Iconological Semantics*

Philippe Schlenker[a], Jonathan Lamberton[b]

December 16, 2023

Abstract. We argue that sign language requires a radical extension of formal semantics. It has long been accepted that sign language employs the same logical machinery as spoken language (occasionally making its abstract components overt), and simultaneously makes extensive use of iconicity. But the articulation between these two modules has only been discussed piecemeal. To capture it, we propose an 'iconological semantics' that combines standard logical semantics with a pictorial semantics in the Greenberg/Abusch tradition. We start by reanalyzing from this perspective earlier data on iconic loci, which are simultaneously variables and simplified depictions of their denotations. We then analyze new data on ASL classifier predicates, constructions that are lexically specified as being iconic. Their behavior argues for a very expressive system, possibly one in which the object language contains viewpoint variables. These can be left free or they may depend on quantifiers, and distinct viewpoint variables can co-occur in a given sentence; this gives rise to an extraordinary interaction between depictions and logical operators. We then sketch an adaptation of pictorial semantics to the dynamic 3D representations used in sign language. Finally, we suggest that iconological semantics might also illuminate the interaction between logical operators and pro-speech gestures in spoken language. In the end, the standard view of language as a discrete compositional system must be revised: it also has a tightly integrated depictive component, and 'textbook semantics' should be revised to capture this fact.

Keywords: sign language, ASL, iconicity, loci, iconic loci, classifier predicates, pictorial semantics, visual narratives, gestures, onomatopoeias

---

* **Authors' contributions:** Philippe Schlenker initiated this research, constructed all examples in consultation with Jonathan Lamberton, and developed the analysis. Jonathan Lamberton was the ASL consultant for the initial phase of the work. When it was written, he checked or provided transcriptions and translations, and discussed important aspects of the analysis. Any theoretical discussion among co-authors occurred only after the data were collected and the first version of the article was written. The separation between the two phases of the work was intended to minimize the risk of a 'theoretical contamination' of sign language judgments. In an additional phase of the work, Jonathan Lamberton also collected the judgments of our second consultant, Jason Lamberton, especially by comparison with his own.

Funding:
Schlenker: This research received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788077, Orisem, PI: Schlenker). Research was conducted at DEC, Ecole Normale Supérieure - PSL Research University. DEC is supported by grant FrontCog ANR-17-EURE-0017.

Acknowledgments: Special thanks to Léo Migotti and Matthew Loder for discussion of the gesture videos, to Emmanuel Chemla for critical discussion of some geometrical questions, and to Diane Lillo-Martin and Ora Matushansky for helpful remarks. Thanks also to Lucie Ravaux for help with the bibliography, the videos, and the verification/computation of the averages. The following colleagues kindly helped with the gesture and onomatopoeia survey: Anna Alsop; Naomi Francis; Matthew Loder; Matt Mandelkern; Pritty Patel-Grosz; Shane Steinert-Threlkeld; Lyn Tieu. Finally, we are extremely to two anonymous reviewers and to Managing Editor Regine Eckardt for extraordinarily constructive and thoughtful suggestions.

[a] Institut Jean-Nicod (ENS - EHESS - CNRS), Département d'Etudes Cognitives, Ecole Normale Supérieure, Paris, France; PSL Research University; New York University, New York.

[b] New York, NY, USA.

1 Introduction

1.1 Goals

We argue that sign language requires a radical extension of formal semantics. It has long been accepted that sign languages (i) employ the same types of grammatical and logical structures as spoken languages (occasionally making their abstract components overt, e.g. Lillo-Martin and Klima 1990, Sandler and Lillo-Martin 2006), and simultaneously (ii) make extensive use of iconicity (e.g. Cuxac 1999, Taub 2001, Liddell 2003b, Schlenker 2018b, Kuhn and Aristodemo 2017). But the articulation between these two modules has only been discussed piecemeal.[1] We argue that formal semantics must be extended with a depictive component inspired by Greenberg's and Abusch's pictorial semantics (Greenberg 2013, 2021; Abusch 2020), which offers a transparent treatment of the interaction between depictions, viewpoints and variables. This extension is what we call an 'iconological semantics'.[2]

Our first argument to integrate pictorial semantics to natural language semantics is that some sign language loci (i.e. positions in signing space that correspond to discourse referents) have the semantics of pictorial variables: their value is simultaneously constrained by an assignment function and by a depictive semantics. The existence of iconic loci was discussed in earlier research, but never integrated with a systematic pictorial framework. Our second argument is that some constructions (classifier predicates, or 'classifiers' for short) are lexically specified as having a depictive semantics. On the basis of new data from ASL (American Sign Language), we will show that their behavior argues for a very expressive system, possibly one in which the object language contains viewpoint variables. These can be left free or can depend on quantifiers, and distinct viewpoint variables can co-occur in the same sentence, including in the scope of quantifiers. This yields an extraordinary interaction between logical operators and iconic conditions. To complete our proposal, we will sketch an adaptation of pictorial semantics to the dynamic 3D representations used in sign language. We wil then show that iconological semantics has applications outside of sign language: in speech, it can help illuminate the interaction between logical operators and gestures, especially speech-replacing ones ('pro-speech gestures'). In the end, our analysis suggests that the view of language as a discrete compositional system must be revised: it also has a tightly integrated depictive component, and 'textbook semantics' should be revised to capture this fact.

To make our goals more concrete, we provide in (1)b the simplified Logical Form of part of the sentence in (1)a (a more literal translation would be: *In all class breaks, there is a student who…*; *there is* is expressed by *HAVE*).

(1) a. CLASS BREAK ALL ALWAYS HAVE STUDENT PERSON-walk-cl.
'In all classes, during the break, there is always a student who leaves along the right side [from the teacher's standpoint] to the far right.'
(ASL, 35, 2254b; anonymized video: https://youtu.be/2vnBm1UJkqI)
b. … always $\exists\pi$ there-is student **person-walk-cl**$^{\pi}$

The sentence in (1)a states, about a variety of classes held in different rooms, that during the break there is always a student that leaves. *Leave* is realized by a classifier predicate, representing a person moving toward an exit; we gloss it as *PERSON-walk-cl*. In the simplified Logical Form in (1)b, this expression carries a viewpoint variable π which simultaneously tells the interpretive system that the expression is iconic, and specifies the viewpoint with respect to which it is interpreted. The classifier moves on the signer's right toward the far right, and thus attributes a similar movement to the departing student: the classifier is interpreted iconically. Relative to which viewpoint (or 'camera position') is this movement interpreted? Plausibly, relative to the teacher's position in the front of the classroom. But the sentence

[1] For instance, the analyses discussed in Schlenker 2018b postulate *ad hoc* lexical rules to account for various iconic effects. This remark also applies to Kuhn and Aristodemo's analysis (2017) of iconic pluractionals, and to Schlenker and Lamberton's (2019, 2022) account of iconic plurals.

[2] The spelling *icono-logical* might convey more precisely what we have in mind, but lexical hyphens tend to have a short life expectancy, so we forego their use in this case. For an unrelated use of *iconological* and *iconology,* see: https://en.wikipedia.org/wiki/Iconology.

quantifies over a variety of classrooms in different configurations, and thus different viewpoints must be involved. In other words, the choice of the viewpoint must be dependent on the relevant classes. This is captured in (1)b because π is bound by the existential quantifier $\exists\pi$ (which comes with an implicit domain restriction), which is itself in the scope of *always*.

More generally, our proposal is that sign language semantics is in essence iconological semantics, or in other words the union of logical semantics and iconic semantics, identified here to pictorial semantics (*modulo* some important adaptations). There is much interesting action at the intersection between these two components, and thus both should be subject to the standards of explicitness that define contemporary formal semantics.

1.2 Structure

After briefly describing our elicitation methods and transcription conventions (in Section 1.3), we present the main conceptual and formal background (Section 2), pertaining to logical semantics in sign language, pictorial semantics in the Greenberg/Abusch tradition, and the integration we propose in this piece. We re-analyze from this new perspective the dual nature of sign language loci, which can both behave as logical variables and as simplified pictorial representations of their denotations (Section 3). We then turn to the iconic nature of classifier predicates, which is encoded by the presence of viewpoint variables in Logical Forms (Section 4). Viewpoint variables may be dependent on quantifiers (Section 5), and several viewpoint variables may co-occur in the same sentence, giving rise to a mix of perspectives (Section 6). To complete our account of sign language classifiers, we explain how the Greenberg/Abusch pictorial semantics can be replaced with 3D animations (Section 7). Finally, we argue that iconological semantics can also illuminate the behavior of pro-speech gestures in spoken language (Section 8), before drawing some conclusions (Section 9).

1.3 Elicitation methods and transcription conventions

The consultant (and co-author) is a Deaf, native signer of ASL (of Deaf, signing parents).[3] Elicitation was conducted using the 'playback method', described for instance in Schlenker et al. 2013, Schlenker and Lamberton 2019, 2022. It involved repeated quantitative acceptability judgments (on a 7-point scale with 7 = best), as well as inferential judgments (when the latter were quantitative, they were also on a 7-point scale, with 7 = strongest inference). Raw scores appear when there was more than a 2-point difference on the same question across different sessions. References such as (e.g. *ASL, 35, 1916, 4 judgments*) at the end of paradigms cross-reference the ASL video (here video 35, 1680) and indicate the number of iterated judgment tasks (on different days, here four). For clarity, we provide links to anonymized versions of the source ASL videos, and specialists are invited to consult the raw judgments in the Supplementary Materials B when relevant. After the data were collected, Jonathan Lamberton informally checked the judgments reported for original data (in numbered examples) with his brother Jason Lamberton, who is also a Deaf native signer; the main points of disagreements are noted as we go (in footnotes when they are inessential), and a full description appears in the Supplementary Materials B. Unless noted otherwise, the two consultants agreed with respect to the contrasts we report.

Transcription conventions are standard for sign language. Loci are alphabetized from dominant to non-dominant side from the signer's perspective (here: from right to left). A suffixed locus, as in *WORD–i,* indicates that the word points toward locus i (a position of signing space associated with a discourse referent). *IX-i* (for 'index') is an index pointing sign toward locus *i*. *EXPRESSION*$_i$ is used for some words associated with locus *i* by virtue of being signed in (rather than by pointing to) the corresponding area of signing space. Agreement verbs include loci in their realization—for instance, the verb *a-HIT-b* starts out from locus *a* and targets locus *b*. We put *-cl* at the end of classifier predicates, and describe their movement with affixed words (e.g. *PLANE-veer_right-cl*) and/or with iconic specifications (e.g. *PERSON-move-cl-c--->|<--- PERSON-move-cl-a+rotation* represents two person

[3] We use the term *consultant* to refer to a collaborator that assesses sentences, including if this person is also a contributor to the article.

classifiers moving toward each other from positions *a* and *c*, and then effecting a rotation).[4] The anonymized videos should be consulted for further details.

2 Logical vs. Pictorial Semantics

To clarify the conceptual issues, we start from simplified examples that will illustrate the difference between a logical semantics, a pictorial semantics, and the 'iconological' mix we advocate for sign language.

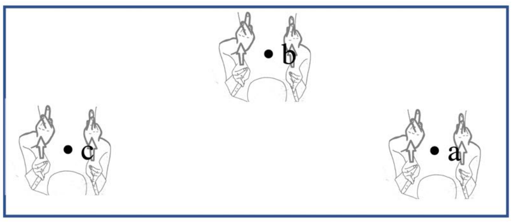
2.1 Logical semantics in sign language

While logical semantics is usually stated for spoken or written forms, it has also been applied to sign languages, with the general (if still debated) assumption that loci may be the visible instantiation of logical variables (Lillo-Martin and Klima 1990, Schlenker 2011b; see also Kuhn 2016 and Schlenker 2016 for alternative or expanded analyses). To make things concrete, let us consider the sentence in (2), a greatly simplified version of examples we will discuss in greater detail below. The word *PERSON* is a standard noun (not a classifier predicate), illustrated in (3). The iterations of *PERSON* introduce three arbitrary loci *a*, *b*, *c*, distinguished by their positions in signing space, as illustrated in (4). The trial pronoun *THE-THREE-a,b,c* then forms the sum of these variables, and the modified verb *PAINT TOGETHER* is predicated of this sum.

(2) PERSON$_a$ PERSON$_b$ PERSON$_c$ THE-THREE-a,b,c PAINT TOGETHER. (simplified sentence[5])
'Three individuals (literally: a person and a person and a person) paint together.'

(3) *PERSON* in ASL[6] (addressee's viewpoint)

(4) Three occurrences of *PERSON* co-occurring with loci (signer's viewpoint)



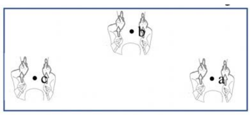
The truth conditions are derived in (5) by a standard semantics that includes quantifiers and variables, and is relativized to an assignment function s, in addition to a context c, a time t and a world w. We simplicity and perspicuity, we treat indefinites as introducing variables bound by unselective

[4] When citing examples from the earlier literature, we have preserved the original glossing conventions, with the result that they may differ in minor respects from those of the present piece; this is unlikely to cause confusions.
[5] While greatly simplified, this example comes closest to (i) below, discussed in (150)c in Supplementary Materials A-1.

(i) POSS-1 CLASS ALWAYS HAVE FOREIGN PERSON-cl$_a$ PERSON-cl$_c$ AMERICAN PERSON-cl$_b$, ALWAYS END THE-THREE-a,b,c PAINT TOGETHER.
'In my class, there are always two foreigners standing in front of American. The three of them always end up painting together.' (ASL, 35, 2098c)
[6] Picture credit: Valli et al. 2005 p. 333.

quantifiers, *à la* Heim 1982, hence the simplified Logical Form in (5)a', and the truth conditions in (5)b (we also introduce covert—and barred—words, e.g. ~~*and*~~, to keep the analysis maximally simple).

(5) **Logical semantics in sign language**

a.  THE-THREE-a,b,c, PAINT-TOGETHER.

a'. $\exists_{a,b,c}$ [person$_a$ ~~and~~ person$_b$ ~~and~~ person$_c$ ~~and~~ pro$_{a,b,c}$ argue]

b. $[[(a')]]^{c, s, t, w} = 1$

iff for some d, d', d", $[[\text{person}_a \text{ and person}_b \text{ and person}_c \text{ and pro}_{a,b,c} \text{ argue}]]^{c, s[a\rightarrow d, b\rightarrow d', c\rightarrow d''], t, w} = 1$,

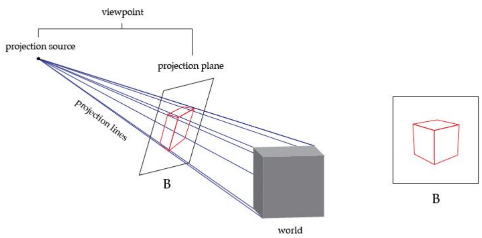
iff for some d, d', d", $\text{person}'_{t,w}(d) = \text{person}'_{t,w}(d') = \text{person}'_{t,w}(d'') = 1$ and $\text{paint-together}'_{t,w}(d+d'+d'') = 1$

In words: There are individuals d, d', d", and the group d+d'+d" plays together.

2.2 *Greenberg's pictorial semantics*

Pictorial semantics produces semantic information in a completely different way. The basic intuition is that a picture is true of all situations that can geometrically project onto it in view of a given projection method. Let us take the example of perspective projection. Following Greenberg 2013, 2021, we start from the notion of a viewpoint, made of a projection source (the center of the projection) and a projection plane (the picture), as is illustrated in (6). A picture is true of those worlds that can project onto the plane.

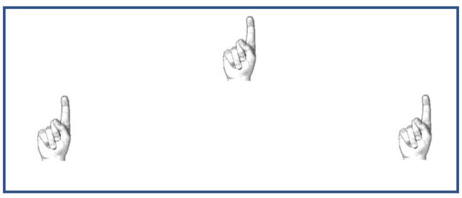
(6) **An example of a projection method: perspective projection (Greenberg 2021)**



Let us apply this idea to the example in (7)a, viewed as a pure picture. It is true of those worlds that contain three raised index fingers arranged as a triangle, in a position that allows them to project onto the picture. The detailed truth conditions obtained depend on the 'marking rules'—for instance, to the effect that the edges of objects are marked on the projection plane, as in (6). If π is the relevant viewpoint, and t and w are the time and world of evaluation, the truth conditions of the picture can be given as in (7)b, where $proj(\pi, t, w) = (a)$ means that relative to viewpoint π (and to the specified method of projection), w projects to the picture in (a) at time t.

(7) **Pictorial semantics (Greenberg)**

a.



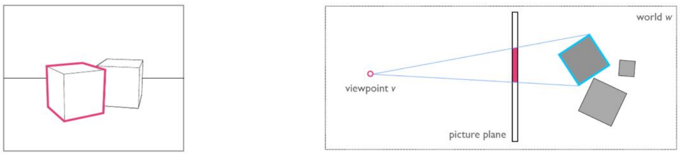
b. $[[(a)]]^{\pi, t, w} = 1$ iff $proj(\pi, t, w) = (a)$

iff at t in w three index fingers form a triangle that projects onto (a) from π

In words: There are three index fingers that form a triangle as depicted.

While in (7) only the picture as a whole has a semantic content, Greenberg 2014 defines a further notion of reference for a picture *part*—a notion we will soon make use of. The intuition is illustrated in (8): the picture in (8)a is true of the situation in (8)b, but in addition the left-most shape in (8)a denotes the top-most cube in (8)b.

(8) **The picture in a. is true of the situation in b., and the left-most shape in a. denotes the top shape in b.**
a. Picture b. Situation



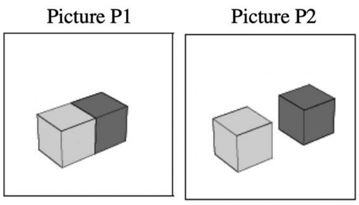
Without giving a fully specified definition (discussed in Greenberg 2014), we will initially take the notion of the projection of an object d onto a picture part p to be primitive (we will modify this point in Section 7). Extending our earlier *proj* notation by adding to it an individual argument, *proj(**d**, π, t, w) = **p*** means that that the *object* d (boldfaced) projects to the picture part p (boldfaced as well) relative to π, t, w, as specified in (9).

(9) If d is an object, π is a viewpoint, t is a time and w is a world, and p is a distinguished picture part (called a 'grapheme' in Greenberg 2014),
proj(d, π, t, w) = p if and only if d projects to p at t in w from viewpoint π

2.3 Abusch's's pictorial semantics with variables

This is not the end of the pictorial story, however. Abusch (e.g. 2013, 2020) has argued that in pictorial narratives, there are ambiguities of cross-reference that can insightfully be captured by positing that pictures can in effect contain variables. To see her motivation, consider the 2-picture sequence in (10), which represents "a short comic of two cubes moving apart".

(10) **Two cubes moving apart** (Abusch and Rooth 2017)

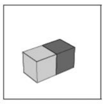
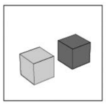


With pictorial semantics alone, and thanks to its relativization to times, the semantics in (7) yields in essence the following truth conditions:

(11) If <P1, P2> is a sequence of two pictures, and if π, t, w are a viewpoint, a time and a world respectively,
$[[<P1, P2>]]^{\pi, t, w} = 1$ iff for some time t' such that t' > t, $[[P1]]^{\pi, t, w} = 1$ and $[[P2]]^{\pi, t', w} = 1$

But this is too weak to capture the most salient reading of this little comic, as Abusch argues. Intuitively, one naturally understands P2 to involve the same cubes as appear in P1, but the semantics in (11) implies no such thing: it just requires that the cubes in P2 *resemble enough* those in P1 to yield the same projection, *modulo* the change of place. One might object that the enrichment whereby the two cubes in the second picture are the same as those in the first picture is just due to common sense reasoning, a process that can apply to the output of any expressive form to strengthen its truth conditions. But things are not so simple: the enrichment can be computed in the scope of operators, as is the case in (12), which is plausibly true if the sequence is understood with the intended coreference, possibly false otherwise.

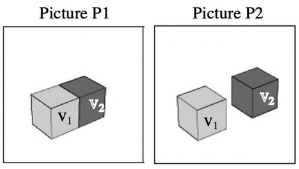
(12) At the beginning of this abstract comic, two cubes led a happy life together. But not all is well – tensions have been mounting between them… what will happen next?

If what happens next is that  , a cube counselor will be needed.

Since this enrichment is computed in the scope of the *if-clause*, it weakens the global truth conditions (because *if*-clauses are a downward-monotonic environment; see also Schlenker 2019b for a similar discussion).

Abusch's solution is to enrich pictures with variables: enrichment boils down to the resolution of an anaphoric ambiguity. Following the spirit of Abusch's analysis, but the implementation of Schlenker 2019b, 2022, we can identify variables to picture parts, and represent the natural reading of (10) and (12) as in (13). The variable v_1 enforces coreference between the light cubes of P1 and P2, and v_2 enforces coreference between the dark cubes of P1 and P2 (see also Maier and Bimpikou 2019 for a related implementation).

(13) Two cubes moving apart – with variables added



The key idea is that for a picture such as P1 to be true, it is not enough that appropriate edges project onto it; in addition, these should be the edges of the objects specified by the assignment function, as is schematically illustrated in (14)a for the interpretation of P1. The extension to pairs of pictures construed as narratives is immediate, as seen in (14)b. Here we make use of the Greenbergian notion of an object projecting onto a picture part (for Greenberg, a 'grapheme'), as in (9); and we take variables to just be picture parts, a point we will refine below.

(14) a. If π, s, t, w are a viewpoint, an assignment function (assigning values to picture parts), a time and world:
$[[P1]]^{\pi, s, t, w} = 1$ iff $proj(\pi, t, w) = P1$ and **in P1, $proj(s(v_1), \pi, t, w) = v_1$ and $proj(s(v_2), \pi, t, w) = v_2$**
b. $[[<P1, P2>]]^{\pi, s, t, w} = 1$ iff for some time t' such that t' > t, $[[P1]]^{\pi, s, t, w}$, and= 1 and $[[P2]]^{\pi, s, t', w} = 1$

When no assignment function is initially specified, it is natural to take a picture or pictorial narrative to be true (at a time and world) just in case it is true on *some* assignment function. This yields the desired coreference relations for (13), as illustrated in (15), where the boldfaced parts enforce coreference.

(15) <P1, P2> is true relative to π, t, w
iff for some assignment function s, $[[<P1, P2>]]^{\pi, s, t, w} = 1$,
iff for some assignment function s, for some t' > t, $[[P1]]^{\pi, s, t, w} = 1$ and $[[P2]]^{\pi, s, t, w} = 1$,
iff for some objects d and d', (i) $proj(\pi, t, w) = P1$ and **in P1, $proj(d, \pi, t, w) = v_1$ and $proj(d' \pi, t, w) = v_2$**, and (ii) $proj(\pi, t', w) = P2$ and **in P2, $proj(d, \pi, t', w) = v_1$ and $proj(d', \pi, t', w) = v_2$**

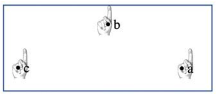
2.4 Pictorial semantics with variables in sign language

At this point, Greenberg's and Abusch's innovations are entirely motivated by pictures, and have nothing to contribute to the semantics of the ASL sentence in (5). But things radically change when we consider a minimal variation of (5) involving a classifier predicate, i.e. a construction whose semantics is highly iconic, such as the person classifier in (16). It is the very shape we used in (7), but it has a conventional meaning: it usually represents a person in upright position, not a finger.[7] However the position and orientation of the classifier are interpreted iconically; specifically, the position of the finger in signing space tracks that of the denoted person, and the finger front represents the person's front.

[7] If this classifier represents a person, the latter must be standing. This classifier is also used to depict various items that are tall and slim, such as posts or pencils.

(16) A classifier in ASL (e.g. representing an upright person)[8]

By replacing the three occurrences of *PERSON* in (2) with three person classifiers in the same configuration, as in (17), we obtain a different meaning, which provides geometric information about the relative position of the three individuals that painted together.

(17) 

THE-THREE-a,b,c, PAINT-TOGETHER. (simplified sentence[9])
'Three individuals arranged as a triangle [as shown] painted together.' (ASL)

Since the finger classifier has a conventional and an iconic component, we will need a mixed notion of projection. In a nutshell, it will resemble the Greenbergian notion in being based on geometric transformations, but the marking rules will be different (we will just require that the face of the person project onto a designated part of the classifier, a point that will be made precise in a 3D framework in Section 7). Importantly, the three finger classifiers in (17) are associated with loci that play the role of variables. This combination will end up making the same kind of contribution as Abuschian pictorial variables, with both a variable and an iconic component.

Specifically, we take the finger classifier to be an iconic predicate, taking as argument the variable it is associated with. To avoid inserting pictures in Logical Forms, we represent the finger classifier as *person-cl*, combined with a viewpoint variable π to indicate that it must be interpreted iconically, and relative to the viewpoint denoted by π.[10] If the classifier is associated with a locus *a*, the latter is written as a subscript, hence: $person\text{-}cl^{\pi}_{a}$. This combination will in the end be interpreted as having a lexical component, a variable component and a projective component, as stated in (18) (which will later be derived from interpretive rules for its component parts $person\text{-}cl^{\pi}$ and *a*).[11]

Notation: We adopt the standard convention of writing as *word'* the semantic value of an expression *word*. When it is evaluated with respect to a time t and a world w, we write is as: $word'_{t,w}$.

(18) [[]]$^{c, s, t, w}$ = $[[\text{person-cl}^{\pi}_{a}]]^{c, s, t, w} = 1$ iff $\text{person}'_{t,w}(s(a)) = 1$ **and proj(s(a), s(π), t, w) = person-cl**

We will think of $person\text{-}cl^{\pi}_{a}$ as an indefinite in the system of Heim 1982, and its variable *a* will be bound by an unselective existential quantifier, yielding the Logical Form in (19)a' for the sentence in (19)a. Cutting some corners, this produces appropriate truth conditions, as sketched in (19).

(19) a. PERSON-cl$_a$ PERSON-cl$_b$ PERSON-cl$_c$ THE-THREE-a,b,c, PAINT-TOGETHER.
a'. $\exists_{a,b,c}$ [person-cl$^{\pi}_{a}$ ~~and~~ person-cl$^{\pi}_{b}$ ~~and~~ person-cl$^{\pi}_{c}$ ~~and~~ pro$_{a,b,c}$ argue]
b. $[[(a')]]^{c, s, t, w} = 1$
iff for some objects d, d', d", [[person-cl$^{\pi}_{a}$ ~~and~~ person-cl$^{\pi}_{b}$ ~~and~~ person-cl$^{\pi}_{c}$ ~~and~~ pro$_{a,b,c}$ argue]]$^{c, s[a\to d,\, b\to d',\, c\to d''], t, w} = 1$,
iff at t in w, some objects d, d', d" are people that project to positions corresponding to a, b, and c respectively from viewpoint s(π), and d+d'+d" paint together

[8] Picture credit: Bill Vicars, https://www.lifeprint.com/asl101/pages-signs/classifiers/classifiers-main.htm
[9] This example comes closest to the far more complicated example in (150)a in the Supplementary Materials A-1.
[10] We reserve capitals for sign language glosses (or gesture glosses, with a different font), and use lowercase words for Logical Forms.
[11] For variable-free reinterpretations of systems with variables, see for instance Quine 1960, Cresswell 1990, Jacobson 1999. See Kuhn 2016 for a variable-free interpretation of loci. For a broader discussion of mixes of iconic and symbolic systems, see Greenberg, to appear.

The crucial observation is that ASL has constructions that have a purely logical semantics, as in (2), and constructions that have both a logical and a pictorial component, as in (17). The existence of the latter requires a pictorial turn in natural language semantics: the question is not so much whether sign language has a pictorial component (it does), but rather how it is integrated with the logical component.

2.5 Viewpoint variables in the object language

In (19)a', including viewpoint variables in the Logical Form was just a convenient way to signal to the interpretive system that the classifier predicates have a pictorial component. But we will discuss data that suggest that these viewpoint variables are crucial in two kinds of cases. First, there are instances of existential quantification over viewpoints, including in the scope of other operators, as was the case in the Logical Form of (1)a, repeated in (20)a. Second, there are cases in which two classifiers appear in the same sentence but carry different viewpoint variables, as is illustrated in (20)b, where the signer's position plausibly corresponds to the teacher's viewpoint in different rooms (hence different 'camera positions' for the interpretation of the classifier).

(20) a. Existential quantification over viewpoints
always $\exists\pi$ there-is student **person-walk-cl**$^{\pi}$
≈ there is always a student who, from some salient viewpoint, walks as shown
b. Multiple viewpoint variables
there-is student λi always$_T$ t_i **person-walk-cl**$^{\pi}$ and always$_{T'}$ t_i **person-walk'-cl**$^{\pi'}$
≈ there is a student who, in certain situations, always walks as shown from viewpoint π, and in other situations always walks as shown from viewpoint π'

We will develop our account in two steps. We will start with the simplest case of an interaction between iconic and pictorial semantics, namely iconic loci, which were argued in earlier work to be simultaneously variables and simplified pictures of their denotations; we will reinterpret earlier results in terms of our iconological semantics. We will then turn to new data about classifier predicates, which re-establish their depictive character and show that they can sometimes make use of existential quantification over viewpoints and of multiple viewpoint variables.

3 Loci in Iconological Semantics

As described in earlier literature, loci can arguably behave as the overt realization of variables, and they can sometimes simultaneously function as simplified pictorial representations of their denotations. Let us briefly go through each property in turn, using iconological semantics as a unifying framework.[12]

3.1 Arbitrary loci as logical variables

Lillo-Martin and Klima 1990, followed by a long line of research, took loci to be, in some cases at least, the overt instantiation of logical variables. Schlenker 2011b further argued that these should be seen as variables within a dynamic semantics (e.g. Kamp 1981, Heim 1982), one in which a variable may be bound by an existential quantifier without being in its c-command domain. To illustrate the simplest (non-dynamic) cases, we can consider (21) and (22) (from ASL and LSF respectively), where two loci *a* and *b* are introduced by antecedents and retrieved by pronominal forms, realized as indexes (pointing signs). The upper limit on the number of loci seems to be due to considerations of performance.[13]

---

[12] A further complication, which we disregard here, is that arbitrary loci can stand in subpart relation with other loci. For instance, if a plural locus a is a subpart of a plural locus b, the denotation of a is understood to be a mereological part of the denotation of b. Since we are interested in spatial rather than mereological relations, we disregard this complication here. See for instance Schlenker et al. 2013 and Kuhn 2021 for examples and discussion. For applications of the notion of loci to a completely different visual domain, dance, see Patel-Grosz et al. 2018, 2019.
[13] An example with 7 loci is given in Schlenker 2017c.

(21) a. IX-1 KNOW BUSH$_a$ IX-1 KNOW OBAMA$_b$. IX-b SMART BUT IX-a NOT SMART.
'I know Bush and I know Obama. He [= Obama] is smart but he [= Bush] is not smart.'
b. IX-1 KNOW PAST PRESIDENT IX-a IX-1 KNOW NOW PRESIDENT IX-b. IX-b SMART BUT IX-a NOT SMART.
'I know the former President and I know the current President. He [= the current President] is smart but he [=the former President] is not smart.'
(ASL, 4, 179; Schlenker 2011b)

(22) DEPUTY$_b$ SENATOR$_a$ CL$_b$-CL$_a$ IX-b a-TELL-b IX-a / IX-b WIN ELECTION
'An MP$_b$ told a senator$_a$ that he$_a$ / he$_b$ (= the deputy) would win the election.' (LSF; 4, 233; Schlenker 2018b)

The similarity between the behavior of loci and that of logical variables can, depending on the account, be interpreted more or less literally, as summarized in (23).

(23) **Variable Visibility**
a. **Weak Version**
In sign language, loci in signing space can be associated both to an arbitrary number of pronouns and to their antecedents to mark their dependency. Furthermore, deictic pronouns that refer to different objects may be associated to different loci.
b. **Strong Version**
In sign language, some uses of loci display the behavior of logical variables, both in their bound and in their free uses.
(Schlenker 2018b)

The close association between loci and pronominal reference has made the statement in (23)a uncontroversial. The stronger view in (23)b is that loci sometimes literally behave as logical variables—a more debated idea. Importantly, nobody claims that variables *must* be overt in sign language: there are numerous null pronouns, and there is no reason at all to assume that the phenomenon of null variables isn't as common as it is in spoken language.

We henceforth adopt the loci-as-variables theory (see Kuhn 2016 for an alternative in which loci are treated as agreement markers, and Steinbach and Onea 2016 for one in which discourse referents can be mapped to areas of space). Correspondingly, loci carried by pronouns such as the pointing sign *IX* are interpreted by the rule in (24).

(24) If c, s, t, w are a context, assignment function, time and world respectively, then:
$[[\text{IX-i}]]^{c, s, t, w} = [[\text{pro}_i]]^{c, s, t, w} = s(i)$

Thus the end of the sentence in (21)a, repeated in (25)a, can be assigned the simplified Logical Form in (25)a', where pro_b is a pronoun carrying an index *b*. It can be interpreted by standard rules of a type-theoretic semantics (e.g. Heim and Kratzer 1998), as illustrated in (25)b.

(25) a. IX-b SMART (…BUT IX-a NOT SMART).
a'. pro$_b$ smart
b. $[[(a')]]^{c, s, t, w} = [[\text{smart}]]^{c, s, t, w} ([[\text{pro}_b]]^{c, s, t, w}) = \text{smart}'_{t,w}(s(b))$

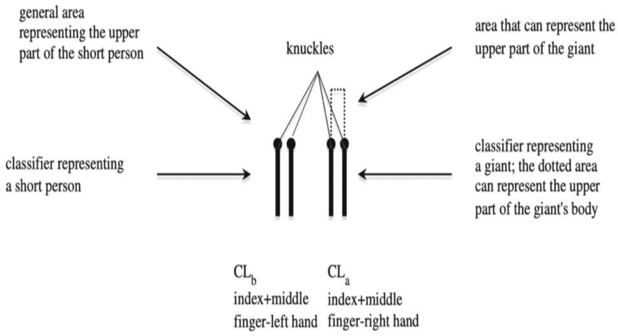
3.2 Iconic loci as pictorial variables

Numerous cases have been described in which loci simultaneously function as variables and as simplified pictures of their denotations (Schlenker et al. 2013, Schlenker 2014, Schlenker 2015). A radical case involved loci introduced by person classifiers, as in (27): two classifiers (each realized with two merged upright fingers, as illustrated in (26)a) served to represent two astronauts training in various rotated positions (illustrated in (26)b). More specifically, CL_a is a classifier on the right, representing a tall astronaut, while CL_b is a person classifier on the left representing a short astronaut.[14] Pointing targeted the position of the classifier that depicted the person's head. This had two consequences: with the astronauts represented in upright position, pointing was higher for the tall than

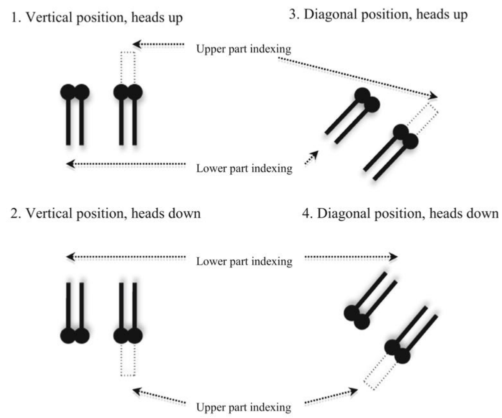
[14] Here the high locus is transcribed as *upper_part* because, after rotation of the classifiers, it finds itself low; but it is the same thing as what is otherwise transcribed as *high*.

for the short astronaut; and as the astronauts were represented in various rotated positions, the pointing followed the position of their heads (the elided part of the sentence was intended to show that these iconic specifications can be disregarded under binding in ellipsis, just like pronominal features can be in English—but we set this point aside in the present discussion).

(26) a. Giant and short astronaut: schematic representation of two classifiers (the signer's perspective)



b. Tall vs. short person rotation - schematic representation (signer's perspective)



(27) HAVE TWO ROCKET PERSON [ONE HEIGHT]$_a$ [ONE SHORT]$_b$. THE-TWO-a,b PRACTICE DIFFERENT VARIOUS-POSITIONS [positions shown].
IX-a HEIGHT IX-b SHORT, **CL_a-[position]-CL_b-[position].**
'There were two astronauts, one$_a$ tall, one$_b$ short. They trained in various positions [positions shown]. They were in [description of the position] position.

IX-a_upper_part LIKE SELF-a_upper_part. IX-b_lower_part NOT.
The tall one liked himself. The short one didn't (like himself).'
(ASL, 17, 178; Schlenker 2014)

Importantly, several examples have been described in the literature in which high and iconic loci are not introduced by classifier predicates, showing that the phenomenon is semantically rather than morphologically determined; in other words, it is not classifier predicates alone that are responsible for the appearance of iconic loci (Schlenker et al. 2013, Schlenker 2015).

Earlier analyses were *ad hoc*. Schlenker et al. 2013 just stipulated that the height of the locus should stand a proportional relation to the height of the denoted object. Schlenker 2014 came close to the present account, as it invoked Greenbergian geometric projections, but without an articulated account of viewpoints or viewpoint variables. From the present perspective, the iconic loci in (27) are simultaneously constrained by an assignment function and by a geometric condition: they are pictorial variables. To be concrete, in order to provide a full explication of the bound reading of (27), we will need to analyze the target part, repeated in (28)a, as in (28)a', where we treat *IX-a* and *SELF-a* alike as occurrences of a simple pronominal *pro*. We just write *a* for the locus they index, assuming it is high, and we represent the bound reading as in (28)a'; if we aimed to represent a mere coreferential reading, without binding, we could posit the simpler Logical Form in (28)b.

(28) a. IX-a-upper_part LIKE SELF-a_upper_part
a'. pro_a^π $\boxed{\lambda a\ t_a \text{ like } pro_a^\pi}$
b. pro_a^π like pro_a^π

Unlike the simplified Logical Form in (25), the Logical Forms in (28) do not include pronouns of the form *pro_a,* but rather *pro_a^π*, where π is a viewpoint variable that indicates that an expression is interpreted iconically, and specifies the relevant viewpoint. The key interpretive rule is stated in (29)a: it makes reference to the notion of an object projecting onto a picture part. We also need an auxiliary assumption, stated in (29)b, to the effect that in this case one considers the projection of the person's head (this will fall under a more general framework in Section 7).[15]

(29) **Iconic loci (initial version)**

a. $[[pro_a^\pi]]^{c, s, t, w}$ = # iff proj(s(a), s(π), t, w) ≠ a
s(a), otherwise

b. Auxiliary assumption: if a is a locus and x is a person, proj(x, s(π), t, w) = a if and only if the head of x projects to a.

In the pronominal case (not in general), we take the iconic component to make a presuppositional contribution, just like English gender features on a pronoun: in both cases, a referential failure (encoded as #) is obtained when the denotation of the variable does not satisfy the specification, and when no failure arises, the value of the pronoun is given by the assignment function[16] (e.g. Cooper 1983, Heim and Kratzer 1998; see Appendix I-A for a further unification between iconic and gender specifications).

For high pointing and gender features alike, the argument for a presuppositional treatment lies in the interaction between pronouns and logical operators, notably negation. Thus in (30)a,b, pronouns in the scope of negation trigger an inference that the denoted person satisfies the high specification or the feminine feature of the pronoun, as the case may be.

(30) a. YESTERDAY IX-1 SEE R [= body-anchored proper name]. IX-1 NOT UNDERSTAND $IX\text{-}a^{high}$.
'Yesterday I saw R [= body-anchored proper name]. I didn't understand him.'
(ASL, 11, 24; Schlenker et al. 2013))
=> R is tall, or powerful/important
b. I don't understand her.
=> the person denoted by *her* is female

With the rule in (29), we can derive the truth conditions of the Logical Form in (28)b following textbook semantics (e.g. Heim and Kratzer 1998, but using # for partiality). The key assumption is that a predicate denotation returns # as soon as one of its argument is #. Together with the lexical entry in (29)a, this guarantees that the entire sentence presupposes that the agent is in a position that corresponds to that of the locus.

(31) $[[(28)b]]^{c, s, t, w} = [[pro_a^\pi \text{ like } pro_a^\pi]]^{c, s, t, w}$
$= like'_{t,w}([[pro_a^\pi]]^{c, s, t, w})([[pro_a^\pi]]^{c, s, t, w})$
= # if proj(s(a), s(π), t, w) ≠ a (where a is placed high)
$like'_{t,w}(s(a))(s(a))$, otherwise

The Logical Form with λ-abstraction yields the same result but in a more complicated way, needed to capture the bound reading of the elided VP. Here we must first compute the value of the λ-abstract, as in (32), after which it can be fed its argument $[[pro_a]]^{c, s, t, w}$, as in (33).

---

[15] An anonymous reviewer notes that we could adopt a weaker view on which loci are just closely associated with standard variables, in which case an index i would appear in addition to a locus a, with the interpretive rule in (i):

(i) $[[pro^{i\pi}_a]]^{c, s, t, w}$ = # if proj(s(i), s(π), t, w) ≠ a
s(a), otherwise

On this view, a variable such as *i* is never overt, but it may co-occur with a locus *a*, in which case s(i) should project to *a*. This analysis could be extended to Abuschian pictures (and might be in the spirit of Abusch's original implementation). But one would need to explain how arbitrary sign language loci work on this theory, since these loci are not iconic and do not carry the viewpoint variable π, but still display a variable-like behavior. We leave this alternative implementation for future research.

[16] In greater detail: the rule in (29) is parallel to that in (i) below, where *fem* is a feminine feature appearing on a pronoun (the fact that we write it as a superscript should not suggest that it is treated as a viewpoint variable, of course). On the right-hand side, *fem'* stands for the semantic interpretation of the feminine feature *fem*.

(i) $[[pro_a^{fem}]]^{c, s, t, w}$ = # if $fem'_{c, t, w}(s(a)) \neq 1$
s(a), otherwise

(32) On the assumption that $\text{like}'_{t,w}$ (the value denotation of *like*) returns # just in case one of its arguments is #:

$[\![\lambda a\ t_a \text{ like } pro_a{}^{\pi}]\!]^{c, s, t, w}$

$= \lambda d.\ \text{like}'_{t,w}([\![pro_a{}^{\pi}]\!]^{c, s[a\to d], t, w})([\![t_a]\!]^{c, s[a\to d], t, w})$

$= \lambda d.$ # if d = # or $[\![pro_a{}^{\pi}]\!]^{c, s[a\to d], t, w} = \#$, i.e. d = # or proj(s(a), s(π), t, w) $\neq$ a
$\text{like}'_{t,w}(d)(d)$, otherwise

$= \lambda d.$ # if d = # or proj(s(a), s(π), t, w) $\neq$ a;
$\text{like}'_{t,w}(d)(d)$, otherwise

(33) If s(a) $\neq$ #,

$[\![(28)a']\!]^{c, s, t, w} = [\![pro_a{}^{\pi}\ \ \lambda a\ t_a \text{ like } pro_a{}^{\pi}]\!]^{c, s, t, w} = [\![\lambda a\ t_a \text{ like } pro_a{}^{\pi}]\!]^{c, s, t, w}([\![pro_a]\!]^{c, s, t, w})$

= # if proj(s(a), s(π), t, w) $\neq$ a
$\text{like}'_{t,w}((s(a))(s(a))$, otherwise

It should be added that the projective condition leaves it to the context to determine the value of the viewpoint variable π, as is standard for free variables in general. For instance, *She$_i$ is tall* uttered out of the blue requires that the context specify an initial assignment function determining the value of the variable *i* (see for instance Heim and Kratzer 1998). There are certainly contextual constraints on such values, including for viewpoint variables. For instance, viewpoints should presumably correspond to imaginable positions of a viewer in normal position (if the viewer were upside down, the representation of scenes would be turned upside down as well). As a result, if a locus is positioned high in signing space, this will yield the inference that the head of the denoted person is positioned higher than the salient viewpoint provided by the context. This seems appropriate for the examples discussed here. (A further question is why the elided part of (27) does not give rise to a presupposition failure; see Schlenker 2014 for an analysis in terms of feature deletion under agreement.)

3.3 Iconic loci moving in space

If Abuschian variables are identified with picture parts, we expect them to move from one picture to the next in examples such (13), since the denotation of the variable v_1 must project to different parts of P1 and P2, and similarly for v_2. Strikingly, under certain conditions, sign language loci can also 'move in signing space', a phenomenon that is sometimes called Locative Shift (Schlenker 2018a). We restrict attention to iconic loci, and focus for simplicity on examples with classifier predicates (examples in which the moving loci are not introduced by classifier predicates are discussed in Schlenker 2015).

In (34), a crocodile movement toward a ball is represented by a classifier predicate. The initial crocodile locus, namely *a*, can be used to effect coreference by way of a pointing sign, as shown in (34)a. But one can also express coreference by targeting the final position of the crocodile, corresponding to locus *b*: (34)b thus has a reading on which the crocodile got sick (rather than: the ball got sick, as might be expected in view of the fact that *b* is initially associated with the ball). This phenomenon is semantically conditioned: this possibility becomes degraded if one *denies* that the crocodile moved toward the ball, as in (35)b.

(34) YESTERDAY CROCODILE$_a$ BALL$_b$
a-CRAWL-SWALLOW-cl-b. FINISHED
a. $^{6.7}$ IX-a SICK.
b. ^{7} IX-b SICK.
'Yesterday, a crocodile went to a ball and swallowed it. And in the end it [= the crocodile] got sick.'
(ASL, 35, 2330; 3 judgments; https://youtu.be/I9ZyK_9zewk; Schlenker et al., to appear)

(35) MOST CROCODILE LOVE EAT BALL. YESTERDAY CROCODILE$_a$ BALL$_b$ BUT NOT a-CRAWL-SWALLOW-cl-b. SO FINISHED
'Most crocodiles love to eat balls. Yesterday there was a crocodile and a ball, but the crocodile didn't go to the ball to swallow it. So in the end
a. ^{7} IX-a FINE.
it [= the crocodile] was fine.'
b. $^{5.3}$ IX-b FINE.
it [= the ball [2/3 judgments] or the crocodile [1/3 judgment]] was fine.'
(ASL, 35, 2334, 3 judgments; Schlenker et al.t, to appear)

To capture these readings, we must take into account the fact that one and the same variable is associated with different picture parts at different points in time—something that was already the case in our pictorial example in (10). So rather than taking a variable to be a designated picture part, we should take a variable to be a function from signing times to picture parts.[17] To take this point into account, we will write these more abstract loci (technically, functions from times to loci) with capital letters, hence, A, B, C instead of a, b, c. It is to these more abstract loci that assignment functions will assign values, but projective conditions will make reference to variable tokens realized in a particular part of signing space. If abstract locus A always corresponds to position a, this won't make a difference relative to our earlier examples. But if locus A 'moves' from position a to position b, the choice between the a token and the b token will make a difference to the projective conditions. We will thus write as A_a an instantiation of A in position a and as A_b for an instantiation of A in position b. These choices are defined in (36).

(36) **Loci as variables: time-dependent version**
a. A sign language locus-as-variable A is a function from signing times to positions in signing space.
Notation: if there is no ambiguity, instead of writing a realization of A as $A(t)$, where t is signing time, we write as A_i an instantiation of A in position i.
b. Assignment functions assign values to loci-as-variables, e.g. $A, B, C, \dots$.
c. Projective conditions make reference to instantiations of loci-as-variables, hence $A\_i, B\_i, C\_i, \dots$.

The value of a pointing sign toward an arbitrary (= non-iconic) locus A_i solely depends on the value that the relevant assignment function s assigns to A, hence only on s(A), as stated in (37)a. By contrast, iconic pointing toward A_i depends both on s(A) and on a projective condition that makes reference to i, the part of space in which the locus is instantiated, as stated in (37)b.

(37) **Arbitrary and iconic loci** (final version)
If A is a locus instantiated in part i of signing space:
a. Arbitrary loci
$[[pro_{A\_i}]]^{c, s, t, w} = s(A)$
b. Iconic loci
$[[pro_{A\_i}{}^{\pi}]]^{c, s, t, w} =$ # if proj(s(A), s(π), t, w) ≠ i
s(A), otherwise

These definitions make it possible to posit for (34)a and (34)b the Logical Forms in (38)a and (38)b respectively. Crucially, (34)b comes with a presupposition that the crocodile projects to position b, which explains the deviance of (35)b, where it is denied that the crocodile moved to the corresponding spatial position.[18]

(38) a. IX-a SICK.
a'. pro_{A_a} sick
a''. $[[(a')]]^{c, s, t, w} = sick'_{t, w}(s(A))$
b. IX-b SICK.
b'. $pro_{A\_b}{}^{\pi}$ sick
b''. On the assumption that s(A) is a crocodile and $sick'_{t, w}(x) = \#$ if and only if x = #,
$[[(b')]]^{c, s, t, w} = sick'_{t, w}([[pro_{A\_i}{}^{\pi}]]^{c, s, t, w})$
= # if proj(s(A), s(π), t, w) ≠ b (i.e. the crocodile doesn't project to position b from viewpoint s(π) at t in w)
sick'(s(A), otherwise

[17] To forestall misunderstandings, our proposal is that variables themselves, not variable *denotations*, are functions from signing times to picture parts. This functional dependency on times captures the intuition that loci 'move in signing space', or in other words that their position varies with time. It is to these time-dependent loci that assignment functions will assign a value, but it will still be the case that an assignment function assigns a single value to a time-dependent (i.e. functional) locus.

[18] As will see below, classifier predicates come with iconic conditions irrespective of whether they are associated with iconic variables, and as a result both sentences in (34) will come with the implication that the crocodile is initially in a position that projects to a.

In sum, instances of Locative Shift further highlight the similarity between iconic loci and Abuschian pictorial variables as implemented above: both kinds of variables 'move' from one representation to the next.[19]

In all the cases discussed up to this point, a single viewpoint variable π appeared in the Logical Forms, and its value was determined (through an assignment function) by the context. One could in principle have done everything with the context parameter without going through an assignment function.[20] We will now extend our investigation in three steps. First, we will show that viewpoint variables can help give an explicit account of the iconic component of classifier predicates (Section 4). While initial cases will involve a single viewpoint variable per sentence, we will then see that viewpoint variables are sometimes dependent on quantifiers (Section 5), and then that several viewpoint variables can co-occur in the same sentence (Section 6). Thus the expressive power afforded by viewpoint variables won't be idle.

4 Classifier Predicates in Iconological Semantics

The iconic loci discussed so far provided pictorial information about the orientation of a single object. We turn to more sophisticated cases in which signing space serves to create a map of the relative positions of different objects and sometimes of their movement. In the cases we considered above, loci were sometimes specified as iconic. We now turn to predicates that are specified as iconic, with the result that their arguments will be subject to iconic conditions as well. Specifically, we claim that classifier predicates always come with an iconic condition—technically, they always carry a viewpoint variable. As a result, the loci they may take as arguments will be (indirectly) subject to iconic conditions as well. We will thus argue for the following generalization:

(39) **Classifier-Loci Generalization**
a. Loci introduced by standard nominals may but need not be interpreted iconically.
b. Loci introduced by classifier predicates must be interpreted iconically.

(For simplicity, we will disregard issues of Locative Shift and correspondingly use lowercase loci, i.e. *a*, *b*, *c*, … . In our 'official' system they correspond to constant time-sensitive loci *A*, *B*, *C*… which remain in positions *a*, *b*, *c*, … respectively.)

4.1 Background on classifier predicates

Since classifier predicates will be our main focus, we should say a word about their defining properties. In a nutshell, classifier predicates (i) have a lexically specified form, but (ii) their position in signing space, their orientation and their movement (if there is one) are unconstrained, and interpreted iconically. In Zwisterlood's (2012) words, classifier predicates "are reported to occur in almost all sign

---

[19] This discussion does not purport to be a complete account of Locative Shift as discussed in Schlenker 2018a, for three reasons. First, some instances of Locative Shift pertain to cases in which no locus is directly interpreted iconically. For instance, one can assign an arbitrary position to *John*, a second arbitrary position to an indefinite meaning *a French city*, and a third arbitrary position to an indefinite meaning *an American city*, and under certain conditions the John locus 'moves' to the French city locus or to the American city locus (Schlenker 2018a). Second, we have no attempted to *constrain* locus movement, and it is thus unclear why there couldn't be an abstract locus *A* that starts out in position *a* and moves to position *b* without giving rise to an iconic interpretation (since the addition of a viewpoint variable π on a pronoun is optional). Third, Schlenker 2018a discusses instances of Locative Shift that apply to temporal and modal loci, and these are hard to handle within a pictorial semantics.

[20] Concretely, we could have replaced the variable π with a diacritic • that just specifies that an expression is interpreted iconically. The rule in (29)a would then have been restated as (i), with a functional term *viewpoint(c)* replacing *s(π)*:

(i) $[\![pro_a{}^{\bullet}]\!]^{c,s,t,w}$ = # if proj(s(a), viewpoint(c), t, w) ≠ a
s(a), otherwise

See also Appendix II for another alternative to the present analysis, without viewpoint variables.

languages researched to date". We will henceforth concentrate on what are called in the literature 'whole entity classifiers'.[21]

A textbook example appears in (40)a: a vehicle classifier represents a car, and the movement of the predicate in signing space iconically depicts the movement of the car.

(40) a. CAR CL-vehicle-DRIVE-BY. (ASL, cited and illustrated from Valli and Lucas 2000, cited in 2017)

b. 'A car drove by *like this*', where the information contributed by *like this* is produced by the movement of the classifier predicate in signing space (after Zucchi 2011)

The gradient and iconic character of the information conveyed was displayed with experimental means in Emmorey and Herzig 2003. While there is much debate about the morphological analysis of classifier predicates (Zwisterlood 2012), we will be agnostic on this point in the present piece, which focuses on their semantics. The syntax is more directly relevant: across sign languages, the syntax of classifier predicates is often different from that of other constructions (e.g. Pavlič 2016). Schlenker et al. (to appear) seek to derive this fact from the iconic character of classifier predicates: these create visual animations and their arguments appear in the order in which their denotations would typically be seen in an event (or in a cartoon), thus giving rise to an 'iconic syntax'.[22]

Regarding the semantics, we will largely follow the insights of Liddell 2003a,b, but we will seek to integrate them to a logical semantics, something that wasn't his goal:

"I am proposing that the classifier predicates analyzed in this paper are fixed lexical verbs. These verbs become full signs by placing and directing them in analogical, gradient ways. This will always include placing the hand at an analogical location, and will sometimes include directing the hand's orientation analogically." (Liddell 2003a)

Zucchi 2011, 2017 and Davidson 2015 take classifier predicates to literally have a demonstrative component, as suggested by Zucchi's paraphrase in (40)b; we use the paraphrase for perspicuity but do not adopt the analysis, for reasons explained in Appendix I-B.

4.2 The Classifier-Loci Generalization: initial examples

A minimal contrast between classifier predicates and normal nouns is displayed in (41). It builds on the fact that the word *PLANE* comes in two varieties. One is a standard noun, produced with a slight repetition, characteristic of nominal marking in ASL;[23] it is realized at a neutral height in signing space.

---

[21] As Zwisterlood (2012) writes, the class of whole entity classifiers "contains classifiers that directly represent referents, by denoting particular semantic and/or shape features". By contrast, the class of 'handling classifiers' includes "classifiers that represent entities that are being held and/or moved; often (but not exclusively) by a human agent". Supalla (1982, 1986) (followed by Davidson 2015) proposed an additional category, the 'body classifier', in which the signer's body represents another person. This is otherwise analyzed as Role Shift, also called Constructed Action (see Zwisterlood 2012 for similar remarks). While Role Shift interacts in interesting ways with viewpoints, this construction lies outside the scope of the present study.

[22] To illustrate, ASL has the basic word-order SVO, but in classifier constructions meaning *a crocodile ate up a ball,* the preferred word order involves a preverbal object (SOV or OSV). By contrast, in classifier constructions meaning *a crocodile spit out a ball*, SVO order is regained. The proposed explanation is that in eat-up-type events, one sees the ball before it is ingested, whereas in spit-out-type events, one only sees the ball after the ejection (Schlenker et al, to appear).

[23] This is different from plural-based repetition, which involves a larger number of iterations, usually with movement.

The other is a classifier version, without the nominal repetition; we gloss it as *PLANE-cl*, and here is it signed low, as befits an airplane on the ground. In (41), both forms introduce loci, but in the classifier case, they obligatorily provide positional information, whereas in the nominal case, they needn't provide such information.

(41) HERE ANYTIME __ SAME-TIME TAKE-OFF, DANGEROUS.
'Here, whenever ___ take off at the same time, there is danger.'
a. ^{7} PLANE$_a$ PLANE$_b$
two planes
b. ^{7} PLANE-cl$_a$ PLANE-cl$_b$
two planes side by side/next to each other
(ASL, 35, 1916, 4 judgments; https://youtu.be/Wyt1AsP6ASk)

The difference in positional information can be brought out in two ways: by way of open-ended questions pertaining to the difference in meaning between the two sentences (see the Supplementary Materials B), and by way of questions of inferential strength, as in (42), with judgments in (43).

(42) Is there risk (i) whenever two planes simultaneously take off here, or (ii) whenever two planes *in a certain configuration* (which?) simultaneously take off here? (Judgments were on a 7-point scale, with with 1 = no inference; 7 = strongest inference)

(43) **Inferential judgments for (41)** (3 judgments)

There is a risk…	(i) whenever two planes simultaneously take off here	(ii) whenever two planes *in a certain configuration* simultaneously take off here
a. PLANE$_a$ PLANE$_b$	6.7	2.3
b. PLANE-cl$_a$ PLANE-cl$_b$	2	6

The facts are similar in (44), which makes reference to three planes rather than two: with the normal nouns in (44)a, the loci may but need not be interpreted iconically, whereas with the classifier predicates in (44)b, an iconic interpretation is obligatory.

(44) HERE ANYTIME
'Here, whenever
a. $^{6.8}$ PLANE$_a$ PLANE$_b$ PLANE$_c$
three planes (or: three planes in parallel)
b. $^{6.5}$ PLANE$_a$ PLANE$_b$ PLANE$_c$
three planes in parallel positions[24]
SAME-TIME TAKE-OFF, THE-THREE$_{a,b,c}$ DANGEROUS.
take off at the same time, all three are in a dangerous situation.
(ASL, 35, 1922; 4 judgments;[25] https://youtu.be/tOpM8yFPJlQ)

(45) **Inferential judgments for (44)** (3 judgments)

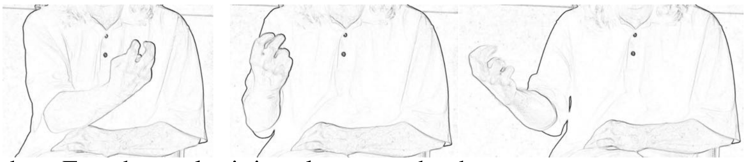
There is a risk whenever…	(i) three planes simultaneously take off here	(ii) three planes *in a certain configuration* simultaneously take off here
a. PLANE$_a$ PLANE$_b$ PLANE$_c$	6.3	2.3
b. PLANE$_a$ PLANE$_b$ PLANE$_c$	2	6.3

[24] As noted by Jason Lamberton, using three rather than two airplane classifiers reduces the strength of the 'parallel take-off' inference: it suggests similar directions but doesn't require strictly parallel take-offs.
[25] In (41) and (44), quantitative questions pertaining to inferential strength were added after the first session, which is why there are 4 acceptability judgments but only 3 inferential judgments.

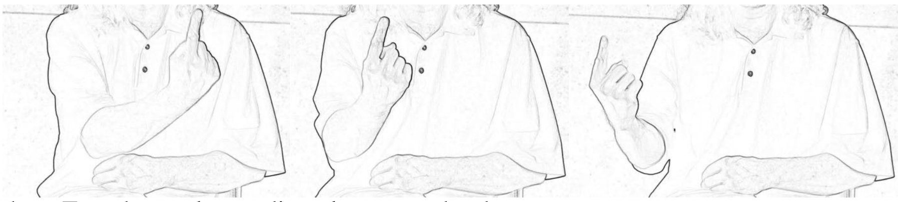
4.3 The Classifier-Loci Generalization: more complex examples

The paradigm in (46) displays a minimal contrast between two classifier predicates and a nominal. (46)c makes use of the normal noun for *PERSON*, iterated three times, and the temporal clause quantifies over situations in which any three Frenchmen are simultaneously in class. By contrast, (46)a,b involve classifier predicates, and correspondingly the temporal clause ('any time…') quantifies over situations in which three Frenchmen are arranged in a particular way: sitting together in (46)a, which involves the 'sitting person' classifier (two clawed fingers, glossed as *SIT*); and standing close to each other in (46)b, which involves the finger classifier representing an upright person.

(46) CLASS ANY-TIME HAVE FRENCH
'In class, any time there are
a. ^{7}SIT-cl$_c$ SIT-cl$_b$ SIT-cl$_a$



three French people sitting close to each other,
b. ^{7} PERSON-stand-cl$_c$ PERSON-stand-cl$_b$ PERSON-stand-cl$_a$



three French people standing close to each other,
c. ^{6}PERSON$_c$ PERSON$_b$ PERSON$_a$

three French people,
IX-arc-a,b,c ARGUE.
they argue.'
(ASL, 35, 1994, 3 judgments; https://youtu.be/2IeoXj09Wg0)

(47) **Inferential judgments for (46)** (3 judgments)

There are arguments whenever…	(i) there are three French people	(ii) there are three French people *in a certain position*
a. SIT-cl$_c$ SIT-cl$_b$ SIT-cl$_a$	2	6.7
b. STAND-cl$_c$ STAND-cl$_b$ STAND-cl$_a$	2.7	6
c. PERSON$_c$ PERSON$_b$ PERSON$_a$	6	3

A variant of the preceding paradigm involves classifier predicates that are simultaneously produced with the two hands, as in (48), where we write *WORD$_1$^WORD$_2$* when two words are signed simultaneously. In (48)a, two occurrences of the sitting person classifier (already seen in (46)a) are used to represent two individuals facing each other, while in (48)b they are behind each other (or maybe just next to each other). Similarly, the upright person classifier of (46)b is used to represent two individuals facing each other in (48)c, and facing each other in (48)d. By contrast, in (48)e there are two occurrences of *PERSON*, whose loci are preferably interpreted non-iconically. Iconic inferences are very strong with all the classifier predicates. Inferential judgments are mixed with the nominal *PERSON*, reflecting the fact that iconic interpretation is optional in this case.

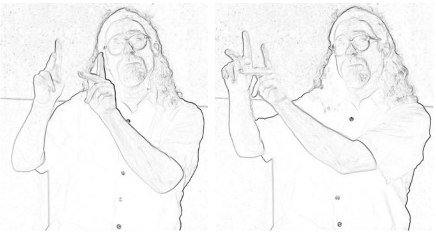
(48) CLASS ANYTIME HAVE FRENCH
'In class, whenever there are
a. SIT-cl_a^SIT-cl_b [mirror]

two French people sitting, facing each other,
b. SIT-cl_a^SIT-cl_b [same orientation]

two French people sitting next to each other [one behind the other],
c. STAND-cl_a^STAND-cl_b [mirror]

two French people standing, facing each other,
d. STAND-cl_a^STAND-cl_b [same orientation]

two French people standing one behind the other,
e. $PERSON_a$ $PERSON_b$



two French people,

THE-TWO ARGUE.

they argue.'
(ASL, 35, 2002; 3 judgments; https://youtu.be/zg6D6DNPavY)

(49) **Inferential judgments for (48)** (3 judgments)

There are arguments whenever…	(i) there are two French people	(ii) there are three French people *in a certain position*
a. [7] SIT-cl$_a$^SIT-cl$_b$ [mirror]	1	7
b. [7] SIT-cl$_a$^SIT-cl$_b$ [same orientation]	1.3	7
c. [7] STAND-cl$_a$^STAND-cl$_b$ [mirror]	1	7
d. [7] STAND-cl$_a$^STAND-cl$_b$ [same orientation]	1.3	6.3
e. [6.7] PERSON$_a$ PERSON$_b$	5.7 (7, 6, 4)	3.3 (3, 2, 5)

4.4 Capturing arbitrary and iconic readings

Focusing on the simplest example, namely (41), we propose that the contrast between classifier predicates and normal nouns can be captured by the simplified Logical Forms in (50).

(50) a. Nouns: anytime$_{a,b}$ plane$_a$ plane$_b$ a+b simultaneously-take-off, danger
b. Classifiers: anytime$_{a,b}$ plane$^{\boldsymbol{\pi}}_a$ plane$^{\boldsymbol{\pi}}_b$ a+b simultaneously-take-off, danger

We treat *anytime* as an unselective binder, indexed with the variables *a* and *b* in this case. The nominal and the classifier case are identical except for the presence of the viewpoint variable π **(boldfaced)** in the classifier case. We informally state in (51) the truth conditions we will derive for the two cases; they are identical except for the boldfaced part of (51)b, which results from the iconic component of the classifier predicates.

(51) a. Nouns: ≈ For any time t and objects a and b such that a and b are planes and a+b simultaneously take off at t, there is danger at t
b. Classifiers: ≈ For any time t and objects a and b such that a and b are planes **and a projects to *plane$_a$* and b projects to *plane$_b$*** and a+b simultaneously take off at t, there is danger at t

On the analytical side, the key step lies in the interpretive rule for the classifier predicate *PLANE-cl* (written in our Logical Forms as *plane-cl*), given in (52)b. It has the same lexical component as the noun *PLANE* in (52)a, but adds to it a requirement (boldfaced) that any object x in its extension should project to this very token of the sign relative to the viewpoint denoted by π.

(52) a. $[[\text{plane}]]^{c,s,t,w} = \lambda x_e\,.\ \text{plane}'_{t,w}(x)$
b. $[[\text{plane-cl}^{\pi}]]^{c,s,t,w} = \lambda x_e\,.$ # if x = #
1 iff $\boxed{\text{plane}'_{t,w}(x) = 1}$ and $\mathbf{proj(x, s(\boldsymbol{\pi}), t, w) = plane\text{-}cl}$, otherwise

The rule in (52)b need not be seen as primitive; rather, it results from the more general interpretive rule in (53), stated for any 1-place predicate (it could be generalized to the case of an n-place predicate).

(53) If P is a 1-place predicate,
$[[P^{\pi}]]^{c,s,t,w} = \lambda x_e\,.$ # if $[[P]]^{c,s,t,w}(x) = \#$
1 iff $\boxed{[[P]]^{c,s,t,w}(x) = 1}$ and $\mathbf{proj(x, s(\boldsymbol{\pi}), t, w) = P}$, otherwise

We will see in Section 7 that the boxed lexical conditions are in the end redundant with the projective condition (as the latter will have a conventional component as well). But things will be clearer if we keep the boxed part separate for the moment.

Since we treat indefinites as predicates with variables, we need a rule to interpret such structures. It is given in (54), and states that an indefinite such as *plane$_i$* is a proposition that is true just in case the denotation of the variable *i* satisfies the predicate *plane* (this is in essence the view of indefinites in Heim 1982).

(54) If P is a predicate, $[[P_i]]^{c,s,t,w} = [[P]]^{c,s,t,w}(s(i))$

When applied to the two lexical entries in (52), the interpretive rule in (54) yields different results: the predicate classifier comes with a projective condition that is absent from the normal noun, as shown in (55).

(55) If $s(i) \neq \#$,

a. $[[plane_i]]^{c, s, t, w} = [[plane]]^{c, s, t, w}(s(i))$
$= 1$ if $plane'_{t,w}(s(i)) = 1$
0, otherwise

b. $[[plane\text{-}cl^{\pi}_i]]^{c, s, t, w} = [[plane\text{-}cl^{\pi}]]^{c, s, t, w}(s(i))$
$= 1$ if $plane'_{t,w}(s(i)) = 1$ **and** $\mathbf{proj(s(i), s(\pi), t, w) = plane\text{-}cl_i}$
0, otherwise

We slightly abuse notation in writing: *proj(x, s(π), t, w) = plane-cl*$_i$, as our 'official' rule in (53) would lead us to expect: *plane-cl*, without the locus *i*. But here it must be remembered that *plane-cl* is a *token* of a sign, not a type; and we need to distinguish different tokens of the same sign. We do so by indexing them with the locus in which they are signed.

Already at this point, we have reached one of our main goals: the boldfaced part of (55)b is almost identical to the variable-related part of the Abuschian rule we used in (14)a. A superficial difference is that in the latter case, we notated the picture part as v_i, whereas in the ASL case the picture part is notated as *plane-cl*$_i$, corresponding to the sign token *plane-cl* appearing in locus *i*.

In the iconic loci seen in Section 3.2, the depictive condition was presuppositional. Quite generally, constraints on pronominal denotations are presuppositional: if an expression is of individual type, any predicative constraints on it (e.g. gender specifications) are presuppositional, and iconic loci follow this pattern. By contrast, classifiers are of predicative type and thus do not fall under this generalization: their iconic component can make an at-issue contribution (and additional non-at-issue contributions in some cases, see for instance Schlenker 2021).

To complete the derivation of the truth conditions of (50)b, we need lexical entries for the remaining words, given in (56) (in a., $R_{c,w}$ is an accessibility relation among times, relativized to the context c and world of evaluation w).

(56) a. $[[anytime_{a,b}\ F, G]]^{c, s, t, w} =$ # if for some t' such that $tR_{c,w}t'$ and for some objects d, d', $[[F]]^{c, s[a\to d, b\to d'], t', w} = 1$ and $[[G]]^{c, s[a\to d, b\to d'], t', w} = \#$
1 if for all t' such that $tR_{c,w}t'$ and for all objects d, d', if $[[F]]^{c, s[a\to d, b\to d'], t', w} = 1$, $[[G]]^{c, s[a\to d, b\to d'], t', w} = 1$, otherwise

b. $[[a+b]]^{c, s, t, w} = s(a)+s(b)$

d. $[[simultaneously\text{-}take\text{-}off]]^{c, s, t, w} = simultaneously\text{-}take\text{-}off'_{t,w}$

e. $[[danger]]^{c, s, t, w} = danger'_{t,w}$

With these lexical entries in hand, the derivation of the truth conditions of (50)b is routine, and leads to the result in (57). The boldfaced part is the iconic contribution of the classifiers.[26]

(57) Assume $s(a) \neq \#$ and $s(b) \neq \#$. Then $[[(50)b]]^{c, s, t, w} \neq \#$, and

$[[(50)b]]^{c, s, t, w} =$ 1 if for all t' such that $tR_{c,w}t'$ and for all objects d, d', if $plane'_{t',w}(d) = plane'_{t',w}(d') = 1$ **and** $\mathbf{proj(d, s(\pi), t, w) = plane\text{-}cl_a}$ **and** $\mathbf{proj(d', s(\pi), t, w) = plane\text{-}cl_b}$ and $simultaneously\text{-}take\text{-}off'_{t',w}(d+d') = 1$, $danger'_{t',w} = 1$
0 otherwise

It is clear that by removing the boldfaced part, we obtain truth conditions for (50)a.

The last point to explain is why a normal noun can optionally introduce an iconically interpreted locus. The simplest answer is that the noun *plane* can optionally come with the viewpoint variable π, as in (58). The general rule in (54) guarantees that the expression *plane*$^{\pi}$ can be interpreted, and yields the same result as the classifier predicate *plane-cl*$^{\pi}$.

(58) **Nouns—iconic interpretation**
$anytime_{a,b}\ plane^{\pi}_a\ plane^{\pi}_b$ a+b simultaneously-take-off, danger

[26] We might want the time of the take-off to be a bit after the time at which the planes are represented as being on the ground. If so, a covert temporal operator (e.g. akin to *then*) should appear in (50)b right after the classifiers.

Thus what is special about classifier predicates is that they obligatorily come with viewpoint variables, whereas this just an option for normal nouns.[27]

4.5 Summary and consequences for loci

It is worth pausing to summarize our findings. In Section 3, we argued that loci can optionally carry a viewpoint variable and thus be iconic. This iconic contribution was taken to be presuppositional, just like the contribution of gender features on pronouns. In the present section, we focused instead on classifier predicates, which are lexically specified as carrying a viewpoint variable; normal predicates may carry a viewpoint variable as well, but this is just an option for them. When iconic predicates 'touch' certain areas of signing space, these become iconic as well, simply because of the semantics of classifier predicates. This also applies when a classifier predicate 'touches' a locus it takes as an argument: the locus must be interpreted iconically, but without carrying a viewpoint variable itself. These two sources of iconicity for loci are summarized in (59).[28]

(59) **Two sources of iconicity for loci**

	Iconic loci	**Loci that are arguments of iconic predicates**
Representation (examples)	pro_i^{π}	PLANE-cl^{π}_i (classifier predicates) $PLANE^{\pi}_i$ (normal predicates with the option of a viewpoint variable)
Source of iconicity	Direct: it is part of the lexical specification of pronouns	Indirect: it is due to the semantics of the iconic predicate
Status of the iconic contribution	Presuppositional	At-issue (or as specified by the classifier predicate)

An additional consequence is worth highlighting. In this section, we only considered examples in which all loci were arbitrary or all loci were iconically interpreted. But there can be mixed cases, and the Classifier-Loci Generalization makes a prediction: any locus that is 'touched' by a classifier predicate becomes *ipso facto* iconic. This prediction appears to be borne out, but the data are subtle and complex, in essence because normal predicates may optionally carry a viewpoint variable as well. Relevant paradigms are discussed in the Supplementary Materials A-1.

Finally, we should ask how the present theory compares to the demonstrative analysis of classifier predicates developed by Zucchi 2011, 2017 and Davidson 2015. We argue in Appendix I-B that the demonstrative analysis either fails to deliver truth conditions, or presupposes an analysis like present one to deliver similarly explicit truth conditions, but in a more roundabout way.

---

[27] Two remarks should be added.
1. Another possibility would be to assume that the loci, rather than the predicates, come with a viewpoint variable π. On this view, the Logical Form in (i) would replace that in (58) (we write *a_π* and *b_π* for a^{π} and b^{π} when these expressions appear as subscriptions).

(i) $anytime_{a,b}$ $plane_{a\_\pi}$ $plane_{b\_\pi}$ $a^{\pi}+b^{\pi}$ simultaneously-take-off, danger

An important difference is that (i) should trigger a presupposition that the planes are in certain positions, whereas this condition is at-issue in (58). We know of no evidence for a presuppositional treatment—but presuppositions can be locally accommodated (especially within restrictors, as is the case here), which makes the question complex. We leave it for future research.
2. There are certainly further conditions on iconically interpreted normal nouns, for instance that their position in signing space should be morphologically sufficiently free to be interpretable iconically; we will not try to explore these additional constraints here.

[28] When loci 'move in signing space', functional loci must be used. Thus to be rigorous, we should write $pro_{A\_i}^{\pi}$; PLANE-$cl^{\pi}_{A\_i}$; $PLANE^{\pi}_{A\_i}$.

5 Quantification Over Viewpoints

In the foregoing discussion, viewpoints played a dual role: they told the interpretive system that an expression had to be interpreted iconically, and they specified the viewpoint relative to which the geometric projection was assessed. But we did not use the full power of viewpoint variables, as there was only a single variable in each sentence, and it was not bound by anything. We could thus have stated our analysis in purely contextual terms, with the assumption that any context c makes available a certain viewpoint, π_c. Specifically, we could have replaced viewpoint variables with a diacritic •, and the rule in (53) with that in (60), which makes reference to the viewpoint of the context.[29]

(60) **A contextual analysis of viewpoint dependency** (insufficient for our data)
If P is a 1-place predicate,
$[[P^{\bullet}]]^{c,s,t,w} = \lambda x_e .$ # if $[[P]]^{c,s,t,w}(x) = \#$
1 iff $[[P]]^{c,s,t,w}(x) = 1$ and $proj(x, \pi_c, t, w) = P$, otherwise

We will now consider two respects in which the power of viewpoint variables can be used. In this section, we show that viewpoint variables may be dependent on quantifiers, with existential quantification being a likely mechanism, as is schematically represented in (61)a. In Section 6, we will argue that several (free) viewpoint variables may co-occur in the same sentence, as sketched in (61)b, where the first classifier predicate is evaluated relative to viewpoint variable π while the second classifier predicate is evaluated relative to viewpoint variable π'.

(61) a. $\exists\pi$ … $classifier^{\pi}_{i}$ …
b. … $classifier^{\pi}$ … $classifier^{\pi'}$…

In both types of configurations ((61)a and (61)b), there will be a striking interaction between viewpoint choice and logical structure, and appeal to a single, fixed contextual viewpoint will be insufficient. (Owing to the insufficiency of fixed contextual viewpoints, we develop our analysis with viewpoint variables. But we consider in Appendix II a third possibility, based on a shiftable viewpoint parameter.)

5.1 Basic analysis

A simple example of existential quantification over viewpoints appears in (62). The context involves four classrooms, each assigned to a single teacher, and the sentence asserts that in each class a student always leaves during the break, with a specific movement iconically depicted by the classifier predicate in (62)b,c, but not in the control (involving a normal noun) in (62)a; as expected, the inferential judgments in (63) display contrasts in iconic specificity.[30] It is clear that a single visual animation couldn't represent movements in four different classrooms unless it involves different 'camera positions', and thus quantification over viewpoints is needed.

(62) *Context:* This school has 4 classrooms, one for each of 4 teachers (each teacher always teaches in the same classroom).
CLASS BREAK ALL ALWAYS HAVE STUDENT
In all classes, during the break, there is always a student that
a. [7] LEAVE.
leaves.
b. [7] PERSON-walk-back_right-cl.
leaves toward the the back, to the right.'

[29] This measure would have yielded the same result as the Zucchian analysis developed in Appendix I-B (see in particular (110)).

[30] Here our second consultant, Jason Lamberton, noted that (62)c is more flexible than (62)b in the location of the exits used by the student. This observation is partly mirrored in our main consultant's inferential judgments above (including their instability for (62)c). The reason for the slight contrast might be that the classifier movement in (109)c is straight (from right to left), and thus easier to produce and more general, whereas the movement in (109)b involves a change of direction and is more marked. We do not seek to account for this fine-grained contrast in what follows but come back to it in the conclusion.

c. [7] PERSON-walk-front_left-cl.
leaves toward the front, to the left.'
(ASL, 35, 2254; 3 judgments; https://youtu.be/h7EnRK6poAE)

(63) **Inferential judgments for (62)** (3 judgments)

One infers that…	(i) at least one student leaves, without further specification	(ii) at least one student leaves, with a specific movement
a. [7] LEAVE	7	1
b. [7] PERSON-walk-back_right-cl	1.7	6.3
c. [7] PERSON-walk-front_left-cl	2.7 (1, 5, 2)	5.3 (7, 3, 6)

We propose to analyze the examples in (62)b,c by way of the simplified Logical Form in (64), which abstracts away from some aspects of the discourse. The crucial assumption is that the viewpoint variable that appears on the classifier *person-walk-cl* is existentially quantified.

(64) $always_T\ \exists\pi$ there-is student person-walk-cl$^\pi$

Natural language quantifiers come with implicit restriction, and we take this to also apply to the existential quantifier over viewpoints in (64), hence the interpretive rule in (65). Relative to a context c, a time t and a world w, existential quantification is restricted to viewpoints in a set $D_{c, t, w}$ (boldfaced in (65)); we interpret it as being the set of 'salient' viewpoints.[31]

(65) If c, t and w are a context, a time and a world, $D_{c, t, w}$ is a set of viewpoints, and we can define:
$[[\exists\pi\ F]]^{c, s, t, w} =$ 1 if for some viewpoint v such that $\mathbf{D_{c, t, w}v = 1}$, $[[F]]^{c, s[\pi\rightarrow v], t, w} = 1$
0, otherwise

The analysis of the classifier *person-walk-cl*$^\pi$ is just an application of the more general rule in (53) above (we assume for simplicity that *person-walk* is not itself a presupposition trigger):

(66) $[[\text{person-walk-cl}^\pi]]^{c, s, t, w} = \lambda x_e$. # if $[[\text{person-walk}]]^{c, s, t, w}(x) = \#$
1 iff $[[\text{person-walk}]]^{c, s, t, w}(x) = 1$ and $proj(x, s(\pi), t, w) =$ *person-walk-cl*, otherwise
$= \lambda x_e$. # if $x = \#$
1 iff $\text{person-walk}'_{t, w}(x) = 1$ and $proj(x, s(\pi), t, w) =$ *person-walk-cl*, otherwise

In the simplified analysis in (64), *always* carries a domain restriction variable and has the semantics in (67)a.[32] And we take *HAVE*, analyzed as *there-is*, to have the semantics of an existential quantifier, as in (67)b.

(67) a. $[[always_T\ F]]^{c, s, t, w} = 1$ iff for all t' such that $s(T)(t') =1$, $[[F]]^{c, s, t', w} = 1$
b. $[[\text{there-is}]]^{c, s, t, w} = \lambda f_{<e,t>}\ \lambda g_{<e,t>}$. # if for every x, $f(x) = \#$ or for every x such that $f(x) = 1$, $g(x) = \#$
1 iff for some object x, $f(x) = g(x) = 1$, otherwise

With these lexical entries, we can complete the derivation of the truth conditions of (64), copied in (68)a, as in (68)b.

(68) a. $always_T\ \exists\pi$ there-is student person-walk-cl$^\pi$
b. $[[(a)]]^{c, s, t, w} \neq \#$. Furthermore,
$[[(a)]]^{c, s, t, w} = 1$ iff for all t' such that $s(T)(t') =1$, $[[\exists\pi \text{ there-is student person-walk-cl}^\pi]]^{c, s, t', w} = 1$,
iff for all t' such that $s(T)(t') =1$, for some viewpoint v such that $D_{c, t', w}v = 1$, for some object x, $\text{student}'_{t', w}(x) = 1$ and $\text{person-walk}'_{t', w}(x) = 1$ and $proj(x, s(\pi), t', w) =$ *person-walk*

[31] For simplicity, we assume that the existential quantifier over viewpoints does not project presuppositions, and thus yields purely bivalent meanings. This is certainly incorrect, but issues of presupposition projection are not crucial to the present analysis, except for the iconic loci of Section 3.2 (and these did not involve existential quantification over viewpoints).
[32] We could have decided to provide a domain restriction in the metalanguage only, without an object-language variable. But in later examples (e.g. (77)) involving two occurrences of *always*, it will be useful to allow them to have different domain restrictions, hence the benefit of an object-language variable.

5.2 Further examples

It is not hard to construct further examples displaying the same pattern of dependency between viewpoints and quantifiers. We mention them to buttress the Classifier-Loci Generalization, but without offering a full analysis, as it would be similar to that of the preceding example.

In (69)b,c, universal quantification over four airport runways (oriented in different ways) co-occurs with an airplane classifier displaying a rolling movement on the right side (in (69)b) and a dipping movement on the left side (in (69)c). These provide relatively precise iconic information, unlike the control with a normal verb in (69)a.[33] In (69)b and (69)c, for our main consultant the viewpoint is at the back of the runaway, as can be established in a separate question (see the Supplementary Materials B); but multiple airways are involved, hence multiple viewpoints as well.

(69) *Context:* This airport has 4 runways oriented in different ways.
WHEN WEATHER BAD, STRIP ALL HAVE AIRPLANE
'When the weather is bad, all runways have an airplane that
a. TAKE-OFF CLUMSY.
takes off clumsily.'
b. PLANE-FLY-roll-right-cl.
rolls right then back left after take-off.'
c. PLANE-FLY-dip-left-cl.
dips right after take-off before stabilizing.'
(ASL, 35, 2426; 3 judgments; https://youtu.be/QYivyug5UBY)

(70) **Inferential judgments for (69)** (3 judgments))

One infers that each runway has	(i) at least one airplane that takes off clumsily, without further specification	(ii) at least one airplane that takes off with a specific movement
a. ^{7} TAKE-OFF	7	1
b. ^{7}PLANE-FLY-roll-right-cl	1	7
c. ^{7} PLANE-FLY-dip-left-cl	1	7

The control in (69)a can be analyzed as in (71)a, without viewpoint variables. The two classifier cases in (69)b,c can be analyzed with a version of (71)b (with different shapes for *fly-cl*), which involves quantification over viewpoints.[34]

(71) a. [all runway] λx there-is airplane [take-off-clumsily [in x]]
b. [all runway] $\lambda x\ \exists\pi\ t_x$ there-is [airplane plane-fly-cl$^{\pi}$ [in x]]

A related paradigm appears in (72), involving a plane that takes off and then turns back. The normal verb version in (72)a provides no iconic information, whereas the classifier versions in (72)b,c do, as they depict a movement to the right or to the left, followed by a loop to the original position. In both classifier cases, for our main consultant the 'camera position' is at the beginning of the relevant runway, but there are multiple runways and hence multiple viewpoints as well, which can be captured by existential quantification over viewpoints.

---

[33] Our second consultant noted that (70)b is a bit more flexible than (70)c, and could involve dips, contrary to the 'rolls right then back left after take-off' translation we give above, in line with our first consultant's judgments. We do not seek to derive this slight contrast in what follows.

[34] In (71), *there-is* can be interpreted with the lexical rule given above in (67)b. For greatest simplicity, we take *in* to be a transitive predicate, as displayed in (i), and for any verb *V*, *V [in x]* is interpreted by intersective modification, as illustrated in (ii).

(i) $[[in]]^{c,\,s\,t,\,w} = in'_{t,w}$ ($\approx \lambda d_e \lambda d'_e$. # if d = # or d' = #
1 iff at t in w d' is located in d, otherwise)

(ii) $[[V[in\ x]]]^{c,\,s\,t,\,w} = \lambda d_e$. # if $[[V]]^{c,\,s\,t,\,w}$ (d) = # or $in'_{t,\,w}(s(x))$ = #
1 iff $[[V]]^{c,\,s\,t,\,w}$ (d) = $in'_{t,\,w}(s(x))$ = 1, otherwise

(72) *Context:* The Boston airport has 20 runways oriented in different ways. The following is signed in New York:
BOSTON WHEN WEATHER BAD, STRIP ALL HAVE AIRPLANE
'In Boston, when the weather is bad, all runways have an airplane that
a. ^{7}TAKE-OFF OH-NO BACK.
takes off and has to fly back.'
b. ^{7}PLANE-FLY-loop-right-cl.
takes off, turns (right[35]) and loops around to come back.'
c. ^{7}PLANE-FLY-loop-left-cl.
takes off, turns left and loops around to come back.'
(ASL, 35, 2430; 3 judgments; https://youtu.be/MMR-hWDEgro)

(73) Inferential judgments for (69) (3 judgments)

One infers that each runway has	(i) at least one airplane that takes off and returns to the runway, without further specification	(ii) at least one airplane that takes off with a specific movement and returns to the runway
a. ^{7}TAKE-OFF OH-NO BACK	7	1
b. ^{7}PLANE-FLY-loop-right-cl	3.7 (1, 5, 5)	3.7
c. ^{7}PLANE-FLY-loop-left-cl	3.3 (1, 4, 5)	4.7

It should be added that the iconic interpretation is a bit more precise in the case in which the classifier looping movement is toward the left (in (72)c) than toward the right (in (72)b); this point was also made by our second consultant. The reason for the contrast is presumably that the looping hand movement made in (73)b is easier to produce than that in (73)c, and correspondingly that the former has a more general meaning than the latter. We do not seek to derive this contrast in what follows but come back to it in the conclusion.

5.3 Intermediate conclusion

In sum, there are clear cases in which the viewpoint of evaluation of classifiers is dependent on quantifiers. We have developed an analysis in terms of (contextually restricted) existential quantification over viewpoints, thus beginning to justify the power afforced by the presence of viewpoint variables in the object language.

This analysis raises two questions. First, all the cases we considered involved narrow scope existential quantification over viewpoints. Are there cases of wide scope existential quantification over viewpoints? There is nothing in the analysis to force viewpoint variables to be bound, of course, and in fact we will see in the next section examples in which a classifier appears in the scope of an operator but does not depend on it; this was foreshadowed in (20)b, which involves free viewpoint variables in the scope of *always*. But as a reviewer asks, since we have examples with existential quantification in the scope of *always*, as in (74)a, can we also display examples with wide scope existential quantification, as in (74)b? We expect this case to arise, but proving its existence is complex because it must be distinguished from that in which the variable is free, as in (74)c (which is itself related to (20)b). If there is uncertainty about the signer's intentions, one might get existential-like truth conditions for the semantics and pragmatics combined—in essence: *for some plausible referential intentions, (74)c is true*. To obtain a definitive argument, we would need to have the existential quantifier appear in the scope of a further operator *Op* (e.g. *not*), as in (74)d; as a result, the crucial examples will be somewhat complex.[36]

[35] Our consultant noted in one session ([JL 22.05.05] in the Supplementary Materials B) that the rightward classifier movement might stand for an unspecified airplane movement, whereas the leftward movement, which is more awkward, is a bit more likely to be interpreted literally.
[36] An additional issue will be to distinguish this case from (74)a combined with an assumption that all the situations quantified over by *always* involve the same salient viewpoint.

(74) a. $always_T$ **∃π** there-is student person-walk-cl^{π}
b. **∃π** $always_T$ there-is student person-walk-cl^{π}
c. $always_T$ there-is student person-walk-cl^{π}
d. **Op ∃π** $always_T$ there-is student person-walk-cl^{π}

Second, restricting attention to narrow scope readings, does the analysis have to be in terms of existential quantification? Couldn't one have instead functional viewpoint terms that can be dependent on quantifiers? We discuss this possibility in some detail in Appendix I-C; as we argue, embedding under negation makes this analysis challenging under some assumptions.

6 Multiple Viewpoint Variables

Up to this point, the pictorial component of classifier predicates could be analyzed by way of a single viewpoint variable, although in some cases it had to be bound by a quantifier. But as we will now see, sometimes one arguably needs two viewpoint variables, as was schematically represented in (61)b, repeated as (75).

(75) … $classifier^{\pi}$ … $classifier^{\pi'}$…

The existence of such cases is an important argument for treating viewpoints in terms of variables (or possibly shiftable viewpoint parameters, as discussed in Appendix II).

6.1 Basic analysis

In (76), it is asserted that there is a student that walks in certain ways toward the teacher or toward the exit during the breaks of the philosophy class and of the linguistics class, which are held in different buildings. Each sentence includes two classifiers, one representing the student's movement in the philosophy class, and the other, their movement in the linguistics class. Both movements are represented from the teacher's perspective, with the signer corresponding to the teacher's location.[37] But since the classes are held in different buildings, there are two locations and thus two viewpoints involved.

(76) *Context*: The signer teaches a philosophy class and a linguistics class in different buildings.

HAVE ONE STUDENT $[\text{PHILOSOPHY CLASS}]_a$ BREAK ALWAYS __, $[\text{LINGUISTICS CLASS}]_b$ BREAK ALWAYS … .
'There is one student who during philosophy class breaks always ___ and during linguistics class breaks always … .'

a. [7] __ = a-PERSON-WALK-slow-cl-1 slowly approaches the teacher from the right	… = b-PERSON-WALK-slow-cl-1 slowly approaches the teacher from the left
b. [7] __ = 1-PERSON-WALK-fast-cl-a quickly leaves the classroom toward the right	… = 1-PERSON-WALK-fast-cl-b quickly leaves the classroom toward the left
c. [7] __ = a-PERSON-WALK-slow-cl-1 slowly approaches the teacher from the right	… = 1-PERSON-WALK-fast-cl-b quickly leaves the classroom toward the left
d. [7] __ = 1-PERSON-WALK-fast-cl-a quickly leaves the classroom toward the right	… = b-PERSON-WALK-slow-cl-1 slowly approaches the teacher from the left

(ASL, 35, 2228, 3 judgments; https://youtu.be/IXP4iPPsoZU)

These sentences can be analyzed by way of the Logical Form in (77), where distinct viewpoint variables, namely π and π', appear on the first and on the second classifier predicate (respectively written as *walk-cl* and *walk'-cl*, to indicate that they have different shapes).

(77) there-is student λi [$always_T$ t_i person-walk-$cl^{\boldsymbol{\pi}}$ and $always_{T'}$ t_i person-walk'-$cl^{\boldsymbol{\pi'}}$]

[37] For our second consultant, the viewpoint is the one that corresponds to the addressee viewing the 3D animation, i.e. from the back of the class in each relevant classroom.

Without getting into a full derivation of the truth conditions, the crucial lexical rule appears in (78): it is just an application of the more general rule in (53) to the predicate classifier *walk-cl* (or *walk'-cl,* of course).

(78) $[[\text{person-walk-cl}^{\pi}]]^{c, s, t, w} = \lambda x_e$. # if $\text{person-walk}_{t, w}(x) = \#$
1 iff $\text{person-walk}_{t, w}(x) = 1$ and **proj(x, s(π), t, w) = *person-walk-cl***, otherwise

When computing the final truth conditions, one projective condition will take the form proj(x, s(π), t, w) = *person-walk-cl*, with a viewpoint s(π) and a classifier movement *person-walk-cl*, and the other will take the form proj(x, s(π'), t, w) = *person-walk-cl'*, with a different viewpoint s(π') and a different classifier movement. The values of the viewpoint variables are determined by the context, and in this case we must assume that s(π) corresponds to the position of the teacher in the philosophy class, while s(π') corresponds to the position of the teacher in the linguistics class.

It should be noted that the two clauses that appear in (77) are in the scope of the same existential quantifier (and λ-abstractor). This makes it unlikely that the context (and thus possibly the viewpoint) shifts from one clause to the next because of a discourse phenomenon: in the absence of context-shifting operators (e.g. Schlenker 2003, 2011a), it is usually assumed that the context remains fixed sentence-internally. On this assumption, the contextual analysis of viewpoint-dependency stated in (60)is clearly insufficient. On the other hand, viewpoint variables are not the only possible alternative to the contextual view, as one could posit instead that iconic elements are evaluated with respect to a shiftable viewpoint parameter. The debate between variable-based and parameter-based analyses is a traditional one in the analysis of time, world and context dependency (e.g. Cresswell 1990, Deal 2020), and we expect that viewpoints will give rise to related debates. A comparison between the present account and a parameter-based view is sketched in Appendix II.

6.2 Further examples

To strengthen our conclusions, we turn to further examples with two distinct viewpoint variables in the scope of a single individual quantifier.

In (79), two take-offs in different cities are described in each sentence, one with a classifier movement away from the signer, the other with a classifier movement toward the signer. In both cases, the viewpoint is that of the air traffic control tower, but these are of course located in different points of space, hence two viewpoints are needed. A Logical Form such as (80) can adequately capture the data.[38]

(79) *Context*: The signer is a part-time air traffic controller in New York and in Boston.
HAVE ONE PILOT NEW-YORK$_a$ ALWAYS __, BOSTON$_b$ ALWAYS … .
'There is one pilot who in New York always __, and in Boston always … . '
a. __ = 1-PLANE-TAKE-OFF-2 … = 2-PLANE-TAKE-OFF-1
takes off away from the control tower takes off toward the control tower
b. __ = 2-PLANE-TAKE-OFF-1 … = 1-PLANE-TAKE-OFF-2
takes off toward the control tower takes off away from the control tower
'Camera position': in all cases, it is the relevant air traffic control tower.[39]
(ASL, 37, 2552; 3 judgments; https://youtu.be/C64QO-pTnHQ)

(80) there-is pilot λi [always_T t_i [plane-take-off-cl$^{\pi}$ in NYC] and $\text{always}_{T'}$ t_i plane-take-off'-cl$^{\pi'}$ in Boston]

[38] This example should be interpreted with a bit of caution because the signer moved his body towards locus *a* for the New York take-offs and towards locus *b* for the Boston take-offs; this might conceivably be interpreted as an instance of Role Shift, an operation we sought to not to mix with our main data points. The same remark applies to (81) below, but far less so to the structurally identical example in (156) in the Supplementary Materials A-2.
[39] Here our second consultant also thought the camera angle was from the air traffic control tower (whether Role Shift might have played a role in his judgment is unclear).

In (79)-(80), the signer is plausibly recounting his own visual experiences, since the context specifies that he is a part-time air traffic controller in the two relevant cities. One might think that this is crucial in order to obtain readings with two viewpoints in the same sentence (the same worry might arise in the context of (76)). But this is not so. In the following variation of (79), the beginning of the target sentence makes clear that the signer does not have direct experience with the events he describes, but this does not affect the acceptability and meaning of the sentences. On the other hand, it is plausible that the signer's status as an air traffic controller working at both airports helps make salient the air traffic control towers as the two viewpoints used in the sentence.

(81) *Context*: The signer has just been hired as part-time air traffic controller in New York and in Boston.
BEFORE MY TIME, HAVE ONE PILOT NEW-YORK$_a$ ALWAYS __, BOSTON$_b$ ALWAYS … .
'Before my time, there was a pilot who in New York always __, and in Boston always … . '

a. __ =	1-PLANE-TAKE-OFF-2 took off away from the control tower	… =	2-PLANE-TAKE-OFF-1 took off toward the control tower
b. __ =	2-PLANE-TAKE-OFF-1 took off toward the control tower	… =	1-PLANE-TAKE-OFF-2 took off away from the control tower

'Camera position': in all cases, it is the relevant air traffic control tower.[40]
(ASL, 37, 2560; 3 judgments; https://youtu.be/g_odIruSp00)

Related examples are discussed in the Supplementary Materials A-2 (with additional complexities).

Since we argued in Section 5 that existential quantification over viewpoints is needed, one might wish to reduce our examples with two viewpoint variables to this quantificational mechanism. This possibility is discussed in Appendix I-D.

7 3D Dynamic Representations: the Puppet Model

Our analysis would be incomplete if we didn't offer a more realistic analysis of the depictions offered by classifier predicates. Pictorial semantics is insufficient in two major respects: first, classifier predicates create 3D representations, whereas pictures are by definition two-dimensional; second, classifier predicates may create dynamic and continuous representations in addition to static ones.[41] We address each problem in turn.

7.1 The Puppet Model

We propose an analysis that may be called the "Puppet Model"; the intuition is by no means new (see for instance Liddell 2003a, Perniss 2007)[42], but to our knowledge the semantic formalization is. In a nutshell, the signer-as-puppeteer selects a viewpoint, and recreates in front of them a simplified and scaled representation of the denoted scene viewed from that point.

To see the general idea, suppose the signer wishes to represent Obama sitting across the Dalai Lama, as in (82)a. To convey the intuitive idea, we can think of the static 3D representation process as being constructed in steps. First, the signer selects a viewpoint from which the scene is perceived; in (82)a, this is the position of the smiley face (☺), between the two paintings on the wall behind the

[40] Our second consultant also took the camera angle to be from the air traffic control tower.
[41] Greenberg's semantics is static, and Abusch's semantics is designed for visual narratives made of discrete picture sequences. It should be added that dynamic 3D representations are also used in dance. Migotti 2021 explicitly addresses this problem, and offers a different solution (based on a kind of 3D version of perspective projection, rather than on scaling). We leave a comparison for future research.
[42] Using the puppet metaphor, Perniss 2007 writes: "The signer is like a puppeteer, manipulating the characters that appear on stage and constructing the event on a reduced scale in the area of space in front of the body." Liddell 2003a also emphasizes the importance of 3D orientation: "some aspects of the orientation of the hands must also be treated as analogical. That is, while the 'upright person' handshape is oriented vertically with respect to a horizontal surface, the direction the palm faces is variable. Similarly, the 'vehicle' classifier is oriented with the ulnar side of the hand down, while the direction of the fingertips is variable."

Dalai Lama. Second, the signer recreates in front of him a scaled version of this scene, with classifier predicates appearing in the position and with the orientation of scaled Obama and scaled Dalai Lama (the signer can't pick the size of the classifiers-qua-puppets, and as a result these need not be to scale). So if Obama and the Dalai Lama are separated by .5 meter and the signer chooses a 1/50 scaling, the corresponding classifiers will be separated by just 1 cm, with the Obama person classifier on the signer's right and the Dalai Lama person classifier on the signer's left, and with the position of the classifiers (facing each other) corresponding to that of the individuals they denote.

(82) **Obama sitting across from the Dalai Lama: from the scene to its classifier representation**[43]

a. Original scene b. Scaling c. Marking with classifiers d. Result

When classifier representations are dynamic, the position of the puppets will change accordingly, although we will need to allow for scaling of time as well (an airplane classifier will typically represent a flight between Paris and New York in 2 seconds rather than 8 hours).

A remark should be added about the notion of 'camera position', used in several inferential questions above. It makes clear sense in photographs or videos, but in puppet-like representations the notion is somewhat ambiguous: one can take the 'camera position' to be the world position in which one would be if one viewed the scene in the same way as the signer, or in the same way as the addressee. In the representation above, we selected in (82)a a real world position (= the smiley face ☺) that corresponds to the way the signer views the scene. But it would make equal sense to select a real world position corresponding to the way the addressee views the scene, in which case the selected viewpoint would be opposite that depicted by the smiley face in (82)a. Our main consultant, who was also the signer in all videos, understood the term 'camera position' to refer to the position of the smiley face. But our second consultant sometimes took it instead to refer to position facing it (see for instance fn. 37, Appendix I-C, Supplementary Materials A-2). This shows that one must be more careful than we have been with the term 'camera position', which is ambiguous with 3D representations in a way that it isn't with pictures or videos. But this issue needn't affect the semantic analysis: whether one chooses a signer viewpoint or an addressee viewpoint, the operations of scaling and marking in (82)b and (82)c can be performed in the same way. We will henceforth stick to definitions based on a viewpoint that corresponds to that of the signer.

7.2 Static projection

To define things more precisely, we will need some assumptions and definitions. For concreteness and simplicity, we will develop the analysis within elementary Cartesian geometry.

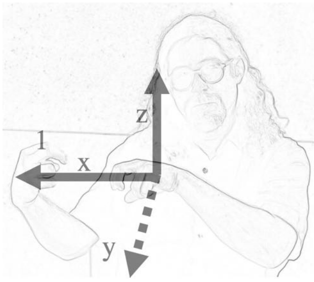
First, the signer is associated with a Cartesian coordinate system (or 'frame of reference') ρ*, with three orthogonal axes x, y z, and numerical coordinates on each axis. Its origin lies in the center of signing space, as is illustrated in (83). We assume that the signer's position is fixed in this coordinate system, here with his chest in position (0, -1, 0), meaning: the x and z coordinates are 0, and the y coordinate is -1 because the signer is 'behind' the coordinate system (the signer's head would be in a higher position, say (0, -1, 1)). For the signer, the x axis extends rightwards, the y axis toward his front, the z axis upwards (we restrict attention to 'right-handed' coordinate systems, with the x, y and z axes

---

[43] Picture credit (Obama and the Dalai Lama): Pete Souza, Public domain, via Wikimedia Commons (we use a mirror-image version).

in the same relative position as the thumb, index and middle fingers of the right hand). We could define the very same coordinate system while starting from the addressee's position, say with coordinates (0, 1, 0), and with the x axis extending toward the addressee's left, the y axis toward their back and the z axis upwards. This equivalence captures the fact that the notion of 'viewpoint' is underspecified in this system, as two salient choices correspond to the signer's position and to the addressee's position.

(83) A Cartesian coordinate system in signing space



This coordinate system will make it possible to situate the position of classifier predicates and loci , and also to encode their orientation by way of vectors (a vector can be thought of as an arrow that starts at the origin of the coordinate system, and reaches a certain point, whose coordinates define that vector).

Second, we will use the term 'viewpoint' with a new meaning, different from what it was in pictorial semantics. A viewpoint π will now determine a real-world coordinate system $\rho(\pi)$. For perspicuity, we will assume that all coordinate systems use the same units. Correspondingly, scaling will be implemented by way of a multiplicative factor applied to all positional coordinates; this scaling factor, $\sigma(\pi)$, will also be determined by π (and we will need a temporal scaling factor $\tau(\pi)$ as well).

Third, classifier predicates as well as real world objects will be assumed to come with a distinguished point, which we will call their 'center', and a distinguished 'orientation', encoded by way of three orthogonal vectors of unit length (i.e. an orthonormal basis; here too, we restrict attention to right-handed ones).

With these tools in hand, we can state that, relative to viewpoint π, a real-world object d projects to a classifier predicate *WORD* just in case (i) d is of the right kind (e.g. a person if this is a person classifier), and (ii) the position and orientation of d relative to $\rho(\pi)$ correspond (in terms of coordinates) to the position and orientation of *CL* relative to the signer's frame of reference ρ^*, *modulo* a scaling factor for position.[44]

These assumptions and definitions are stated in (84), and make it possible to define projection as in (85).

(84) **Assumptions and definitions**

a. For d any real-world object or sign language classifier, d is associated with:

(i) a distinguished point, its center, which encodes the position of d. Relative to a a frame of reference ρ, its coordinates are given by center(d, ρ), of the form <x, y, z>, a triple of real numbers;

(ii) a (right-handed) triple of orthogonal vectors of unit length that encode the orientation of d. Relative to a frame of reference ρ, the coordinates of this triple of vectors are given by orientation(d, ρ), of the form

[44] Two remarks should be added.

1. It might be that one and the same object, say a person, has different points that count as its 'center' depending on the classifier one considers. For instance, the sitting and standing person classifiers in (46)a,b might involve projections from different points—maybe the middle of the body for the sitting person classifier and the head for the standing person classifier. If so, the *center* function would need to have an additional argument, namely the word relative to which the center is computed—with the result that center(SIT-cl, d, ρ) might be different from center(STAND-cl, d, ρ). The same measure (namely the addition of an argument) might conceivably be necessary for the *orientation* function. We leave these refinements for future research.

2. Just like puppets, classifiers can sometimes be articulated, which would require an extension of the analysis.

<u, v, w>, where u is a triple of coordinates of a unit-length vector, and similarly for v and w.[45]

b. Any viewpoint π specifies a Cartesian frame of reference $\rho(\pi)$, a spatial scaling factor $\sigma(\pi)$ **(and for dynamic projection, a temporal scaling factor $\tau(\pi)$).**

c. The signer is associated with a distinguished frame of reference ρ^*, whose center (namely the point <0, 0, 0>) is in front of the signer, in the center of signing space.

d. Any classifier *WORD* (e.g. *PERSON-cl*) is associated with a set of objects that are of the appropriate type to appear in its extension. Relative to a time t and world w, we write this value (= the lexical content of the classifier) as $word'_{t,w}$.

We can now provide a revised definition of projection, adapted to the 3D sign language case.

(85) **Projection**
Let ρ^* be the Cartesian frame of reference associated with the signer, and let *WORD* be a word (e.g. a predicate classifier). For any viewpoint π, time t, world w, and object d,

proj(d, π, t, w) = WORD iff

(i) $\boxed{word'_{t,w}(d) = 1}$, and

(ii) modulo the scaling factor $\sigma(\pi)$ (> 0), the position and orientation of d relative to $\rho(\pi)$ correspond (in terms of coordinates) to the position and orientation of *CL* relative to ρ^*, or in formal terms:

a. center(d, $\rho(\pi)$) = $\sigma(\pi) \bullet$ center(WORD, ρ^*)[46]
b. orientation(d, $\rho(\pi)$) = orientation(WORD, ρ^*)

Let us illustrate. We start by assuming that the viewpoint of the Obama/Dalai Lama scene corresponds to someone observing the scene from behind, between the two paintings; this is the frame of reference notated as ρ^* in (85). With this frame of reference, the Dalai Lama is at the center, and thus has coordinates (0, 0, 0), while Obama has coordinates (50, 0, 0).[47]

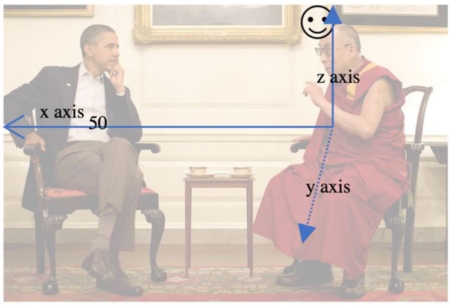
---

[45] As an anonymous reviewer notes, it would be more standard and elegant to define orientation as a triple of angles, as this suffices to characterize the triple of orthogonal vectors associated with the object (see for instance https://en.wikipedia.org/wiki/Orientation_(geometry)). Orientation would then be defined by a triple of angles, rather than by a triple of triples of vector coordinates.

[46] Here $\bullet$ represents scalar multiplication. Specifically, center(WORD, ρ^*) is a triple of coordinates <x, y, z>, and $\sigma(\pi) \bullet$ center(WORD, ρ^*) yields its pointwise scaling, namely $\langle\sigma(\pi)x, \sigma(\pi)y, \sigma(\pi)z\rangle$.

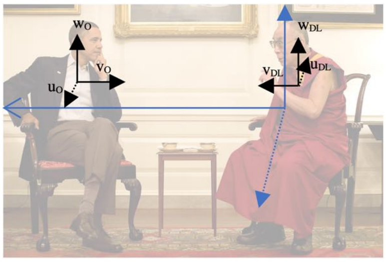
[47] While only the coordinates of the characters matter, the viewer, behind the Dalai Lama, can for instance be taken to be in position (0, -10, 0): the viewer's chest is behind the center of the frame of reference, and a bit higher than it. This is the viewer position corresponding to that of the signer, but we could equally define a viewer position corresponding to that of the addressee, which would look at the same scene from a position facing the smiley face in (86), say with the addressee's chest in position (0, 10, 0).

(86) A viewpoint and coordinate system for the depicted characters



Next, the orientation of Obama and the Dalai Lama is given by two triples of orthogonal unit vectors. We will call these vectors $<u_O, v_O, w_O>$ for Obama and $<u_{DL}, v_{DL}, w_{DL}>$ for the Dalai Lama. Each of these six vectors is defined by three coordinates (think of an arrow whose tail is at the origin of the coordinate system, and whose head extends by one unit in a certain direction; the coordinates of this head fully define the vector). Simplifying a bit, we take these vectors, represented in (87)[48], to have the coordinates in (88).

(87) Vectors encoding the orientation of the characters relative to a frame of reference



(88) Orientations of the two individuals

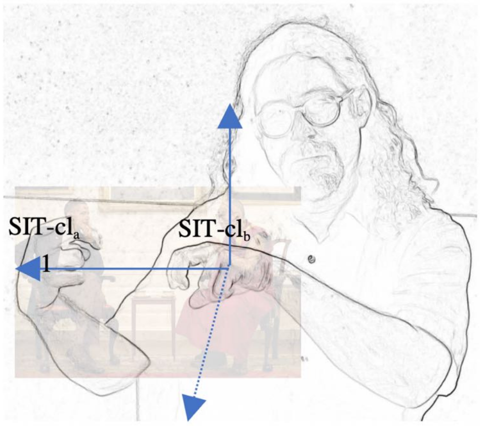
Obama:	$<u_O, v_O, w_O>$	=	<(0, 1, 0), (-1, 0, 0), (0, 0, 1)>
Dalai Lama:	$<u_{DL}, v_{DL}, w_{DL}>$	=	<(0, -1, 0), (1, 0, 0), (0, 0, 1)>

Finally, Obama and the Dalai Lama project to the classifiers $SIT\text{-}cl_a$ and $SIT\text{-}cl_b$ just in case in the signer's frame of reference (notated as $\rho(\pi)$ in (85)), the classifiers have the same coordinates modulo the relevant scaling factor as well as the same orientation as the objects they depict. We will assume that the scaling factor is 1/50. Since we conveniently picked our example so that all coordinates are 0 except for the Obama's x-coordinate, which is at x = 50, we only need to revise the latter, putting the Obama classifier at (1, 0, 0) in the signer's frame of reference ρ^*, as illustrated in (89)a. No scaling factor is applied to the vectors that define orientation, since all vectors are of unit length, and we just need to ensure that the very same triples of vectors define the orientation of the classifiers as the orientation of the individuals they depict, as illustrated in (89)b.

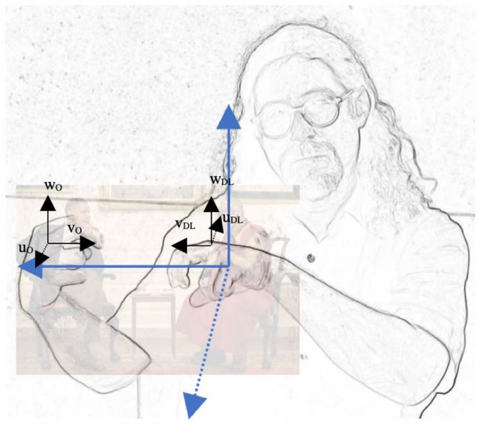
[48] Since orientation vectors are of unit length 1, they are not represented to scale in (87), as they would be too small to be visible (on the assumption that the x distance between Obama and the Dalai Lama is 50).

(89) **A scaled representation of the scene in (86) by way of classifier predicates**

a. Without orientation vectors



b. With orientation vectors



This general framework can also be applied to the semantics of iconic loci, with minor adjustments that are discussed in Appendix I-E.

It should be added that the condition in (85)(i) is lexical in nature: for *proj(d, π, t, w) = WORD* to be satisfied, d should satisfy the lexical meaning of *WORD* at t in w. In our example, for Obama to be represented by *SIT-cl$_a$* at t in w, it should be the case that at t in w Obama is a sitting person (and to this a specifically projective condition must be added). The result is that, as announced, the lexical conditions we had in (53) become redundant with the definition of projections in (85).[49]

7.3 Dynamic projection

While classifier predicates may be static, they may also move in signing space, in which case their movement is iconically interpreted; we will put an index τ on these temporally dynamic classifiers. We must thus extend our analysis to account for this dynamic character. We will view a dynamic classifier as a function from times to static classifiers. So if δ is the duration of the movement of *word-cl$_\tau$* in signing space, then for each δ' ≤ δ, *word-cl$_\tau$(δ')* is a static classifier.

It would be simplest to posit that at evaluation time t, *word-cl$_\tau$* is true of an object d just in case the classifier movement tracks the object movement starting at time t, or in other words: for each δ' ≤ δ, at t+δ' d projects to *word-cl$_\tau$(t+δ')*. But this implies that 1 second of the classifier movement corresponds to 1 second of the denoted object movement, which would lead to absurd results (to represent a flight from Paris to New York, the classifier movement would need to take as long, say 8 hours). It is clear that durations are scaled, but there is no reason to assume that the scaling factor is the same for time and for space. This is why we posited that a viewpoint π doesn't just make available a spatial scaling factor σ(π), but also a temporal scaling factor τ(π). We can thus require that for each δ' ≤ δ, at t+(τ(π))δ', d projects to *word-cl$_\tau$(δ')*. The scaling factor could be huge: if an 8-hour flight is represented by a 2-second movement, the scaling factor is… 14,400 (= (8*60*60) / 2).

This leads us to the definition in (90).

(90) **Dynamic projection**

Let ρ* be the Cartesian frame of reference associated with the signer, and let *word-cl$_\tau$* be a dynamic

[49] Two remarks should be added. First, we suggested above that the redundancy could be eliminated by doing without the boxed lexical condition in (53). An alternative would be to do without the boxed part of (85), but the result would be unintuitive: within the Puppet Model, it seems be part and parcel of the depiction that the sitting person classifier can only represent people in sitting position. Second, although the two versions are identical in case *WORD* is lexical (simplex), there could in principle be cases in which *WORD* is a complex expression containing variables. The boxed part of (53) would have no trouble dealing with this case, but the boxed part of (85) would, for lack of access to an assignment function.

predicate classifier. If δ is the duration of the classifier movement, we view *word-cl*$_\tau$ as a function from the interval $[0, \delta]$ to static classifiers. Then:
$proj(\pi, t, w, d) = word_\tau$ iff
(i) $word'_{t,w}(d) = 1$, and
(ii) for each $\delta' \leq \delta$, $proj(d, \pi, t+(\tau(\pi))\delta', w) = word\text{-}cl_\tau(\delta')$.

To illustrate, if the scene in (86) is modified so that Obama and the Dalai Lama are initially both standing, and then Obama moves toward the Dalai Lama to shake his hand, we may use two 'standing person' classifiers as in (16), facing each other, and initially in the same position as the sitting classifiers discussed above. Both will be construed as dynamic classifiers, so their representations will be *[PERSON-cl*$_\tau$*]*$_a$ for Obama on the signer's right and *[PERSON-cl*$_\tau$*]*$_b$ for the Dalai Lama on the signer's left. Restricting attention to the Obama classifier, if the expression *[PERSON-cl*$_\tau$*]*$_a$ moves toward b for a duration of 1 second, and the scaling factor is 5, the condition in (90)(ii) will translate into (91), and will require that for 5 seconds after t, Obama moves toward the Dalai Lama in a way that mirrors the classifier movement.

(91) $proj(\pi, t, w, Obama) = CL\text{-}person_\tau$

This concludes our sketch of the main revisions needed to adapt pictorial projection to sign language 3D representations. The revised notion can be plugged into the truth conditions we discussed informally in earlier sections. Importantly, the analysis is highly modular, reflecting the fact that depictive semantics and compositional semantics are very different systems, although both are needed to account for sign language meaning. As long as some notion of projection and viewpoint can be specified, it can be embedded within the iconological system we defined in earlier sections; but some explicit notion of projection is needed if explicit truth conditions are to be derived.

8 Iconological Semantics for Gestures and Onomatopoeias

While one might think that the iconic mechanisms we described and their integration to Logical Forms are unique to sign language, this is not so. We will now discuss data that suggest that they have counterparts with pro-speech gestures and even with pro-speech onomatopoeias (which can be analyzed as vocal gestures). For pro-speech gestures, the Puppet Model sketched in the preceding section could be applied with little modification;[50] for pro-speech onomatopoeias, a full analysis of auditory representations would be needed, something we leave for future research.

The existence of gestural counterparts of iconic loci was already discussed in print (Schlenker 2020), and thus we focus on gestures that might resemble classifier predicates. Importantly, the literature distinguishes between the case of object viewpoint gestures, where the scene is enacted in front of the speaker, and that of character viewpoint gestures, where the scene is enacted by the speaker herself (e.g. McNeill 1992, Parill 2009, 2010). An illustration appears in (92) (= Parill 2009, Fig. 1a,b,c): the cartoon scene in (92)a is represented with an object viewpoint gesture in (92)b and with a character viewpoint gesture in (92)c.

(92) a. Cartoon scene: hopping b. Object viewpoint gesture c. Character viewpoint gesture

[50] We say 'little modification' rather than 'no modification' because the lexical component of pro-speech gestures is far less clear than that of classifier predicates.

Our ASL classifier examples involve object viewpoint and thus we will restrict attention to object viewpoint gestures (a possible counterpart of character viewpoint gestures would involve Role Shift [Davidson 2015]; we leave its interaction with iconological semantics for future research).

8.1 Iconological semantics for pro-speech gestures

On the syntactic side, a remarkable finding is that in pantomimes, subjects tend to produce SOV orders irrespective of the word order of their native language (Goldin-Meadow et al. 2008). Schlenker et al. (to appear) replicate the main finding with sequences of some pro-speech gestures embedded in French sentences, and connect this behavior to that of ASL classifier predicates: although the basic word order of ASL is SVO, standard classifier predicates override this order and preferably yield SOV orders as well.[51] Following earlier insights, they propose that 'iconic syntax' is at the root of both phenomena: with classifier predicates and pro-speech gestures alike, highly iconic constructions create a visual animation, with the arguments appearing in the order that would be natural in a simplified cartoon. The main motivation for this analysis is that the traditional preference for preverbal objects only holds if the patient is presented as visible before the action, for instance in a scenario in which a crocodile swallows a ball (one sees the ball before the swallowing). When the patient only becomes visible after the action, as in a crocodile spitting out a ball, an SVO order is regained. This generalization holds both with ASL classifier predicates and with (French) pro-speech gestures, which highlights the similarity between the two case.

On the semantic side, pro-speech gestures are usually iconic in nature, just like classifier predicates; in addition, some authors have explicitly suggested that classifier predicates have a gestural component (Schembri et al. 2005, concurring with Liddell 2003b).[52] We might thus expect that they can mirror the behavior of classifier predicates not just in their syntactic behavior, but also in their interaction with viewpoints. Starting from our own introspective judgments, we designed a small survey with 7 consultants, all of them linguists (one of them had provided feedback on a draft of the survey), with both acceptability and (multiple-choice) inferential questions.[53] Since our goal is just to show that some examples mirror mechanisms found in sign language, we only discuss clear cases below. Detailed results can be found in the Supplementary Materials B (this includes examples that we thought were clear, but were not to our consultants). We indicate as superscripts our consultants' average acceptability scores on a 7-point scale, with 7 = best, and we provide links to anonymized videos.

Transcriptions:
FLY-╰$_{front}$ transcribes a flat hand moving horizontally toward the speaker, and then abruptly upwards.
FLY-╰$_{right}$ transcribes the same movement on the speaker's right.
FLY-//// transcribes a flat hand, representing an airplane, moving upwards in a bumpy fashion.

In (93), the dominant reading has the pilot flying toward the control tower, and vertically; the viewpoint would thus be that of someone in the control tower (and this person couldn't be the speaker, since he is recounting events that took place before his time).

(93) **Salient viewpoint**
Context: The speaker has just been hired as an air traffic controller at JFK airport.

$^{5.3}$ Before my time, there was a pilot who upon take-off would always FLY-╰$_{front}$.

---

[51] This is a simplification, as Schlenker et al. (to appear) find in gestures and classifier predicates alike not just SOV but also OSV orders.
[52] Specifically, Schembri et al. 2005 argue that "classifier constructions are blends of linguistic and gestural elements". Following Okrent 2002, they assume that "gestures, unlike morphemes, are relatively unconventionalized; they exhibit a relationship between form and meaning that may be gradient, and they are used to express imagistic aspects of thought through forms specifically created to reflect aspects of that imagery."
[53] In some cases, some consultants wrote that they didn't get any of the inferences they had to choose from; full survey results should be consulted for details.

https://youtu.be/glZ9PFxOaig

Dominant reading: The pilot always took off toward the air control tower, and vertically.

A simplified Logical Form appears in (94). It involves a gestural predicate FLY-\\$^{\pi}$ that includes a viewpoint variable π. To yield the target reading, the context must specify that π denotes a viewpoint associated with the JFK control tower.

(94) there-is pilot λi always$_T$ t_i FLY-\\$_{front}^{\pi}$
where the context specifies that the value of π is a viewpoint associated with the JFK control tower.

There are more complex examples in which viewpoint choice is dependent on a universal quantifier, as illustrated in (95) .

(95) **Viewpoints dependent on universal quantifiers**
Context: The speaker is an airline pilot with prior experience in Crete, which has several airports.

[5.7] In Crete, when the weather is bad, every hour there is a plane that FLY-////.
https://youtu.be/wmrNPppEW04

Dominant reading: Every hour, there is some point or other of space in which there is a plane that has a bumpy ascent.

There are also cases in which an existential quantifier over viewpoint takes scope under negation, as in (96) (such cases are important for the discussion of Appendix I-C).

(96) **Existential quantification over viewpoints in the scope of negation**
Context: The speaker is a pilot talking to another pilot.

[5.4] Tomorrow the weather will be nice, so your plane won't FLY-////.
https://youtu.be/qHxsyRD74B4

Dominant reading: There is no point of space in which the plane will have a bumpy ascent.

(95)a can be analyzed with a modified version of the Logical Form discussed for classifier predicates in (64), adapted in (97).

(97) always$_T$ ∃π there-is plane FLY-///$^{\pi}$

Disregarding tense, the negative case in (96) can be analyzed with the Logical Form in (98): the existential quantifier over viewpoints takes scope under negation (for a structurally related example with a classifier predicate in ASL, see Appendix I-C, (113)).

(98) not ∃π [your airplane] FLY-///$^{\pi}$]

Cases that involve two free viewpoint variables can be replicated as well, as in (99), where the first gesture is evaluated relative to a viewpoint involving the New York control tower, while the second gesture is evaluated relative to the Boston control power.

(99) **Evaluation relative to two viewpoint variables**
Context: The speaker works as a part-time air-traffic controller in New York and in Boston.

[4.7] Before my time, there was a pilot who upon take-off in New York would always FLY-\\$_{front}$, and in Boston would always FLY-\\$_{right}$.
https://youtu.be/xbPxD9gfXh8

Dominant reading: 1. In New York, the pilot would take off toward the local air traffic control tower from the front. 2. In Boston, he would take off toward the local air traffic control tower from the right.

A simplified Logical Form appears in (100), which involves distinct viewpoint variables, π and π'. To obtain the target reading, the context must guarantee that they denote viewpoints associated with the New York and with the Boston control tower respectively.

(100) there-is pilot λi [$always_T$ t_i FLY-$\backslash_{front}^{\pi}$ in NYC] and $always_{T'}$ t_i [FLY-$\backslash_{right}^{\pi'}$ in Boston]
where the context specifies that the value of π is a viewpoint associated with the NYC control tower, and he value of π' is a viewpoint associated with the Boston control tower

If this sketch is on the right track, pro-speech gestures are similar to sign language classifier predicates not just in their syntax, but also in their viewpoint-sensitivity. The difference is of course that pro-speech gestures don't come with lexical entries and thus have a far less precise meaning than classifier predicates. Still, the analogy could be more generally useful: *modulo* the absence of clear lexical entries, some pro-speech gestures might make it possible to 'import' the behavior of classifier predicates into spoken language.

8.2 Iconological semantics for pro-speech onomatopoeias

Onomatopoeias are sometimes called 'vocal gestures' to highlight their similarity with manual gestures (e.g. Grenoble et al. 2015, Schlenker 2018e, Migotti and Guerrini, to appear). With a focus on musical stimuli, Pasternak 2019 and Migotti and Guerrini, to appear use ascending or descending scales to evoke an increase or decrease in light intensity, or a movement up or down. By analogy with analyses of semantic correlates of loudness in music (e.g. Schlenker 2017d, 2022), we can also use increasing or decreasing loudness to represent an object approaching or moving away. This has the advantage of involving a clear notion of auditory perspectival point, which can be seen as an auditory analogue of viewpoint in pictorial semantics.

Such vocal gestures make it possible to replicate the patterns we found with pro-speech gestures, as can be seen in the simple examples in (101), where sh<>sh transcribes a fricative noise becoming louder and then softer.

(101) **Evaluation relative to a single auditory point**
Context: The speaker lives in France.
[5] On Bastille Day, I spend my time on my balcony. There are always fighter jets that sh<>sh.
https://youtu.be/AuDkVtAX9zM

Dominant reading: There are always jets that approach the speaker's balcony with increasing noise and move away with decreasing noise.

Auditory points also can also be dependent on universal quantifiers, as seen in (102).[54]

(102) **Auditory points dependent on universal quantifiers**
Context: The speaker is a journalist who used to be a correspondent in France.

[5.1] In France, on Bastille Day, there are always fighter jets that sh<>sh.
https://youtu.be/7htSTCEYhas

Dominant reading: There are always jets that approach some point or other with increasing noise and move away with decreasing noise.

There are also cases in which two vocal gestures are evaluated relative to different auditory points in the same sentence, as illustrated in (11), where *shshsh* transcribes a soft continuous fricative sound.

---

[54] In order to test existential quantification in the scope of negation, one would need to assess examples such as (i):

(i) *Context:* The speaker is talking to a fighter jet pilot who always takes part in the Bastille Day parade. There are real concerns about noise pollution in the country, so on Bastille Day next year, your jet won't sh<<>>sh.

(103) **Evaluation relative to two auditory points**

Context: The speaker lives in France.

[4.7] On Bastille Day, in my neighborhood fighter jets always sh<>sh, whereas over the Champs-Elysées they $_{\text{shshsh}}$.

https://youtu.be/JmoKUc6HX3M

Dominant reading: 1. Fighter jets fly over the speaker's neighborhood with considerable and increasing noise before moving away. 2. Fighter jets fly over the Champs Elysées with less noise.

These cases can be analyzed with Logical Forms that are strikingly similar to ones we postulated before, with viewpoint/auditory point variables, for instance (104) for existential quantification in the scope of universal quantifiers, and (105) for evaluation of two vocal gestures relative to two distinct auditory points (further examples would be needed to exclude alternative analyses).

(104) $\text{always}_T\ \exists\pi$ there-is military-jet sh<>sh^{π}

(105) [always_T fighter-jets [sh<>sh^{π} in my-neighborhood] and [$\text{always}_{T'}$ fighter-jets[$_{\text{shshsh}}{}^{\pi'}$ in Champs-Elysées]]

Needless to say, the depictive component of our semantic analysis would need to be adapted to apply to iconic sounds, something we leave for future research.

9 Conclusion

9.1 Main results

Sign language linguistics has traditionally been divided between the 'formal camp' (e.g. Sandler and Lillo-Martin 2006 and references therein) and the 'iconic camp' (e.g Cuxac 1999; Cuxac and Salandre 2007; Taub 2001; Liddell 2003b). But it is clear that iconicity can and must be integrated to formal approaches. Earlier formal analyses did one of two things: they offered piecemeal integrations of iconic effects (e.g. Schlenker et al. 2013, Kuhn and Aristodemo 2017, Schlenker 2018b), with *ad hoc* rules for each construction; or they analyzed iconic expressions as making demonstrative reference to their own form, but without any explicit treatment of their truth conditions (e.g. Zucchi 2011, 2017, Davidson 2015). By contrast, for classifiers at least, iconological semantics offers a systematic way of integrating sign language iconicity to a formal approach: sign language semantics is in essence the union of standard compositional semantics and of pictorial semantics. The glue between them lies in viewpoint manipulation; while evaluation under a time and world works in similar ways for words and pictorial representations, the latter also depend on a viewpoint, which in our account is systematically provided by viewpoint variables. These can be free or bound. In the latter case, there can be an extraordinary interaction between logical operators and pictorial semantics. While pictorial semantics in the Greenberg/Abusch tradition offers a good point of departure to offer an explicit account of sign language depictions, it must be refined. In particular, the Puppet Model of sign language classifiers can be made precise by replacing the notion of projection onto a plane with scaling to a 3D representation, with further adjustments to obtain dynamic notions.

While it might initially seem that the depictive component of iconological semantics is solely relevant to sign language, this is not so: pro-speech gestures appear to display a viewpoint-sensitivity that is similar to that of classifier predicates. There remains an important difference between classifier predicates and pro-speech gestures: the former have lexical entries, the latter usually don't.

More broadly, the 'textbook view' of natural language semantics ought to be expanded: besides the compositional component, human language has a depictive component which is tightly connected to it. In our terms, logical semantics ought to be extended into an iconological semantics.

9.2 Open questions

On an empirical level, we have only discussed two salient cases of interaction between the logical skeleton of sign language and depictions, namely iconic loci and classifier predicates. There are many further iconic constructions that ought to be revisited from an iconological perspective, including but not restricted to: (i) iconic plurals (e.g. Schlenker and Lamberton 2019, 2022), (ii) iconic pluractionals (e.g. Kuhn and Aristodemo 2017), (iii) iconic modulations of verbs (including telic verbs such as LSF *UNDERSTAND*, atelic verbs such as LSF *REFLECT* or ASL *GROW;* Schlenker 2018b), (iv) visible degrees (e.g. Aristodemo and Geraci 2018), and (v) Locative Shift (Schlenker 2018a). Role Shift is also known to have a strong iconic component (e.g. Schlenker 2017a,b, Davidson 2015), and an iconological account has yet to be developed for it. On the gestural front, we solely considered object viewpoint pro-speech gestures. Generalizing the analysis to character viewpoint gestures will be important; and the analysis of pro-speech gestures should extend to co-speech gestures, with additional complexities (because they modify words).

On a theoretical level, there are three types of open questions. One type pertains to the depictive component. The Puppet Model is only a first stab at an explicit account, and it should be made more precise and sophisticated. We wrote as if all the details of the position and movement of classifier predicates are semantically interpreted, but this is probably not so, as seen in fine-grained contrasts (in (62)and (72)) between 'easy to produce' and 'hard to produce' movements: the former may be interpreted as generic, the latter tend to be interpreted precisely. An account has yet to be offered. In addition, it is an open question how more abstract notions of iconicity discussed in the literature should be unified with the present account; these include the diagrammatic use of plural loci (Schlenker et al. 2013) and ordering by size of domain restrictions (Davidson and Gagne 2022).

A second type of question pertains to our Logical Forms. In this initial attempt, we dealt with viewpoint manipulation by way of object language variables, in part for reasons of perspicuity. This provides considerable expressive power, and in some cases it is indeed useful. But a less expressive alternative is to do everything with an implicit but shiftable viewpoint parameter. This is a more constrained framework because one can think of a parameter as the value of a unique viewpoint variable.[55] A related debate exists in the analysis of time and world dependency (Cresswell 1990), and especially in the analysis of context-shifting constructions: initial accounts of 'monsters' (= constructions that shift the context of evaluation of indexicals) were developed in terms of object language context variables (Schlenker 2003); later accounts argued for more restrictive frameworks with context parameters and shifting operators (e.g. Anand and Nevins 2004, Anand 2006, Deal 2020). A similar debate will arise in the analysis of viewpoint manipulation; see Appendix II for a parameter-based restatement of the present, variable-full analysis.

A third but related type of question pertains to the relation between viewpoints for iconic representations and viewpoints for perspectival expressions such as *come* and *go*.[56] Sudo 2015 develops an account of the latter in terms of viewpoint parameters that comes very close to the parameter-based version of the present analysis. A key question is to determine whether one and the same notion can or should account for both kinds of constructions; here too, see Appendix II for some remarks, and some examples with co-speech gestures in which the viewpoint of evaluation of *come* and *go* is preferentially aligned with that of the corresponding co-speech gestures.

---

[55] To sketch the idea in greater detail: If we can only make use of a single viewpoint variable π, we can define things as in (i)a, with the understanding that the assignment function s can only make reference to viewpoint variable π (although it may assign values other variable types, notably individual-denoting variables). We can also rewrite this as in (i)a, where we have divided s into two parts: an assignment function s- that assigns a value to all variables except π; and a specification of the value of π. But if there is only one viewpoint variable to take care, we can treat its value as a parameter, as in (i)c, where v alone replaces $\pi \rightarrow v$.

(i) If the only viewpoint variable is π and $s(\pi) = v$, and if s- is an assignment function identical to s except that it does not assign a value to π, we can write things near-equivalently as:

a. $[[F]]^{c, s, t, w} = \ldots$

b. $[[F]]^{c, s-, \pi \rightarrow v, t, w} = \ldots$

c. $[[F]]^{c, s-, v, t, w} = \ldots$

[56] Thanks to Ora Matushansky (p.c.) for insisting that we explore the connection between viewpoints for sign language and viewpoints in the semantics of *come* and *go*.

Finally, a more systematic comparison with other depiction-friendly frameworks should be offered in the future. Outside of formal semantics, Liddell 2003a,b and Dudis 2004 analyzed some of the same data (with insights we followed) within cognitive grammar; Huenerfauth 2004 did as well within a framework for machine translation. Within formal semantics but with very different data, Rooth and Abusch 2019 and Maier 2019 develop frameworks that allow for interactions between pictures and texts in comics.[57] The present study should be seen as a contribution to this broader movement that seeks to integrate grammatical and depictive phenomena.

[57] In addition, Maier and Bimpikou (2019) and Schlöder and Altshuler (to appear) develop versions of DRT for purely pictorial sequences.

Appendix I. Details and Alternatives

A. Unifying phi features and iconic loci [supplement to Section 3.2]

Our analysis of iconic specifications developed in (29) can be further unified with that of gender features by assuming that pronouns can generally include predicates that constrain their denotation. These predicates include phi-features, and iconic predicates of the form P^{π} (see also Esipova 2019 for an analysis in which phi-features are treated as modifiers). On this view, both (106)a and (106)b have the general form in (106)c. We can then posit that *P* is in essence an area of space that serves as a pictorial representation, and is interpreted as in (107).

(106) a. $[[pro_a^{fem}]]^{c, s, t, w} =$ # if $[[fem]]^{c, s, t, w}(s(a)) \neq 1$
s(a), otherwise

b. $[[pro_a^{P\_\pi}]]^{c, s, t, w} =$ # if $[[P^{\pi}]]^{c, s, t, w}(s(a)) \neq 1$
s(a), otherwise

c. $[[pro_a^{F}]]^{c, s, t, w} =$ # if $[[F]]^{c, s, t, w}(s(a)) \neq 1$
s(a), otherwise

(107) $[[P^{\pi}]]^{c, s, t, w} = \lambda x_e.$ # if x = #
1 iff proj(x, s(π), t, w) = P, otherwise

B. The demonstrative analysis of classifier predicates [supplement to Section 4.5]

Zucchi 2011, 2017 and Davidson 2015 take classifier predicates to literally have a hidden demonstrative component, as suggested by Zucchi's paraphrase in (40)b. In Zucchi's words (2017),

> the intuitive idea behind these proposals is that movement in classifier predicates of motion is a gesture which fixes the referent of a hidden demonstrative (i.e., a demonstration), in the same way in which the reference of [a] demonstrative may be fixed by a gesture: *The car moved in a way similar to this*.

From the present perspective, there are major objections to this analysis. First, it does not offer an explicit semantics (Schlenker 2018c). If a pictorial semanticist had stated that the semantics of a picture is just: *the world is like this*, where *this* refers to the very form of the picture, nobody would have thought that an *analysis* had been offered. The hidden demonstrative analysis does not offer an analysis of the pictorial component either; rather, it presupposes one, but leaves it implicit.[58]

Second, even if one is willing to supply the missing pictorial component, one must ask whether there is any evidence for the demonstrative component. If classifier predicates contained an explicit word akin to *this*, there would be such an argument. But as things stand, nothing in the form of classifier predicates necessitates such a decompositional analysis. The analogy with *moves in a way similar to this* is deceptive: in English, *this* could refer to any salient representation, whereas the ASL classifier predicate can only refer to its own form (this difference is acknowledged in Zucchi 2011).[59]

---

[58] Zucchi 2011 appeals to a contextually specified 'mapping' between signing space and real space to obtain explicit truth conditions.

[59] Let us mention three further objections. First, without an explicit treatment of the pictorial component, it is difficult for the demonstrative analysis to handle the sophisticated interactions between iconic representations and viewpoints discussed in later sections. Second, as noted in Schlenker 2018c, when one considers in detail the inferences triggered by classifier predicates, there are differences between them and constructions that come close to the 'like this' paraphrases. Specifically, classifier predicates sometimes trigger presuppositions in cases in which 'like this'-like paraphrases don't (Schlenker 2018c, building on data from Schlenker 2021). Third, the demonstrative analysis might be faced with an Occam's razor problem. In three respects, the semantics of classifier predicates is similar to that of pro-speech gestures: (i) in terms of the inferential types they trigger, such as at-issue content and presuppositions (Schlenker 2021); (ii) in terms of their syntax (both types of constructions give rise to unexpected preverbal objects in a range of cases, as discussed in a comparative fashion in Schlenker et al., to appear); and (iii) in terms of their interaction with viewpoints (see Section 8 of the present piece). It is clear that a propositional/predicative and iconic semantics must be offered for pro-speech gestures, and furthermore

Still, if one wishes to make the demonstrative analysis explicit, one can in the end rely on iconological semantics. While Zucchi 2011 assumes an event semantics, his ideas can be adapted to the present framework. A Zucchian analysis yields the truth conditions in (108)a, where the boldfaced condition specifies that at t in w, the location of object x (in real space) is similar to the location of the signed token of P (in signing space). This can be compared to the iconological analysis, repeated in (108)b (to avoid confusion, we write as $[[P]]_Z$ the Zucchian lexical entry).

(108) If P is a (token of a) 1-place classifier predicate of position,
a. Demonstrative analysis (in the spirit of Zucchi 2011)
$[[P]]_Z^{c,s,t,w} = \lambda x_e$. # if $[[P]]^{c,s,t,w}(x) = \#$
1 iff $[[P]]^{c,s,t,w}(x) = 1$ and **similar(location(x), t, w, P)**, otherwise
b. Iconological analysis (present piece)
$[[P^{\pi}]]^{c,s,t,w} = \lambda x_e$. # if $[[P]]^{c,s,t,w}(x) = \#$
1 iff $[[P]]^{c,s,t,w}(x) = 1$ and **proj(x, s(π), t, w) = P**, otherwise

Crucially, the Zucchian analysis still needs to explain what 'similar' means (a point granted in Zucchi 2011). One way to do so is by way of the notion of a projection, as in (109).

(109) **A specification of the Zucchian analysis in (108)a**
For every context c (specifying a viewpoint π_c), object x, time t, world w, and classifier predicate of position P, similar(location(x), t, w, P) if and only if proj(x, π_c, t, w) = P

Roughly put, the location of x is similar to P at t in w just in case x projects to P at t in w. But this still requires that a notion of projection be specified, as was done in Section 7. In addition, a projection involves a viewpoint, and it is natural to take it to be specified by the context c; we write this contextual viewpoint as π_c in (109).

When the Zucchian theory is made explicit in this way, it becomes a special case of the analysis advocated in this piece. Specifically, the Zucchian analysis as made precise by (109) yields the same meaning as the present analysis *on the assumption* that the variable π denotes the contextually specified viewpoint π_c, as is stated in (110).

(110) If $s(\pi) = \pi_c$,
$[[P]]_Z^{c,s,t,w} = \lambda x_e$. # if $[[P]]^{c,s,t,w}(x) = \#$
1 iff $[[P]]^{c,s,t,w}(x) = 1$ and similar(location(x), t, w, P), otherwise
$= \lambda x_e$. # if $[[P]]^{c,s,t,w}(x) = \#$
1 iff $[[P]]^{c,s,t,w}(x) = 1$ and proj(x, π_c, t, w) = P, otherwise
$= [[P^{\pi}]]^{c,s,t,w}$

The iconological analysis has the advantage of being more explicit, of avoiding the detour through demonstrative reference, and also of allowing for a manipulation of the viewpoint variable π, as we saw in Sections 5 and 6.

C. Viewpoints dependent on quantifiers: alternative analyses [supplement to Section 5.3]

In Section 5, we argued that there can be existential quantifiers over viewpoints with scope under other operators. But it could be objected that viewpoint dependency on quantifiers need not imply the presence of a narrow scope existential quantifier over viewpoints. Instead of positing the Logical Form in (111)a, we could posit that in (111)b. It has no existential quantifier over viewpoints, but it contains a functional viewpoint-denoting term, $\Pi(x)$, which is dependent on the variable x introduced by a universal quantifier over runways.

that it is unlikely to be due to a lexical demonstrative component, since these expressions usually don't have a predetermined lexical form in the first place. Once an iconic propositional/predicative semantics is available for pro-speech gestures, one can use it for classifier predicates as well (the only difference is that marking rules have a conventional component in classifier predicates, but this can occur across pictorial constructions, and is not at all an obstacle to a projection-based analysis, as was detailed in Section 7 above). Proponents of the demonstrative analysis might have to posit that, on non-lexical grounds, pro-speech gestures have a demonstrative meaning as well.

(111) a. [all runway] $\lambda x\ \exists\pi\ t_x$ have [airplane plane-fly-cl$^{\pi}$]
b. [all runway] $\lambda x\ \ t_x$ have [airplane plane-fly-cl$^{\Pi(x)}$]
c. $\exists\Pi$ [all runway] $\lambda x\ \ t_x$ have [airplane plane-fly-cl$^{\Pi(x)}$]

The reason this analysis can be given is that there is presumably some pragmatic leeway in determining the value of the function term *Π*, with the result that the semantics combined with the pragmatics is close to what is represented in (111)c. But the latter is just an instance of 'Skolemization', the process by which existential quantifiers in the immediate scope of a universal quantifier, as in (112)a, is equivalent to wide scope existential quantification over appropriate functions, as in (112)b.[60]

(112) a. $Qx_e\ \exists y_e\ P(x, y)$
b. $\exists f_{<e,e>}\ Qx_e\ \ P(x, f(x))$

Importantly, this equivalence fails in non-upward-monotonic environments. This is easiest to see with a simple negation, as in (113). If the existential quantifier over viewpoints *∃π* has a sufficiently large domain restriction, no matter what the functional dependency *Π* is, (113)a entails (113)b,c but the converse isn't true. Importantly, we assume for the moment that *Π* is only functional on its individual argument *x*, not on the implicit time argument with respect to which any sentence is evaluated.

(113) a. not $\exists\pi$ [your airplane] plane-fly-cl$^{\pi}$]
b. not [[your airplane] plane-fly-cl$^{\Pi(x)}$
c. $\exists\Pi$ [not [your airplane] plane-fly-cl$^{\Pi(x)}$]

The problem with (113)b is that it is too weak. Since the variable *x* isn't bound, and since *Π* doesn't depend on further arguments, *Π(x)* selects a viewpoint, say at the back of the plane as it takes off, and the claim is just that the airplane doesn't fly as depicted from that viewpoint. But from this it doesn't follow that there isn't another viewpoint—say one that it behind the plane 5 minutes after take-off—from which the plane would fly as shown. Since (113)c is even weaker than (113)b, it too is overly weak.

Turning to ASL, it is clear that the classifier examples in (114)b,c involve an iconic representation and thus a viewpoint, but crucially not one that corresponds to a fixed point of space, as this would make the meaning overly weak.[61] So (113)a is a good analysis of (114)b,c but (113)b,c aren't. (A related paradigm is discussed in Supplementary Materials A.)

(114) *Context:* The signer is a pilot chatting with another pilot.
TOMORROW WEATHER GOOD. POSS-2 PLANE WON'T
'Tomorrow, the weather will be good. Your plane won't
a. [7] FLY ROUGH.
have a rough flight.'
b. [7] PLANE-FLY-move-upwards-bumpy-cl.
have a bumpy ascent.'
c. [7] PLANE-FLY-move-downwards-bumpy-cl.
have a bumpy descent'.
(ASL 37, 2568 ; 3 judgments; https://youtu.be/0oQY9ba4G-s)

---

[60] The equivalence holds more generally if *Q* is an upward-monotonic quantifier. Let us write **E** for the semantic value of an expression *E*, and let us prove the entailment in both directions.
–If (112)a is true, the set D = {d: ∃y **P**(d, y) = 1} is in **Q** (seen as a set of sets). For each d, if d is in D, one can choose a y such that P(d, y) = 1 and define f(d) = y, making arbitrary choices if d isn't in D. It is then clear that {d: **P**(d, f(d)) = 1} = D, witnessing the truth of (112)b.
—Suppose now that (112)b is true. Then for some f, D' = {d: **P**(d, f(d)) = 1} is in **Q**. Furthermore, D = {d: ∃y **P**(d, y) = 1} is a superset of D', so by upward monotonicity of *Q*, D is in **Q**, which shows that (112)a is true.
[61] The advantage of using a simple negation rather than a quantifier is that it trivializes the functional dependency determined by ∏ and thus simplifies the discussion.

Here it is worth mentioning two of the questions we used to elicit data (see the Supplementary Materials B for full data). One pertains to the fixed or non-fixed nature of the spatial viewpoint from which the flight is depicted, and the other to the location of the viewpoint, as can be seen in (115).

(115) Inferential questions for (114)

	One infers that tomorrow's the addressee's plane won't…			If some movement is represented iconically, from which perspective is it represented, i.e. where is the 'camera position'?
	(i) fly roughly, without further specification	(ii) fly as iconically depicted, from a fixed spatial viewpoint	(iii) fly as iconically depicted, no matter what the viewpoint	
a. FLY ROUGH	7	1	1	n/a
b. PLANE-FLY-upwards-bumpy-cl	2.3 (1, 2, 4)	3	6.7	from behind the plane
c. PLANE-FLY-downwards-bumpy-cl	2	2 (4, 1, 1)	6.7	from behind the plane

The question about the viewpoint, i.e. the position of the 'camera position', would be unambiguous if we were talking about a picture or a video. But as we saw in Section 7, this is a simplification: classifiers are really 3D representations, and behave very much like simplified puppets enacting a scene. Accordingly, the question can be interpreted in two ways. (i) Where in the world would one be if one viewed the scene in the same way as the signer? (ii) Where in the world would one be if one viewed the scene in the same way as the addressee? Our main consultant, who also signed the videos, systematically went for interpretation (i). Our second consultant often went for interpretation (ii) instead. As we argued in Section 7, the disagreement does not affect the semantic analysis once pictorial representations in the Greenberg/Abusch tradition are replaced with *bona fide* 3D representations.

Be that as it may, it is clear that the reading obtained can be analyzed with the Logical Form in (113)a (with narrow scope existential quantification over viewpoints), and not with the functional analyses in (113)b,c, where Π is the functional viewpoint-denoting term. Importantly, the dialectical situation changes if Π is sensitive to enough arguments. In (114)b and c alike, the viewpoint is right behind the plane.[62] This spatial location changes as the plane moves, so in order for this position to be picked out by $\Pi(x)$, we must make its value sensitive not just to the (fixed) value of the variable *x*, but also to the time parameter. This can be achieved in a richer temporal framework that countenances a future operator, as in (116), with an appropriate lexical stipulation for Π: in (117), we just state that $\Pi(x)$ denotes a position right behind the position occupied by the denotation of *x at the time of evaluation*.

(116) $[[\text{FUT } P]]^{c, s, t, w} =$ 1 if for some $t' > t$, $[[\text{FUT } P]]^{c, s\, t', w} = 1$
0, otherwise

(117) $[[\Pi(x)]]^{c, s, t, w} = \mathbf{\Pi}_{t,w}(s(x)) =$ the position right behind s(x) at t in w

With these assumptions in hand, with can compare an existential and a functional analysis of viewpoints, as in (118)a,b.

(118) a. not FUT $\exists\pi$ [your airplane] plane-fly-cl$^{\pi}$]
b. not FUT [your airplane] plane-fly-cl$^{\Pi(x)}$

As before, the analysis with an existential quantifier in the scope of negation yields appropriate truth conditions, as seen in (119). But with the additional time dependency encoded in (117), we do obtain plausible truth conditions, as seen in (120).

(119) On the assumption that
(i) the domain restriction of $\exists\pi$ is universal (i.e. $\mathbf{D}_{c, t, w}$ is true of all viewpoints for any c, t, w), and
(ii) $s(x) = [[\text{your airplane}]]^{c, s, t, w} =$ your-airplane'

[62] For our second consultant, the viewpoint is right in front of the plane.

$[[(118)a]]^{c, s, t, w} = 1$ iff for no time t' > t, for no viewpoint v, $[[[\text{your airplane}]\ \text{plane-fly-cl}^{\pi}]]^{c, s[\pi\to v], t', w} = 1$,
iff for no time t' > t and no viewpoint v, $[[\text{plane-fly-cl}^{\pi}]]^{c, s[\pi\to v], t', w}$(your-airplane') = 1,
iff for no time t' > t and no viewpoint v, plane-fly$_{t', w}$(your-airplane') and proj(your-airplane', v, t', w) = *plane-fly-cl*

(120) With the assumption in (117),
$[[(118)b]]^{c, s, t, w} = 1$ iff for no time t' > t, $[[[\text{your airplane}]\ \text{plane-fly-cl}^{\Pi(x)}]]^{c, s, t', w} = 1$,
iff for no time t' > t, $[[\text{plane-fly-cl}^{\Pi(x)}]]^{c, s, t', w}$(your-airplane') = 1,
iff for no t' > t, plane-fly$_{t', w}$(your-airplane') and proj(your-airplane', $\prod_{t',w}$(s(x)), t, w) = *plane-fly-cl*
iff for no t' > t, plane-fly$_{t', w}$(your-airplane') and proj(your-airplane', the-position-right-behind-s(x)-at-t'-in-w, t', w)

In sum, by introducing viewpoint functional terms that are dependent on time arguments, we can in this case derive plausible truth conditions without existential quantification over viewpoints. Whether this holds in general remains to be seen. But in any event, be it by existential quantification or by functional dependency, it is clear that viewpoint choice can give rise to non-trivial quantificational dependencies.

D. Free viewpoints variables: alternative analyses [supplement to Section 6]

In this section, we ask whether the analysis with free viewpoint variables developed in Section 6 could be replaced with existential quantification over viewpoints, a mechanism justified in Section 5.

On this alternative analysis, the examples in (62) should not include two (distinct) free viewpoint variables, as in (121)a, but rather two occurrences of a single viewpoint variable bound by different existential quantifiers, as illustrated in (121)b or (121)c, depending on the desired scope of the existential quantifier relative to *always*.

(121) a. there-is student λi [always$_T$ t_i person-walk-cl$^{\boldsymbol{\pi}}$ and always$_{T'}$ t_i person-walk'-cl$^{\boldsymbol{\pi}'}$]

b. there-is student λi [always$_T$ **∃π** [t_i person-walk-cl$^{\boldsymbol{\pi}}$] and always$_{T'}$ **∃π** [t_i person-walk'-cl$^{\boldsymbol{\pi}}$]

c. there-is student λi **∃π** [always$_T$ t_i person-walk-cl$^{\boldsymbol{\pi}}$ and **∃π** [always$_{T'}$ t_i person-walk'-cl$^{\boldsymbol{\pi}}$]

These existential analyses won't work unless the domain restriction D mentioned in our interpretive rule for ∃π, copied in (122), is very small indeed and is essentially reduced to a singleton.

(122) $[[\exists\pi\ F]]^{c, s, t, w} = 1$ iff for some viewpoint v such that $D_{c, t, w}v = 1$, $[[F]]^{c, s[\pi\to v], t, w} = 1$

Without this 'singleton' assumption, the existential analysis might be too liberal. Consider the Logical Form in (121)c (the problem of excessive liberality will be worse when the existential quantifier over viewpoints has narrow scope, as in (121)b). If, relative to a given time t', $D_{c, t', w}$ is true of a large set of viewpoints, we will fail to obtain the inference that the student's movement is represented from the teacher's perspective. World knowledge is unlikely to help, as one could in principle adopt another viewpoint, for instance one corresponding to the back of the class.

The problem disappears if the domain restriction D in (122) is true of a single viewpoint, or if existential quantifiers over viewpoints come equipped with additional restrictions, realized as *V* and *V'* in (123). If their values are contextually determined, they could be reduced to singleton sets, which would give rise once again to a referential-like reading.

(123) λi [**∃π: Vπ**] [always$_T$ t_i walk-cl$^{\boldsymbol{\pi}}$ and [**∃π: V'π**] always$_{T'}$ t_i walk'-cl$^{\boldsymbol{\pi}}$]

This solution can be adapted to the analysis in (122), with implicit rather than explicit restrictions. But it comes with additional constraints. The Logical Form in (121)c would impose the same implicit restriction on the two occurrences of ∃π, which would have to be evaluated with respect to the same context, time and world. Here (121)b would fare better provided it is accompanied with the pragmatic assumption in (124), which posits that the times of the break in the philosophy class are all associated with the teacher's position in that class, and similarly for the times of the break in the linguistics class.

(124) **Pragmatic assumption for (121)b**

Let $v_{philosophy}$ and $v_{linguistics}$ be the viewpoints corresponding to the teacher's position in the philosophy and in the linguistics class respectively. Assume that s(T) and s(T') do not overlap. We posit:
For every time t such that s(T)(t) = 1, $D_{c, t, w} = \{v_{philosophy}\}$, and for every time t such that s(T')(t) = 1 $D_{c, t, w} = \{v_{linguistics}\}$.

The complexity of the choice between a referential analysis (with free variables) and an existential analysis (with singleton restrictions) is nothing new: in logic, the formulas in (125) are equivalent, despite the fact that one is apparently existential and the other isn't. The heart of the matter is that singleton restrictions make it possible to 'imitate' referential readings, and this is as true for viewpoint variables as for anything else; it is correspondingly difficult to fully adjudicate among two alternatives that are nearly equivalent.

(125) a. Px
b. [∃y: y = x] Px

E. Projective Conditions for Iconic Loci [supplement to Section 7.2]

In this section, we sketch an application of the Puppet Model of Section 7 to iconic loci. Two adjustments are needed. First, we must specify that for iconic person-denoting loci, the center of the locus projects, roughly, to the head of the denoted person—as illustrated in (26)b in the main text.[63] Second, unlike person classifiers, which have a clear 3D orientation (with a front and a back, for instance), the 3D orientation of iconic loci might be ambiguous: when one points toward a person-denoting iconic locus, one might not specify how this person's head is oriented; if so, part of the conditions given in (85) will have no 'bite'.

The definition of the semantics of iconic loci from the main text (in (37)a) is repeated in (126). The projective conditions of the main text (in (85)) are copied in (127). Projection rules for iconic loci require (a version of) the special assumptions in (128), including the stipulation that orientation conditions are trivialized (= (128)(ii)b).

(126) **Iconic loci**

$[[pro_{A\_i}{}^{\pi}]]^{c, s, t, w} =$ # if proj(s(A), s(π), t, w) ≠ i
s(A), otherwise

(127) **Projection**

Let ρ* be the Cartesian frame of reference associated with the signer, and let *WORD* be a word (e.g. a predicate classifier). For any viewpoint π, time t, world w, and object d,
proj(d, π, t, w) = WORD iff
(i) $word'_{t,w}(d) = 1$, and
(ii) modulo the scaling factor σ(π) (> 0), the position and orientation of d relative to ρ(π) correspond (in terms of coordinates) to the position and orientation of *CL* relative to ρ*, or in formal terms:
a. center(d, ρ(π)) = σ(π) • center(WORD, ρ*)
b. orientation(d, ρ(π)) = orientation(WORD, ρ*)

(128) **Assumptions for iconic loci**

Our assumptions apply to (127) in case *WORD* is a locus *i*.
(i) Lexical content: instead of $word'_{t,w}(d) = 1$, we use: $entity'_{t,w}(d) = 1$: d has to be an entity.
(ii) Centers and orientations
a. Center: We take the center of an iconic locus to be whatever part of it the signer points to. If an iconic locus denotes a person, we take the latter's center to be their head.
b. Orientation: We take the orientation of a locus to be ambiguous, in the sense that one can pick any orientation one wishes. For present purposes, we can take this condition to be trivialized, i.e. always satisfied.

[63] As mentioned, this example involves iconic loci introduced by classifier predicates, but similar facts hold of iconic loci that are not introduced in this way, as discussed in Schlenker 2015.

To illustrate, the crucial part of (129)a (copied from (30)a in the main text) is given the Logical Form in (129)a'. Locus *a* is positioned high and is iconic, and we can apply (127)-(128).

(129) a. YESTERDAY IX-1 SEE R [= body-anchored proper name]. IX-1 NOT UNDERSTAND IX-a^{high}.
'Yesterday I saw R [= body-anchored proper name]. I didn't understand him.'
(ASL, 11, 24; Schlenker et al. 2013))
=> R is tall, or powerful/important
a'. not I understand $pro_{A\_a}^{\pi}$
b. By (126), $[[pro_{A\_a}^{\pi}]]^{c, s, t, w}$ = # if proj(s(A), s(π), t, w) ≠ a
s(A), otherwise

Writing ρ* for the signer's referential, by (128)(i), $[[pro_{A\_a}^{\pi}]]^{c, s, t, w}$ = # unless
1. $entity'_{t, w}(s(A)) = a$;
2. center(s(A), ρ(s(π))) = σ(s(π)) • center(a, ρ*)
3. orientation(s(A), ρ(s(π))) = orientation(a, ρ*)
By (128)b(i), condition 2 requires that the coordinates of the head of s(A) in the referential corresponding to s(π) (i.e. ρ(s(π)))) should be those of (the center of) locus *a* in the signer's referential ρ*, modulo the scaling factor σ(s(π)).
By (128)(ii)b, condition 3 is trivialized (disregarded).

To be concrete, suppose that locus *a* has a z-coordinate of 30cm above the center signing space (of coordinates (0, 0, 0) in the signer's referential ρ*), and suppose that the center of the referential corresponding to s(π) (the point with coordinates (0, 0, 0) in the referential ρ(s(π)))) is at the level of a person's chest (say at 1m20). If the scaling factor σ(s(π) is 3, the head of s(A) should have a z-coordinate of 3 * 30cm = 90cm, and thus it should be 90cm above the level of a person's chest, so at 2m10—which indeed qualifies as very tall.

Importantly, this framework only derives the 'tall' part of the inference in (129)a, not the 'powerful or important' part. Additional measures are needed for the latter, possibly based on more metaphorical notions of projection.[64]

[64] Independently of the issue of loci, the present geometric framework could definitely be improved in the future. First, as noted in fn. 45, our definition of orientation could be replaced with a more standard and elegant one. Second, as noted by E. Chemla (p.c.), we have in effect encoded real-world objects and sign language classifiers alike by way of points and frames. We could then take a classifier representation to be true of an object just in case there is a certain geometric transformation (based for instance on translations, rotations and scaling) from one to the other. We leave these refinements for future research.

Appendix II. Viewpoint Variables, Viewpoint Parameters, and Perspective Parameters

In this appendix, we briefly introduce Sudo's analysis of the perspectival dependency of *come*, based on parameters rather than variables. We then restate iconological semantics within a parameter-based framework, and we ask how Sudo's perspectival parameter relates to our viewpoint parameter.[65]

❑ *The parameter-based analysis of* come *(Sudo 2015)*

Building on Percus 2011, Sudo 2015 proposes an analysis in which the perspectival component of *come* is evaluated relative to a new parameter, which is identical in kind but distinct from the context parameter.[66] Adapted to the framework of the present piece, a Sudoian entry is given in (130), and illustrated with one of Sudo's examples in (131). We will call *c* the context parameter and π the new perspectival parameter. It will be convenient to take both c and π to be quadruples of the form <speaker, addressee, time, world>, with $c = <c_s, c_a, c_t, c_w>$ and $\pi = <\pi_s, \pi_a, \pi_t, \pi_w>$.

(130) $[[\text{come}]]^{c, \pi, s, t, w}(y)(x) =$ # if at the time and in the world of π, y isn't a home base of the agent or addressee of π
1 iff x goes to y at t in w, otherwise

(131) a. Sergei came to London.
b. $[[(a)]]^{c, \pi, s, t, w} = [[\text{come}]]^{c, \pi, s, t, w}(\text{London'})(\text{Sergei'})$
= # if at the time and in the world of π, London' isn't a home base of the agent or addressee of π
1 iff Sergei' goes to London' at t in w, otherwise

It is crucial that the perspectival parameter π should be distinct from the context parameter c because one can shift π without shifting c, as is illustrated in (132)a: *comes* is evaluated relative to Gianni's perspective, but the context isn't changed, or else one would expect (132)b to be acceptable with a meaning on which *my department* refers to Gianni's department.

(132) a. Gianni meets with any linguist who comes to Milan.
b. Gianni meets with any linguist who comes to my department in Milan.
cannot mean: Gianni meets with any linguist who comes to Gianni's department in Milan.

Building on Bylinina et al. 2015, Sudo posits a perspective-shifting operator with the lexical entry in (133); as desired, it shifts the perspectival parameter without shifting the context parameter (Sudo posits further perspective-shifting operators for other cases).

(133) For any expression F of type <e, t> (relative to the parameters of the interpretation function),
$[[\text{M F}]]^{c, \pi, s, t, w} = \lambda x_e \,.\, [[F]]^{c, \pi[x], s, t, w}(x)$, where $\pi[x] = <x, \pi_a, \pi_t, \pi_w>$.

One can thus analyze (132)a, which has the form in (134)a, as in (134)b. The result is that the predicative expression *F* is evaluated relative to a perspective π[Gianni'] which has as its speaker coordinate Gianni, as is desired. Shifting the value of this perspectival parameter has no effect on the context parameter c, which remains the same outside and inside the scope of M.

(134) a. Giannni M F.
b. $[[(a)]]^{c, \pi, s, t, w}$ $= [[\text{M F}]]^{c, \pi, s, t, w}(\text{Gianni'})$
$= [\lambda x_e \,.\, [[F]]^{c, \pi[x], s, t, w}(x)](\text{Gianni'})$
$= [[F]]^{c, \pi[\text{Gianni'}], s, t, w}(\text{Gianni'})$

[65] We leave for future research a comparison with Schein's use of quantification over perspectives in the analysis of plurals (Schein 1993).
[66] Sudo's analysis of *go* is more complex because it involves competition with *come*: *come to x* triggers a presupposition about the value of x, *go to x* involves the corresponding anti-presupposition, namely that *come to x* would yield a presupposition failure.

❑ *Replacing viewpoint variables with viewpoint parameters within Iconological Semantics*

The relation between viewpoints for iconic representations and perspectives for *come* is an open question. But we can restate our variable-based analysis of viewpoint-dependency within a parameter-based framework, thus making it closer to Sudo's analysis of the perspective-dependency of *come*.

To offer a restatement, the following steps can be followed. When a Logical Form involves a single viewpoint variable π:
(i) replace π with a diacritic • to indicate that the expression on which π appears is interpreted iconically;
(ii) if there is any quantifier $\exists\pi$, replace it with an operator $\exists$•,[67]
(iii) adjust interpretive rules by adding a viewpoint parameter to the interpretation function, and redefine viewpoint-sensitive rules as follows (where • was introduced by rule (i)):

(135) $[[pro_{A\_i}{}^{\bullet}]]^{c, s, \pi, t, w} =$ # if proj(s(A), π, t, w) $\neq$ i.
s(A), otherwise

(136) If P is a 1-place predicate,
$[[P^{\bullet}]]^{c, s, \pi, t, w} = \lambda x_e$. # if $[[P]]^{c, s, \pi, t, w}(x) = \#$
1 iff $[[P]]^{c, s, \pi, t, w}(x) = 1$ and **proj(x, π, t, w) = P**, otherwise

(137) If c, t and w are a context, a time and a world, $D_{c, t, w}$ is a set of viewpoints, and we can define:
$[[\exists\bullet\ F]]^{c, s, \pi, t, w} =$ 1 iff for some viewpoint π such that $D_{c, t, w}\pi = 1$, $[[F]]^{c, s, \pi, t, w} = 1$;
0, otherwise

To see an application, consider (64) from the main text, copied below as (138)a. The syntactic transformations in (i)-(ii) above yield the Logical Form in (138)b. It is interpreted as in (139), which yields appropriate results.

(138) a. $always_T$ $\exists\pi$ there-is student leave-cl^{π}
b. $always_T$ $\exists$• there-is student leave-$cl^{\bullet}$

(139) $[[(138)b]]^{c, s, t, \pi, w} \neq \#$. Furthermore,
$[[(138)b]]^{c, s, \pi, t, w} = 1$ iff for all t' such that s(D)(t') =1, $[[\exists\bullet$ there-is student leave-$cl^{\bullet}]]^{c, s, \pi, t', w} = 1$,
iff for all t' such that s(D)(t') =1, for some viewpoint π' such that $D_{c, t', w}\pi' = 1$, $[[F]]^{c, s, \pi', t', w} = 1$,
iff for all t' such that s(D)(t') =1, for some viewpoint π' such that $D_{c, t', w}\pi' = 1$, for some object x, $student'_{t', w}(x) = 1$ and $person\text{-}walk'_{t', w}(x) = 1$ and **proj(x,π', t', w) = *person-walk-cl***

Things are trickier when a Logical Form contains two viewpoint variables, as is the case in (77), copied as (140)a. In this special case, we can posit the Logical Form in (140)b, in which the value of the variable π is given by the initial viewpoint parameter, while the variable π' is translated with an existential operator *∃•*, which would have to range over a singleton domain to yield the target reading. The necessary assumption is stated in (140)c, and applied in (140)d to the boxed part of (140)b.

(140) a. there-is student λi [$always_T$ t_i person-walk-cl^{π} and $always_{T'}$ t_i person-walk'-$cl^{\pi'}$]

b. there-is student λi [$always_T$ t_i person-walk-$cl^{\bullet}$ and $\boxed{\exists\bullet\ always_{T'}\ t_i\ person\text{-}walk'\text{-}cl^{\bullet}}$]

c. Assumption if (b) is evaluate relative to c, s, π, t, w:

$D_{c, t, w} = \{\pi^*\}$, where π^* is the value of the viewpoint variable π' in (140)a.

d. With the assumption in (c), with F = $always_T$ t_i person-walk'-$cl^{\bullet}$,

$[[\exists\bullet\ F]]^{c, s, \pi, t, w}$ = 1 if for some viewpoint π' such that $D_{c, t, w}\pi' = 1$, $[[F]]^{c, s, \pi', t, w} = 1$
0, otherwise
= 1 iff $[[F]]^{c, s, \pi^*, t, w} = 1$
0, otherwise

[67] We could call this operator anything we want, e.g. *Op*. We use the symbol *∃•* because it keeps the translation transparent.

In the general case, it is clear that the variable-full approach is more liberal than the parameter-based approach. This should give rise to the same kind of debate that took place in the literature about theories of context shifting (see Schlenker 2011a for a survey, and also Schlenker 2003, Anand and Nevins 2004, Anand 2006, and especially Deal 2020).

❑ *Relation between viewpoints and perspectives*

Is perspective-dependency for *come* the same mechanism as viewpoint dependency for classifier predicates and especially pro-speech gestures? A natural way to investigate this question in the case of pro-speech gestures is to explore paradigms such as (141). In line with the characterizations of *come* in terms of movement toward the 'home base' of the perspectival center (Sudo 2015), (141)a is acceptable because the speaker is associated with his office. Thus the perspectival center is in this case the speaker and/or his office. The question is what happens when a co-speech gesture representing a person walking is added: in (141)b, the depiction is represented from the speaker's viewpoint—finger movement is toward the speaker; in (141)c, it is represented from another viewpoint—finger movement is away from the speaker. We provide the average judgments of the 7 consultants we surveyed (see the Supplementary Materials B).

Notation: *a-WALK-1* glosses a finger movement from the speaker's right to his chest. *1-WALK-a* glosses a finger movement from the speaker's chest to his right. A co-speech gesture precedes the expression it modifiers, which is boldfaced.

(141) *Context:* The speaker is at home and is talking to his 14-year-old daughter about his colleague Ann. Ann usually works remotely but lives close to the office.
Ann normally works remote, but she'll
a. $^{6.7}$ come to the office tomorrow.
https://youtu.be/mTCMHHFtsZY
b. ^{5} a-WALK-1 **[come to the office]** tomorrow.
https://youtu.be/0fcL9VAwwjo
c. $^{3.7}$ 1-WALK-a **[come to the office]** tomorrow.
https://youtu.be/n8MZV3bpZFM
a'. $^{6.9}$ go to the office tomorrow.
https://youtu.be/3nwSQlPdjvA
b'. $^{3.4}$ a-WALK-1 **[go to the office]** tomorrow.
https://youtu.be/E4_tytkA2VE
c'. $^{5.6}$ 1-WALK-a **[go to the office]** tomorrow.
https://youtu.be/JkPdW8HvfBQ

Overall, our consultants prefer a gesture toward the speaker than away from the speaker with *come*: (141)b is preferred to (141)c. This can be explained if there is a preference to have the depictive viewpoint match the value of the perspectival parameter for *come*. One possibility is that the preference for matching stems from the fact that there is a single parameter for both constructions, but of course one cannot conclude this on this very limited empirical basis.

Since in (141), the speaker is not at the office, *go* can also be used, as in (141)a'. This makes it possible to test co-speech gestures co-occurring with *go*. Here the pattern is reversed: (141)c' is preferred to (141)b', and again this suggests a preference for matching the depictive viewpoint with the value of the perspectival parameter for *go*. More work will be needed to assess the generality of this phenomenon.

Supplementary Materials A. Additional Examples: Classifier-Loci Generalization and Viewpoint Dependency

We provide below additional data about (1) the Classifier-Loci Generalization, and (2) the behavior of viewpoint variables (with cases of existential quantification, and examples in which two non-coreferring viewpoint variables co-occur in the same sentence).

1. The Classifier-Loci Generalization

The Classifier-Loci Generalization predicts that if a classifier predicate 'touches' arbitrary loci, these become iconic. As a result, in the presence of three arbitrary loci, a classifier that connects just two of them iconically situates them with respect to each other, but not necessarily with respect to the third arbitrary locus. On the other hand, a classifier predicate that connects three loci that are initially introduced as arbitrary turns them all into iconic loci evaluated with respect to a single viewpoint. It is this dual prediction that we test in the paradigms below.

❑ *Partial iconicity: two loci out of three*

In (142), we consider sentences with three loci, one to represent an American student, and two to represent foreign students. There is no classifier in (142)a: the three loci are arbitrary, and the sentence just implies that the two foreign students paint together. In (142)b, the three loci are initially arbitrary, but two classifiers are then added to represent the two foreign students coming together and rotating toward the front. Information about the position of the American student appears to be optional, giving rise only to a weak inference that the American student is behind the foreign students. This sharply contrasts with (142)c,d, where the three loci are introduced by classifier predicates and are thus interpreted iconically: they provide strong iconic information to the effect that "the foreign students always sit apart toward the front while the American is between and behind them".[68]

(142) *Context:* In the presence of American students, foreign students tend to bond together—even with hesitations due to Covid.

POSS-1 CLASS ALWAYS HAVE
'In my class, there are always
a. ^{7}FOREIGN PERSON$_a$ PERSON$_c$ AMERICAN PERSON$_b$. ALWAYS THE-TWO-a,c PAINT TOGETHER.
two foreigners and an American. The two foreigners always paint together.'
b. ^{7}FOREIGN PERSON$_a$ PERSON$_c$ AMERICAN PERSON$_b$. ALWAYS THE-TWO-a,c PERSON-move-cl-c--->|<--- PERSON-move-cl-a+rotation.
two foreigners and an American. The two foreigners always come together in the middle of the room and turn to face the front of the class.'
c. ^{7}FOREIGN PERSON-cl$_a$ PERSON-cl$_c$ AMERICAN PERSON-cl$_b$, ALWAYS THE-TWO-a,c PAINT TOGETHER.
two foreigners apart and an American behind them. The two foreigners always paint together.'
d. ^{7}FOREIGN PERSON-cl$_a$ PERSON-cl$_c$ AMERICAN PERSON-cl$_b$, ALWAYS THE-TWO-a,c PERSON-move-cl-c--->|<--- PERSON-move-cl-a+rotation.
two foreigners apart and an American behind them. The two foreigners always come together in the

[68] One might object that the framework developed here leads us to expect that in (142)b there should be precise iconic information about the initial position of the two foreigners relative to each other, since classifier predicates start from their initial positions and meet halfway between them. But without the specification of a viewpoint and a scale, this is very unspecific, in the sense that when two individuals move towards each other, it's always possible to find a viewpoint from which the scaled movement would develop as displayed by the classifiers. By contrast, when three individuals are situated relative to each other, the information is much richer. To make an analogy: without a specification of the scale, a map that just situates two points relative to each other is rather uninformative; by contrast, a map that situates three points relative to each other is far more informative.

middle of the room and turn to face the front of the class.'
(ASL, 35, 2092; 3 judgments; https://youtu.be/nJatQIuM8Cs)

(143) **Inferential judgments for (142)** (3 judgments)

It can be inferred that…	(i) the class always has two foreign students and an American student, **without further information about their initial position.**	(ii) the class always has two foreign students and an American student, **and the three are initially in a certain configuration**
a. 3 normal nouns - 1 normal verb	6.7	1.7
b. 3 normal nouns - 2 classifiers	5.7	3
c. 3 classifiers - 1 normal verb	2	6
d. 3 normal nouns - 2 classifiers	1	7

Here it is worth noting that our second consultant partly disagreed with these judgments: he found that (143)b was iconically stronger than indicated here, and (143)c weaker. Importantly, he still agreed that (143)c was iconically stronger than (143)b. From the perspective of the present theory, such disagreements need not be surprising because (some) normal predicates have the *option* of being interpreted iconically, and so it could be that our second consultant treats the normal nouns *PERSON* in (142)c as carrying viewpoint variables in this case.[69]

The Logical Forms in (144) can account for our main consultant's judgments; for simplicity, we assume that some repeated elements are elided (the assumption that conjunction can be unexpressed is standard in ASL).

(144) a. $\exists_{a,b,c}$ [[foreign person]$_a$ ~~and [foreign~~ person]$_c$ ~~and~~ [American person]$_b$ ~~and~~ pro$_{a,c}$ paint-together]
b. $\exists_{a,b,c}$ [[foreign person]$_a$ ~~and [foreign~~ person]$_c$ ~~and~~ [American person]$_b$ ~~and~~ **[person-move-cl$_\tau^\pi$]$_a$** ~~and~~ **[person-move-cl'$_\tau^\pi$]$_c$]**
c. $\exists_{a,b,c}$ [[foreign person-cl$^\pi$]$_a$ ~~and [foreign~~ person-cl"$^\pi$]$_c$ ~~and~~ [American person-cl'$^\pi$]$_b$ ~~and~~ pro$_{a,c}$ paint-together]
d. $\exists_{a,b,c}$ [[foreign person-cl$^\pi$]$_a$ ~~and [foreign~~ person-cl"$^\pi$]$_c$ ~~and~~ [American person-cl'$^\pi$]$_b$ ~~and~~ [person-move-cl$_\tau^\pi$]$_a$ ~~and~~ [person-move-cl'$_\tau^\pi$]$_c$]

They crucial example is in (144)b: it involves a mixed representation in which loci *a* and *c* are initially introduced by normal nouns, but then receive an iconic interpretation because they are arguments of the classifier predicates. By contrast, the third locus *b* (corresponding to the American person) can remain arbitrary. Without going into full details, the key is that relative to an assignment function s', whose details are determined by the unrestricted quantifier $\exists_{a,b,c}$, the conjunct *[person-move-cl$_\tau^\pi$]$_a$* will impose the following condition:

(145) Assuming that s'(a) ≠ #:
[[[person-move-cl$_\tau^\pi$]$_a$]]$^{c, s' t, w}$ = [[person-move-cl$_\tau^\pi$]]$^{c, s' t, w}$(s'(a))
= [λx$_e$. # if x = #
1 iff person'$_{t,w}$(x) = 1 **and proj(x, s(π), t, w) =** walk-cl$_\tau$](s'(a)), otherwise](s'(a))
= 1 iff person'$_{t,w}$(s'(a)) = 1 **and proj(s'(a), s(π), t, w) =** walk-cl$_\tau$
0, otherwise

Now the crucial formal observation is that the beginning of the movement of the classifier *person-walk-cl$_\tau$* originates in locus *a*, and as a result the (dynamic) projection rule will require that at the time of evaluation t and in the world of evaluation w, s'(a) should be in a position that projects to that locus.

By parity of reasoning, an iconic condition will equally be imposed on locus *c*, as it is the origin of the movement of the second classifier. By contrast, no such condition is imposed on the third locus, *b* (associated with the American), as it is not 'touched' by any classifier predicate. If the three loci *a*, *b*,

[69] This might make sense on pragmatic grounds: if the signer goes through the trouble of using classifier predicates to situate the movement of the two foreign students, this might be with the general intention of conveying information about the positions of the individuals involved in the scene. Such pragmatic pressure might be absent in (142)a, which does not involve classifier predicates.

c are initially arbitrary, the classifiers touching *a* and *c* will situate their denotations with respect to each other,[70] but not with respect to the American. (If the normal nouns *PERSON* carry a viewpoint variable, they will in essence behave like classifier predicates and will thus be interpreted iconically. But for them this is just an option, whereas for classifier predicates this is an obligation.) Unsurprisingly, in (144)c,d, the three person classifiers that appear at the beginning of the sentence force all three loci to be interpreted iconically.

A shorter paradigm appears in (146):

(146) *Context:* In the presence of American students, foreign students tend to bond together – even with hesitations due to Covid.

CLASS WHEN HAVE _____ , SOON THE-TWO-a,c PERSON-move-cl-c--l-><-l-- PERSON-move-cl-a[71].
'In class, whenever there are _____, the two foreigners soon come together.'
a. normal nouns
$^{6.7}$ FOREIGN PERSON$_a$ PERSON$_c$ AMERICAN PERSON$_b$
two foreigners and an American
b. finger classifiers facing the addressee (student's perspective on the scene)
^{7} FOREIGN PERSON-cl$_a$ PERSON-cl$_c$ AMERICAN PERSON-cl$_b$
two foreigners behind an American
c. finger classifiers facing the signer (teacher's perspective on the scene)
^{7} FOREIGN CL-person$_a$ CL-person$_c$ AMERICAN CL-person$_b$
two foreigners in front of an American
(ASL, 35, 2078; 3 judgments; https://youtu.be/8M-CtR-xVuA)

(147) Inferential judgments for (146) (3 judgments)

It can be inferred that the two foreign students move closer together whenever…	(i) the foreign students are in the presence of an American student.	(ii) the foreign students are in the presence of an American student **and the three are in a certain configuration.**
a. 3 normal nouns	6	2.7
b. 3 classifiers facing the addressee	1.7	6.3
c. 3 classifiers facing the signer	1.7	6.3

In (146)a, the three loci representing the three students can be construed as arbitrary, and thus the *when*-clause quantifies over all situations in which there are (at least) two foreign students and an American student in class. Nonetheless, due to the presence of a classifier predicate, the main clause provides iconic information about the two foreign students: they move toward each other, hesitate, and then end up next to each other. By contrast, in (146)b,c, the three students are represented by three person classifiers, and correspondingly the *when*-clause quantifies over situations in which two foreign students and an American student stand in a particular spatial relation. In (146)b, the three finger classifiers are seen 'from the back' (as they are facing the addressee), so that the viewpoint is probably that of a student, and the situations quantified over involve two foreigners behind an American. In (146)c, the three person classifiers face the signer and thus the viewpoint is probably that of the teacher, and now the *when*-clause quantifiers over situations in which two foreigners are in front of an American. (Our second consultant agreed with our main consultant on this paradigm.)

Let us add that the mixing of iconic and arbitrary loci can be found in further examples, sometimes with less clear results. Be it across signers or examples, this variation is unsurprising since normal nouns can optionally carry viewpoint variables and thus behave like classifiers with respect to iconic interpretation. In (148), the main clause uniformly involves two classifier predicates originating in loci *a* and *c*. The *when*-clause includes either three normal nouns for an airplane (in (148)a,b), or

[70] As noted in fn. 68, if the viewpoint is underdetermined, this provides little information, e.g. that they should be sufficiently far apart that they could move closer together.
[71] The classifier predicate involves a movement of the two indexes towards each, with a pause before they finally meet.

three minimally different airplane classifier predicates (in (148)c, d). The latter clearly provide positional information about the three airplanes. For (148)a,b, positional information about the American plane, corresponding to locus *b*, is far more open, but there is variation across (148)a and (148)b, and across judgment tasks. Still, a contrast between (148)a,b on the one hand and (148)c,d on the other hand is preserved: the latter are more iconically specific than the former. (Here our second consultant agreed with our main consultant's judgments given below.)

(148) *Context:* Space is severely limited at DC airport.

DC AIRPORT WHEN HAVE
'At the DC airport, when there are
a. ^{7}FOREIGN PLANE$_a$ PLANE$_c$ AMERICAN PLANE$_b$, CONTROLLER ASK THE-TWO-a,c PLANE-cl-c--->| |<--- PLANE-cl-a.
two foreign planes and an American plane [2/3 judgments]
or: two foreign planes behind an American plane [1/3 judgment],
the controller asks the two foreign planes to move closer together.'
b. ^{7}FOREIGN PLANE$_a$ PLANE$_c$ AMERICAN PLANE$_b$, CONTROLLER ASK THE-TWO-a,c PLANE-cl-c---\ \--- PLANE-cl-a.
two foreign planes behind an American plane,[72] the controller asks the two foreign planes to line up behind each other.'
c. ^{7}FOREIGN PLANE-cl$_a$ PLANE-cl$_c$ AMERICAN PLANE-cl$_b$, CONTROLLER ASK THE-TWO-a,c PLANE-cl-c--->| |<--- PLANE-cl-a.
two foreign planes behind an American plane, the controller asks the two foreign planes to move closer together.'
d. ^{7}FOREIGN PLANE-cl$_a$ PLANE-cl$_c$ AMERICAN PLANE-cl$_b$, CONTROLLER ASK THE-TWO-a,c PLANE-cl-c---\ \--- PLANE-cl-a.
two foreign planes behind an American plane, the controller asks the two foreign planes to line up behind each other.'
(ASL, 35, 2084; 3 judgments; https://youtu.be/DGI36jPTx4w)

(149) **Inferential judgments for (148)** (3 judgments)

The two foreign airplanes move closer together in the airport whenever…	(i) the foreign airplanes are in the presence of an American plane.	(ii) the foreign airplanes are in the presence of an American plane **and the three are in a certain configuration.**
a. 3 normal nouns - 2 classifiers moving closer	5 (3, 6, 6)	3 (5, 2, 2)
b. 3 normal nouns - 2 classifiers lining up behind each other	3.7	4.3
c. 3 classifiers - 2 classifiers moving closer	1.7	6.3
d. 3 classifiers - 2 classifiers lining up behind each other	1.3	6.7

❑ *Regaining full iconicity*

Following the same logic, we expect if the three arbitrary loci that are introduced in the first clause are all 'touched' by classifier predicates in a later clause, they will *ipso facto* acquire an iconic interpretation. This is the case in (150)b, which minimally differs from (142)b: in the second clause, classifier predicates start from loci *a* and *c* and move toward the center (as was the case in (88)b), but in addition a classifier predicate moves from the *b* locus toward the center. Since now all three loci have been 'touched' by a classifier predicate, they all acquire an iconic interpretation, and the sentence provides information about the original position of the two foreign students relative to the American student.

[72] The inferential judgment about (148)b should be qualified. As seen in the Supplementary Materials B, in 2 judgment tasks out of 3, the consultant wrote that the foreign planes were probably/likely behind the American plane. And upon proofreading the glossed examples he mentioned that this is what he would 'guess' if asked, a weaker inference than in (148)b,c.

(Here our second consultant finds that (150)b is iconically stronger than (150)c: he would give a higher endorsement of the inference in (151)(ii) for b than for c.[73])

(150) *Context:* Some students tend to bond together – even with hesitations due to Covid.

POSS-1 CLASS ALWAYS HAVE
'In my class, there are always
a. ^{7}FOREIGN PERSON$_a$ PERSON$_c$ AMERICAN PERSON$_b$. ALWAYS END THE-THREE-a,b,c PAINT TOGETHER.
two foreigners and an American. The three of them always end up painting together.'
b. ^{7}FOREIGN PERSON$_a$ PERSON$_c$ AMERICAN PERSON$_b$. ALWAYS END THE-TWO-a,c PERSON-move-cl-a+c->center PERSON-move-cl-b->center.
two foreigners in front of American. In the end, it's always the case that the two foreigners come together in the middle and are then joined by the American. '
c. ^{7}FOREIGN PERSON-cl$_a$ PERSON-cl$_c$ AMERICAN PERSON-cl$_b$, ALWAYS END THE-THREE-a,b,c PAINT TOGETHER.
two foreigners standing in front of American. The three of them always end up painting together.'
d. ^{7}FOREIGN PERSON-cl$_a$ PERSON-cl$_c$ AMERICAN PERSON-cl$_b$, ALWAYS END THE-TWO-a,c PERSON-move-cl-a+c->center PERSON-move-cl-b->center.
two foreigners standing in front of American. In the end, it's always the case that the two foreigners come together in the middle and are then joined by the American. '
(ASL, 35, 2098; 3 judgments; https://youtu.be/4AcnqwYUP9E)

(151) **Inferential judgments for (150)** (3 judgments)

It can be inferred that…	(i) the class always has two foreign students and an American student, **without further information about their initial position.**	(ii) the class always has two foreign students and an American student, **and the three are initially in a certain configuration**
a. 3 normal nouns - 1 normal verb	6	2
b. 3 normal nouns - 3 classifiers	2	6
c. 3 classifiers - 1 normal verb	2	6
d. 3 normal nouns - 3 classifiers	1	7

The contrast between (142)b and (150)b is captured by the Logical Forms in (152); the key point is that only two loci are 'touched' by classifier predicates in the former case, whereas all three loci are in the latter case (owing to the addition of the boldfaced classifier).

(152) a. Logical Form of (142)b
$\exists_{a,b,c}$ [[foreign person]$_a$ ~~and [foreign~~ person]$_c$ ~~and~~ [American person]$_b$ ~~and~~ [person-move-cl$_\tau^\pi$]$_a$ ~~and~~ [person-move-cl'$_\tau^\pi$]$_c$]
b. Logical Form of (150)b
$\exists_{a,b,c}$ [[foreign person]$_a$ ~~and [foreign~~ person]$_c$ ~~and~~ [American person]$_b$ ~~and~~ [person-move-cl$_\tau^\pi$]$_a$ ~~and~~ [person-move-cl'$_\tau^\pi$]$_c$ ~~**and**~~ **[person-move-cl"$_\tau^\pi$]$_b$]**

In sum, in some cases at least, there are mixed cases in which some loci introduced by normal nouns become iconic because they have been 'touched' by classifier predicates, while other loci remain arbitrary. More work would be needed, however: all we have done is provide an argument that *sometimes* one can mix arbitrary and iconic loci. Since any locus can in principle be interpreted as iconic, there should be variation across examples, as we saw in (148).

[73] As seen in the Supplementary Materials B, "he feels b is stronger than c and would perhaps make c 2;5 instead of 2;6".

2. Viewpoint Variables

We provide below additional examples that can be analyzed with Logical Forms involving (i) existential quantification over viewpoints (to complement Section 5), and (ii) distinct viewpoint variables (to complement Section 6). The same alternative theoretical analyses could be considered as in Appendix I-C, D.

- *Existential quantification over viewpoints*

In (153)b,c, a plane classifier veers right or left, giving rise to an iconic meaning in the scope of negation. What is understood isn't that there is a viewpoint relative to which the plane won't have a similar movement (with existential quantification scoping over negation); rather, it is that there is no viewpoint relative to which the plane will have such a movement (with existential quantification scoping under negation). There is a complication relative to (114) in Appendix I-C, however: movement to the right or to the left need not be interpreted literally, but may be taken to encompass diverse ways for the plane to veer away from its course (see the Supplementary Materials B for details). Still, in both cases there is a strong intuition that the movement is viewed from the back of the plane. Our second consultant mostly agreed with these judgments, except that he saw the camera angle as being in front of the plane, not behind the plane. As explained in Section 7 of the main text, within the Puppet Model such disagreements are unsurprising as one may equally view the 3D scene from the signer's or from the addressee's position.

(153) *Context:* The signer is a pilot chatting with another pilot.
TOMORROW WEATHER GOOD. POSS-2 PLANE WON'T
'Tomorrow the weather will be nice. Your plane won't
a. ^{7}FLY ROUGH.
have a rough flight.'
b. ^{7}PLANE-veer_right-cl.
veer off course.'
c. ^{7}PLANE-veer_left-cl.
veer off course.'
(ASL, 37, 2564; 3 judgments; https://youtu.be/La4NWeG7p5c)

(154) Inferential judgments for (153) (3 judgments)

	One infers that tomorrow's the addressee's plane won't…			If some movement is represented iconically, from which perspective is it represented, i.e. where is the 'camera position'?
	(i) fly roughly, without further specification.	(ii) fly as iconically depicted, from a fixed spatial viewpoint.	(iii) fly as iconically depicted, no matter what the viewpoint.	
a. FLY ROUGH	7	1	1	n/a
b. FLY-upwards-bumpy-cl	2	1	7	from behind the plane
c. FLY-downwards-bumpy-cl	2	1	7	from behind the plane

As was the case for (114)b,c and (118)a (in Appendix I-C), (153)b,c can be analyzed by way of the simplified Logical Form in (155).

(155) not FUT $\exists\pi$ [your airplane] plane-move-cl$^{\pi}$

- *Multiple viewpoint variables*

The paradigm in (156) parallels that in (76) in the main text, but with entirely different lexical choices, pertaining to airports and airplanes rather than classrooms and students. Here too, our second consultant sees the camera angle to be from the front of the plane rather than from behind the plane (the same remark applies as in our earlier discussions: within the Puppet Model, the 3D scene may be viewed from the signer or from the addressee position).

(156) *Context*: The signer is a part-time air traffic controller in New York and in Boston. Depending on air traffic, pilots use different runways, i.e. no pilot is assigned to a specific runway.

HAVE ONE PILOT NEW-YORK$_a$ ALWAYS __, BOSTON$_b$ ALWAYS … .
'There is one pilot who in New York always __ and in Boston always … . '

a. ^{7} __ = a-PLANE-roll_right-cl rolls to the right after take-off	… = b-PLANE-climb_steeply-cl climbs steeply after take-off
b. ^{7} __ = a-PLANE-climb_steeply-cl climbs steeply after take-off	… = b-PLANE-roll_right-cl rolls to the right after take-off
c. ^{7} __ = a-PLANE-roll_right-cl rolls to the right after take-off	… = b-PLANE-roll_left-cl rolls to the left after take-off
d. ^{7} __ = a-PLANE-roll_left-cl rolls to the left after take-off	… = b-PLANE-roll_right-cl rolls to the right after take-off

(ASL, 35, 2436; 3 judgments; https://youtu.be/_KofNCdwyNU)

In all examples, the take-offs are represented from the beginning of the runway (or from the front of the plane for our second consultant). But these correspond to different positions, as some are in New York while others are in Boston, and in any event each airport has several runways. Paralleling our analysis of (76)-(77), we can posit the Logical Form in (157), which involves two different viewpoint variables π and π'.

(157) there-is pilot λi [always$_T$ t_i plane-fly-cl$^{\pi}$ and always$_{T'}$ t_i plane-fly'-cl$^{\pi'}$]

There are potential objections, however. First, our analysis is correct only if we take the New York viewpoint to be fixed, and similarly for the Boston viewpoint. This is reasonable if we assume that the signer, who the context says is an air traffic controller, is always in exactly the same position while working in New York, and while working in Boston. But the context of (156) makes clear that pilots change runways, so it is unlikely that the representation of the New York take-off, or of the Boston take-off, could involve a fixed 'camera position'. Thus existential quantification might be needed for one or both viewpoint variables, as is represented in (158)a (where both variables are quantified).

(158) there-is pilot λi [always$_T$ $\exists\boldsymbol{\pi}$ t_i fly-cl$^{\pi}$ and always$_{T'}$ $\exists\boldsymbol{\pi}'$$t_i$ fly'-cl$^{\pi'}$]

A second objection is that the data reported in (156) show that (for our main consultant) the 'camera position' is understood to be at the beginning of the runway, and thus one might think that an existential analysis is insufficiently specific. But given what is known about take-offs, it is clear if *some* viewpoint allows for the representations given, that has to be a viewpoint corresponding roughly to the beginning of the runway.

Supplementary Materials B. Raw Data and Survey Results

A. Raw data

Raw data can be found in the following folder: https://osf.io/3qnem/

They contain:

ASL data: 1. Raw data with the main consultant's judgments. 2. Averages for the main consultant's judgments. 3. Summary of the second consultant's judgments. 4. Anonymized ASL videos.

Gestures and onomatopoieas: 1. Anonymized version of the survey in pdf format. 2. Anonymized version of version of the results, in pdf format. 3. Anonymized version of the results, in xlsx format. 4. Anonymized gesture and onomatopoeia videos.

B. Detailed results of the survey on pro-speech gestures and onomatopoeias

Here we provide further information about the judgments given by our 7 consultants (all of them linguists[74]) in the survey on pro-speech gestures and onomatopoeias mentioned in Section 8, and made available above. (Survey results pertaining to co-speech gestures co-occurring with *come* and *go* were discussed in Appendix II and are not repeated here.)

We describe our paradigms, with (i) average acceptability judgments as superscripts, and (ii) inferential questions and choices made by our consultants, as (rounded) percentages of consultants who selected the various options. We comment below on dominant or intended readings; in the latter case, some examples were unclear to our consultants and were not discussed in Section 8.

Transcriptions:
FLY-╰$_{front}$ transcribes a flat hand moving horizontally toward the speaker, and then abruptly upwards.
FLY-╰$_{right}$ involves the same movement on the speaker's right.
FLY-//// glosses a flat hand, representing an airplane, moving upwards in a bumpy fashion.
WALK-$^{\uparrow\leftarrow}$ transcribes a two-fingered gesture representing a person walking toward the speaker and then away on the speaker's left.
WALK-↓$^{\leftarrow}$ transcribes the same movement going toward the speaker, but then away on the speaker's right.

Pro-speech gestures

(159) **Evaluation relative to a single viewpoint: airplane**

a. *Context:* The speaker has just been hired as an air traffic controller at JFK airport.

$^{5.3}$ Before my time, there was a pilot who upon take-off would always FLY-╰$_{front}$.
https://youtu.be/glZ9PFxOaig

Dominant reading: The pilot always took off toward the air control tower, and vertically.

b. In view of the context, what is the most accurate description of what the relevant pilot always did upon take-off?
0% The pilot always took off away from the air control tower, and vertically.
71% The pilot always took off towards the air control tower, and vertically.
29% The pilot always took off towards some point or other, and vertically.

[74] As noted in Section 8, one colleague who took the survey also gave feedback on an earlier draft of it.

(160) **Evaluation relative to a single viewpoint: student**

a. *Context:* The speaker is the Chair of the linguistics program, and he has just talked to the syntax instructor about a recurring problem in the syntax class. This class always takes place in the same classroom.

[5.1] I have just talked to the syntax instructor. Apparently, there's a student who in the middle of class always WALK-↑←.

https://youtu.be/ScXW4z-rSaA

Intended reading, but unclear from the data: The student always walks toward the relevant teacher and then leaves from the left [from the teacher's perspective].

b. In view of the context, what is the most accurate description of what the relevant student always does in the middle of the syntax class?
29% The student always walks towards the teacher and then leaves from the left [from the teacher's perspective].
29% The student always walks away from the teacher and then leaves from the right [from the teacher's perspective].
43% The student always moves towards some point or other and then takes a right and leaves.

(161) **Viewpoints dependent on universal quantifiers: airplanes**

a. *Context:* The speaker is an airline pilot with prior experience in Crete, which has several airports.

[5.7] In Crete, when the weather is bad, every hour there is a plane that FLY-////.

https://youtu.be/wmrNPppEW04

Dominant reading: Every hour, there is some point or other of space in which there is a plane that has a bumpy ascent.

b. In view of the context, what is the most accurate description of what always happens in Crete when the weather is bad?
0% There is some point of space in which, every hour, there is a plane that has a bumpy ascent.
100% Every hour, there is some point or other of space in which there is a plane that has a bumpy ascent.

(162) **Viewpoints dependent on universal quantifiers: students**

a. *Context:* The speaker is the chair of the linguistics program. Each 1st year student takes phonology, syntax and semantics, taught by three separate instructors in three separate classrooms.

[5.1] I have just talked to our three instructors. There's a student who in the middle of class always WALK-↑←.

https://youtu.be/3S0NqsqqIok

Intended reading, but unclear from the data: The student always walks toward the relevant teacher and then leaves from the left [from the teacher's perspective].

b. In view of the context, what is the most accurate description of what the relevant student always does in the middle of class?
43% The student always walks towards the relevant teacher and then leaves from the left [from the teacher's perspective].
29% The student always walks away from the relevant teacher and then leaves from the right [from the teacher's perspective].
29% For some fixed location at school, the student always moves towards that location and then takes a right.

(163) **Existential quantification over viewpoints in the scope of negation**

a. *Context:* The speaker is a pilot talking to another pilot.

[5.4] Tomorrow the weather will be nice, so your plane won't FLY-////.

https://youtu.be/qHxsyRD74B4

Dominant reading: There is no point of space in which the plane will have a bumpy ascent.

b. In view of the context, what is the most accurate description of what will happen tomorrow?
0% There is some point of space in which the plane won't have a bumpy ascent [leaving open whether or not it will have a bumpy ascent in other parts of space].
100% There is no point of space in which the plane will have a bumpy ascent.

(164) **Evaluation relative to two viewpoint variables**
a. *Context:* The speaker works as a part-time air-traffic controller in New York and in Boston.

[4.7] Before my time, there was a pilot who upon take-off in New York would always FLY-↳$_{\text{front}}$, and in Boston would always FLY-↳$_{\text{right}}$.
https://youtu.be/xbPxD9gfXh8

Dominant reading: 1. In New York, the pilot would take off toward the local air traffic control tower from the front. 2. In Boston, he would take off toward the local air traffic control tower from the right.

b. In view of the context, what is the most accurate description of what the relevant pilot always did upon take-off?
0% 1. In New York, he would take off towards the local air traffic control tower from the right. 2. In Boston, he would take off towards the local air traffic control tower from the front.
86% 1. In New York, he would take off towards the local air traffic control tower from the front. 2. In Boston, he would take off towards the local air traffic control tower from the right.
0% Viewed from a single, fixed point of space, 1. in New York, he would take off towards that point from the right, 2. in Boston, he would take off towards that point from the front.
14% Viewed from a single, fixed point of space, 1. in New York, he would take off towards that point from the front, 2. in Boston, he would take off towards that point from the right.

(165) a. *Context:* The speaker is the chair of the linguistics program. Each 1st year student takes phonology and syntax, taught by two separate instructors in two separate classroms.

[4.9] I have just talked to our instructors. There's a first-year student who in the middle of the phonology class always WALK-$^{\uparrow}$←, and in the middle of the syntax class always WALK-↓$^{\leftarrow}$.
https://youtu.be/fi-q6O9JmR0

Intended reading, but unclear from the data: 1. In the phonology class, the student always walks toward the teacher and then leaves from the left [from the teacher's perspective]. 2. In the syntax class, the student always walks toward the teacher and then leaves from the right [from the teacher's perspective].

b. In view of the context, what is the most accurate description of what the relevant student always does in the middle of class?
43% 1. In the phonology class, the student always walks towards the teacher and then leaves from the left [from the teacher's perspective]. 2. In the syntax class, the student always walks towards the teacher and then leaves from the right [from the teacher's perspective].
29% 1. In the phonology class, the student always walks away from the teacher and then leaves from the right [from the teacher's perspective]. 2. In the syntax class, the student always walks away from the teacher and then leaves from the left [from the teacher's perspective].
29% Viewed from a single, fixed point of space, 1. in the phonology class, the student always walks towards that point and then takes a right, 2. in the syntax class, the student always walks towards that point and then takes a left.

Pro-speech onomatopoeias

(166) **Evaluation relative to a single auditory point**

a. *Context:* The speaker lives in France.

[5] On Bastille Day, I spend my time on my balcony. There are always fighter jets that sh<>sh.
https://youtu.be/AuDkVtAX9zM

Dominant reading: There are always jets that approach the speaker's balcony with increasing noise and move away with decreasing noise.

b. In view of the context, what is the most accurate description of what happens on Bastille Day?
86% There are always jets that approach the speaker's balcony with increasing noise and move away with decreasing noise.
14% There are always jets that approach some point or other with increasing noise and move away with decreasing noise.

(167) **Auditory points dependent on universal quantifiers**
Context: The speaker is a journalist who used to be a correspondent in France.

[5.1] In France, on Bastille Day, there are always fighter jets that sh<>sh.
https://youtu.be/7htSTCEYhas

Dominant reading: There are always jets that approach some point or other with increasing noise and move away with decreasing noise.

b. In view of the context, what is the most accurate description of what happens on Bastille Day?
86% There are always jets that approach the speaker's past French neighborhood with increasing noise and move away with decreasing noise.
14% There are always jets that approach some point or other with increasing noise and move away with decreasing noise.

(168) **Auditory points dependent on two auditory points**

a. *Context:* The speaker lives in France.
[4.7] On Bastille Day, in my neighborhood fighter jets always sh<>sh, whereas over the Champs-Elysées they shshsh.
https://youtu.be/JmoKUc6HX3M

Dominant reading: 1. Fighter jets fly over the speaker's neighborhood with considerable and increasing noise before moving away. 2. Fighter jets fly over the Champs Elysées with less noise.

b. In view of the context, what is the most accurate description of what always happens on Bastille Day?
0% 1. Fighter jets fly over the Champs Elysées with considerable and increasing noise before moving away. 2. Fighter jets fly over the speaker's neighborhood with less noise.
100% 1. Fighter jets fly over the speaker's neighborhood with considerable and increasing noise before moving away. 2. Fighter jets fly over the Champs Elysées with less noise.
0% Heard from single, fixed point, 1. fighter jets that fly over the Champs Elysées make considerable and increasing noise before moving away, and 2. fighter jets that fly over the speaker's neighborhood make less noise.
0% Heard from single, fixed point, 1. fighter jets that fly over the speaker's neighborhood make considerable and increasing noise before moving away, and 2. fighter jets that fly over the Champs Elysées make less noise.

References

Abusch, Dorit: 2020, Possible worlds semantics for pictures. In L. Matthewson, C. Meier, H. Rullmann, and T. E. Zimmermann (eds.), *The Wiley Blackwell Companion to Semantics*, Wiley.

Abusch, Dorit and Rooth, Mats: 2017, The formal semantics of free perception in pictorial narratives. In Alexandre Cremers, Thom van Gessel & Floris Roelofsen (eds), *Proceedings of the 21st Amsterdam Colloquium*.

Anand, Pranav: 2006. *De De Se*. PhD dissertation. Santa Cruz: University of California.

Anand, Pranav and Nevins, Andrew: 2004, Shifty Operators in Changing Contexts. In R. Young (ed.), *Proceedings of SALT XIV*, 20-37, Ithaca, NY: Cornell University.

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

Bylinina, Lisa, Eric McCready & Yasutada Sudo. 2015. Notes on perspective sensitivity. In Peter Arkadiev, Ivan Kapitonov, Yuri Lander, Ekaterina Rakhilina & Sergei Tatevosov (eds.), *Donum Semanticum: Opera Linguistica Et Logica In Honorem Barbarae Partee A Disscipulis Amicisque Rossicis Oblata*, 67–79. Moscow: LSC.

Cooper, Robin: 1983, *Quantification and Syntactic Theory*. Reidel Publishing Company, Dordrecht.

Cresswell, Max J.: 1990, *Entities and Indices*. Kluwer.

Cuxac, Christian. 1999. French sign language: Proposition of a structural explanation by iconicity. In *Gesture-based communication in human–computer interaction*, ed. A. Braort, et al., 165–184. Berlin: Springer.

Cuxac, Christian and Sallandre, Marie-Anne: 2007, Iconicity and arbitrariness in French Sign Language: Highly iconic structures, degenerated iconicity and diagrammatic iconicity. In E. Pizzuto, P. Pietrandrea & R. Simone (eds.), *Verbal and signed languages: Comparing structures, constructs and methodologies*, 13–33. Berlin: Mouton de Gruyter.

Davidson, Kathryn: 2015, Quotation, Demonstration, and Iconicity. *Linguistics & Philosophy* 38(6), 477-520.

Davidson, Kathryn and Gagne, Deanna: 2022, "More is up" for domain restriction in ASL. *Semantics & Pragmatics* Volume 15 https://doi.org/10.3765/sp.15.1

Deal, Amy Rose: 2020, *A Theory of Indexical Shift: Meaning, Grammar, and Crosslinguistic Variation*. MIT Press.

Dudis, Paul G.: 2004, Body partitioning and real-space blends. *Cognitive Linguistics* 15,2:223–238

Emmorey, Karen and Herzig, Melissa: 2003. Categorical versus Gradient Properties of Classifier Constructions in ASL. In K. Emmorey (Ed.), *Perspectives on Classifier Constructions in Signed Languages*, Lawrence Erlbaum Associates, Mahwah NJ, pp. 222–246

Esipova, Maria: 2019, *Composition and Projection in Speech and Gesture*. PhD dissertation, New York University.

Grenoble, Lenore A., Martina Martinovic, and Rebekah Baglini. 2015. Verbal gestures in Wolof. In *44th Annual Conference on African Linguistics (ACAL)*, eds. Ruth Kramer et al., 110–121. Somerville: Cascadilla Proceedings Project.

Goldin-Meadow, Susan, So, Wing Chee, Özyürek, Asli and Mylander, Carolyn: 2008, The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences* 105(27): 9163-9168.

Greenberg, Gabriel: 2013. Beyond Resemblance. *Philosophical Review* 122:2, 2013

Greenberg, Gabriel: 2014, Reference and Predication in Pictorial Representation. Handout of a talk given at the London Aesthetics Forum (February 19, 2014).

Greenberg, Gabriel: 2021, Semantics of Pictorial Space. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-020-00513-6

Greenberg, Gabriel: to appear, The Iconic-Symbolic Spectrum. *Philosophical Review*.

Heim, Irene: 1982, *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts.

Heim, Irene: 2008, Feature on bound pronouns. In D. Harbour, D. Adger and S. Béjar (eds.) *Phi Theory*. Oxford: Oxford University Press. 35–56.

Heim, Irene and Kratzer, Angelika: 1998, *Semantics in Generative Grammar*. Oxford: Blackwell.
Hinterwimmer, Stefan; Patil, Umesh; Ebert, Cornelia: 2021, On the Interaction of Gestural and Linguistic Perspective Taking. *Front. Commun*. 6:625757. doi: 10.3389/fcomm.2021.625757
Huenerfauth, Matt: 2004, Spatial representation of classifier predicates for machine translation into american sign language. In: *Workshop on Representation and Processing of Sign Language, 4th Internationnal Conference on Language Ressources and Evaluation* (LREC 2004), Lisbon, Portugal, pp. 24–31.
Jacobson, Pauline 1999, Towards a variable free semantics. *Linguistics and Philosophy* 22:117–84
Kamp, Hans: 1981, A Theory of Truth and semantic Representation. In J. A. G. Groenendijk, T. M. V. Janssen and M. B. J. Stokhof (eds.): *Formal Methods in the Study of Language*, (pp. 277-322). Amsterdam: Mathematical Centre.
Kuhn, Jeremy. 2016, ASL Loci: Variables or Features? *Journal of Semantics* 33(3). 449–491.
Kuhn, Jeremy: 2021, Disjunctive reference in French Sign Language. In Dreier, Kwon, Darnell, and Starr (eds.), *Proceedings of Semantics and Linguistic Theory 31 (SALT 31)*, 104-121.
Kuhn, Jeremy and Aristodemo, Valentina: 2017, Pluractionality, iconicity, and scope in French Sign Language. *Semantics & Pragmatics* 10(6).
Liddell, Scott K. 2003a, Sources of Meaning in ASL Classifier Predicates. In: Emmorey, Karen (ed.), Perspectives on Classifiers in Sign Languages. Mahwah, NJ: Lawrence Erlbaum, 199220.
Liddell, Scott K.: 2003b, *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge Cambridge University Press.
Lillo-Martin, Diane and Klima, Edward S.: 1990, Pointing out Differences: ASL Pronouns in Syntactic Theory. Fischer, Susan & Siple, Patricia (ed.), *Theoretical Issues in Sign Language Research*, Volume 1: Linguistics, 191–210. Chicago: University of Chicago Press.
Maier, Emar: 2019, Picturing Words: the Semantics of Speech Balloons. In Julian Schlöder, Dean McHugh, and Floris Roelofsen (eds), *Proceedings of the 22nd Amsterdam Colloquium* pp. 584–592. ILLC, University of Amsterdam.
Maier, Emar and Bimpikou, Sofia: 2019, Shifting Perspectives in Pictorial Narratives. In M. Teresa Espinal et al. (eds), *Proceedings of Sinn und Bedeutung 23*, pp. 91–105. Universistat Autonoma de Barcelona, Bellaterra.
McNeill, David: 1992, *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
Migotti, Léo: 2021, From Picture Semantics to Dance Semantics. Slides of a talk given at the 15th International Symposium of Cognition, Logic and Language, November 26, 2021, University of Latvia, Riga.
Migotti, Léo and Guerrini, Janek: to appear, Linguistic inferences from pro-speech music. To appear in *Linguistics & Philosophy*.
Okrent, Arika: 2002. A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (or at least give us some criteria to work with). In R. P. Meier, K. Cormier, & D. Quintos-Pozos (eds.), *Modality and structure in signed and spoken languages*, pp. 175–198. Cambridge, UK: Cambridge University Press.
Parrill, Fey: 2009, Dual viewpoint gestures. *Gesture* 9, 271 –289.
Parrill, Fey: 2010, Viewpoint in speech–gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25,5: 650-668
Pasternak, Robert: 2019, The projection of co-speech sound effects. Manuscript, ZAS Berlin. https://ling.auf.
Patel-Grosz, P., Grosz, P. G., Kelkar, T. & Jensenius, A. R. (2018). Coreference and disjoint reference in the semantics of narrative dance, In Uli Sauerland & Stephanie Solt (ed.), *Proceedings of Sinn und Bedeutung* 22, (vol. 2, pp.199-216). ZASPiL 61. Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS).
Patel-Grosz, P., Grosz, P. G., Kelkar, T. & Jensenius, A. R. (to appear). Steps towards a formal semantics of dance. *Journal of Semantics*. https://doi.org/10.1093/jos/ffac009
Pavlič, Matic: 2016, The word order parameter in Slovenian Sign Language Transitive, ditransitive, classifier and locative constructions. PhD dissertation, University of Venice.

Percus, Orin: 2011, *Indexicality and Compositional Semantics*. Lecture notes, ESSLLI 2011.
Perniss, Pamela M.: 2007, Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective, *Lingua* 117, 7:1315-1338.
Quine William V: 1960, Variables explained away. *Proceedings of the American Philosophical Society* 104: 343–347
Rooth, Mats and Abusch, Dorit: 2019, Indexing Across Media. In Julian J Schlöder, Dean McHugh, and Floris Roelofsen (eds), *Proceedings of the 22nd Amsterdam Colloquium* pp. 612–624. ILLC, University of Amsterdam.
Sandler, Wendy and Lillo-Martin, Diane: 2006, *Sign Language and Linguistic Universals*. Cambridge University Press.
Schein, Barry: 1993, *Plurals and Events*. MIT Press.
Schembri, Adam; Jones, Caroline; Burnham, Denis: 2005, Comparing Action Gestures and Classifier Verbs of Motion: Evidence From Australian Sign Language, Taiwan Sign Language, and Nonsigners' Gestures Without Speech. *Journal of Deaf Studies and Deaf Education*, 10, 3:272–290. https://doi.org/10.1093/deafed/eni029
Schlenker, Philippe: 2003, A Plea for Monsters, *Linguistics & Philosophy* 26: 29-120.
Schlenker, Philippe: 2011a, Indexicality and De Se Reports. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning,* Volume 2, Article 61, pp. 1561-1604, Mouton de Gruyter.
Schlenker, Philippe: 2011b, Donkey Anaphora: the View from Sign Language (ASL and LSF). *Linguistics and Philosophy* 34, 4: 341-395 DOI: 10.1007/s10988-011-9098-1
Schlenker, Philippe: 2014, Iconic Features. *Natural Language Semantics* 22, 4: 299-356.
Schlenker, Philippe: 2015, Gradient *and* Iconic Features in ASL (squib). *Snippets* (Issue 29) doi: 10.7358/snip-2015-029-schl
Schlenker, Philippe: 2016, Featural Variables. *Natural Language and Linguistic Theory*.
Schlenker, Philippe: 2017a, Super Monsters I: Attitude and Action Role Shift in Sign Language. *Semantics & Pragmatics,* Volume 10, 2017.
Schlenker, Philippe: 2017b, Super Monsters II: Role Shift, Iconicity and Quotation in Sign Language. *Semantics & Pragmatics*. Volume 10, 2017.
Schlenker, Philippe: 2017c, Sign Language and the Foundations of Anaphora. *Annual Review of Linguistics*, 3:149–77
Schlenker, Philippe: 2017d, Outline of Music Semantics. *Music Perception: An Interdisciplinary Journal* 35, 1: 3-37 DOI: 10.1525/mp.2017.35.1.
Schlenker, Philippe: 2018a, Locative Shift. *Glossa: A Journal of General Linguistics*, 3(1), 115. DOI: http://doi.org/10.5334/gjgl.561
Schlenker, Philippe: 2018b. Visible Meaning: Sign Language and the Foundations of Semantics. *Theoretical Linguistics* 44(3-4): 123–208
Schlenker, Philippe: 2018c, Sign Language Semantics: Problems and Prospects [replies to peer commentaries]. *Theoretical Linguistics* 44(3-4): 295–353.
Schlenker, Philippe: 2018e, Iconic Pragmatics. *Natural Language & Linguistic Theory* 36, 3:877–936
Schlenker, Philippe: 2019b, What is Super Semantics? *Philosophical Perspectives* 32, 1: 365-453 https://doi.org/10.1111/phpe.12122
Schlenker, Philippe: 2020, Gestural Grammar. *Natural Language & Linguistic Theory* 38:887–936.
Schlenker, Philippe: 2021, Iconic Presuppositions. *Natural Language & Linguistic Theory* 39:215–289.
Schlenker, Philippe: 2022, Musical Meaning within Super Semantics. *Linguistics and Philosophy* 45:795–872.
Schlenker, Philippe; Bonnet, Marion; Lamberton, Jonathan; Lamberton, Jason; Chemla, Emmanuel; Santoro, Mirko; Geraci, Carlo: to appear, Iconic Syntax: Sign Language Classifier Predicates and Gesture Sequences. Accepted with minor revisions in *Linguistics & Philosophy*.
Schlenker, Philippe and Lamberton, Jonathan: 2019, Iconic Plurality. *Linguistics & Philosophy* 42: 45-108
Schlenker, Philippe and Lamberton, Jonathan: 2022, Meaningful Blurs: the Sources of Repetition-based Plurals in ASL. *Linguistics & Philosophy* 45: 201–264.

Schlenker, Philippe, Lamberton, Jonathan and Santoro, Mirko: 2013, Iconic Variables. *Linguistics & Philosophy* 36, 2, 91-149.
Steinbach, Markus and Onea, Edgar: 2016, A DRT Analysis of Discourse Referents and Anaphora Resolution in Sign Language. *Journal of Semantics* 33(3). 409–448. doi:10.1093/jos/ffv002.
Sudo, Yasutada : 2015, Come and Go. Handout of a talk given at the 'Perspectives in Context' workshop, Georg-August Universität Göttingen, July 10, 2015.
Supalla, Ted: 1982, *Structure and Acquisition of Verbs of Motion and Location in American Sign Language*. PhD Dissertation, University of San Diego.
Supalla, Ted: 1986, The Classifier System in American Sign Language. In: Craig, Colette (ed.), *Noun Classes and Categorization*. Amsterdam: Benjamins, 181-214.
Taub, Sarah F. 2001. *Language from the Body*. Cambridge: Cambridge University Press.
Valli, Clayton; Lott, Peggy Swartzel; Renner, Daniel; and Hills, Rob: 2005, *The Gallaudet Dictionary of American Sign Language*, Gallaudet University Press.
Valli, Clayton, and Lucas, Ceil: 2000. *Linguistics of American Sign Language: an Introduction*. 3rd ed. Washington, DC: Gallaudet University Press.
Zucchi, Sandro: 2011, Event descriptions and classifier predicates in sign languages. Presentation at FEAST (Formal and Experimental Advances in Sign language Theory) in Venice, June 21, 2011.
Zucchi, Sandro: 2017, Event Categorization in Sign Languages. In: Cohen H. and Lefebvre C. (eds.) *The Handbook of Categorization in Cognitive Science*, pp. 377-396. Elsevier.
Zucchi, Sandro: 2018, Sign language iconicity and gradient effects. *Theoretical Linguistics* 44(3-4): 283–294.
Zwitserlood, Inge: 2012, Classifiers. In: Roland Pfau, Markus Steinbach and Bencie Woll (eds) *Sign Language: An International Handbook*. De Gruyter Mouton, pp. 158-186. https://doi.org/10.1515/9783110261325.158