provides several contrastive examples. To further clarify the distinction between depictive iconicity and descriptive iconicity, we turn briefly to two highly iconic phenomena (one found in spoken language, one in signed language) that each illustrate both types of iconicity and their dissociability, as well as the way that descriptive iconicity supports depictive iconicity, in two different language modalities.

2.1. DEPICTION VS. DESCRIPTION IN SPOKEN LANGUAGE IDEOPHONES. One can find beautifully illustrative examples of both kinds of iconicity in spoken language ideophones in work that has emphasized the depiction/description distinction (Kita 1997; Dingemanse 2012; Dingemanse et al. 2015; Dingemanse 2015; Dingemanse & Akita 2017; Dingemanse 2019). Let's begin with an example from Kita (1997) from Japanese, which includes the ideophone/mimetic expression *gorogoro*, which is used to convey the movement of a heavy round object with a continuous rotation.

(1) **Mimetic/ideophone (Japanese)** (Kita 1997)

tama ga gorogoro to korogat-ta no o mi-ta
ball NOM Mimetic roll-Past Nominalizer ACC see-Past

'(One) saw a ball rolled *gorogoro*.'
(*gorogoro* = movement of a heavy round object with continuous rotation)

In one sense, *gorogoro* is iconic: there is a non-arbitrary relationship between the form and its meaning in that it sounds enough like a ball rolling and/or it's easy to connect the meaning to its

form, such that it seems that *gorogoro* is an especially well motivated form to use for that meaning. However, notably this still involves a conventionalized form-meaning mapping that must be learned by language users, and that vary from language to language: a Japanese speaker could not have used just any series of sounds in this sentence position that evokes a ball rolling, but rather must use this specific conventionalized form. In this sense it is descriptive and symbolic, just like the English word *roll*. English sound-symbolic onomatopeia illustrate the same phenomenon, as in (2); the choice of form may be motivated by its meaning (i.e. *knock* somewhat reminds one of the sound of knocking on a door), but it still must be learned as a conventionalized mapping. Crucially, it also could have taken a different form and still describe the same class of events of knocking, e.g. French *frappe* 'knock'.

(2) **Onomatopeia (English)**
She knocked on the door
(*knock* = similar to the round of a fist hitting a door)

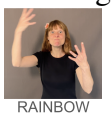
When used in their most conventionalized forms, iconic words like *gorogoro* or *knock* are often highly integrated into the linguistic system as modifiers, verbs, etc, and present an example of descriptive iconicity: these involve a somewhat non-arbitrary form , but it is a form that provides meaning through conventional symbolic mapping. Although their iconicity may play an important role in organization of the lexicon and semantic access and perhaps also in acquisition (see Section 6 below), for a particular sentence token these contribute to truth conditional meaning in the very same way as non-iconic words in any language, that is, the iconicity present in the conventionalized form is ignored for interpretation of a proposition conveyed by this sentence.

However, these same forms, because they are descriptively iconic, i.e. have a motivated connection between form and meaning, can easily be used to further create a depiction, where the interpretation of an additional aspect of meaning comes not only via convention but by a more direct perceptual experience, i.e. from the length of the ideophone's production to the length of the action, or via several small repetitions to depict the duration of several repeated actions. Kita (1997) in fact notes that *gorogoro* tends to be used with gestures and exaggerated expression to somehow also **show**, and not just **tell**, how the rolling happened. Dingemanse (2015) notes the example of the Siwu ideophone *gelegele* 'shiny', a motivated base form which can then be used to provide meaning via depiction, as in *gelegele-gelegelegelegelegele*. Ideophones show an inverse relationship between grammatical integration and depiction: those that combine directly with arguments are less depictive, while ideophones that are supporting depictions are typically set off with markers of reported speech/quotation (Dingemanse & Akita 2017). We can similarly see this kind of difference in English between (3-a) and (3-b), where we can observe a truth-conditionally inert iconicity in the movitated form of the symbolic word *knock* in (3-a), which then makes an

excellent basis for a depiction in (3-b). The proposition conveyed by the sentence in (3-a) could be conveyed by an entirely non-iconic lexicon, but that is not the case for the meaning of the sentence in (3-b).

(3) a. *She knocked on the door* (**Descriptive iconicity in *knock***)
b. *She was like, "knock-knock!"* (**Descriptive iconicity supporting an iconic depiction**)

In sum, the form of an ideophone or onomatopeia can exhibit descriptive iconicity without involving depiction; they still contribute meaning via a conventionalized symbolic mapping. In addition, these same forms are also especially ripe for use in supporting additional depictions/depictive iconicity, with the result that the same form can be conveying meaning via symbolic description (the conventionalized use of a motivated form) and depiction (modification of that form to resemble a referent or event in some way) (Clark 1996; Kita 1997; Dingemanse 2012).

2.2. DEPICTION VS. DESCRIPTION IN SIGN LANGUAGE. Ferrara & Hodge (2018) extend semiotic observations about description and depiction to the manual/visual modality via several rich examples from both sign languages and speech plus gesture. They note that many forms in sign languages have non-arbitrary form-meaning mappings but are ultimately symbols that participate in description, such as the verb INFORM in American Sign Language which begins at the forehead/source of knowledge, BABY in American Sign Language which resembles the shape of rocking an infant in one's arms, or RUN in Norwegian Sign Language which resembles the way one swing's one's arms while running. Like *knock* or *chirp*, these forms have some resemblence to their meaning that presumably may aid in word learning or aiding lexical access, but that resemblence is inert when it comes to the proposition the sentence conveys, i.e. any iconicity need not be used to comprehend a particular sentence token, and in fact signers need not recognize any iconicity of the sign to use it fluently. For example, the Norwegian SL sign RUN can be used even in the case when one doesn't swing one's arms, the American SL sign BABY is used to discuss a baby sleeping in a cradle (not in one's arms), and the sign 

can be used for a rainbow that doesn't have the same arc as the conventionalized form. All of these are thus examples of motivated symbols/descriptive iconicity, i.e. iconicity within the symbolic lexicon.

However, just as we saw with ideophones and onomatopeia in spoken languages, descriptively iconic forms in sign languages can be used to support iconic depictions. One example comes from depicting classifiers, which do so in a highly parallel manner to spoken language ideophones. These signs, found in most of the world's sign languages (Zwitserlood 2012), have a descriptively iconic component, the classifier handshape, which comes from a limited set of conventionalized, iconic forms, such as the  handshape for vehicles or the  handshape for a line of upright

individuals. Ferrara & Hodge (2018) point out that the handshapes can be descriptively iconic, and participate in regular grammatical alternations and hierarchical structures in the language at the morphological and phrase-structure level (Benedicto & Brentari 2004; Zwitserlood 2012), similar to onomatopeia and ideophones. On top of the basic descriptive iconicity in the handshape, they are also used to support iconic depictions in their use of movement and location to convey spatial relationships, as in (4), in which the classifier CL-4(line) depicts a line of students behind a library desk.

(4)

'Ten students stood in a line behind/along the library desk.'

Like the iconic depictions found in ideophones and onomatopeia (e.g. *knock! knock! knock!*), the classifier both **tells** and **shows** how the individuals were arranged. The depictive component is not semantically inert, but rather important to understanding the utterance: it is not possible to convey exactly this information via a symbolic lexicon, exactly as we saw for (3-b) above. Intriguingly, this distinction between the discrete categories of sign language classifier handshapes and their analog meanings conveyed by their depictive movements and locations has been supported experimentally (Emmorey & Herzig 2003) and forms the basis for formal semantic analyses of classifiers as handshapes conveying lexical content and movements and locations as demonstrations (Davidson 2015; Zucchi 2012), again supporting the idea that these complex forms illustrate in their conventionalized forms (handshapes) a descriptive iconicity, which is especially ripe for use in supporting additional depictive iconicity (spatial locations and movements) (Ferrara & Hodge 2018).

2.3. DISSOCIATING MODALITY AND SEMIOTICS: THE MEANING OF "GESTURES". As illustrated by spoken language ideophones and sign language classifiers, the semiotic distinction between depiction and description critically cuts across language modalities: speech can involve depiction or description, and so can signing, writing, etc. There is a close kinship between this discussion and the argument within sign language and gesture studies that expressive modality is a poor cue for distinguishing "language" from "gesture" (with "gesture" often but confusingly not always taken to mean depiction), as highlighted by Goldin-Meadow & Brentari (2017) and the related replies, and of course decades of research on the linguistic structure of sign languages (Stokoe et al. 1976; Hill et al. 2018; Sandler & Lillo-Martin 2006; Padden 1988). This brings the well-studied question of the relationship between "language and gesture" to the forefront. As it happens, approaches to the question of compositionality in language vs. gesture have tended to

fall into two camps, both of which tend to err on the side of collapsing the depictive/descriptive distinction when it comes to semantic contribution, despite early work (e.g. McNeill 1992) noting how gesture often contributes in some sense to a different kind of imagistic meaning component than non-gestural language.

Traditionally, most research focused on gestural meaning has tended to come from cognitive linguistics, assuming that the language and gesture systems, however one divides them, both contribute to meaning by adding details of different sorts to a meaning constructed to be experienced by conversational participants (Enfield 2009; Taub 2001; Liddell et al. 2003; Kendon 2004; Perniss et al. 2010). Such meanings are typically understood to be a construct of a particular experience, allowing gestures to naturally add depictive detail. In fact, from this perspective it is often curious why language is not *more* iconic, given the direct route to adding details of a particular available in iconic depiction, but certainly the presence of depiction in language is expected in this perspective. On the other hand, a well known challenge to these views of meaning from the perspective of compositionality is how logical linguistic components like *not*, *or*, etc. contribute meaning. These clearly convey or affect meaning, but how do we account for their effect if the meaning that they act on is something like one particular representation of the world? Semantically, what does it mean to depict something and then negate it, or to understand logical operators like negation and conditionals of particular experiences? Integrating the logical properties of these operators and the effects they have on meaning, especially the meaning that may be conveyed by gestures, is a challenge, and one that has not been the focus of work on gesture and language in cognitive linguistic frameworks.

In contrast, recent work within formal semantics instead approaches the question of how we account for the meaning of (depictively) "iconic gestures" by making use of truth conditions and treating their content as propositional. For example, Lascarides & Stone (2009a,b) relate the content of gestures to the content of accompanying speech via coherence relations that are already familiar from formal analyses of spoken discourse. In a more direct comment on the pictoral nature of gestural depictions, a formal semantics for pictures proposed by Greenberg (2013) states that a picture is an accurate depiction if there is a projection from a viewpoint that maps the picture into the referent; if not, the picture fails to depict. Schlenker (2018a) builds on this semantics for depiction to provide a truth-conditional semantics for such phenomena as the use of space for discourse referents and expressing plurality in both sign languages and gesture. Within this kind of semantics, integration of depiction with negation and other logical operators is more straightforward. However, in taking gestures to have the same compositional properties as speech, it runs into problems accounting for their differences. For example, many co-speech gestures seem to not be able to compose directly with negation (Kita 1997; Ebert & Ebert 2016; Ebert 2018), and so that difference must be explained either by prosody, information structure, modality, appeal to syntax

(Esipova 2019b), etc. On top of some compositional questions, though, a more general issue of two systems arises: why do we ever use depiction to communicate, if its interpretation requires a different mechanism (viewpoint projection) and why is it part of the same system? And does a truth-conditional semantics for pictures really capture the meaning of a picture as something that can be true, or not?

Ultimately, a downside of both cognitive and truth-conditional approaches to addressing the language vs. gesture question is that they both seem to be overly unificational. By assuming that an utterance "meaning" is ultimately of the same type whether it comes from description or depiction, the different pathways afforded by language and gesture actually become surprising: why do we so consistently take two different routes if we're ultimately trying to arrive at the same destination? The suggestion here is that when we communicate we have the potential to be contributing to two separate kinds of representations/two kinds of meaning, one that constructs representations of particulars to be experienced by participants in the conversation, and another which allows us to convey propositional content and resulting entailments via a compositional symbolic system. This is very much in the spirit of "multi-dimensional" approaches to depictive and descriptive meaning (Kita 1997), but now with the goal to explain *why* they are separate. To the extent that there are modality differences beyond the semiotic distinction (e.g. difference in why speech puts certain information in the auditory mode vs. manual gestures in the visual mode), the "language/gesture" divide is important and interesting to pursue on top of the semiotic one, but they are often confused, and the argument here is that the semiotic distinction (cutting across all modalities) and not the visual/auditory distinction is what matters for questions of compositionality.

In their overview on the language and gesture question and the "coming of age of sign language and gesture research", Goldin-Meadow & Brentari (2017) cite the non-compositional, analog properties of gesture (across language modalities) and contrast it with the compositional, discrete properties of language (again, across language modalities). They take a view roughly in line with the idea here that regardless of modality, compositionality relies on discrete composition and analog meaning is something separate, although we will use semiotic terminology instead of "language" vs. "gesture" since many things we might call "gestures" (due to their modality) are interpreted as ultimately symbolic descriptions and many people fail to include the spoken modality when they think of "gesture". Notably, Coppola & Senghas (2017) argue for a dissociation of compositionality from the language/gesture question, suggesting that demonstratives like *that one*, pointing signs, and spatial depictions in sign languages are examples of analog content that is compositional, while examples of discrete content that are noncompositional include emblems, exclamations, and onomatopeia. In the next section we sketch a proposal and address these points, arguing that demonstratives, spatial information, etc. are interpreted via *depiction*, while the emblems are all *description*, even when they are iconic, and even when they are non-compositional.

The following section will illustrate an approach of separating two kinds of semantics built on description and depiction which nevertheless cares about this critical compositional question: how exactly do iconcally depictive aspects of language (including depictive gestures) compose with symbolic language in both spoken and signed languages?

3. Parallel representations. Parallel representations are common throughout cognitive science. For example, developmental and cognitive psychology provide strong evidence that number cognition involves two separate representational systems: one of analog magnitude estimations and another of discrete representations strictly limited in size (Feigenson et al. 2002; Feigenson & Carey 2005; Dehaene 2011). Both systems are available to pre-linguistic infants and many non-human animals, although neither system is independently able to represent large exactly quantities like 178, which seems to be a technology developed with the help of a symbolic language, building on these systems (Le Corre & Carey 2007; Frank et al. 2008). Similarly, memory involves analog episodic memory of particular experiences, and seemingly discrete semantic memory of generalized knowledge (Tulving 1983, 1985). Given the regularity in which we see parallel representations in other domains of cognition, it shouldn't be surprising to see linguistic meaning bifurcate in a similar way, and indeed suggestions along these lines have been made, including in recent work arguing for one compositional system and one system without compositionality (Baggio 2021). This section sketches a means by which a formal compositional, symbolic system that derives propositional meaning can interact with representations of particulars coming from depiction.

3.1. PARTICULARS VS. PROPOSITIONS. The implementation suggested here is one in which a symbolic compositional system provides one representation of linguistic meaning as instructions for eliminating alternative possibilities. Symbols by their very nature evoke alternatives: a *rainbow* is defined for a given context by contrast to all of the non-rainbow entities (Saussure 1916). In other words, what it is to know the word *rainbow* is to know what counts (in a given context) as a rainbow and what is not a rainbow, just as what it is to know the meaning of a sentence is to know the situations in which that sentences is true and the situations in which it is not true (Heim & Kratzer 1998). Both functionally and formally, we regularly make use of this notion of contrast to understand entailment, for the purpose of encoding information about the way that things are via propositions (generalizations, facts, and their consequences). In other words, it is a core property of discrete symbolic reasoning that it allows us to make and share very precise claims about sets of circumstances and it allows our interlocutor to know what logically follows from them. This has been productively modeled as a narrowing of possibilities, following insights from truth conditional semantics for natural language, modeling language via logic over possible worlds (Lewis 1986) and a view of conversation in which assertions affect participants' common ground by elim-

inating possibilities under consideration (Stalnaker 1978; Groenendijk & Roelofsen 2009), among other traditions. In this view of meaning, a series of assertions like *I saw a rainbow. It was above the clouds* might first rule out all states of affairs lacking rainbows seen by the speaker, and then further rules out all that lack clouds and those that lack the particular relation between a rainbow and clouds, and then it continues to narrow down the set of circumstances we are considering by the addition of further information. The entire function of such a system is for us to collect information and narrow the circumstances we are considering, i.e. to re-weight our probability mass in a certain way/update our priors with new information, made ever more precise by building our communication system on a logic, which can be aligned with a syntactic structure to model these entailments (Heim & Kratzer 1998; Chierchia & McConnell-Ginet 2000).

In such representations, a symbol like the English word *rainbow* is, roughly, a function which divides all objects into rainbows and non-rainbows in the context-relevant way (e.g. $\lambda x.x$ counts as a rainbows in the relevant context, i.e. is an object we could label in that context with the concept of RAINBOW). In this case we can build a proposition by using this symbol *rainbow* in a larger linguistic context of an existential statement, like *There is a rainbow*, which is also a (complex) symbol, in this case a propositional function which divides all situations into those in which there is and those in which there is not a rainbow (e.g. $\lambda w \exists x.x$ is a rainbow in w). Table 2 provides an example of propositional content provided by various symbols, including several words and phrases; the idea in each is that these contribute to determining what state of affairs one finds themselves in (or wants to be in, or believes themselves to be in, etc. depending on modality, see Von Fintel & Heim 2002 for discussion of this approach for the way that we generalize about alternative possibilities to the actual world).

Discrete **propositional** representations stand in contrast to a separate system for representing **particulars**, in which both description and depiction can be used to add detail, as in the process of creating a painting or using clip art or simply slowly uncovering an object in the real world. A property of these representations is that they must necessarily be of a particular, just as an encoding of an event via episodic memory must be of a particular event, even if it is not an accurate one or even if, in reality, its details are borrowed from an amalgamation of various actual experiences. The resulting representation of the process is a particular event/object/place/etc. which can both be (a) experienced by the conversational participants and (b) be referred to just as we refer to events, objects, places out in the world, i.e. *this person* or *that depiction*. In this view of meaning, the English utterance *There was a rainbow above the clouds* evokes for the interlocutor a particular experience of viewing a rainbow and clouds, and any accompanying co-speech gesture that shows, e.g. a full arc of a rainbow, could be used to adjust aspects of this model to have a full arced rainbow. Whereas propositional meaning represents functions over possibilities (containing various kinds of rainbows), this kind of meaning is a representation of a particular state of the world

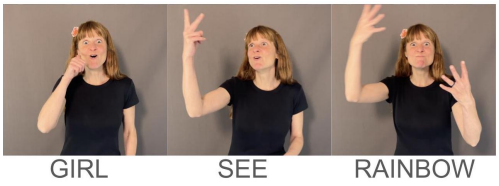
involving a particular rainbow.

Two abstract symbols that contribute their own functions, such as *beating* and *heart*, when used together as *beating heart* have a propositional contribution that may be simple intersective modification (objects that are both beating, and hearts), as well as an entirely non-compositional particular representation, emphasizing the noncompositional nature (i.e. not adding more detail to a shape of a heart, but rather an image-like representation of that particular human organ). Likewise, a dialogue in ASL using the (descriptively iconic/motivated symbol) sign RAINBOW in a non-depictive/purely conventionalized form is pure description and if the interlocutor represents a particular state of afffairs it likely includes a prototypical rainbow, exactly like the noniconic English word *rainbow*. However, both can also support depictions: if in ASL the sign is marked as depictive (set off or signed in a way that seems to intentionally deviate from a citation form, as in an emphatic full 180 degree arc in (5), as opposed to the conventionalized smaller arc for the sign as exemplified in (6)) then it will convey a slightly different particular (but not affect the propositional content), just like a depictive co-speech gesture to accompany the English word might affect the experienced particular in the same way, though not the propositional content.

(5)

'I saw a/the (specific) rainbow' (One infers that it had a full arc)

(6)



'A/the girl saw a/the rainbow' (A more conventional form, could have had any arc)

In contrast to much of the current literature on formal semantics for gestures, here the lack of interactivity between the depictive components and propositional content is a property of the depictive semiotics of gestures and not of the visual/manual modality: both co-speech gestures in spoken languages and signs in sign languages have the potential to be interpreted as symbolic/description, as we just mentioned and saw in (6). Another example of symbolic gesture from spoken language outside of an established lexicon might be size co-speech gestures expressing coffee sizes, e.g. *coffee*  vs. *coffee*  , which we discuss more in the next section. These occur in the gestural/visual medium but because they correspond to stable concepts they can be interpreted as symbols (either independently or perhaps in concert with the accompanying speech), basically equivalent to *espresso* (a standard version of a tiny coffee) and *Americano* (a standard version of a

larger cup of coffee).

Table 2 provides several examples including depictive size gestures providing size details to a representation of a particular lizard, depictive illustrations of a rainbow arc, the use of pure descriptions to modify the representation of a particular (both with words like *beating heart* and co-speech gestures for coffee sizes), classifier predicates, and quotations which depict a speech event. The idea is that some content is only depictive (the co-speech size gesture of a lizard, in contrast to the form without a gesture), some is only descriptive (negation), and many forms contribute both to a representation of a particular and to propositional content, as when the sign RAINBOW is used in a modified way with a full 180 arc, when one has a quotation, or the different contributions of classifier predicates or ideophones. This should also emphasize the independence of semantic contribution and modality: co-speech size gestures can be depictive, as in the length of a pencil or lizard (assuming one does not have natural categories for pencil or lizard sizes - this might turn out to be symbolic/descriptive for biologists who study these things), and they can also be interpreted as symbolic, as in the use of size gestures for coffee sizes. Similarly, spoken languages can use quotations to introduce depictions (e.g. *Alexis was like, "No, David!"*), and those depictions which are particular events out in the world can then themselves (*No, David!*) be interpreted (in the context they were uttered) via description.

3.2. TYPE e EXPRESSIONS OF LANGUAGE VS. PARTICULARS. Note that the distinction between propositional vs. particulars is a distinction in the way that meaning is represented, going beyond the claim that these differ (merely) in formal type. For example, referential expressions like *that rainbow*, or *the rainbow I saw yesterday* or names like *Alex* and pronouns like *she* are expressions in language which can make reference to particular things outside of language. In formal semantics of natural language based on a type-driven lambda calculus (e.g. Heim & Kratzer 1998), we can analyze these expressions as being of type e, but the particulars that they make reference to in the world (i.e. the rainbow, the actual person Alex, etc.) are parts of the world, not expressions of a natural or logical language that have a type. The idea put forward here is that a referential expression (e.g. *that rainbow*) can refer to a particular rainbow via convention of forms and compositional linguistic structure that necessarily evokes alternatives (i.e. non-rainbows); in contrast, a depiction of a rainbow *is* a particular in the world, just like a rainbow or a window or a seeing event. We understand a rainbow depiction in the same way we understand our perceptions of a rainbow itself. Of course, language is also part of the world: the utterance of the words *that rainbow* by a speaker is a particular event which we can also sometimes refer to directly in instances of quotation, and in fact in quotational contexts works in a parallel fashion to depiction, where one event of quotation is related directly to another reported speech event via a demonstration/depiction relation (Davidson 2015). Again, in such a case, we understand the quotation itself via the same mechanisms we understand the original speech event, in that case via our interpretation of language.

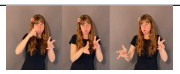
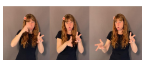
Expressions (*w/depiction)	**Particulars**	**Propositional content**
heart		$\lambda x.x$ is a heart
There is a heart		$\lambda w \exists x.x$ is a heart in w
beating heart		$\lambda x.x$ is a heart and x is beating
lizard		$\lambda x.x$ is a lizard
**lizard*  (co-speech gesture)	 (lizard of that size)	$\lambda x.x$ is a lizard
* *Alexis was like, "No, David!"*	*No, David!*	$\lambda w \exists e.$ agent(e,Alexis) $\wedge$ e demonstrates (that $\leftarrow$)
*  (ASL Classifier predicate)	 (spatial layout)	$\lambda e \exists x.$ theme$(e,x) \wedge$ vehicle(x) $\wedge$ that $\leftarrow$ demonstrates e
rainbow		$\lambda x.x$ is a rainbow
There was a rainbow		$\lambda w \exists x \exists e.$ theme(e,x) $\wedge$ rainbow(x)
RAINBOW		$\lambda x.x$ is a rainbow
* RAINBOW modified w/180°		$\lambda x.x$ is a rainbow
coffee  , *not coffee* 		$\lambda x.x$ is an espresso $\wedge$ x is not an American coffee
English *not*, ASL NOT		$\lambda p.\neg p$

Table 2. Comparison of representations of particulars and contributions to propositions

The consequence of the distinction between propositional content vs. particulars is primarily seen in composition. A referential expression like *the rainbow* may include more detailed information via restrictive modifiers, e.g. *the beautiful rainbow I saw yesterday near the Alewife station*, each of which is represented as their own functions that can evoke alternatives: *beautiful* (as opposed to not beautiful, i.e. $\lambda x.$ beautiful (x)), *[which] I saw yesterday near the Alewife station* (as opposed to not seeing it there, i.e. $\lambda x.$ I saw x yesterday near the Alewife station), etc. Exactly which contrasting alternatives are drawn depends on information structure, i.e. our conversational goals, what is given in the discourse, etc. (Roberts 2012). Different modifiers of course influence which possible rainbows might be referents for the referential expression, based on which actual rainbows satisfy the restrictions, but the variety of particular rainbows out in the world do not themselves form part of the propositional representation. In contrast, understanding increasingly more detailed depictions works entirely differently: the idea is that just as we approach a window and try to get a clearer view of the rainbow itself outside a window, and thus have more perceptual information available about the rainbow, so too do we preceive increasingly detailed aspects of a depiction, and in both cases directly encode more perceptual details about particulars. There is a longstanding debate in cognitive science about the discrete vs. analog nature of this process and/or how much a perception of a rainbow might be encoded with language-like properties; see Quilty-Dunn et al. (2022) for an overview and argument that there is linguistic structure in some aspects of non-linguistic processes. The distinction being made here isn't so much that there could not be discrete features at some level of representation in the perception of particulars. Rather, the idea is that however we perceive and represent particulars, once we represent something as a particular thing in the world, we can refer to it using language but by its very nature our representation of the particular itself does not form part of the propositional representation or compose with symbolic linguistic structures. (See Burge (2018) for how we might more accurately think of constraints gained via perception as accuracy conditions instead of proposition-like truth conditions, for example.)

Finally, saying that a depiction is understood as a particular doesn't mean that it can't be used to convey information about a more general process or lead to inferences that generalize about the world. Consider that I can model the process of egg-cracking in a baking recipe for a child by first cracking an egg neatly on the side of a bowl, modeling the process first, and then letting the aspiring cook take her turn. My action (which actually cracks the egg) demonstrates the process and thus is informative in a general way. I could alternatively choose to not actually crack an egg but merely depict what it should look like, showing the quick motion on the side of a bowl either with or without holding a real egg, and then let the child proceed. Whether the egg actually cracks in my demonstration is immaterial to showing the child how to crack an egg: the action (a particular event) shows how an egg should generally be cracked. So too with other depictions: I

can depict an American football through a drawing or a gesture, and you understand some aspects of the ball (such as its unusual oblong shape) in the same way as if I found an example of such a ball and showed it to you, a general point provided in many further illustrative examples by Clark (1996, 2016). Understanding and representing a communicative act *as a particular* (which we can then refer to) doesn't mean that particular can't be used to communicate generalities. The dissociation between representation and generality of inference happens in the other direction as well: understanding something via a symbolic representation that evokes alternatives doesn't mean it can't be used to refer to a particular, as in the case of referential expressions discussed above, i.e. *the rainbow that I saw, my friend Alex, the mother's swift egg cracking*, etc.

3.3. COMPOSITIONALITY AND DEPICTION: ARE THEY REALLY DISSOCIABLE?. As noted by Coppola & Senghas (2017), symbolic compositionality and analog depictions are not completely dissociable at first blush: there appear to be examples of depiction that compose with symbolic language and examples of symbolic language that seem to not be compositional. This section briefly address each in turn. First, there are indeed many examples of depiction that participate in a larger compositional structure. One of the clearest examples is the use of space in the grammar of sign languages, such as in the use of space for keeping track of discourse referents via "referential loci" and in the use of space in depictive classifier predicates. These involve what seem to be analog uses of space with compositional contents like pronouns, verbs, etc. However, from the perspective of having parallel semantic representations, the analog component need not necessarily be interpreted *within the same meaning component* as the compositional symbolic propositinoal representation, and in fact that is exactly the proposal put forward here.

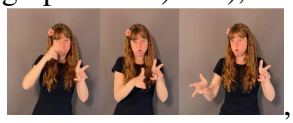
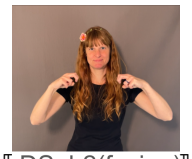
If the analog component contributes only to depiction and is not part of the propositional meaning, how might they interact? One way, alluded to in the previous section, is that we can refer to depictions just as we can refer to any particulars, e.g. via demonstratives that point deictically to particulars external to language: (*this person* (while pointing to a friend), *that place* (while pointing toward a geographic area) etc.), we can also use demonstratives to point to a particular depiction (*like this* →  , *like this* → "No, David!", *like this*  , etc.). External reference to depictions/demonstrations is one analysis of quotation within formal semantics, in which reference is made to a constructed particular (often speech) event via depiction. This kind of analysis is illustrated in (7), where the quote "I'm happy" is a depiction (by the speaker of the whole sentence) of a previous (speech) event in which Alex was a participant (Clark & Gerrig 1990; Potts 2007; Davidson 2015; Maier 2018). Similarly even without a verbatim quote, we can use the same formalism involving demonstrations, as in (8), in which the depiction $depiction_1 =$ *"No, David!"*) is referred to in the descriptive proposition, just like one can refer to any language-

Figure 1. An event of two cats sitting, and an event of demonstrating it

external particular. Finally, sign language classifiers have been handled in exactly the same way, with reference to a depiction, as in (9) (Zucchi 2012; Davidson 2015). The idea is that there are many kinds of events in the world and sometimes we want to refer to these events and relate them to each other by means of a demonstration predicate (which imposes similarity between its two arguments, similar to an expression "like this"), as the event of Alex's talking and the event of another speaker quoting it, or an event of two cats facing each other and an event of demonstrating it, as in Figure 1.

(7) ⟦Alex said, "I'm happy"⟧ =
$\exists v.saying(v) \wedge agent(v, Alex) \wedge demonstrates(\text{I'm happy}, v))$

(8) ⟦Alexis was like, "No David!"⟧ =
$\exists v.agent(v, Alex) \wedge demonstrates(\text{No David!}, v))$

(9) ⟦  ⟧ = $\lambda x \exists v[animal(x)].theme(x, v) \wedge demonstrates($  $, v)]$
lit. a predicate which takes individuals (which have to be animals) and returns TRUE if there is an event in which they are the theme and which the event can be demonstrated by 

This approach of treating depictions as event particulars can be extend to other areas where iconicity seems to integrate compositionality in sign languages, such as the reduplication of verbs and adjectives in sign languages to convey event plurality (Kuhn & Aristodemo 2017): in these cases conventionalized forms can contribute to propositional content but the depictive modifications of these repetitions can be analyzed as depictions, and in fact this is precisely the intuition suggested by Dingemanse (2015) for the very similar possibilities for reduplication seen in spoken language ideophones.

Another way that iconic depictions can affect propositional meaning, beyond reference to the depiction, is as a means of disambiguating polysemous symbols. In work on co-speech gestures,

Fricke (2022), building on Peirce 1974, calls this a "concept oriented" meaning for gestures, in contrast to an object oriented meaning. Let's take, for example, a word like *coffee*, which can denote many different kinds of coffee, such as an American style coffee (a large cup) or European style coffee (a small espresso cup), in a way that depends on the context of speech. A co-speech gesture which is the size of a small cup, such as this (), can as part of a context disambiguate the denotation for coffee, as in (10-b), which is a multi-modal utterance of the English word *coffee* accompanied by an iconic gesture. Just like a caption for a photograph, though, there is no syntactic composition implied in this case, and in fact we might write this more accurately as the interpretation of *coffee* in the context which includes the given depiction of a coffee cup, i.e. (10-c). This is similar to pointing to an example of the kind of coffee one intends, as in (10-d), to provide an exemplification of what one means by *coffee*, which is traditionally also treated as an element of the context (although the broader question of linguistic integration of deictic pointing in both spoken and sign languages is beyond the scope of the current paper). See Ebert & Ebert (2016) for discussion of co-speech gestures as exemplifications, which seems analogous to the distinction here and in Fricke 2022.

(10) a. $[\![\text{coffee}]\!]^C = \lambda x.x$ is coffee in C

b. $[\![\text{coffee}$  $]\!] = \lambda x.x$ is small-cup coffee, i.e. espresso

c. $[\![\text{coffee}\ ]\!]^C$, where C includes the depiction  $= \lambda x.x$ is coffee in C, i.e. espresso

d. $[\![\text{coffee}$  $]\!] = \lambda x.x$ is small-cup coffee, i.e. espresso

e. $[\![\text{coffee}\ ]\!]^C$, where C includes the gesture  $= \lambda x.x$ is coffee in C, i.e. espresso

There are several ways that this concept-oriented use of depictions in a context differs from our earlier object oriented case of reference to depictions. First, the concept-oriented case is generally available only when there are pre-existing separate concepts that the noun (e.g. *coffee*) may refer to. For example, while linguistic modifiers like *small* might just as easily contrast coffee sizes vs. pencil sizes (11-a)-(12-a), and we can make reference to a depiction of a pencil of any size (e.g. like *this*, classifiers, etc.), concept-oriented depictions are strange when there are no obviously pre-existing categories (such as in pencils (12-b)), suggesting a different kind of process than intersective composition. This variation is confirmed quantitatively in Esipova (2019a), where depictively iconic gestures appear to be rated more natural in contrastive contexts like (11) when they express pre-existing concepts (e.g. small pints vs. long glasses of beer, small vs. large dogs,

etc.), in contrast to depictions for less conventionalized distinctions.

(11) a. I'd like an American-style coffee, not an espresso.

b. I'd like a coffee  , not a coffee  .

(12) a. I'd like a long pencil, not a short pencil.

b. ?I'd like a pencil  , not a pencil. 

Importantly, this is not a feature of co-speech gestures specifically, but of the way that iconic depictions can interact with symbolic content. Consider the case of emoji in (13), where a depiction of an animal or a piece sports equipment disambiguates the reference of existing senses of *bat*. Further evidence that both the gesture and the emoji case are not actually integrating analog depictive elements into the propositional content is that analog details are ignored; conceptual disambiguation is the entire purpose of the depiction. In the case of coffee or beer sizes, we merely need to know which kind of drink; in the case of bats, the bat that the sentence is referring to need not have its wings ever in the form of the bat emoji or share its color; in fact it need share no perceptual features at all with the iconic depiction (whether a gesture or an emoji), it merely needs to be in the same category, i.e. in the denotation of the symbol *bat* (the animal).

(13) I found a bat  , not a bat  .

In sum, we do see cases where depictions affect propositional meaning, but they are either cases of reference, or cases of contextual influence for conceptual disambiguation. For further reading on this distinction in the domain of co-speech gestures (where they have been argued by many to integrate more compositionally), see Hunter (2019) about disambiguation of homophones, where she discusses the contrast between *bank/bank* via speech and/or co-speech gesture. Another especially interesting example comes from Fricke (2022), where she notes both uses of depiction in the same sentence, about the same object: she gives the example of a circular gesture accompanying the spoken word *hole* in a case where the hole itself was rectagular (and this was depicted as rectangular within the same utterance), but the circular gesture merely exemplified the fact that holes are prototypically round. Finally, Ebert & Ebert (2014) similarly distiguish the exemplifying use, which we see differing in both seemingly "compositional" properties, and in interpretation in context.

Beyond the use of depictions, Coppola & Senghas (2017) note two other challenges to the notion that compositionality and iconicity are incompatible. One is the use of space for keeping track of discourse referents via pronominal pointing in sign languages. However, it is important to note that across the spectrum from the most cognitive linguistics-based accounts (Liddell et al.

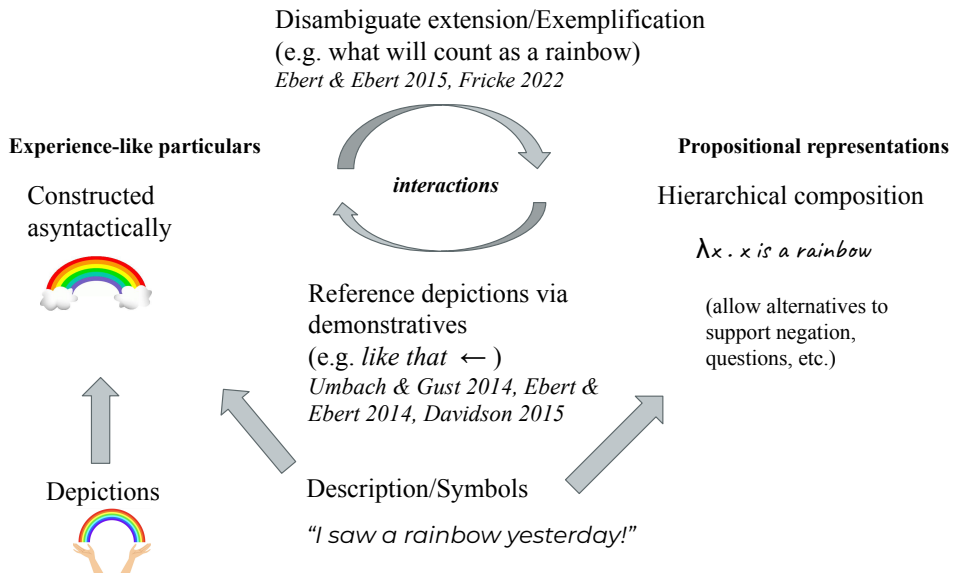


Figure 2. Schema of interactions between map-like depictions and propositional representations

2003) to the most formal semantics accounts (Schlenker et al. 2013) for analyzing this use of space in sign languages, a consistent conclusion seems to be that there are depictive elements overlaid in some sense on top of the same kind of existing symbolic system that spoken languages have for keeping track of discourse referents; the two are separable. Essentially, depictively iconic aspects are interpreted via different mechanisms than descriptive contributions of loci even in formal implementations, such as an additional presupposition (Schlenker et al. 2013; Schlenker 2018a). In these accounts the difference between the depictive aspects of these loci and the descriptive ones (say, gender or noun class) is a depiction that fails to be accurate, crucially evaluated in a component separate from the propositional component. So, whatever the analysis of use of space for loci, the analog depictive nature that concerns Coppola & Senghas (2017) is and should be separable from its propositional contribution. Moreover, recent work highlighting close parallels between demonstratives and many cases of pronominal pointing in sign languages (Ahn 2019; Koulidobrova & Lillo-Martin 2016) further lends weight to the view that while the comparison of pronominal pointing in ASL to English *that* may not always hold, the use of pointing to loci in sign languages should nonetheless consistently separate the (a) pronominal pointing from the (b) use of space to distinguish two alternatives from yet again (c) the use of space to depict; the first two contribute description and the third contributes depiction (and, speaking to Coppola & Senghas (2017)'s point, is the only analog component).

The last examples proposed by Coppola & Senghas (2017) of potential analog compositional forms are in fact demonstratives themselves (e.g. *that*), which indeed seem to "allow" the analog content of depictions to contribute compositionally: Ebert & Ebert (2016) and Ebert et al. (2020)

show that the use of a demonstrative with a depiction (in their case using the German *so*) allows negation and other propositional operators to directly target depictions. But, this is not so much depiction entering the language as depiction providing a representation of a particular, and the demonstrative creating a predicate of similarity to that object or event. We can, for example, say that a cup was green *like this* (pointing to a green plant), which doesn't mean that the plant was part of the content of the language, but rather that the demonstrative can create a linguistic object using similarity to the demonstrated item, while the item itself is entirely external to the propositional component. Umbach & Gust (2014) provide a formal semantic account of how demonstratives can relate language external content to linguistic content, creating a new predicate via a demonstrative and a similarity semantics ("like this"). In the past, the challenge to simple demonstrative accounts of iconic language in sign language and gesture has been to understand what is being demonstrated: often there is not a particular in existence in any of the states of affairs described, which challenges these views in a strict truth-conditional framework. However, if we work via parallel semantic representations that include a depicted representation of a particular, plus the ability to refer to them in a propositional system via demonstratives, then we have a referential target for which we can pick out relevant features (Umbach & Gust 2014) via demonstration of a particular.

By way of clarifying the general stance in this paper on depiction as "particulars" and not as "propositions", let us push the comparison with demonstrative pronouns further. Pointing to a particular plant while saying *This plant needs some water*, it is clear both that pointing to different plants lead to different truth conditions (pointing to different plants will make this utterance true in some cases and false in others), and also that the plant itself does not contribute to the propositional content in any way other than being referred to. Similarly, we can make reference to depictions and in that way depictions may influence truth conditions; depictions might also help disambiguate expressions, just as any element of context might, and in that case also influence truth conditions. But in neither case is the depiction, nor the plant, part of a linguistic representation from which we derive its propositional content.

Finally, there may certainly also be discrete components of language that do not seem to participate in further semantic composition. For example, Coppola & Senghas (2017) provide examples like *Shhh* or the thumbs up emblem 👍, which indeed are noncompositional and seem to involve some iconicity/metaphorical meaning, but it is (iconic) description/symnolic meaning: they access a stored meaning via cultural/linguistic convention. In a parallel semantic representational system, these do not construct a representation of any particular, but rather affect only the propositional component. Of course, descriptive forms sometimes stand in for an entire proposition and simply do not take any further arguments, that is, they are saturated: other examples include *Yes!*, *Bravo!*, etc. (These contrast with symbols like *jumped* which is an unsaturated predicate that requires an argument to be semantically complete.) So, the iconic symbols like *Shhh* are simply fully satu-

rated, and hence do not further "compose" (Coppola & Senghas 2017), while at the same time conveying only propositional content; any iconicity is ignored in composition.

In sum, we have addressed potential counterexamples that had been argued to involve true compositionality of analog, iconic depictions (Coppola & Senghas 2017); these cases involved either an inability to compose (in the case of fully saturated discrete symbols) or are precisely the sort of phenomena that are contributing only via depiction to the representation of particulars. The takeaway is that iconic depictions are represented as particulars, and that these interact with propositional meaning in one of two ways: they can be referred to (via demonstratives), or they can affect propositional meaning via exemplification of a concept or kind, but they do not seem to involve proper symbolic composition. In the next section we will consider these problems from the perspective of pragmatics/information exchange and try to understand why this might be, and also argue that this proposal also makes the right predictions when it comes to pragmatic reasoning.

4. Alternatives require propositional representations. Parallel representations run the risk of proliferating theoretical complexity without any empirical bite. This section focuses on motivating these parallel representations via information theoretic views of propositional content, based on the observation that by their very nature representing via symbols naturally involves contrast (Saussure 1916) and thus supports reasoning over alternatives; in contrast, representing particulars does not. We begin by establishing two tests for whether an aspect of language participates in reasoning over alternative propositions, namely, whether it is able to be used with negation and polar questions, and then fit this to a commonly adopted theory of information exchange and propositional content.

4.1. NEGATION AND QUESTION DIAGNOSTICS. A first diagnotic for propositional content, negation, was in fact used by Kita (1997) to make the argument for two dimensions of semantic content in ideophones, separating the propositional from the expressive content of these spoken language forms (see also Zwicky & Pullum 1987 for an argument that depictive meaning is of a different "expressive" sort). Kita noted that in Japanese, ideophones/"mimetics" like *gorogoro* 'heavy object rolling', which initially appear to be fully integrated into an utterance meaning in a positive context, become unaccepable in negative utterances unless they have a metalinguistic meaning, i.e. not like *this (gorogoro)* but like *this (some other linguistic form)*. This lack of integration contrasts with purely symbolic/descriptive modifiers like *sizukani* 'quietly' which are acceptable and quite natural in the scope of negation (i.e. being interpreted as 'not in a quiet manner', this manner reading being unavailable for the ideophone).

(14) **Depiction, no negation** (acceptable) (Kita 1997)

tama ga gorogoro to korogat-ta no o mi-ta
ball NOM Mimetic roll-Past Nominalizer ACC see-Past

'(One) saw a ball rolled *gorogoro*.'

(*gorogoro* = movement of a heavy round object with continuous rotation)

(15) **Depiction, with negation** (not acceptable)

**tama ga gorogoro to korogat-ta no de wa na-i*
ball NOM Mimetic roll-Past Nominalizer COP Focus Neg Pres

'It was not the case that a ball rolled *gorogoro*.'

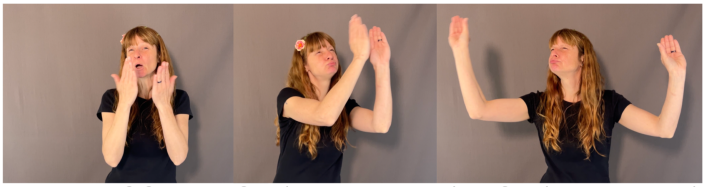
(16) **Descriptive modifier, with negation** (acceptable)

tama ga sizukani korogat-ta no de wa na-i
ball NOM quietly roll-Past Nominalizer COP Focus Neg Pres

'It was not the case that a ball rolled quietly.'

Intruigingly, we see that the very same pattern occurs in depictive classifier predicates in American Sign Language: depictions that are acceptable in positive sentences are unable to be interpreted as part of a negated proposition, and in this they contrast with symbolic modifiers (17)-(19). In both Japanese and ASL, judgments are graded: the more depictive the ideophone/classifier, i.e. the more the conventional form is modified to depict, say with changing the form more to match the perceptual experience by elongating the expression, adding emotional/affective features, etc., the less acceptable they are under negation.

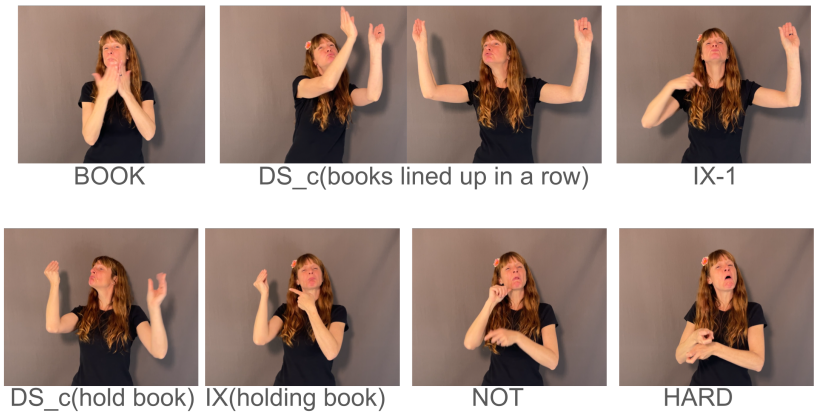
(17) **Depiction, no negation**



'Of all the books in a row, it was difficult to pull one down'

(18) **Depiction, with negation** (not acceptable)
*BOOK DS_c(books lined up), NOT DS_c(pull down w/difficulty)
'Of all the books in a row, it wasn't difficult to pull one down'

(19) **Descriptive modifier, with negation**



'Of all the books in a row, it wasn't difficult to pull one down'

The very same intuition also arises in English onomatopeia modified with depictions, as in the contrast between (20) (depiction, no negation), (21) (depiction, negation, and far less natural unless interpreted as a metalinguistic comment), and finally (22) (symbolic modifier, negation).

(20) **Depiction, no negation**
The bird was chirrrp-chirrping[expressed in a sing-songy manner] on her perch.

(21) **Depiction, with negation** (not acceptable)
*The bird wasn't chirrrp-chirrping[expressed in a sing-songy manner] on her perch.

(22) **Descriptive modifier, with negation**
The bird wasn't chirping loudly on her perch.

Conceptually, negation works as a diagnostic because it is a logical operator that applies only to propositional content, since its semantics explicitely reverses the truth value of a proposition. Note that here we are making a semantic point not a syntactic one: of course negation can sometimes seem to apply to smaller syntactic constitutents, e.g. *Mary enjoys not swimming*, but its function in terms of meaning is to return the opposite truth value of a proposition. What do we mean by a proposition? Formally, we can model propositions as functions from worlds or situations to truth values, such that any proposition, e.g. *I saw a rainbow* is going to return TRUE just for those worlds/situations in which I did see a rainbow, and FALSE for those in which I did not. This is similar to thinking about a predicate as a function over objects, e.g. *blue* returns TRUE for all of the blue things and FALSE otherwise, in some context-dependent way. This way of viewing propositions and subparts (like predicates) is useful primarily for compositionality: we can then quite easily define a propositional operator like negation in such a way that it takes any proposition q and returns the proposition $\neg q$ with the opposite truth value of q, i.e. for any world in which q was false, $\neg q$ is true, and vice versa.

In terms of information, we can think about the function of a propositional assertion as convey-

ing information by eliminating live possibilities Stalnaker (1978), a process schematized in Figure 3. The idea is that conversational participants begin by assuming some possible ways things might be (perhaps a "common ground"), and then narrow down possibilities via assertions, either that q is true (i.e. that *I saw a rainbow*, as in Figure 3), or perhaps that q is false (i.e. that $\neg q$ is true, i.e. that *I did not see a rainbow*). In contrast, the argument in this paper is that depiction seems to be doing something else entirely in these cases, constructing a positive representation of a particular event of chirping, aspects of which participants can experience directly through the depiction but are too fined grained detail and not in the correct (symbolic) format to contribute to propositional meaning. The idea is that a depiction would be too specific, depicting its referent as it is in a particular world/scenario, whereas symbolic language is powerful in creating propositional meaning precisely because it generalizes over particular cases of the ways things might be, e.g. all of the particular chirpings to instead care about the whole event class, much more informative for the function of propositional meaning.

Lupyan & Winter (2018) discuss exactly this grounded particularity of iconic forms in contrast to symbolic forms: they note that a picture of a guitar can't help but commit to whether the guitar is an acoustic or an electric guitar, how many strings it has, a particular color, etc. They note that this contrasts with the linguistic symbol *guitar*, which generalizes more broadly over both classes and configurations of objects that we all categorize as guitars. In the context of negation, we can see why this generality is valuable: saying *She hasn't seen a guitar* generalizes over all kinds of guitars in a way that we can imagine would be helpful in some cases, whereas *She hasn't seen that guitar* can only speak about a particular (familiar/present) guitar; the symbol allows us to refer to different guitars both within a particular world and guitars across worlds/situations, allowing for a highly productive use of logical operators like negation. (Note the same logic goes through for rainbows of a particular orientation, location, etc.)

For all of the same reasons that negation can serve as a diagnostic for propositional content, polar questions also act as a diagnostic, motivated by the view that questions partition possibilities (Groenendijk & Stokhof 1984; Groenendijk & Roelofsen 2009) and request of the interlocutor to help with elimination of possibilities. For example, *Is the bird chirping loudly on her perch?* asks whether we should eliminate possibilities according to the positive and negative propositions, i.e. whether it is true or false that the bird is chirping loudly on her perch. Similarly, the polar question *Did you see a rainbow?* asks whether we should eliminate possibilities according to whether the interlocutor did or did not see a rainbow, all and any details of particular chirpings and rainbows aside.

(23) **Depiction, in question** (not acceptable)
*Was the bird chirrrp-chirrping[expressed in a sing-songy manner] on her perch?

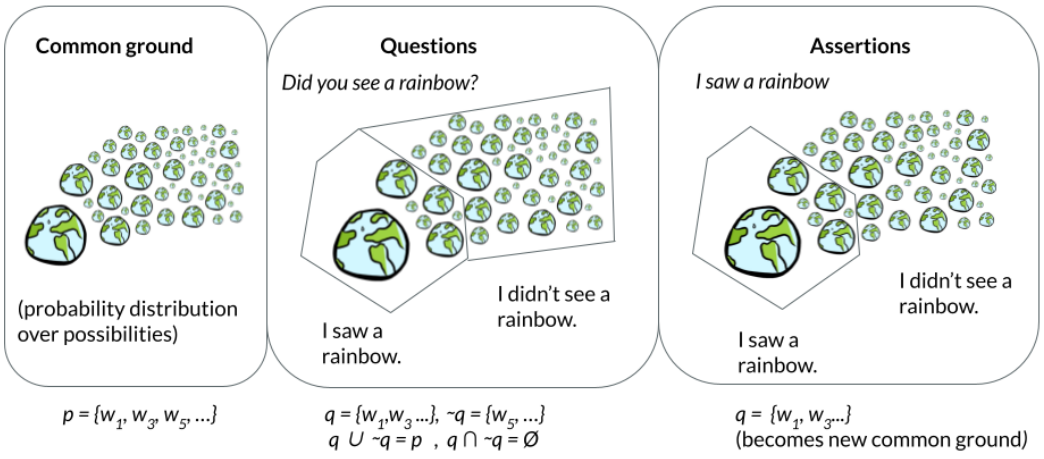


Figure 3. Propositional content eliminating possibilities via questions and assertions

(24) **Descriptive modifier, in question**
Was the bird chirping loudly on her perch?

Quantitative data comparing acceptability judgments of descriptive and depictively iconic elements in spoken language in both negation and questions are reported in Davidson (2023) and confirm the informal judgments reported here: depiction via either vowel lengthening and quotation results in significantly lower acceptability judgments under forms of negation than in positive sentences, in contrast to non-depictive (but still descriptively iconic) language.

Even when there is no overt negation or form of a question, we are often implicitely considering alternatives through information structure. In fact, modeling all discourse as a series of questions and their answers has many advantages including an integrated theory of focus placement and information structure (Roberts 2012; Rooth 1992); in such a framework, information is modeled as constantly partitioning a common ground and then answering with propositions, with focus placement constraining the kind of partitions or alternatives that we consider. We might, for example, add narrow focus on the phrase *on her perch* as in (*Was the bird chirping loudly on her PERCH?)* which partitions the possibilities into those in which she was chirping loudly on her perch, and those in which she was chirping loudly somewhere else. Linguistic forms (utterances including focus, etc.) map to partitions of possibilities via propositional representations; linguistic content then is propositional if it can raise or answer a specific question by partitioning possibilities under consideration. Examples include *It is going to rain today?*, *Is that (pointing to a window) a rainbow?*, or *Have you ever seen a rainbow?*

In contrast to propositional content, a picture/depiction may bring to mind many questions or convey quite a lot of information about how things are, but the idea is that depictive content cannot create a crisp partition of the set of possibilities in the way that propositional content can. As we

saw with depictions in ideophones and classifiers, this is graded: the more details in an image, the more it is taken to be a depiction of a particular, and the less it supports a reasoning over a class and contrasts with an alternative; in terms of consequences for acceptability, the less depictive detail there is, the more it can be interpreted as a symbol/description and the more it supports use in a question. In fact, the same patterns seen within the traditional domain of language (ideophones, depictive classifiers, even co-speech gesture) extend to more novel domains where there are a range of semiotic possibilities. Consider, for example, emoji: a highly abstracted rainbow form 🌈 could marginally/potentially be used to ask question: '🌈?' might be interpreted as asking about the existence of a rainbow, or anything that rainbows stand as a symbol for, and in both of these cases would be interpreted as symbolic/description. But with more detail, an image becomes harder to understand as symbolic description and must be taken as depiction, as in the nearly impossible to interpret question '  ?'. A highly detailed image (in this case, a photograph) is simply taken as a depiction offered for the interlocutor to experience, and can't naturally form a question. Both rainbow images occur in the same "modality" (visual images), but provide different semiotic contributions. Like highly depictive ideophones, classifiers, etc., highly depictive images can only contribute to propositional meaning through pronominal/demonstrative reference to the depiction, e.g. *Was it like this (→*  *)?*, just as in the same way that we can point to events and things in the world around us to ask questions via demonstratives, e.g. *Was it like this* (→ the view of a rainbow outside a window)? Again, this doesn't make the window or the scene part of the propositional contribution, anymore than it makes depictions part of the propositional contribution; both are particulars referred to with a demonstrative, with the demonstrative providing the link to the propositional meaning. At the same time, the depictive meaning complements, in parallel, the propositional meaning.

An alternative interpretation of the interaction of negation and question operators with depictions is that depictions simply include an abundance of information, and that it is not their depictive nature (analyzed here as being represented by a particular) but rather their informational load which is incompatible with negation and questions. One might predict that we find the same degredation of very long descriptions under negation and in question formation that we see for depictions. Testing this is beyond the scope of this paper, but an excellent next step. This would be surprising in light of previous discussion in the co-speech gesture literature on the general insability of gestures to be targeted by negation (which takes the negation issue to be one of modality, not information content or semiotic status), and given that there is not, to my awareness, any previous discussion in the literature on negation that heavy information load is difficult to negate.

Another avenue for further empirical investigation concerns cases of alternative degrees that are conveyed via depiction. For example, an anonymus reviewer provides the example of a measuring stick depicting the height for which one has to exceed to participate on a rollercoaster. In this case,

the height is understood via a depiction, yet the alternatives are trivial to encode for purposes of reasoning: it divides heights of participants by those who are able to go on the rollercoaster and those who cannot. The idea is that there should be some possibility to negate and form questions given the clear set of alternatives, but whether and how these degree depictions differ from other cases of depictions (which we analyze here as particulars) is worth further examination.

4.2. ADVANTAGES OF PARALLEL REPRESENTATIONS. What is keeping depictive linguistic forms from composing in propositional meaning, as measured by negation and question diagnostics? We alluded to this explanation as one of alternatives: for example, using the polar question *Did you dance?* can be viewed as contributing an enumerable set of propositional alternatives, in this case {You danced, You didn't dance} (Hamblin 1976). The ability to refer to concepts with discrete symbols seems to support the generation of these alternatives in a way that contrasts with depictions, which for all of the content that they can convey about the world do not seem to be structured in a way that is amenable to alternative formation. A symbol like the word *rainbow* and perhaps also sometimes an abstract emoji 🌈 has some potential to pick out rainbows as compared to some other alternatives (different weather patterns, perhaps), whereas a more detailed image like  doesn't clearly contrast with any particular other alternatives, it simply presents an experience. Or consider an emoji like 💃, which appears to have semiotic potential as either as symbol for description or, frequently as a depiction, as argued by (Maier 2023). An emoji generally seems to be unable to support the generation of any alternatives if interpreted as a depiction of a particular person dancing in a particular way in a particular moment, and is less likely to do so the more depictive details are added. To the extent that it supports alternatives, it seems highly restricted: it might do so only as a symbol of, say, general dancing (as opposed to other activities: {dancing, singing, ... }), or of a person (as opposed to other people: {Alex, Beth, Cindy, ... }), but not, say, as that person only during that moment featured in the depicted particular (even more if it is a photograph of someone dancing in a particular moment). Note in this way a *depiction of an individual* differs from the *name of an individual*, which in the latter case contributes a particular individual to the propositional content (who might contrast with other individuals), whereas the depiction might contrast with other people, or perhaps that person in other moments, etc, it's not clear. A symbol permits contrast with a set of alternatives; a depiction cannot support alternatives except when (re-)interpreted as a symbol.

A similar reasoning applies to an example of negation and co-speech gesture in Lascarides & Stone (2009b) (their (6)).

(25) D: And um I thought not too edgy and like a box, more kind of hand-held more um ...not as uh [computery] and organic, yeah, more organic shape I think.
When D says "computery", her right hand has fingers and thumb curled downward (in a

5-claw shape), palm also facing down, and she moves the fingers as if to depict typing.

Lascarides & Stone (2009b) use this example to motivate an integrated model of gesture and speech, given the scope of negation over the "computery" gesture: they say that the "overall multimodal act entails *not with a keyboard*" (although what full proposition is entailed that contains this predicate is not clear, perhaps e.g. that the object not have a keyboard). This seems to be the case of concept-oriented gesture Fricke (2022) discussed above, that is, a gesture disambiguating what is meant by "computery", i.e. computery *in the sense of containing a keyboard* (see Hunter (2019) for a similar take on the same data), much like a small pinching gesture during *coffee* clarifies that one wants coffee in the espresso sense, not the Americano sense. So, even here we have a case not of true composition of a depictive gesture with a noun under negation, but rather an exemplifying gesture to disambiguate senses of a word, which makes sense to interpret as symbolic in the sense that it picks out an existing concept that comes with a natural set of contrasts.

Recent work by Esipova (2021) makes highly effective use of prohibition signs to illustrate the way that negation (in the form of prohibition symbols) forces drawn pictures to be interpreted as symbolic. She notes an excellent example in this prohibition sign of feeding birds:  . Although the viewer assumes the sign is intended to be interpreted as "bird feeding is prohibited", Esipova notes that it is ambiguous: one could, in a stretch, interpret it as "feeding of birds in this particular direction is prohibited", or "feeding this particular number of crackers to birds is prohibited". As this illustrates, negation doesn't target the particular event, but negation searches for a symbolic target of a class of events, either events of bird feeding, events of feeding in a particular direction, etc. Saussure (1916) noted over a century ago how (depictive) icons do not exist in structural contrast with anything else, whereas (descriptive/symbolic) words naturally evoke their alternatives. Thus, in a highly nonconventionalized visual image, it's going to be unclear how to build the set of alternative possibilities, as exemplified by the ambiguity in the alternatives related to bird feeding, but it apparently must be done if it is to be interpreted symbolically. In some cases where we try to force negation with a picture, communication break downs since negation forces a symbolic interpretation and yet the options for alternatives are more ambiguous than for conventionalized symbols like words in natural languages like English or ASL, which constrain the possible alternative spaces. Question formation, like negation, relies on alternatives, and pictures are not naturally structured in the right way to support alternatives, and can only support them if reinterpreted as symbolic.

Of course, depictions do tell us about the world, and there is a long tradition in philosophy of engaging with precisely the question of "pictoral meaning" and its similarities or differences to "linguistic meaning". For example, a picture of a rainbow as viewed through a window seems

to entail the existence of the rainbow, and such entailments have been used to argue that pictures (and picture-like content in language such as iconic gestures, etc.) should be modeled through the same kind of inferences as content conveyed via symbolic, conventionalized linguistic forms (Fodor 1975; Greenberg 2013; Tieu et al. 2019). A philosophical view perhaps more naturally consistent with the one taken here can be found by those who separate pictorial from propositional content, e.g. Burge (2018); Camp (2018); Block (2023); Fodor (2007) and others, in which interpreting a picture may be much more complex and structured than simply moving through the world and registering information about it, but that even so this process does not involve propositional representations of the same sort found in symbolic interpretation and composition. In other words, depictive interpretation may involve accuracy conditions that model perception but not truth conditions of the sort that motivate propositional approaches to linguistic meaning. This is consistent with the argument here that unlike propositional content, depictive content cannot serve the function of raising and answering questions through eliminating alternatives, a core function of propositional meaning.

This distinction we have been emphasizing in this paper and in this section is one between depiction being limited to particulars, while description applies to entire classes/sets. This relates to a question many have posed in the iconicity literature: why are languages not more iconic across the lexicon, i.e. acrossed conventionalized form-meaning symbolic mappings? One compelling answer is that iconicity can hinder communicating about generalities: recall Lupyan & Winter (2018)'s point that a picture of a guitar can't help but be about a particular guitar, and thus express a particular color and style (Lupyan & Winter 2018), but often we want to speak of guitars in general, or even the lack thereof (e.g. "Sorry, we don't have any guitars", or "Have you ever seen a rainbow?"). Used in this way, words/symbols serve a useful function of abstracting away from particular features of a particular guitar. However, this insight from Lupyan & Winter (2018) and its surrounding discussion typically miss an important distinction emphasized in this paper, namely the difference between descriptive and depictive iconicity. *Depictive* iconicity necessarily involves details, and thus is ideal for communicating parallel meaning about particulars, but is much less natural in contexts without a referential commitment to a particular. However, not all iconicity has this property: *descriptive* iconicity seems quite beneficial in aiding the processing (retrieval and acquisition) of conventionalized form-meaning pairings. These forms (say, onomatopeia, iconic signs in ASL, sound symbolism in names (as discussed in Shih et al. 2018), etc.) derive their meaning via convention, picking out a stable category independently of their iconic form, and thus the form is not interpreted directly as picking out a particular. Nevertheless, the iconicity may serve as a helpful support to aid conceptual access. Work on the interaction of depiction and description in the same iconic forms and their acquisition seems fertile for these investigations; for an example, see Hou (2013) for detailed exposition of signing children's acquisition of plurality in American

Sign Language verbal forms, which involves both description (conventionalized plurality markers) and opportunities for depiction (in locations, manner, etc.), and thus perhaps unsurprisingly contains different periods in their protracted acquisition process. More generally, a takeaway to the question of "why isn't language more iconic" is that depiction and description contribute different functions in meaning so they appear for different reasons, and investigating iconicity by looking at form distribution only across a lexicon is not going to reflect these differences; attention must also be given to each of their interpretations in the context of utterance.

5. Iconicity within the symbolic lexicon. Beyond compositionality in a stable linguistic system, the question of iconicity in emergence of forms in a language and in an individual's lifespan are intruiging. The picture of child language acquisition as moving from wholistic to eventually more compositional comprehension and production suggest that perhaps the categories of depiction and compositionality may be fluid over time. For example, in a given token a form may be interpreted as non-compositional: a RAINBOW gesture or sign may at some developmental stage both depict a particular rainbow and be making a descriptive claim about the existence of a rainbow, and as the compositional system grows more complex over time, it may begin to lose a propositional meaning (e.g. there is a rainbow) but rather contribute only a nominal ("rainbow") that is looking for (in some languages) a determiner, a predicate, and perhaps increasingly complex structure, as has been the focus on decades of work in language development. It seems equally likely that the depictive system develops in complexity on both the comprehension and production sides of the equation, such that constructed models of particulars include more detail and draw increasingly on both depiction and description as time goes on (again across both an individual and across development).

Although descriptive iconicity may not contribute directly to meaning in a given sentence token, it plays two important roles. First, it may support transmission through language acquisition and/or persistence of the form: if *knock* is a particularly good symbol for the concept of knocking on a door (due to its descriptive iconicity/motivated form) then it will more likely persist as the chosen form in a language community. This may be related to possible advantages for (descriptively) iconic words in sign language acquisition (Caselli & Pyers 2017; Caselli et al. 2021), and the broad persistence of systematicity and iconicity in many areas of the lexicon in many different langugaes (Blasi et al. 2016; Shih et al. 2018; Dingemanse et al. 2015) In addition, descriptive iconicity also supports depiction in a particular sentence token, as we saw for ideophones, sign language classifiers, etc, as in *She went knock! knock! knock!* to fill in depictive detail in the representation of the particular (Dingemanse et al. 2015). So, there are many reasons to see iconicity persist within symbolic, descriptive content in language, yet this need not contribute actively to the propositional content of an utterance token that contains descriptively iconic content.

6. Comparing approaches to iconicity and compositionality.

6.1. CURRENT APPROACH: PARALLEL REPRESENTATIONS. The broad view proposed here takes the way we represent depiction to be of a different sort entirely from propositional representations. Is this reflected in any current approaches to compositional semantics of depictive content? The general approach is consistent with the one illustrated in detail in Davidson (2015) and Maier (2018) for sign language role shift and quotation, as well as sign language depictive classifiers (see also Zucchi 2012) in which the depiction is not something that itself has truth conditions but is rather something that can be created and referred to in the same way we refer to objects and events. Under this view the depiction is itself an event in the world (a "demonstration" event), and the idea is that a depiction does not fail or succeed in its depiction of a state of affairs (i.e. it does not have truth conditions) but rather is something that makes up part of the world. One can predicate of an event that it accurately demonstrates another (i.e. the content of a quotation being a demonstration of another speech event) by using verbs that introduce quotations, classifier predicates, etc., but it is not a property of the depiction itself to have propositional content. The general thrust of such a view is that the interpretation of depictive and descriptive content are quite separate, and depiction comes into play in description only via external reference or concept-oriented meaning that supports disambiguation. Extending this approach in other directions seems potentially fruitful. Consider the interesting claim that sign language verbs iconically encode event structure (Wilbur 2008). We can pull apart two pieces of this phenomenon based on the two kinds of representations we have been discussing: on the one hand, there seems to be a convincing case of descriptive/symbolic iconicity in the motivating mapping between many telic verbs in sign languages (that express bounded events) and forms that have abrupt end points. Evidence that this mapping is motivated comes from studies showing that people who don't already know the forms guess at their telicity values at above chance levels for both signs (with abrupt end points) (Strickland et al. 2015) and spoken words (that end with stop consonants) (Kuhn et al. 2021). This greater-than-chance success seems to be evidence of some descriptive iconicity/motivated forms in the symbolic lexicon, and doesn't seem to play a necessary role in the online interpretation of a particular sentence token since it can be dissociated from the meaning since not all predicates express their telicity in this way (Davidson et al. 2019). That said, this mapping also supports depiction in how we interpret gradability in the form of these signs (Kuhn 2017) roughly in the same way as is seen in spoken language ideophones (Dingemanse 2015): through their iconicity we can easily recognize lengthening, repetition, and interruptions of the verb as events which demonstrate another event in some way (e.g. their temporal features).

A possible criticism of this kind of two-part approach to iconic compositionality is that it doesn't give a complete compositional semantics (Schlenker 2021), and in an important sense this is true: the view here is emphastically not a *truth conditional* semantics for pictures, as depic-

tions are not taken to have truth conditional content, but rather exist as things one perceives and represents as particulars. One committed to a propositional semantics for pictures would prefer to have full truth conditions for the pictures built into those for the language in the case of language with depiction, but the argument here is that there are strong reasons to consider depictions as distinct in compositional contribution from descriptions. Historically, this division is actually often assumed without much ado in formal semantics, where pictures have been largely ignored until research in the last decade highlighted more depictive aspects of language (especialy via sign language, gestures, and ideophones). In fact, in the early wave of recent formal semantic work on iconic gestures, Giorgolo (2010) takes iconic gestures to contribute only to a spatial representation, separate from propositional content, and provided a process for connecting them. But given the variety of iconic phenomena in both co-speech gesture, speech, and sign, a natural move by formal semanticists working with the tools of truth-conditional/propositional semantics was to account for increasingly wide variety of phenomena using roughly the same tools as symbolic content (Schlenker 2018b; Patel-Grosz et al. 2022). The argument here, instead, is that we should more seriously examine that assumption in the case of iconic depictions across all language modalities, a view we contrast with others in the rest of this section.

6.1.1. ONTOLOGICAL DIFFERENCES ARE REFLECTED IN THE SYNTAX. A recent formal proposal in which depictive content is viewed as having an incompatible *type* of content from description, yet in a different way from the current view, comes from Ramchand (2019), who provides a framework for handling precisely this question of compositionality and iconicity based on sentence structure/syntactic hierarchy. Ramchand argues that iconicity is a natural organizing principle of communication but only in a limited place in sentence structure: in the VP domain and below, i.e. only in the syntactic domain in which a verb composes with its arguments, which maps to the semantic domain of event composition. Within this domain, semantic composition is purely conceptual composition, taking cues and insights from cognitive linguistics and related approaches to essentially create a "picture" of the event, seemingly along the same lines as the representation of a particular suggested here. Under her system, it is only at higher syntactic levels that "meaning" as understood as this cognitive model/particular is translated to external/truth conditions, via a quotational move using the notion of a demonstration and quantification over events (Davidson 2015; Champollion 2015). For Ramchand, then, the separation of cognitive/ conceptual combination from truth conditional possible world semantics corresponds to syntactic hierarchy such that the truth conditional content is higher in the syntactic hierarchy than the level of the VP, i.e. at the IP level and above where functional syntactic projections realized event types as instantiated in particular states of affairs, and thus providing a final contribution as "classic" propositional content.

This account has many advantages, especially its acknowledgement of the different roles of

iconicity and depiction from propositional content, in light of the arguments here and in previous work (Clark 1996; Dingemanse 2015; Ferrara & Hodge 2018) that they are distinct. It is not clear that it immediately accounts for the lack of interaction of negation and polar question formation with depiction that we have seen, but it seems like a potentially straightforward route to determine whether this can be modeled as an ontological mis-match at different syntactic heights. All else being equal, it would also be an advantage to have distinctions in the semantics reflected in the syntax, which is the biggest difference from the current account, in which the proposed representations are parallel and not sequential.

On the other hand, in preserving the presence of truth conditional/propositional content as part of the meaning of a verb and its argument composition, the parallel proposal adopted in this paper remains conservative with respect to the decades of work in formal compositional semantics, while Ramchand's sequential account diverges more radically for composition within the VP domain. A possibly more critical failing of the sequential proposal in its current form is that it fails to distinguish descriptive from depictive iconicity within the VP domain. This can erroneously give the expectation of far more depiction within the VP than there is with respect to descriptively iconic lexical items. For example, the expectation that a verb and its arguments constructs an event particular seems to predict that descriptively iconic lexical items like BABY or RAINBOW in ASL or *chirp* or in English would affect the event representation with their iconic components, since they occur in the VP domain. In contrast, a major argument of this paper has been that depictive and descriptive iconicity make separate contributions, i.e. their iconicity does not have to be interpreted when used in descriptions, it is only interpreted when it is used to support an additional depiction, so that the particular arc of the rainbow need not match the arc in descriptive uses of the ASL sign RAINBOW, only when they are specially marked as depictions. Similarly, the manner of cradling in the ASL sign BABY, or the sound of the English word *chirp* are all simply symbols and thus need not necessarily reflect any aspects of the events conveyed by VPs in which they are components, unless they are intentionally modified to include depictive iconicity (in which case, then they must). Presumably the sequential account could be adjusted to account for special marking of depictions as indicating that the content contribute to the construction of the event representation in contrast to conventionally appearing lexical items, whose form can be comfortably ignored for purposes of constructing the event particular, but then the power of the generalization of the syntactic locus for iconic content is significantly reduced.

6.1.2. EVERYTHING IS PROPOSITIONAL. In a third approach to the iconic compositionality question, a prolific line of research in recent years models depiction within formal frameworks as contributing truth conditions just as descriptions do, except the truth conditions come about in different ways: truth conditions for pictioral content involve a mapping projected from a particular viewpoint, as in work on formal truth-conditional approaches to picture semantics by Greenberg

(2013), in the tradition of philosophical approaches to pictoral content as propositional. Under this view, a picture is an accurate depiction if there is a projection from a particular viewpoint that maps the picture onto a state of affairs; if not, the picture fails to depict the state of affairs. This semantics for depiction has been applied to such phenomena as the spatial layout of discourse referents in sign languages (Schlenker et al. 2013), the use of plurals in both sign languages and gesture, and depictive co-speech gestures (Schlenker 2018a, 2019). A similar approach is extended to pictorial anaphora, using the notion that discourse referents can occur in the picture/comic medium as well, and thus continue or fail in reference (Abusch & Rooth 2017).

In theory a unificatory approach with depiction and description contributing to truth conditions allows for the integration and interaction of logical operators like negation and depictions, but special modifications are required when depictions pattern differently from descriptions. As we have seen, many depictions seem to not be able to compose directly with negation, both in speech (Kita 1997) and in gesture (Ebert & Ebert 2016; Ebert 2018), and so this difference must be explained either by prosody (they cannot take prosodic focus?), information structure (they aren't questioned), or appeals to structural differences (Schlenker 2019; Esipova 2019b), since it cannot be due to depiction. Schlenker (2021) proposes a general analysis in terms of presupposition: when a depiction is not accurate, there is a presupposition failure, and negation does not tend to target presupposition content, and detailed analyses are worked out for a wide variety of depictively iconic phenomena in gesture, sign, and speech. The contribution of the current paper is to raise doubt about whether depictions are even the sort of thing that should have truth conditional content in linguistic utterances. We have argued that they should not be, in light of their inability to support polar questions, negations, etc. This could be handled as a property of the presuppositional nature of this content, given that presuppositions classically "project" through questions, but being presuppositional is hardly explanatory, especially for content that cannot possible be lexically stipulated, as emphasized by Tieu et al. (2019). Instead, we can consider parallel semantic representations as a possible explanation for why depictive content so often appears to pattern with linguistic presuppositions in not being targeted by negation, questions, etc, due to its inability to support discrete partitions of worlds/alternatives.

One related distinction that has been highlighted in this approach to iconic depiction and compositionality is prosody and timing: Schlenker (2018a) distinguishes "co-speech" from "pro-speech" gestures, the latter occuring in their own time slot and the former co-occuring with spoken content. Schlenker notes that the latter are much more easily able to be targeted by negation. Later experimental work by Esipova (2019a) illustrates that there is variation within these classes: co-speech gestures can sometimes support alternatives, and this seems to be especially the case when they are interpreted symbolically, e.g. size gestures for espresso vs. American-style coffee. There's clearly more work to be done, but the suggestion arising from the current proposal would

be to more clearly distinguish the semiotic contributions (depiction vs. symbolic interpretation) in addition to modality and timing, and the prediction is that pro-speech gestures are much more often actually interpreted symbolically/as description (to stand in for a concept, with only as much depictive detail as is helpful to convey that concept and not a particular). At first blush this appears to be supported given that the most natural examples of negation targeting co-speech gestures in Esipova (2019a) were of this sort, e.g. two conventionalized beer glass sizes, as well as in research showing that depictive co-speech gestures do not lead to contrastive inference in the same way as descriptive adjectives (Alsop et al. 2018). In sum, the fact that these gestures occur in the visual modality along with speech doesn't rule out interpretation as a symbol/description, but it is simply the case that co-speech gestures are far more likely to be depictive than descriptive, and consequently lead to an inability to support contrastive inferences.

Both the purely propositional view and the parallel representation view proposed in this paper are conservative with respect to the semantics of descriptive content. Given this, the fundamental point of difference comes down to the view of depiction, namely, the argument for using truth conditional semantics to interpret pictures, and arguments in favor and against this view (as in the interaction with negation, question, etc.). Someone who is committed to a truth conditional theory of meaning for both descriptive and depictive phenomena can use an propositional account of pictures (e.g. Greenberg 2013) for the depictive content, and apply many of the observations of this paper. However, an argument in the current paper at a different level is that the interpretation of a picture is independent and different in kind from the interpretation of the propositional content, by which we mean the kind of thing which is able to narrow our possibility space and can be described by a discrete function based on contrast that partitions possible worlds/states of affairs. Rather, while depictions can be referred to (as can any parts of the world, which we represent as particulars), they do not themselves either refer and/or predicate (Burge 2018; Camp 2018).

6.1.3. COMPUTATIONAL APPROACHES TO ICONIC SPATIAL GESTURES. The problem of iconicity and composition has also arisen recently within computational approaches to multi-modal language, i.e. modeling speech plus gesture. Computational approaches most often share a spirit with the "everything in propositional" view, given the explicit computational/engineering goal to integrate the meanings via formal computations and the success of the propositional view of meaning in providing a basis for computational models of symbolic descriptive language. For example, Lawler et al. (2017) takes as given that iconic spatial gestures denote as words do (e.g. that articulations that approximate roundness should be interpreted similar to the linguistic symbol/word *round*), and implements a procedure for "modeling the meaning of gestures *qua* linguistic signs" in which the meaning of a gesture is dependent on the surrounding speech (with this difference being primarily how gestures differ from speech). Alahverdzhieva & Lascarides (2010) similarly assume that the target interpretation for a gesture is something like a (linguistic) predicate, but one which

might be (as for Lawler et al. 2017) more underspecified than a linguistic communication of that predicate. The goal of having meaning from gestures contribute to the kind of propositional representations provided by language is taken as a given, seemingly because of the long literature noting that speech and gesture are highly coorelated in production and the fact that gestures clearly affect interpretation, e.g. from McNeill (1992); Kendon (2004), etc. For example, Rieser & Lawler (2020) specifically note that in their multi-modal model "gesture meaning and speech meaning combine to a single complex proposition" without necessarily focusing on the distinction between depiction and description in gesture. That said, the goals that motivate this unified approach (integrated production, influence of gesture on interpretation) can be achieved in many ways without assuming iconic depictive gestures necessarily contribute the kind of propositional meaning that linguistic symbols do, and so there are also aspects of the current view that appear within the computational literature. For example, Giorgolo (2010) and Lücking (2016) both implement a system based on similarities in modeling gestures via processes similar to general mechanisms of spatial perception of (non-communicative) actions and events. Giorgolo (2010) further shares with the current view the notion that there are two kinds of representations at some level, one propositional and another spatial (McNeill 1992), but ultimately in that system they compose to a single informational content of the same type as language. This view might be the most amenable to the one in this paper that takes iconic gestures as particulars, though, and it would be interesting to amend it in this way. (A starting point would be to separate depictive spatial gestures from descriptive spatial gestures.) In contrast, Lücking (2016) takes gestures to convey the kind of percept-related information that might be available in the intensions of predicates (i.e. that a can has a cylindrical shape), so while being grounded in the view that gestures are somehow "closer" to perception than symbols, ultimately the content composes directly with linguistic content in a propositional representation of meaning.

Not all computational approaches treat gestural content as compositional at the level of syntax: notably, Lascarides & Stone (2009a) provide a foundational framework for considering the contribution of gestures to discourse relations, which shares in common with the view here that gestures are not contributing to compositional semantics at the level of syntax. Similarly, Hunter (2019) makes several arguments that engage with the formal accounts in the previous section arguing against a compositional syntactic account and in favor of the relation between gesture and speech being modeled via discourse relations instead. Rieser & Lawler (2020) follow similar goals of integrating gestural and speech meaning at the discourse level, in order to account for asynchrous dependencies across a discourse. However, these accounts do generally share with the "everything is propositional" view the notion that we want to handle gestures via formal frameworks intended to model language, not in frameworks intended to handle perception of particulars out in the world. For Lascarides & Stone (2009a), a motivation for this is the scope interaction between logical op-

erators like negation and iconic spatial gestures, but we have seen in Section 4.2 that this seems to happen specifically in the case where a spatial gesture is used for an existing concept, i.e. as a symbol, and doesn't extend to true depictive cases that one needs to interpret via their form.

In general the current view diverges from these computational accounts viewed here, whether they treat gestures at the level of syntax or discourse, in explicitly rejecting the idea that gestural depictions should inherently be treated via frameworks intended to handle language (and its symbolic discrete nature). Rieser & Lawler (2020) include a summarizing overview of different features of many of these accounts in terms of their syntactic vs. discourse level composition, alignment, meaning independence, and other features. It seems though, that the modality (visual/manual gesture vs. speech or text) has been the foundational distinction that these accounts attempt to capture. In contrast, the current account argues, in agreement with semiotic and psychological accounts (Clark 1996, 2016; Dingemanse 2015; Ferrara & Hodge 2018) that there should be different treatments for depictive vs. descriptive gestures (not a modality specific difference) and that many iconic gestures are depictions which do not "integrate" into the propositional symbolic system any more than other aspects of the world which we may want to refer to. Of course these gestures may, however, communicate quite a lot by providing a way for the interlocutors to represent particulars of the world, i.e. a rainbow (even if it is not actually present in the context), a method for cooking (even if we are not currently making food), etc. In these cases, computational implementations via simulation will be more amenable to the current view of depictive iconic spatial gestures (and other depictive elements in language, spoken and signed), in contrast to symbolic/propositional representations. Although an implementation is clearly beyond the scope of this paper, a computational theory of cognition committed to modeling depictions in a way that is aligned with this paper would begin by representing depictions in the same way that it represents events and objects in the world. Communicative utterances involving depictions may involve dependencies in either the sense of context-based disambiguation and reference, just as in any non-linguistic elements in the world, pictures adjacent to text, etc.

6.1.4. EVERYTHING IS A REPRESENTATION OF A PARTICULAR. A final family of approaches to iconicity in language gives up on propositional meaning altogether. In fact, for many decades, research focusing on depictive meaning in language and gesture overwhelmingly originated from a cognitive linguistic or psycholinguistic perspective, assuming that the language and gesture systems, however one divides them, both contribute to meaning by adding details to a model of the world constructed as a particular, much like one creates a picture by adding increasingly more detail (Enfield 2009; Taub 2001; Liddell et al. 2003; Kendon 2004; Johnson-Laird 1980), although constructed representations need not always be imagistic, but rather could involve mental models with more abstract symbols, etc. The key is that they deal in particulars that are experienced, not generalizations over possibilities represented as propositions. However, a well known challenge

to such views of meaning from the perspective of compositionality is its inability to account for the informational content/entailments of language, and how logical linguistic components like *not*, *or*, etc. contribute meaning. What is the negation of a picture or a model that we interpret via perception/experience (Kaup et al. 2006)? This approach would see to have an advantage when we see depiction integrated into grammatical structures (Clark 2016), and is challenged by limitations to this integration, such as when we see increasing depiction inversely related to grammatical integration (Dingemanse & Akita 2017; Kita 1997). Finally, an additional downside of a construction-based/cognitive semantics approach to the semantics of depiction and description is that it is overly unificational in assuming that description and depiction contribute meaning in the same way: why do we so consistently take two different routes if we're ultimately trying to arrive at the same destination? The proposal involving parallel representations provide a motivation for depictive and descriptive content in language in their own ways, drawing on the strengths of truth-conditional approaches to meaning for descriptive language, and taking pictures as contributing to construction of particulars that can be perceived directly and also referred to within descriptive/propositional language just like any other particulars.

7. Conclusions. This paper argues that only by separating the semiotic functions of depiction and description (Clark 2016) can we understand the compositional, and noncompositional, properties of iconic language. It presents the argument in favor of seprating (a) a constructive component of linguistic meaning that deals with particulars (which we can describe and depict, and interpret via perceptual experience), and (b) a propositional component of linguistic meaning that is built via symbolic description, based in notions of structural contrast (Saussure 1916) and interpreted via a logic and functions that link to concepts/generalizations. Where depiction appears to integrate with propositional meaning is better modeled as occuring simply in parallel representations or in some cases making use of (implicit or overt) demonstratives to refer outside of the propositional component to depictions, or exemplifications to disambiguate application of existing concepts.

In addition to the empirical coverage gained by separating the functions of linguistic meaning into a particular type and a propositional type, this separation maps neatly into several similar distinctions outside of language proper. First, other areas of cognition like *number* present evidence for parallel analog and discrete representations (analog magnitudes and object files, respectively) (Feigenson et al. 2002; Feigenson & Carey 2005), illustrating the advantages of some redunancy in the methods for encoding within a broader cognitive domain. Similarly, memory has been traditionally divided into representations of particulars (episodic memory) and representations of generalities (semantic memory) (Tulving 1983, 1985). In both number and memory the systems interact but serve essentially different encoding functions, along the same kind of lines suggested here. Finally, evidence from neuro-linguistics suggests separate pathways for compositionality/propositional representations versus wholistic, imagistic representations (Frankland &

Greene 2020; Baggio 2021), consistent with a parallel encoding and theoretical separation of particular vs. propositional content. In general, parallel processing is rampant throughout different cognitive domains generally so it should be unsurprising to see in language (Jackendoff 2007); the claim here is that parallel semantic representations provide an advantage even within the narrow domain of linguistic meaning, an idea that has been explored within semantics often under the name of multiple "dimensions" of meanings (Kita 1997; Zwicky & Pullum 1987; Potts 2007), but here cashed out within a cognitive science view as separate representational formats. Evidence here in favor of this view comes from investigating the compositionality of iconic depictions, while preserving much of the insights of propositions as symbolic and of symbolic meanings as conveying information. A key takeaway here is that it is a core property of symbols to evoke alternatives (Saussure 1916), and so symbolic description (unlike depiction) supports reasoning over alternatives, supported by longstanding views of the role of questions and answers in this kind of meaning (Roberts 2012) and the central role of alternatives throughout language (Hamblin 1976; Horn 1989; Rooth 1992; Chierchia et al. 2012; Fălăuş 2016; Repp & Spalek 2021).

Of course, there are many important issues closely related to the depiction/propositional distinction in meaning made here that fall outside the scope of this paper. Among these are other kinds of meaning conveyed by language that go beyond description and depiction, one obvious example being social meaning (Eckert 2012, 2018). We've also not yet addressed iconicity at levels above the meaning of an utterance, such as organization of a narrative (Cohn 2013). Finally, the main thrust of the proposal here is that understanding the contribution of depictions and descriptions in natural language cuts across language modalities: sign languages (Ferrara & Hodge 2018; Hodge et al. 2019; Davidson 2015; Maier 2018), spoken languages (Clark 1996; Kita 1997; Dingemanse 2015), and even written language (Clark & Gerrig 1990; Davidson 2015; Maier 2023) all make use of both description and depiction as mechanisms for conveying meaning. This makes highly salient further questions, such as how it is that language users make use of multiple languages and modalities for various semiotic purposes; see Kusters et al. (2017) for overview. The hope is that the current paper provides a framework from the perspective of semantic compositionality for addressing the two kinds of meaning that iconic linguistic forms can convey, and contributes to clarity both within the study of iconicity and in the study of compositionality.

References

Abusch, Dorit & Mats Rooth. 2017. The formal semantics of free perception in pictorial narratives. *Proceedings of the 21st Amsterdam Colloquium, ILLC, University of Amsterdam* .

Ahn, Dorothy. 2019. *THAT thesis: A competition-based mechanism for anaphoric expressions*. Cambridge, MA: Harvard University dissertation.

Alahverdzhieva, Katya & Alex Lascarides. 2010. Analysing language and co-verbal gesture in constraint-based grammars. In *Proceedings of the 17th international conference on head-driven phase structure grammar (hpsg)*, 5–25.

Alsop, Anna, Elaine Stranahan & Kathryn Davidson. 2018. Testing contrastive inferences from suprasegmental features using offline measures. *Proceedings of the Linguistic Society of America* 3(1). 71–1.

Aristodemo, Valentina & Carlo Geraci. 2018. Visible degrees in Italian sign language. *Natural Language & Linguistic Theory* 36(3). 685–699. https://doi.org/10.1007/s11049-017-9389-5.

Baggio, Giosuè. 2021. Compositionality in a parallel architecture for language processing. *Cognitive Science* 45(5). e12949–e12949.

Benedicto, Elena & Diane Brentari. 2004. Where did all the arguments go?: Argument-changing properties of classifiers in ASL. *Natural Language & Linguistic Theory* 22(4). 743–810.

Blasi, Damián E, Søren Wichmann, Harald Hammarström, Peter F Stadler & Morten H Christiansen. 2016. Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences* 113(39). 10818–10823.

Block, Ned. 2023. *The border between seeing and thinking*. Oxford University Press.

Burge, Tyler. 2018. Iconic representation: Maps, pictures, and perception. In *The map and the territory*, 79–100. Springer.

Camp, Elisabeth. 2018. *Why maps are not propositional*, vol. 10. Oxford: Oxford University Press. doi.

Caselli, Naomi K, Karen Emmorey & Ariel M Cohen-Goldberg. 2021. The signed mental lexicon: Effects of phonological neighborhood density, iconicity, and childhood language experience. *Journal of Memory and Language* 121. 104282.

Caselli, Naomi K & Jennie E Pyers. 2017. The road to language learning is not entirely iconic: Iconicity, neighborhood density, and frequency facilitate acquisition of sign language. *Psychological science* 28(7). 979–987.

Champollion, Lucas. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38(1). 31–66.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. *Semantics: An international handbook of natural language meaning* 3. 2297–2332.

Chierchia, Gennaro & Sally McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics*. MIT press.

Clark, Herbert H. 1996. *Using language*. Cambridge University Press.

Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological review* 123(3). 324.

Clark, Herbert H & Richard J Gerrig. 1990. Quotations as demonstrations. *Language* 764–805.

Cohn, Neil. 2013. *The visual language of comics: Introduction to the structure and cognition of sequential images.* A&C Black.

Coppola, Marie & Ann Senghas. 2017. Is it language (yet)? The allure of the gesture-language binary. *Behavioral and Brain Sciences* 40.

Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38(6). 477–520.

Davidson, Kathryn. 2023. Semiotic distinctions in compositional semantics. In *to appear in Proceedings of the 58th Chicago Linguistic Society*, .

Davidson, Kathryn, Annemarie Kocab, Andrea D Sims & Laura Wagner. 2019. The relationship between verbal form and event structure in sign languages. *Glossa: a journal of general linguistics* 4(1).

Dehaene, Stanislas. 2011. *The number sense: How the mind creates mathematics*. OUP USA.

Dingemanse, Mark. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics compass* 6(10). 654–672.

Dingemanse, Mark. 2015. Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 39(4). 946–970.

Dingemanse, Mark. 2019. "Ideophone" as a comparative concept. *Ideophones, mimetics, expressives* 13–33.

Dingemanse, Mark & Kimi Akita. 2017. An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to japanese 1. *Journal of Linguistics* 53(3). 501–532.

Dingemanse, Mark, Damián E Blasi, Gary Lupyan, Morten H Christiansen & Padraic Monaghan. 2015. Arbitrariness, iconicity, and systematicity in language. *Trends in cognitive sciences* 19(10). 603–615.

Ebert, Christian, Cornelia Ebert & Robin Hörnig. 2020. Demonstratives as dimension shifters. In *Proceedings of Sinn und Bedeutung*, vol. 24 1, 161–178.

Ebert, Cornelia. 2018. A comparison of sign language with speech plus gesture. *Theoretical Linguistics* 44(3-4). 239–249.

Ebert, Cornelia & Christian Ebert. 2014. Gestures, demonstratives, and the attributive/referential distinction. *Handout of a talk given at Semantics and Philosophy in Europe (SPE 7), Berlin* 28.

Ebert, Cornelia & Christian Ebert. 2016. The semantic behavior of co-speech gestures and their role in demonstrative reference. presentation, LSCP LANGUAGE seminar talk at the Institut Jean-Nicod, Département d'études Cognitives, école normale supérieure, Paris, Dec. 2016.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology* 41. 87–100.

Eckert, Penelope. 2018. *Meaning and linguistic variation: The third wave in sociolinguistics*. Cambridge University Press.

Emmorey, Karen. 2014. Iconicity as structure mapping. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1651). 20130301.

Emmorey, Karen & Melissa Herzig. 2003. Categorical versus gradient properties of classifier constructions in ASL. *Perspectives on classifier constructions in signed languages* 222. 246.

Enfield, Nicholas J. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances* 8. Cambridge University Press.

Esipova, Maria. 2019a. Acceptability of at-issue co-speech gestures under contrastive focus. *Glossa: a journal of general linguistics* 4(1).

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

Esipova, Maria. 2021. On not-at-issueness in pictures. *Glossa: a journal of general linguistics* 6(1). 83.

Fălăuş, Anamaria (ed.). 2016. *Alternatives in semantics*. Springer.

Feigenson, Lisa & Susan Carey. 2005. On the limits of infants' quantification of small object arrays. *Cognition* 97(3). 295–313.

Feigenson, Lisa, Susan Carey & Marc Hauser. 2002. The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological science* 13(2). 150–156.

Ferrara, Lindsay & Gabrielle Hodge. 2018. Language as description, indication, and depiction. *Frontiers in Psychology* 9. 716.

Fodor, Jerry. 2007. The revenge of the given. *Contemporary debates in philosophy of mind* 105–116.

Fodor, Jerry A. 1975. *The language of thought*, vol. 5. Harvard university press.

Frank, Michael C, Daniel L Everett, Evelina Fedorenko & Edward Gibson. 2008. Number as a cognitive technology: Evidence from pirahã language and cognition. *Cognition* 108(3). 819–824.

Frankland, Steven M & Joshua D Greene. 2020. Concepts and compositionality: in search of the brain's language of thought. *Annual review of psychology* 71. 273–303.

Fricke, Ellen. 2022. The pragmatics of gesture and space. *Pragmatics of Space* 14. 363.

Giorgolo, Gianluca. 2010. A formal semantics for iconic spatial gestures. In *Logic, language and meaning: 17th amsterdam colloquium, amsterdam, the netherlands, december 16-18, 2009, revised selected papers*, 305–314. Springer.

Goldin-Meadow, Susan & Diane Brentari. 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences* 40.

Greenberg, Gabriel. 2013. Beyond resemblance. *Philosophical review* 122(2). 215–287.

Groenendijk, Jeroen & Floris Roelofsen. 2009. Inquisitive semantics and pragmatics .

Groenendijk, Jeroen Antonius Gerardus & Martin Johan Bastiaan Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*: Univ. Amsterdam dissertation.

Hamblin, Charles L. 1976. Questions in Montague English. In *Montague grammar*, 247–259. Elsevier.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*, vol. 1185. Blackwell Oxford.

Hill, Joseph C, Diane C Lillo-Martin & Sandra K Wood. 2018. *Sign languages: Structures and contexts*. Routledge.

Hodge, Gabrielle, Lindsay N Ferrara & Benjamin D Anible. 2019. The semiotic diversity of doing reference in a deaf signed language. *Journal of Pragmatics* 143. 33–53.

Horn, Laurence. 1989. A natural history of negation .

Hou, Lynn YS. 2013. Acquiring plurality in directional verbs. *Sign Language & Linguistics* 16(1). 31–73.

Hunter, Julie. 2019. Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics and Philosophy* 42(4). 317–332.

Jackendoff, Ray. 2007. A parallel architecture perspective on language processing. *Brain research* 1146. 2–22.

Johnson-Laird, Philip N. 1980. Mental models in cognitive science. *Cognitive science* 4(1). 71–115.

Kaup, Barbara, Jana Lüdtke & Rolf A Zwaan. 2006. Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics* 38(7). 1033–1050.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge University Press.

Kita, Sotaro. 1997. Two-dimensional semantic analysis of japanese mimetics. *Linguistics* 379–415.

Koulidobrova, Elena & Diane Lillo-Martin. 2016. A 'point'of inquiry: The case of the (non-) pronominal IX in ASL. *The impact of pronominal form on interpretation* 221–250. https://doi.org/10.1515/9781614517016-009.

Kuhn, Jeremy. 2017. Telicity and iconic scales in asl. *Semantics Archive* .

Kuhn, Jeremy & Valentina Aristodemo. 2017. Pluractionality, iconicity, and scope in French Sign Language. *Semantics and Pragmatics* 10. https://doi.org/10.3765/sp.10.6.

Kuhn, Jeremy, Carlo Geraci, Philippe Schlenker & Brent Strickland. 2021. Boundaries in space and time: Iconic biases across modalities. *Cognition* 210. 104596.

Kusters, Annelies, Massimiliano Spotti, Ruth Swanwick & Elina Tapio. 2017. Beyond languages, beyond modalities: Transforming the study of semiotic repertoires. *International Journal of Multilingualism* 14(3). 219–232.

Lascarides, Alex & Matthew Stone. 2009a. Discourse coherence and gesture interpretation. *Gesture* 9(2). 147–180.

Lascarides, Alex & Matthew Stone. 2009b. A formal semantic analysis of gesture. *Journal of Semantics* 26(4). 393–449.

Lawler, Insa, Florian Hahn & Hannes Rieser. 2017. Gesture meaning needs speech meaning to denote-a case of speech-gesture meaning interaction. In *Fadli@ esslli*, 42–46.

Le Corre, Mathieu & Susan Carey. 2007. One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition* 105(2). 395–438.

Lewis, David K. 1986. *On the plurality of worlds*, vol. 322. Blackwell Oxford.

Liddell, Scott K et al. 2003. *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press.

Lücking, Andy. 2016. Modeling co-verbal gesture perception in type theory with records. In *2016 federated conference on computer science and information systems (fedcsis)*, 383–392. IEEE.

Lupyan, Gary & Bodo Winter. 2018. Language is more abstract than you think, or, why aren't languages more iconic? *Philosophical Transactions of the Royal Society B: Biological Sciences* 373(1752). 20170137.

Mahowald, Kyle, Anna A Ivanova, Idan A Blank, Nancy Kanwisher, Joshua B Tenenbaum & Evelina Fedorenko. 2023. Dissociating language and thought in large language models: a cognitive perspective. *arXiv preprint arXiv:2301.06627* .

Maier, Emar. 2018. Quotation, demonstration, and attraction in sign language role shift. *Theoretical Linguistics* 44(3-4). 265–276.

Maier, Emar. 2023. Emojis as pictures. *Ergo* .

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Meir, Irit. 2010. Iconicity and metaphor: Constraints on metaphorical extension of iconic forms. *Language* 86(4). 865–896.

Padden, Carol A. 1988. *Interaction of morphology and syntax in american sign language*. Routledge.

Patel-Grosz, Pritty, Salvador Mascarenhas, Emmanuel Chemla & Philippe Schlenker. 2022. Super linguistics: an introduction. *Linguistics & Philosophy* .

Peirce, Charles Sanders. 1974. *Collected papers of charles sanders peirce*, vol. 1. Harvard University Press.

Perniss, Pamela, Robin Thompson & Gabriella Vigliocco. 2010. Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in psychology* 1. 227.

Potts, Christopher. 2007. The dimensions of quotation. *Direct compositionality* 14. 405–431.

Quilty-Dunn, Jake, Nicolas Porot & Eric Mandelbaum. 2022. The best game in town: The re-emergence of the language of thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences* 1–55.

Ramchand, Gillian. 2019. Verbal symbols and demonstrations across modalities. *Open Linguistics* 5(1). 94–108.

Repp, Sophie & Katharina Spalek. 2021. The role of alternatives in language. *Frontiers in Communication* 6. 111.

Rieser, Hannes & Insa Lawler. 2020. Multi-modal meaning: An empirically-founded process algebra approach. *Semantics and Pragmatics* 13. 8–1.

Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5. 6–1.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural language semantics* 1(1). 75–116.

Sandler, Wendy & Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge University Press.

Saussure, Ferdinand. 1916. *Course in general linguistics*. Columbia University Press.

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

Schlenker, Philippe. 2018b. What is super semantics? *Philosophical Perspectives* 32(1). 365–453.

Schlenker, Philippe. 2019. Gestural semantics. *Natural Language & Linguistic Theory* 37(2). 735–784.

Schlenker, Philippe. 2021. Iconic presuppositions. *Natural Language & Linguistic Theory* 39(1). 215–289.

Schlenker, Philippe, Jonathan Lamberton & Mirko Santoro. 2013. Iconic variables. *Linguistics and philosophy* 36(2). 91–149. https://doi.org/10.1007/s10988-013-9129-1.

Shih, Stephanie S, Jordan Ackerman, Noah Hermalin, Sharon Inkelas & Darya Kavitskaya. 2018. Pokémonikers: A study of sound symbolism and pokémon names. *Proceedings of the Linguistic Society of America* 3(1). 42–1.

Stalnaker, Robert C. 1978. Assertion. In *Pragmatics*, 315–332. Brill.

Stokoe, William C, Dorothy C Casterline & Carl G Croneberg. 1976. *A dictionary of american sign language on linguistic principles*. Linstok Press.

Strickland, Brent, Carlo Geraci, Emmanuel Chemla, Philippe Schlenker, Meltem Kelepir & Roland Pfau. 2015. Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences* 112(19). 5968–5973.

Taub, Sarah F. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.

Tieu, Lyn, Robert Pasternak, Philippe Schlenker & Emmanuel Chemla. 2017. Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: a journal of general linguistics* 2(1).

Tieu, Lyn, Philippe Schlenker & Emmanuel Chemla. 2019. Linguistic inferences without words. *Proceedings of the National Academy of Sciences* 116(20). 9796–9801.

Tulving, Endel. 1983. Elements of episodic memory .

Tulving, Endel. 1985. How many memory systems are there? *American psychologist* 40(4). 385.

Umbach, Carla & Helmar Gust. 2014. Similarity demonstratives. *Lingua* 149. 74–93.

Von Fintel, Kai & Irene Heim. 2002. Lecture notes on intensional semantics. *ms., Massachusetts Institute of Technology* .

Wilbur, Ronnie B. 2008. Complex predicates involving events, time and aspect: Is this why sign languages look so similar. *Theoretical issues in sign language research* 217–250.

Winter, Bodo & Marcus Perlman. 2021. Iconicity ratings really do measure iconicity, and they open a new window onto the nature of language. *Linguistics Vanguard* 7(1).

Zucchi, Sandro. 2012. Formal semantics of sign languages. *Language and Linguistics Compass* 6(11). 719–734.

Zwicky, Arnold M & Geoffrey K Pullum. 1987. Plain morphology and expressive morphology. In *Annual meeting of the berkeley linguistics society*, vol. 13, 330–340.

Zwitserlood, Inge. 2012. Classifiers. In *Sign language: An international handbook*, De Gruyter.