Semiotic distinctions and compositional semantics *

Kathryn Davidson
(Draft: January 12, 2022)

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 utterances across all modalities frequently incorporate elements of iconic depiction. How exactly these aspects of language interact is an area of active research on spoken and signed languages and gesture studies within linguistics and across the cognitive sciences. However, questions related to iconicity are rarely studied from the perspective of compositionality: words like red and ball can combine to describe a red ball, but how do meanings from iconic forms compose with non-iconic elements in language? This paper builds upon the established semiotic distinction between depiction and description to argue against accounts that treat depiction as contributing toward propositional linguistic meaning, and in favor of two separate systems for representing meaning in language, one that involves a non-symbolic representation of a particular, and another symbolic/propositional representation, defined via ability to contrast propositional alternatives in order to eliminate possibilities, with an eye toward explaining the pervasiveness of both depiction and description as means of conveying meaning in language.

Keywords. Depiction, description, iconicity, semantics, alternatives, semiotics

1. Introduction. From a very wide lens view in the cognitive sciences, the study of natural language semantics as the study of human language sits in an uncomfortable position, arising from different views about what even is a “meaning” for a piece of natural language, and especially how to model the observation that meanings of utterances are frequently related to the meaning of their components (compositionality, Partee 1984). On the one hand, there is a sense that meanings involve particulars that we construct, building to share with one’s interlocutor an experience or memory about a particular (real or imagined or desired) event. Under this view, more linguistic forms lead primarily to more detail in an image, like adding clip art of colors, shapes, categories, etc, e.g. red (_rectangle) + heart (❤) = red heart (❤). This view encompasses a wide range of perspectives, including much of psychology, communication, cognitive linguistics, and mental models. It differs from a second perspective on meaning in human language taken by other linguists as well as some philosophers and computer scientists interested in the logical structure of language, namely, that meaning involves conveying information by eliminating possibilities. Under this view, linguistic meaning is conceived not as something to be built but rather as instructions for eliminating possibilities.

* Thanks to Mark Dingemanse, Masha Esipova, Ray Jackendoff, Josh Martin, Hande Sevgi, and Carla Umbach for comments on earlier drafts, but this remains a drafty draft, so comments welcome to: kathryndavidson@harvard.edu.
hypotheses about the possibilities under consideration: mentioning that one has some red apples might lead your interlocutor to first eliminate all states of affairs that don’t include apples, arriving at a set of possibilities that all contain apples (e.g. \( w_1, w_3, w_4, w_{16}, w_{27}, \ldots \)), and then from those eliminate the states of affairs in which the apples aren’t red, leading to a subset of the previously live ones that now only have apples that are also red (e.g. \( w_1, w_4, w_{27}, \ldots \)), getting us closer to knowing the way that things are. In this view we understand what a piece of language means by how the information it contains allows us to update a prior set of (infinite) possibilities, and what is entailed as what must hold in the updated set of possibilities. These possibilities might be modeled using truth conditions in a possible world semantics, as probabilistic inferences over possible states of affairs, or other implementations focused on deductions about what facts hold across various considered scenarios. Resolving “constructivist” vs. “eliminativist” views of meaning has been challenging to address directly via empirical means; instead approaches frequently have focused on separate phenomena to illustrate the advantages of each perspective.

The observation of the current paper is that what is often taken as a joint in the social history of approaches to semantics is more accurately seen as a real joint in nature: each type of “meaning” plays its own role in understanding human language. The argument made here for making use of both kinds of semantic representations in language comes from phenomena which fall under the label of “iconicity”, at the classic language/gesture divide. We build on a semiotic distinction emphasized in recent work on iconicity and multi-modality between depiction and description (in the sense of Clark 2016; Ferrara & Hodge 2018; Dingemanse et al. 2015; Dingemanse & Akita 2017) to argue that meaning from iconic depictions can only bear on constructed meanings, while meaning conveyed via description can contribute to both propositional/eliminative meanings and to model-like constructed meanings. A key to diagnosing the compositional difference is in the ability to support alternatives: it’s long been known that depictive content interacts differently with propositional operators like negation (Kita 1997), and we extend this to polar question formation, and argue that this is because depictions are unable to support the generation of alternatives without being (re-)interpreted as symbolic description.

Critically, the distinction between depiction and description is independent from questions of language modality: decades of research in sign language and gesture have shown that the human mind will manifest the full complexity of human language in whatever modality is available, including manual/visual signs, gestures, protactile language, speech, etc, all of which can involve depiction or description and are equally excellent vehicles for compositional language (Padden 1988; Klima & Bellugi 1979; Stokoe et al. 1976; Sandler & Lillo-Martin 2006; Goldin-Meadow & Brentari 2017). We unequivocally agree with this conclusion: the equi-potential of language modality (sign/speech/tactile/etc) is one of the most important findings in linguistics and cognitive science. However, just because language can occur across multiple modalities doesn’t mean
that we should immediately apply the same formal semantic tools that have grown from analyzing primarily written language to the more semiotically varied options available in face-to-face speech, sign, etc. For example, within formal semantic analyses of sign languages, the question of the compositionality of depictive content has arisen in such varied areas as iconic verb forms (Strickland et al. 2015; Kuhn & Aristodemo 2017), gradable adjectives (Aristodemo & Geraci 2018), depictive use of space for anaphora (Schlenker et al. 2013; Kuhn 2020; Abusch & Rooth 2017; Schlenker 2021), and the approach has generally been to model it using the same truth-conditional/“eliminative” tools as non-iconic content, often without fully examining this assumption, or by directly arguing in favor of extending tools from formal semantics to this work (Schlenker 2018b). In contrast, this paper argues that the depictive/descriptive distinction is critical to the tools we use for analysis: symbols (conventionalized signs, words, etc.) link to stored conceptual representations and compose hierarchically and contribute to eliminative/propositional meaning along traditional lines in formal semantics. Conventionalized symbols also contribute to representations of particulars, and iconic depictions can only contribute to representations of these particulars, constructed along traditional lines in cognitive linguistics.

The overall picture is an argument for a parallel architecture for sentence meaning, one of which is constructing models of particulars using non-syntactic perceptual and conceptual representations, and the other which is conveying propositional meaning that supports generalizations like entailment, etc. via hierarchical composition. In terms of how they interact, the picture that emerges in one that favors a traditional view of demonstrative reference for depictive content (Kaplan 1989; Partee 1973; Davidson 1979). This paper concludes by comparing how the distinction in the compositional properties of description and depiction is accounted for in existing proposals for the compositional contributions of iconic content (Schlenker 2021; Ramchand 2019), and briefly to parallel analog/discrete representations in other areas of cognitive science.

2. Depiction vs. Description. In language, the arbitrariness of the sign (Saussure 1916) highlights the independence of many linguistic forms from their meaning, such as the fact that the form of the English word/symbol *cat* has no particular relation to its meaning. Arbitrariness 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 much general interest in the extent of iconicity/non-arbitrariness among expressions across the world’s spoken languages (Blasi et al. 2016; Winter & Perlman 2021; Ohala 1984; Haynie et al. 2014; Shih et al. 2018) and a surge of research on both iconic and noniconic aspects of sign languages and gesture (Goldin-Meadow & Brentari 2017), there has largely been a lack of consideration in that literature for the role that iconicity plays in composing sentence mean-
ing: 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? Within cognitive approaches that involve constructing meanings, 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 eliminative 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 again under the assumption that the iconicity of the content has the same potential for compositionality as non-iconic content.

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 its interpretation from the kind seen in compositional language (Fodor 2008; Burge 2018). 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 rolling 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 iconicity**: these linguistic forms are not arbitrary (they happen to be iconic), but the meaning they convey in their most conventionalized form has the **potential** to be represented by an arbitrary symbol, and is accessed via convention. This same concept has also been categorized as “patterned iconicity” (Hwang et al. 2017; Padden et al. 2013, 2015). This contrasts with **depictive iconicity**, which does not express stable concepts, and 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 (Clark 2016; Dingemanse 2015; Burge 2018). These include modifications of onomatopeia 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); depictive iconicity involves neither but meaning derives from mapping a form directly to a particular pictoral representation. In many areas of linguistics these have both been categorized as “iconicity”, including especially in recent formal work (Schlenker 2021; Ram-
chand 2019), but this difference is 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 as case studies that each illustrate both types of iconicity and their dissociability, in two different language modalities (one found in spoken language, one in signed language).

2.1. Depiction vs. Description in Spoken Language Ideophones. To begin with an example from spoken language, work on iconic ideophones very clearly illustrates the depiction/description distinction (Kita 1997; Dingemanse 2012; Dingemanse et al. 2015; Dingemanse 2015; Dingemanse & Akita 2017; Dingemanse 2019). Ideophones have conventionalized form-meaning mappings that are learned by language users, and that vary from language to language, despite typically being iconic in the sense of having a non-arbitrary relationship to their meaning. When used in their most conventionalized forms, these are often highly integrated into the linguistic system as modifiers, verbs, etc, and present an example of descriptive iconicity: these involve a non-arbitrary form, but it is a form that provides meaning through conventional symbolic mapping. These contribute meaning in the same way as non-iconic words in any language - the iconicity present in the conventionalized form is ignored for interpretation. However, these same forms can very easily be elaborated in a way that supports a depiction, where the interpretation of this aspect of meaning comes not 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 (e.g. Siwu gelegele ‘shiny’ → gelegelegelegelegelegele, Dingemanse 2015). 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). English sound-symbolic onomatopoeia illustrate the same contrast, in the difference between The hen clucked at the pig (cluck is descriptively iconic
but not depictive) and *The hen went, ”cluuck, cluuck!”* (*cluck* is used both for description and for depiction, and the depiction is supported by this iconicity). In sum, the form of an ideophone or onomatopeia can exhibit descriptive iconicity without involving depiction; in addition, the 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 description (the conventionalized use of an iconic form) and depiction (modification of that form to resemble a referent or event in some way) (Clark 1996).

2.2. **Depiction vs. Description in Sign Language Depictive Classifiers.** 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 *cluck*, these forms have some resemblance to their meaning, but that resemblance need not be used for interpretation: 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), etc. On top of this, though, Ferrara & Hodge (2018) show how many signs exhibit both description and depiction (along with indication, see Clark 1996). 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 flat handshape for flat objects or the handshape for vehicles. 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 nearly always also support iconic depictions in their use of movement and location to convey spatial relationships, such as this: ![Vehicle](Vicars 2015), which conveys the spatial relationship of a flat road surface and a vehicle. Like ideophones, ![Vehicle](Vicars 2015) is used both for description and then modified for depiction, a depiction which is especially supported by the descriptive iconicity. This distinction between the discrete categories of sign language depictive 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 depictive iconicity (spatial locations and movements) (Ferrara & Hodge 2018).

2.3. **Comparison to the “Language” vs. “Gesture” Distinction.** 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 here to the now obvious fact that expressive modality is a poor cue for distinguishing “language”, supported by of course decades of research showing how the linguistic structure of sign languages is just as complex as spoken languages (Stokoe et al. 1976; Hill et al. 2018; Sandler & Lillo-Martin 2006; Padden 1988). Nevertheless, this brings different approaches to the well-studied question of compositionality within “language and gesture” to the forefront, since in written/spoken language systems like English the semiotic distinction is frequently mapped to the modality distinction, as written language lends itself much less naturally to depiction, while gesture frequently depicts.

In approaches that take a “meaning as construction” perspective, the distinction between “language” and “gesture” is frequently de-emphasized, with both assumed to 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). Linguistic content contributes to construct of a particular, which allows to gestures to naturally add depictive detail. In fact, from this perspective the question is often raised 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, and “gesture” and “language” are expected to contribute similarly to semantics.

On the other end of the spectrum, a recent line of work within formal semantics has taken both gestural and linguistic content to contribute to truth conditions, under an eliminative approach to meaning. For example, 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 depictive use of space to track discourse referents and expressing plurality in gesture. 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, etc, since “gesture” and “language” convey propositional meaning.
Finally, in their overview on the language and gesture question and the “coming of age of sign language and gesture research”, Goldin-Meadow & Brentari (2017) discuss the non-compositional, analog properties of (what they call) “gesture” (across expressive visual/auditory modalities) and contrast it with the compositional, discrete properties of “language” (again, across expressive visual/auditory modalities). They take a view roughly compatible with the idea here that regardless of modality, compositionality relies on discrete composition, separate from the interpretation of analog content. In a highly relevant reply, Coppola & Senghas (2017) raise precisely the question tackled in this paper: is there evidence for a 1-to-1 relationship between (lack of) compositionality and analog content/“gesture”? They argue against “the allure of the gesture-language binary,” and in favor of a dissociation of compositionality from the language/gesture question, suggesting that demonstratives like that one, pointing signs, and the use of space in discourse in sign languages are examples of analog content that is compositional, while examples of discrete content that are non-compositional include emblems, exclamations, and onomatopoeia. In the next two sections we will address both of these purported dissociations, coming down ultimately on the side of a motivated separation for compositional purposes between “analog”/depictive content and propositional content based on ability to support reasoning over alternatives. The key in addressing their analog compositional cases is to acknowledge that depictive content can indeed contribute to propositional meaning, but only in a very particular way, via either demonstratives or conceptual disambiguation. These are the very same ways that we link propositional content to non-linguistic entities (e.g. people, objects, and other present aspects of a discourse), and so the interpretation of depictive content can and should be separated from interpretation of propositional content. Overall, instead of characterizing the binary as “language” and “gesture”, a more accurate terminology can be borrowed from semiotics, since many forms we might call “gesture” are in fact descriptive, and many forms we might call “language” are in fact depictive; the use of particular modalities is entirely dissociable from potential compositional semantic contributions, although of course it remains an important (but different!) question how and why modalities are chosen to express different aspects of meaning in multi-modal language use (Kusters et al. 2017; Cohn 2016).

3. Alternatives require description. Intuitively, depiction and description seem to serve different functions, but what are the empirical arguments for separating their contributions? This section proposing testing for whether language participates in propositional content by whether it is able to be used with negation and polar questions, motivated by the informational/eliminative function of propositional content. Following pioneering work by Kita (1997), we observe that depiction behaves differently under these operators and argue that this is because they are based on alternatives, while depictions contribute instead to constructed particulars that are intended to be experienced.
3.1. **Negation.** Negation was used by Kita (1997) to make the argument for two dimensions of semantic content in ideophones, separating the propositional from the “expressive” (depictive) content of these spoken language forms (see also Zwicky & Pullum 1987 for an argument that this meaning is of a different “expressive” sort, along the lines of Potts 2007b). Kita shows that in Japanese, ideophones/“mimetics” like *gorogoro* ‘heavy object rolling’, which appear to be integrated into an utterance meaning in a positive context, become unacceptable 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.

(1) a. **Depiction, no negation** (Japanese, Kita 1997)

> tama ga gorogoro to korogat-ta no o mi-ta
> ball Nom Mimetic roll-Past Nominalizaer Acc see-Past
> ‘(One) saw a ball rolled *gorogoro*’
> *(gorogoro* = movement of a heavy round object with continuous rotation).

b. **Depiction, with negation**

> tama ga gorogoro to korogat-ta no de wa na-i
> ball Nom Mimetic roll-Past Nominalizaer Cop Focus Neg
> ‘It was not the case that a ball rolled *gorogoro*.’

c. **Descriptive modifier, with negation**

> tama ga sizukani korogat-ta no de wa na-i
> ball Nom quietly roll-Past Nominalizaer Cop Focus Neg
> ‘It was not the case that a ball rolled quietly.’

One might suppose that the interaction with negation is something that must be learned or stipulated about the lexical class of ideophones, but this is not at all the case: when we look to depictions elsewhere in language, depiction across language is similarly resistant to negation. For example, in depictive classifier predicates in American Sign Language, the very same pattern appears. Data in (2) are based on informal judgments by four Deaf adult signers of ASL. We see that depictions that are acceptable in positive sentences (2a) are unable to be interpreted as part of a negated proposition (2b), and in this they contrast with symbolic modifiers (2c).

(2) a. **Depiction, no negation** (ASL)

> BOOK CL- [(in a row), CL- ] (PULL DOWN WITH STRAINED EFFORT)
> ‘Of all the books in a row, it was difficult to pull one down’

b. **Depiction, with negation**
Of all the books in a row, it wasn’t difficult to pull one down.

c. **Descriptive modifier, with negation**

BOOK CL-(in a row), NOT HARD CL-(PULL-DOWN)

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

In both Japanese and ASL, judgments are highly graded, in a very important way: 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 exaggerated emotional/affective features, etc.) the less acceptable they are under negation. The very same intuition also arises in English onomatopeia modified with depictions, as in (3), which reflects judgements provided by four informally consulted adult English speakers.

(3) a. **Depiction, no negation**

The bird was chirrrp-chirrping[expressed in a sing-songy manner] on her perch.

b. **Depiction, with negation**

*The bird wasn’t chirrrp-chirrping[expressed in a sing-songy manner] on her perch.

c. **Descriptive modifier, with negation**

The bird wasn’t chirping loudly on her perch.

Conceptually, negation works as a diagnostic because its function in propositional content is to move the common ground to the complement set of possibilities under consideration. For example, if *The bird was chirping loudly on her perch* (symbolic modifier, no negation) eliminates one (infinite) set of possibilities under consideration (*q*), its negation eliminates the complement set, *The bird wasn’t chirping loudly on her perch* (*¬q*). The function of propositional representations is to convey information by eliminating live possibilities; when content contributes directly to that function, it is propositional. In contrast, 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 through the depiction, but which, it is argued here, do not bear directly on propositional content.

3.2. **Polar Questions.** 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) or activate a set of alternatives (Hamblin 1976; Karttunen 1977) and request of the interlocutor to help with elimination of these 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, details of particular chirpings and rainbows aside. Modeling a 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* (*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 *Is that (pointing to a window) a rainbow?, Is it going to rain today?, 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 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 depictive 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 question; the less depictive detail there is, the more it can be interpreted as a symbol/description and the more it supports a question (4a-c).

(4)  
   a. **Depiction** (English)  
      The bird was chirrrp-chirrping[expressed in a sing-songy manner] on her perch.  
   
   b. **Depiction, polar question**  
      *Was the bird chirrrp-chirrping[expressed in a sing-songy manner] on her perch?*  
   
   c. **Descriptive modifier, polar question**  
      Was the bird chirping loudly on her perch? 

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 may have different semiotic potential. Like highly depictive ideophones, classifiers, etc., highly depictive images can only contribute to eliminative meaning by making external/demonstrative reference to it, 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 (Kaplan 1989; Davidson 1979; Partee 1973). At the same time, the depictive meaning is provided by the speaker/signer and complements the propositional meaning.

3.3. DEPICTION IS UNABLE TO SUPPORT ALTERNATIVES. 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 the bird chirp loudly? can be viewed as contributing an enumerable set of propositional alternatives, in this case {She chirped loudly, She didn’t chirp loudly}. 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, the sign RAINBOW or (sometimes) even the emoji have potential to pick out rainbows as compared to some other alternatives. These contrast with a more detailed image like or the enactment chirrrp-chirrrping (expressed in a sing-songy manner) or CL- (pull down with strained effort) which fail to clearly contrast with any particular other alternatives, but rather present an experience of particular objects or events at a given time and place.

As this paper and many others have emphasized, semiotic contribution (here, to supporting alternatives) depends not on the form specifically but rather how it is interpreted; for example, an emoji like seems to have semiotic potential as either as symbol for description or as a depiction. It seems to be unable to support the generation of 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 to 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. As we saw above with rainbows, this
is even more apparent 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 but not in a depictive manner: interpretation via depiction is experiential, whereas interpretation via description can naturally support conceptually linked alternatives (individuals accessed via other names (e.g. {Alex, Beth, Cindy, ... )}).

Esipova (2021) makes highly effective use of prohibition signs to illustrate the way that negation (in the form of prohibition symbols) forces pictures to be interpreted as symbolic. An example is this prohibition sign of feeding birds: you; although intended to be interpreted as “bird feeding is prohibited”, a joke is made of the ambiguity in the information packaging of pictures to force it instead to be interpreted as “feeding of birds in this particular direction is prohibited”. It becomes clear that negation doesn’t target the particular event as depicted, but rather negation searches for a target of a generalization communicated symbolically by this picture, either (a) events of duck feeding, (b) events of feeding in a particular direction, (c) the duck kind generally, etc. In a nonconventionalized visual image, it’s often unclear how to build the set of alternative possibilities, but it must be done if it is to be interpreted symbolically, as in prohibition signs which express negation. Similarly question formation, like negation, relies on alternatives, and the more detailed/picture-like an image is in order to be “experienced”, the less it is going to be structured in the right way to support alternatives.

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; Schlenker 2021). A philosophical view perhaps more naturally consistent with the one taken here can be found in Burge (2018), in which interpreting a picture is much more complex and structured than simply moving through the world and registering information about it, but that this process does not involve propositional representations of the same sort found in symbolic interpretation and composition. To use Burge’s terminology, depictive interpretation may involve accuracy conditions that model perception but not truth conditions of the sort that motivate eliminative approaches to linguistic meaning. As such, depictive content cannot serve the function of raising and answering questions through eliminating alternatives, a core function of propositional meaning; depiction instead serves to provide a particular, which can be experienced or referred to as one experiences and refers to particulars out in the world.
3.4. WHY ISN’T LANGUAGE MORE ICONIC. 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 that has been proposed is that iconicity can hinder communicating about generalities: 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) misses an important distinction emphasized in this paper, namely the difference between descriptive and depictive iconicity. Indeed, 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. But 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 lexical signs in ASL like BABY and INFORM, sound symbolism in names as in Shih et al. 2018, etc.) derive their meaning via convention, picking out a stable meaning independently of their iconic form, and thus the form is not interpreted directly as picking out a particular object/event in time and space. Since it is not focus on particulars, descriptive iconicity may serve as a helpful support to aid conceptual access, and explain patterns found across lexicons; in contrast, depictive iconicity is used to communicate details for the interlocutor to experience, an entirely different goal from the informational content available in description. Thus, a takeaway to the question of “why isn’t language more iconic” is that depiction and description contribute different functions in meaning so they arise for different reasons, and investigating iconicity by looking at large-scale form distribution across a lexicon is not going to reflect these differences; attention must also be given to each of their interpretations in the context of an utterance.

4. Constructed particulars and eliminative propositions. Parallel representations within a single cognitive domain are common throughout cognitive science. For example, developmental and cognitive psychology provides evidence that number cognition involves two separate representational systems: one of analog magnitude estimations and another of discrete representations strictly limited in size, based on both behavioral and neural data (Feigenson et al. 2002; Feigenson & Carey 2005; Hyde & Spelke 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, which seems to be a technology developed with the help of language, building on these systems (Le Corre & Carey 2007; Frank et al. 2008). Similarly, memory involves analog episodic mem-
ory 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 will provide an implementation of such a distinction in linguistic meaning, with one system familiar to formal semanticists based on compositionality and propositions, and the other system familiar within cognitive linguistics based on constructed, simulated particulars.

4.1. PROPOSITIONAL SYSTEM. A symbolic compositional system provides one representation of linguistic meaning as instructions for eliminating possibilities, for the purpose of encoding information about the way that things are (generalizations, facts, and their consequences). A function of discrete symbolic reasoning is 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 (Davidson 1967), modeling language via logic over possible worlds (Lewis 1986) and a view of conversation in which assertions affect participants’ common ground by eliminating possibilities under consideration (Stalnaker 1978; Groenendijk & Roelofsen 2009), among other traditions. In such “eliminative” representations, a series of statements like I saw a rainbow. It was above the clouds might first rules 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 “eliminative” traditions, a symbol like the English word heart is, roughly, a function which divides all objects into hearts and non-hearts (e.g. λx.x is a heart). In this case we can build a proposition by using this symbol heart in a larger linguistic context of an existential statement, like There is a heart, which is also a (complex) symbol, in this case a propositional function which divides all states of affairs into those in which there is and those in which there is not a heart (e.g. λw∃x.x is a heart in w). Table 2 provides an example of several proposals for 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).

4.2. SIMULATED PARTICULARS. The discrete eliminative/propositional representations stand in contrast to a separate system for representing a particular, in which both description and depiction can be used to add detail, as in the process of creating a painting or using clip art. 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 result 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 “constructive” meaning, the English utterance There was a rainbow above the clouds may convey to the interlocutor the 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. Where propositional meaning provides functions over possibilities (containing various kinds of rainbows), this kind of meaning is a representation, or simulation, of a particular rainbow, along the lines for semantics explored in cognitive linguistics and mental model theory (Johnson-Laird 1980).

4.3. DESCRIPTION CONTRIBUTES TO BOTH SYSTEMS, BUT WE ONLY DEPICT PARTICULARS. Two 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, an ASL dialogue using the (descriptively) iconic sign RAINBOW in a non-depictive/purely conventionalized form is pure description and may convey a typical rainbow to the interlocutor, exactly like the noniconic English word rainbow, and for the same conversational participants, both will likely convey the same experiential particular if used in a conventionalized form. 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, as opposed to the conventionalized smaller arc for the sign) then it will convey a slightly different particular, just like a depictive co-speech gesture to accompany the English word will affect the experienced particular in the same way, though not the propositional content.

In contrast to much of the current literature on gesture semantics, the proposal here is that compositionality is a property of a form’s interpretation in context as depictive vs. descriptive and not of the expressive modality: co-speech gestures in spoken languages and modifications to signs in sign languages can be interpreted as symbolic/description if they refer to discrete categories
that support alternatives. An example from spoken language might be size co-speech gestures expressing coffee sizes, e.g. coffee ☕ vs. coffee ☕, which can map easily to existing concepts of espresso-sized coffee and Americano-sized coffee. These occur as gestures but can have a descriptive semiotic contribution, basically equivalent to espresso and Americano, and correspondingly activate two different constructed representations and two different eliminative functions. Indeed, these even support negation: one can say I want a coffee ☕, not a coffee ☕ (Esipova 2019a). The expectation is that co-speech gestures expressing unexpected/unconventionalized sizes (say, the size of someone’s arms or the size of a lizard’s tail, which do not map to any stable concept for most people) would instead necessarily be depictive.

Table 2 provides several examples of contributions by various forms to the system of representation for constructed particulars and the propositional system. These include depictive size gestures providing size details to a map-like representation of a 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. Some content is only depictive (the co-speech size gesture for a lizard’s tail), some is purely propositional (negation), and many forms contribute to both via depiction and description. Examples of both include when the sign RAINBOW is used in a modified way with a full 180 arc (describing with the lexeme and depicting with the arc), when one has a quotation (describing the person and context of an utterance and depicting the quotatin/utterance), or the different contributions of classifier predicates (a descriptive handshape and a depictive movement/location). Again, this should also emphasize the independence of semantic contribution and modality: co-speech size gestures can be depictive, as in the length of a lizard’s tail (assuming one does not have natural categories for lizard sizes - this might be 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 can then themselves (No, David!) be interpreted (in the context they were uttered) via description.
<table>
<thead>
<tr>
<th>Expressions (*w/depiction)</th>
<th>Particulars</th>
<th>Propositional content</th>
</tr>
</thead>
<tbody>
<tr>
<td>heart</td>
<td><img src="heart.png" alt="Heart" /></td>
<td>$\lambda x. x$ is a heart</td>
</tr>
<tr>
<td>There is a heart</td>
<td><img src="heart.png" alt="Heart" /></td>
<td>$\lambda w \exists x. x$ is a heart in $w$</td>
</tr>
<tr>
<td>beating heart</td>
<td><img src="heart.png" alt="Heart" /></td>
<td>$\lambda x. x$ is a heart and $x$ is beating</td>
</tr>
<tr>
<td>lizard</td>
<td><img src="lizard.png" alt="Lizard" /></td>
<td>$\lambda x. x$ is a lizard</td>
</tr>
<tr>
<td>*lizard (co-speech gesture)</td>
<td><img src="lizard.png" alt="Lizard" /> (lizard of that size)</td>
<td>$\lambda x. x$ is a lizard</td>
</tr>
<tr>
<td>* Alexis was like, “No, David!”</td>
<td><img src="AlexisDavid.png" alt="Alexis &amp; David" /></td>
<td>$\lambda w \exists e. \text{agent}(e,\text{Alexis}) \land e \text{ demonstrates (that $\leftarrow$)}$</td>
</tr>
<tr>
<td>* <img src="DepictiveClassifier.png" alt="Depictive Classifier" /> (depictive classifier in ASL)</td>
<td><img src="SpatialLayout.png" alt="Spatial Layout" /></td>
<td>$\lambda e \exists x. \text{theme}(e, x) \land \text{vehicle}(x) \land \text{that $\leftarrow$ demonstrates } e$</td>
</tr>
<tr>
<td>rainbow (in English)</td>
<td><img src="Rainbow.png" alt="Rainbow" /></td>
<td>$\lambda x. x$ is a rainbow</td>
</tr>
<tr>
<td>There was a rainbow</td>
<td><img src="Rainbow.png" alt="Rainbow" /></td>
<td>$\lambda w \exists e. \text{theme}(e, x) \land \text{rainbow}(x)$</td>
</tr>
<tr>
<td>RAINBOW (in ASL)</td>
<td><img src="Rainbow.png" alt="Rainbow" /></td>
<td>$\lambda x. x$ is a rainbow</td>
</tr>
<tr>
<td>* RAINBOW modified w/180°</td>
<td><img src="Rainbow.png" alt="Rainbow" /></td>
<td>$\lambda x. x$ is a rainbow</td>
</tr>
<tr>
<td>coffee, not coffee</td>
<td><img src="Coffee.png" alt="Coffee" />, <img src="NotCoffee.png" alt="Not Coffee" /></td>
<td>$\lambda x. x$ is an espresso $\land x$ is not an American coffee</td>
</tr>
<tr>
<td>English not, ASL NOT</td>
<td></td>
<td>$\lambda p. \neg p$</td>
</tr>
</tbody>
</table>

Table 2. Comparison of representations of particulars and contributions to propositions
4.4. COMPOSITIONALITY AND DEPICTION: PURPORTED POINTS OF INTEGRATION. As noted by Coppola & Senghas (2017), discrete compositionality and analog depictions do appear to non-trivially interact, at least at first blush: there are examples of depiction that exhibit compositionality and examples of description seem to not be compositional. Let us 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, this paper’s argument is that from the perspective of having parallel semantic representations, the analog component need not necessarily be interpreted within the same meaning component as the compositional structure.

If the analog component contributes only to depiction and is not part of the propositional meaning, how does it happen that the depiction seems to be part of the propositional meaning? Consider that just like we can use demonstratives to point to particulars external to language, out in the world (this person → 🧙‍♂️, that place → 🌋 etc.), we can also use demonstratives to point to a particular depiction (this arrangement → 🎤, this utterance → “No, David!”, this big 🎤, etc.). Partee (1973) notes that expressions like She went whoooah! seem to be elided versions of obvious demonstratives like She went like this: whoooooah!, and indeed, external reference to depictions/demonstrations is one analysis of iconic enactments within formal semantics, in which a constructed representation of a particular event via depiction (e.g. depiction₁ = “No, David!”) is referred to in the descriptive proposition (e.g. for the quotation in Table 2, λw.∃e. agent(e,Alexis) ∧ e demonstrates depiction₁) (Clark & Gerrig 1990; Potts 2007a; Davidson 2015; Maier 2018). It is also behind the analysis of depictive classifier predicates by Davidson (2015) and Zucchi (2012), in which the interpretation of the depictive material is separated from the compositional truth-conditional structure as a demonstration, something referred to in the propositional content but not interpreted as containing truth conditions (the depiction is crucially not a proposition itself).

Another challenge is to extend this analysis to the use of space for keeping track of discourse referents via pronominal pointing in sign languages, considered by Coppola & Senghas (2017), Schlenker et al. (2013), and others as a case of analog content that bears on propositional meaning. However, across the spectrum from the most construction-based accounts (Liddell et al. 2003) to the most eliminative-based accounts (Schlenker et al. 2013) for analyzing this use of space in sign language discourse, a consistent conclusion seems to be that there are clearly depictive elements which at most interaction are taken to be overlaid in some sense on top of the same kind of system that spoken languages have for keeping track of discourse referents. For example, within a cognitive linguistic framework, Liddell et al. (2003) shows how the use of space for loci constructs a representation/simulation of a scenario via depiction. Within an eliminative tradition, Schlenker
et al. (2013) and Schlenker (2018a) expand on many of these ideas and implement the depictive component as a presuppositional restriction on locus use (similar to a gender presupposition on, e.g. English or French pronouns): the locus requires a mapping between its depiction and the referents in the state of affairs it is describing. If this presupposition holds, then the interpretation of indices via an assignment function proceeds exactly as in spoken language. Thus, the interpretation of depictive elements is generally taken to be separate and orthogonal to the many other active questions in this area of sign linguistics such as whether loci themselves are a visible way to track discourse referents as in dynamic semantic indices (Lillo-Martin & Klima 1990; Schlenker 2011; Steinbach & Onea 2016; Barberà 2015) which even in spoken languages is discrete and non-depictive, or more like semantic features in restricting but not determining discourse reference (Kuhn 2015; Neidle 2000). Neither are depictive so whatever the analysis, the purported analog component of the use of space in sign language loci that concerns Coppola & Senghas (2017) is and should be separable from its propositional contribution.

The last examples provided by Coppola & Senghas (2017) of analog compositional forms are in fact demonstratives themselves (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 making direct reference to a particular. Moreover, it’s known that the demonstrative can create 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 itself was part of the content of the language, but rather that the demonstrative can create a linguistic object using similarity to the demonstrated item (Umbach & Gust 2014), while the particular item is entirely external to the propositional component.

4.5. NON-COMPOSITIONAL DISCRETE FORMS. Finally, regarding the second dissociation that Coppola & Senghas (2017) propose, there may certainly also be discrete components of language that do not seem to participate in further semantic composition. They provide examples like Shhh or the OK emblem 👍, which indeed involve (iconic) description: 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 saturated, and hence do not further “compose” (Coppola & Senghas 2017), while at the same time conveying only propositional content; any iconicity is ignored in composition. We see then that potential
counterexamples to the requirement of discreteness for compositionality are either expected (these saturated discrete symbols) or are precisely the sort of phenomena that are contributing only via depiction to the representation of particulars or involve analyses via demonstratives (which we suggest make external reference to constructed particulars). Figure 1 presents a view of these possible points of interaction: at the top we see that simulated particulars, which can involve depiction, may influence the choice of functions for propositional content in cases of ambiguity, polysemy, etc. (e.g. in the extreme case, a baseball gesture will disambiguate an animal bat from a baseball bat). At the bottom, the compositional semantic system can refer via demonstratives to constructed particulars just as it can refer to particulars out in the world, and make new predicates (e.g. like that, Umbach & Gust 2014), or in the specific case of event particulars refer to demonstrations (Davidson 2015). Meaning via description affects both constructed particulars (one is chosen to exemplify the symbol) and propositional representations, while meaning via depiction can only affect/depict particulars.

5. Comparison of approaches to the question of iconicity and compositionality.

5.1. Current approach: Parallel representations. The broad view proposed here takes meaning expressed by depiction to be of a different sort entirely from propositional representations, although they can occur simultaneously. Is this reflected in any current approaches to compositional semantics of depictive content? The general approach is consistent with the one taken by Davidson (2015) and Maier (2018) for sign language role shift and quotation, as well as sign lan-
guage 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 actually an event (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. Extending this approach in other directions seems potentially fruitful: sign language verbs are known to display iconicity with respect to their underlying event structure (Strickland et al. 2015) in a way that seems to involve both depiction and description in terms of repetitions and extent of the event (Kuhn & Aristodemo 2017), roughly in the same way as is seen in spoken language ideophones (Dingemanse 2015). It seems then that a separation of depictive and descriptive components may provide more insight than the broader notion of iconicity that has collapsed them in the sign language literature, especially when comparing to hearing non-signers (Strickland et al. 2015) who might make very different uses of the two kinds of iconicity, since they will have access to depictive meaning but presumably not to descriptive meaning via conventionalization.

A criticism of this approach to iconic compositionality is that it doesn’t give a complete compositional semantics (Schlenker 2021), and in an important sense this is true: a parallel representational story for depictions vs. propositions is emphatically not a truth conditional semantics for pictures, as pictures are not taken to have truth conditional content. One committed to a truth conditional 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 semiotically distinct in kind 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 (especially via sign language, gestures, and ideophones), and the natural move by formal semanticists working with the tools of eliminative meaning/propositional semantics was to account for increasingly wide variety of phenomena using the same approach (Schlenker 2018b). This assumption has not been seriously examined in the case of iconic depictions across language modalities, and the argument here, instead, is that there is evidence for keeping them separate once we focus on semiotic contribution (depiction(description) instead of expressive modality (language/gesture).
Ontological differences are reflected in the syntax. Another recent formal proposal in which depictive content is viewed as having an incompatible type of content from description 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 conceptual composition, taking cues and insights from cognitive linguistics and related approaches to create a “picture” of the event along the same lines as the model-like cognitive linguistic representation suggested here; at higher syntactic levels, meaning is translated to external/truth conditions, via a quotational mode using the notion of a demonstration (Davidson 2015) and quantification over events (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 the goals of propositional content. 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 of negation and 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 account proposed in this paper remains conservative with respect to the decades of work in formal compositional semantics, while the sequential account diverges more radically for composition within the VP domain. A 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, along the same lines of cognitive linguists viewing all as contributing to a constructed meaning. 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 of rampant constructive meaning in the combination of a verb and its arguments 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 primary 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 conventionalized form for the ASL sign rainbow need not convey meaning about a particular rainbow RAINBOW, and only does so when they are specially marked as depictions, deviating from the conventionalized form. Similarly, the running manner in the Norwegian sign language sign for RUN or the manner of cradling in the ASL sign BABY, or the sound of the English word chirp need not necessarily reflect any aspects of the events conveyed by VPs in which they are components, since these can be used purely as symbols, unless they are intentionally modified from the conventionalized form to depict (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 contributes 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, but if that amendment is made then the power of the generalization of the syntactic locus for iconic content is significantly reduced.

5.1.2. EVERYTHING CONTRIBUTES TO ELIMINATIVE MEANING. 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 pictorial content involve a mapping projected from a particular viewpoint, following work on picture semantics by Greenberg (2013), in the tradition of philosophical approaches to pictorial 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, 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 under a view in which they all contribute propositional meaning, this difference must be explained either by prosody (they cannot take prosodic focus?), information structure (they aren’t questioned), language modality (“gesture” is always secondary to speech), or appeals to syntactic differences (Schlenker 2019; Esipova 2019b). Schlenker (2021) proposes a very general analysis in terms of presupposition:
when a depiction in gesture or sign language is not accurate, there is a presupposition failure, and negation does not tend to target presupposition content. 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 as presuppositions or assertions in linguistic utterances. We have argued that they should not be, in light of their inability to support the sort of content that motivates eliminative/truth conditional meaning, like polar questions, negations, etc. On the eliminative view, the lack of integration of depictive content with negation is handled as a property of the presuppositional nature of this content, given that presuppositions classically “project” through negation and questions, but being presuppositional is hardly explanatory, especially for content that cannot possible be lexically stipulated (Tieu et al. 2019). Instead, we propose parallel semantic representations which are separated by function (constructed particulars to experience, propositional content to eliminate possibilities) as a possible explanation for why depictive content appears to pattern with linguistic presuppositions in not being targeted by negation, questions, etc, due to its inability to support discrete partitions/alternatives.

A final 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 occurring in their own time slot and the former co-occurring 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. I want a coffee, not a 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 in 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 emphasized in this paper, the fact that such gestures occur in the visual modality along with speech doesn’t rule out interpretation as a symbol/description.

Both the purely eliminative view and the parallel representation view proposed here are conservative with respect to the semantics of descriptive content and its contributions to the elimination of possibilities within that tradition. 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 using an eliminative account of pictures (e.g. Greenberg 2013) for the pictorial 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 can support alternatives.

5.1.3. Everything contributes to construction of representations of particulars. A final family of approaches gives up on eliminative 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 by participants, 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 our metaphors often are visual and much of the work is in visual language, constructed representations need not be visual/imagistic, but rather could involve mental models with more abstract symbols, auditory stimulations, etc. The key is that they deal in particulars that are experienced, not in generalizations over sets of objects or worlds. 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 seem 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).

From the point of view of this paper, another major additional disadvantage of a purely construction-based/cognitive semantics approach to the semantics of depiction and description is that it is overly unificationical in assuming that description and depiction can contribute to the same kind of meaning: why do languages so consistently create two different routes if we’re ultimately trying to arrive at the same destination? The above proposal suggests a purpose to both depictive and descriptive content in language in their own ways, drawing on the strengths of truth-conditional approaches to meaning for descriptive language, while acknowledging the evidence in favor of both depiction and description as contributing to constructed particulars.

6. Conclusions. This paper argues that only by separating the semiotic functions of depiction and description can we understand the compositional, and noncompositional, properties of iconic language. It presents the argument in favor of separating (a) a constructive component of linguistic meaning that deals with particulars (which we can describe and depict, and interpret via perceptual
experience), and (b) an eliminative component of linguistic meaning that is built via description and interpreted via a logic and functions that link to concepts/generalizations. Where depiction appears to integrate with propositional meaning is better modeled as occurring simply in parallel representations or in some cases making use of (implicit or overt) demonstratives to refer outside of the propositional component to depictions.

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 intriguing. 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 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 iconicity in depiction may not contribute directly to meaning in a given particular, iconicity may be supported in the descriptive lexicon if it aids supporting depiction as well as perhaps aiding in language acquisition, cultural transmission, etc.

In addition to the empirical coverage (as in interactions of negation and questions with depictive content) gained by separating the functions of linguistic meaning into a positive particular type and an eliminative 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 redundant methods for information encoding. 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 constructive vs. eliminative content. In general, parallel processing is rampant throughout different cognitive domains generally so it should be unsurpris-
ing 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 2007a), 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 via elimination/generalizations. A key takeaway here is that symbolic propositional meaning comes part and parcel with the notion of supporting alternatives, itself 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; Beltrama 2020). We’ve also not yet addressed iconicity at levels above the meaning of an utterance, such as organization of a narrative (Cohn 2013; Abusch & Rooth 2017). 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), spoken languages (Clark 1996; Kita 1997; Dingemanse 2015), and even written language (Clark & Gerrig 1990; Davidson 2015; Maier 2018) 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 choose among 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


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.


Potts, Christopher. 2007b. The expressive dimension.


