Steps towards a semantics of dance

Pritty Patel-Grosz, Patrick G. Grosz, Tejaswinee Kelkar, Alexander Refsum Jensenius
(University of Oslo)

Abstract

As formal theoretical linguistic methodology has matured, recent years have seen the advent of applying it to objects of study that transcend language, e.g., to the syntax and semantics of music (Lerdahl & Jackendoff 1983, Schlenker 2017a; see also Rebuschat et al. 2011). One of the aims of such extensions is to shed new light on how meaning is construed in a range of communicative systems. In this paper, we approach this goal by looking at narrative dance in the form of *Bharatanatyam*. We argue that a semantic approach to dance can be modeled closely after the formal semantics of visual narrative proposed by Abusch (2013, 2014, 2021). A central conclusion is that dance not only shares properties of other fundamentally human means of expression, such as visual narrative and music, but that it also exhibits similarities to sign languages and the gestures of non-signers (see, e.g., Schlenker 2020) in that it uses space to track individuals in a narrative and performatively portray the actions of those individuals. From the perspective of general human cognition, these conclusions corroborate the idea that linguistic investigations beyond language (see Patel-Grosz et al. forthcoming; Schlenker 2019a) can yield insights into the very nature of the human mind and of the communicative devices that it avails.

Keywords: dance, formal semantics, visual narrative, pictorial semantics, situation semantics

1. Introduction

In recent years, the formal theoretical linguistic methodology that has been developed in connection with natural language phenomena has matured to a stage where scholars have raised the question of whether such methodology can also be applied in a fruitful way *beyond language*, constituting a sub-field of *Super Linguistics* (where *super* is used in its original Latinate meaning, 'beyond'; Patel-Grosz et al. forthcoming). Building on the pioneering work of Lerdahl & Jackendoff (1983), recent applications of linguistic methodology to music are instantiated by Katz (2017) and Schlenker (2017a, 2019b), amongst others (see also Rebuschat et al. 2011). The aim of linguistic investigations of non-standard objects is fourfold: first, to understand what unifies natural language with other human competencies (such as music, dance, or visual narrative); second, to clearly delimit what counts as language proper *vs*. what is a language-like system that should not count as a language (e.g., music); third, building on these first two goals, to achieve a better understanding of the unique *vs.* non-unique *features* of human language; and, fourth, to contribute new insights, based on linguistic methodology, to the very study of the human mind and what sets it apart from the minds of other animals.

The present investigation is part of the larger *Super Linguistics* program, in that it involves the application of linguistic methods to non-standard (and non-linguistic) objects, i.e., objects of inquiry outside the realm of natural language. As such, we can outline our methodological assumptions as follows. The analytical framework of proposing a precise and predictive semantic analysis, and, more generally, our focus on the semantics of potentially meaningful objects, qualify as an extension of linguistic methodology. One of the consequences of this endeavour is showing that well-established semantic formalism can be applied to meaning in narrative dance, which is a new object of study from a linguistics perspective. In doing so, we build on Abusch's (2013, 2014, 2021) formal semantics of visual narrative.

Building on such a foundation allows us to use Bharatanatyam dance (our object of study) as a window into the question of how the human body and visual space can be

used to communicate (i) the tracking of individuals in a narrative, and (ii) perspective taking with regards to such an individual – to perform the individual's actions in the narrative. A better understanding of such communicative mechanisms in the non-linguistic medium of narrative dance can inform our theory of semantics, and, more broadly, linguistic theorizing, while also contributing to our understanding of general human cognition.[1]

Empirically, we apply the method of controlled elicitation, which is a well-established methodology in traditional linguistics; the only difference is that we elicit dance sequences from a trained dancer, rather than natural language examples from a native speaker consultant.

In our formal implementation, we draw on commonalities between narrative dance and speech-accompanying gesture (see, e.g., McNeill 1992, 2012, Kendon 2004, Abner et al. 2015, Schlenker 2020),[2] and introduce concepts such as *indexical bases*, which we define as positions in space that are associated with a given referent in the narrative, and *action-performance* (in a technical sense), which we define as the acting out of a referent's actions – from that referent's perspective – by the dancer. Our notions of base and action-performance are inspired by two concepts from sign language research, namely referential *loci* (see, e.g., Lillo-Martin & Klima 1990, Liddell 1990, Schlenker 2017b), and so-called *action role shift* or *action reports* (see, e.g., Padden 1986, Lillo-Martin 1995, Quer 2005, Sandler & Lillo-Martin 2006, Herrmann & Steinbach 2009, 2012, Davidson 2015). However, we do not argue that bases *are* loci, or that action-performance *is* role shift. Loci and role shift are grammatical and linguistic in a narrow sense. By contrast, we propose that the establishment of *bases* for reference tracking, and the subsequent use of *action-performance* for perspective taking, are non-grammatical (and non-linguistic) mechanisms that are plausibly rooted in general cognition, and which underlie phenomena in different domains and modalities, including dance, silent comics, speech-accompanying gestures, and so forth. It is an interesting question for future research whether loci and role shift in sign languages may have their origins in such general cognitive mechanisms, from where they have been recruited as linguistic components of sign language grammars. The study of bases and action-performance in narrative dance may thus open lines of inquiry with regards to cognitively grounded commonalities between dance and sign language.

This paper is roughly divided into two parts: section 2 presents the empirical investigation, and section 3 the semantic analysis. Subsequently, section 4 explores an alternative analysis that relies on visual iconicity alone, and section 5 concludes.

2. A super linguistic approach to narrative dance – methodology

2.1 The object of study: Bharatanatyam

Given the broad range of different musical genres and dance forms, linguistic investigations that venture into music or dance can adopt one of the following

[1] We are grateful to an anonymous reviewer for proposing the framing in this paragraph, and for suggesting some of the wording that we used in this paragraph.

[2] For the purpose of this paper, we define *gesture* as communicative body movements (i.e., meaning-bearing bodily actions), which is the definition commonly used in linguistic research (see, e.g., McNeill 1992, 2012, Kendon 2004, Abner et al. 2015). This contrasts with definitions in other fields, such as human-computer interaction and music/dance research, where *gesture* is frequently used to refer to body motion in general (see Jensenius et al. 2010 and Jensenius 2017 for discussion).

approaches. They can either try to establish generalizations across genres (e.g., Napoli and Kraus 2017) or focus on case studies (see Katz and Pesetsky 2011 and Schlenker 2017a, 2019b, who zoom in on Western art music as instantiated by the works of Bach, Mozart, Saint-Saëns and Strauss; see also Charnavel 2016, 2019, who focuses on ballet and modern dance). In our study, we choose the second route, focusing on *Bharatanatyam*,[3] a classical South Indian dance that originates in Tamil Nadu (see Puri, 1986, 2004; Williams, 2003; Ramesh, 2013, 2014); Bharatanatyam is a type of figurative (narrative) dance that typically serves to tell a story. As a figurative dance, it is more similar to language (and silent visual narrative) than other dance forms (such as ballet, contemporary or street dance), yet more conventionalized than pantomime (which can be viewed as an extreme form of figurative dance; see Charnavel 2016). We thus expect it to share properties of silent visual narratives. Note that, while Bharatanatyam is typically accompanied by music and or spoken word (e.g., singing of the narrative), it is not necessarily accompanied by music, and we recorded our stimuli without music.

Traditionally, Bharatanatyam is used to articulate religious narratives, but it also allows for secular and modern stories in contemporary dance productions. As outlined by Puri (1986), the dance has a rich inventory of conventionalized gestures, including approximately 31 types of single hand gestures (*hasta mudras*) and 27 types of double hand gestures, which have received some attention in the semiotic literature (see Puri, 1986:271-276; see also Ikegami, 1971). The double hand gestures are combinations of two single hand gestures. Gesture inventories and their sizes vary, depending on the source material, since this is a 2000-year-old dance form. Hand gestures are semantically underspecified; for instance, the *patāka* ('flag') gesture, which involves a flat hand with fingers touching (similar to the hand position when 'high-fiving') can be interpreted as one of the entities from the set in (1) (from Ikegami, 1971:373).

(1) *possible meanings associated with the patāka ('flag') mudra*
'clouds, a forest, things, bosom, might, peace, a river, heaven, prowess, moonlight, strong, sunlight, wave, entering, silence, an oath, the sea, sword, a palmyra leaf'

This underspecification is resolved by the context, i.e., the eventual meaning of a *patāka* mudra depends on factors such as the position of the arm, the accompanying movement, and so forth.

In addition to hand gestures, Bharatanatyam makes gestural use of the entire body; Puri (1986:251) identifies whole body gestures as "larger action sign units", which subsume a dancer's eyes, face, neck, torso, limbs and feet. We can thus differentiate between "local" gestures such as hand-and-arm combinations, and "global" full-body gestures. In our study, we focus on such "global gestures", since we take hand gestures to have symbolic meanings, which are conventional in the sense that they may be rote learned (requiring a trained audience to correctly interpret them). Global gestures are a phenomenon that we may also expect to find in non-conventionalized dance by untrained participants, which is relevant for future studies that build on our findings.[4]

[3] We follow the convention in the literature and capitalize the first letter of *Bharatanatyam*.

[4] Note that facial expressions are also used as part of the Bharatanatyam sign system; given the nature of our study, our dancer aimed to minimize the use of facial expressions and compensate for it with other gestures.

From a big-picture perspective, cognitively interesting findings would include the existence of meanings that can be inferred without explicit teaching, and possibly by non-specialists. Such findings would clearly further our understanding of human cognition. By contrast, the existence of conventional meanings that are inaccessible to audience members who have not been instructed in a dance form would not be enlightening.[5] To move away from low-level symbols such as hand gestures (which may simply have a sign-based semantics that is rote-learned by trained dancers), our strategy was to look at more abstract and global types of meaning such as the coreference/disjoint reference distinction, which is central to reference tracking. We now proceed with describing the setup of our exploratory production study.

2.2 Motivation of our study

When we investigate the semantics of dance, we naturally aim to look for any phenomena that may reflect properties similar to those found in natural language semantics. Inspired by Abusch's (2013, 2014, 2021) seminal work on the semantics of visual narrative, which builds on Greenberg's (2011, 2013) pictorial semantics, we carried out an exploratory production study of Bharatanatyam. Our investigation focused on the encoding of coreference *vs*. disjoint reference in this dance form, to explore the very tools available to a dancer with the intention of encoding such contrasts. Coreference vs. disjoint reference is a very basic and fundamental distinction in natural language semantics. While Bharatanatyam is highly conventionalized, coreference *vs.* disjoint reference is abstract enough to raise the expectation that it may be encoded through strategies that involve less conventional symbolism.

The encoding of coreference and disjoint reference between noun phrases is illustrated (very coarsely) in (2) and (3), respectively. Note that we do not aim to contribute to the large body of literature on how exactly such sentences should be analyzed (e.g., Heim, 1982), i.e. we gloss over the difference between truth-conditional and presuppositional content in (2) and (3), and we take (2a) to roughly have the truth conditions in (2b), whereas (3a) roughly has the truth conditions in (3b). The difference between (2) and (3) that is at the center of our exploration is that (2) introduces a single discourse referent whereas (3) introduces two separate discourse referents (see also Kamp and Reyle, 1993).

(2) *coreference*
   a. *A man* is sitting on the ground and *that man* is holding a spear.
   b. true iff $\exists$ x[x is a man & x is sitting on the ground & x is holding a spear]

(3) *disjoint reference*
   a. *A man* is sitting on the ground and *another man* is holding a spear.[6]

[5] Our approach thus follows the strategy of Napoli & Kraus, to focus on the non-cultural physical aspects of dance, as captured by the following two statements (Napoli & Kraus 2017:468): "Dance and language are produced and performed by the body and governed by cognitive faculties. […] thus applying linguistic methods grounded in biology to the study of dance might reveal insights."

[6] There is a non-trivial question of how the sentence in (3) relates to the (seemingly simpler) sentence with two indefinites in (i.). Given the novelty inference that arises from indefinites, it would seem superfluous to use *another man* instead of *a man*. In our study, we opted for *another man*, as this is often perceived to be more natural – which was relevant for constructing items for the production study in section 2.3-2.4. See Grønn & Sæbø (2011) for discussion of *a, the* and *another*.

i. *A man* is sitting on the ground and *a man* is holding a spear.

b. true iff $\exists$ x[x is a man & x is sitting on the ground
& $\exists$ y[y is a man & y is holding a spear & y ≠ x]]

Abusch (2013) investigates comics without words (French *sourds*), i.e. purely visual narratives.[7] She focuses on mangas such as Masashi Tanaka's *Gon*, which tell the story of Gon, a small dinosaur that interacts with real life animals. The question that Abusch raises is as follows: in a comic (Episode 4) that contains a number of eaglets, a reader can establish coreference across panels, i.e. if, in Abusch's example in Figure 1, we see an eaglet depicted in panels 32 (top right), 33 (top center), 34 (top left), and 36 (bottom left), we generally infer that this is the same eaglet (as opposed to one of the others that have been introduced earlier). The central question for Abusch is how coreference across panels is established in such comics, i.e., what is the cognitive mechanism behind such identity inferences? In the absence of words and pointing gestures, Abusch takes this to be a non-trivial question.

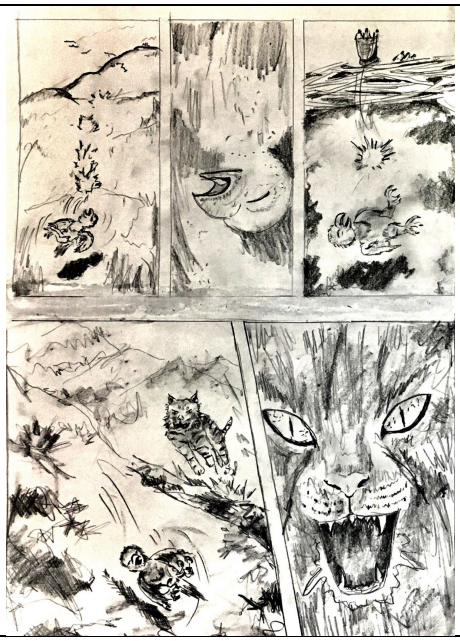


Figure 1: coreference across panels (from Abusch 2021:8), image from Tanaka (1992), redrawn for Abusch (2021) by artist Milka Green, courtesy of Dorit Abusch. Due to the right-to-left orientation of manga, the narrative excerpt starts in the top right corner, from where it continues leftwards, and eventually concludes in the bottom left corner.

In line with Discourse Representation Theory (Kamp and Reyle, 1993), Abusch proposes that the referents in comic panels are existentially quantified, (4a-c), and coreference arises from post-semantic identification of discourse referents in the pragmatics (which is a type of pragmatic enrichment), (4d). Such existential quantification is plausible in visual narratives, as there are no definite descriptions comparable to *the eaglet* in natural language.

(4) *coreference in comics without words (Abusch 2013:13)*
a. top-left panel: "[*an eaglet*]$_1$ bounced down a cliff face"
b. bottom-right panel: "[*a bobcat*]$_2$ looked and opened its mouth"

[7] See also Cohn (2018, 2019, 2020) on the cognitive relevance of comics, and specifically on the syntax (hierarchical structure) of comics.

c. bottom-left panel:
"[*a bobcat*]$_3$ jumped toward [*an eaglet*]$_4$ that was bouncing down"

d. *pragmatic enrichment*
→ "[*the bobcat*]$_{3=2}$ jumped toward [*the eaglet*]$_{4=1}$ that was bouncing down"

The contribution of Abusch's that is most central to addressing dance semantics – and thus most relevant for our investigation – includes: [i.] a generalized possible worlds model of information content (Abusch 2021:2), which allows us to define truth in visual narratives and which we discuss in section 3.1, as well as [ii.] the introduction of discourse referents into the semantic analysis of comics without words, which becomes useful for analyzing co-reference and disjoint reference in a visual narrative (compare Abusch 2013:12-15), as discussed in section 3.2.

Our point of departure is that, crucially, the questions and insights that Abusch addresses for comics without words carry over to any type of silent visual narratives, including narrative dance and pantomime. This motivates our case study of Bharatanatyam as presented in the remainder of this paper.

2.3 Design of our study

We recorded dance sequences based on a set of items that we constructed in order to probe for coreference vs. disjoint reference. We carried out two exploratory studies with a professional Bharatanatyam dancer; the first study (henceforth *Study 1*) took place in November 2016 and had the aim to probe and establish core generalizations; it yielded 30 short dance sequences (see Appendix), of which we analyze a set of 12 sequences (*Set 1*) both qualitatively and quantitatively in section 2.4 (the remaining 18 sequences exploring other phenomena, irrelevant to the present study).[8] The second study (henceforth *Study 2*) took place in December 2017 and had the aim to replicate patterns from the first study, and probe questions that arose from the earlier analysis; it yielded 24 short dance sequences, which we analyzed qualitatively (as there was no motivation for further quantitative analysis). Our focus was on the communicative strategies that a dancer can employ in order to encode coreference vs. disjoint reference.

Since the same dancer participated in both recording sessions (separated by one year), overlaps between the strategies that were employed in the two separate recording sessions can be taken to be systematic, indicating a stable communicative strategy of this dancer. In sections 2.3–2.5, we focus on the design and results from Study 1; we comment on the design of Study 2 in section 2.6. The complete stimuli for both studies are included in the Appendix.

We designed our stimuli as short narrative texts. The items were designed in a way that aims to utilize meanings conventionally encoded in Bharatanatyam hand gestures, such as the ones illustrated in (1) in section 2.1 (including objects such as 'palmyra leaf', cf. (7)). The context for all items in Study 1 is given in (5); this context (an artist having designed a statue for a temple) was chosen to be as natural as

[8] In Study 1, we also recorded 6 mini-narratives (12 dance sequences) on the distinction between reflexive and non-reflexive reference, and 3 mini-narratives (6 dance sequences) on the encoding of de se vs. (non-de-se) de re readings in Bharatanatyam narratives. This yielded 30 dance sequences in total. While we found clear strategies for encoding coreference vs. disjoint reference, which we report in this paper, the strategies that the dancer employed for encoding reflexive as opposed to non-reflexive reference (or de se readings vs. non-de-se readings) are less systematic and were thus not explored further at this point.

possible, with the aim of limiting artificial components in the narrative that are solely due to the study design. What is crucial for our setup is the idea that there are several possible referents in the context (here: 'the room is full of people'); this allows us to freely introduce discourse referents.

(5) *Context:* An artist has designed a statue for a temple. She is at the temple, watching how people interact with the statue; the room is full of people.

To probe for coreference vs. disjoint reference in Study 1, we recorded 6 mini-narratives in 2 conditions, i.e. 12 dance sequences in total. Two sample narratives are given in (6) and (7) (with the remainder in (10)-(13) in section 2.4.1). This setup allows us to elicit pairs of dance sequences in our production study, which semantically only differ in whether the agent of the last-mentioned event has been introduced earlier (*the same man* / *that woman*) or not (*another man* / *another woman*).[9] In each item, both dance sequences start the same, e.g. in (6a-b), the artist sees a strong man sitting on the ground. Then they differ in terms of whether the same individual is involved in another action, or a different individual. The embedding in perception contexts ('the artist sees…') aims at fixing a perspectival center for the narrative; in follow-up elicitation as part of Study 2, we included unembedded variants (e.g. 'A woman is sitting on the ground. […]').[10] The resulting dance sequences do not reflect this difference.

(6) *Item 1*
   a. The artist sees a strong man sitting on the ground.
      Then she sees that *the same man* is holding a spear. (coreference)
   b. The artist sees a strong man sitting on the ground.
      Then she sees that *another man* is holding a spear. (disjoint reference)

(7) *Item 2*
   a. The artist sees a woman waving a palmyra leaf in the sunlight.
      Afterwards *that woman* is pointing at the clouds in the sky. (coreference)
   b. The artist sees a woman waving a palmyra leaf in the sunlight.
      Afterwards *another woman* is pointing at the clouds in the sky. (disj. ref.)

In terms of possible manipulations, Bharatanatyam is relatively flexible. It is typically accompanied by music and chanting (see Puri 2004:52-53), but it can also be danced

[9] Our approach is inspired by the construction of 'minimal pairs' in natural language elicitation. We aimed at the pairwise comparison of expressions that only minimally differ, which can be taken to be a central part of linguistic methodology, going back to structuralist analyses in the first half of the 20th century (see Hockett 1942:7 for a discussion of "the traditional term 'minimal pair'"). Such an approach has, in later years, seen extensions to semantic and syntactic minimal pairs (e.g. Fodor et al. 1980:301), and minimal pairs beyond language (as part of Super Linguistics; see Schlenker 2017a:5). As we will see, the pairs of dance sequences that we elicited are not minimal in that they systematically differ with regards to four properties (or *cues*). Nevertheless, we consider it worth emphasizing that our approach to probing meaning in dance is thus inspired by linguistic methodology, a central feature of super linguistic research.

[10] The intention behind fixing a perspectival center in Study 1 was to make the narrative more natural (based on consultation with the dancer), but also to favor a dance sequence where the dancer aims to *narrate* by means of dance, rather than *acting out* the narrative by playing the different characters on stage. This methodological choice ended up being irrelevant, as there is no difference with regards to the relevant findings between the studies that had a perspectival center, and the follow-up study where we removed the perspectival center, (17) and (18).

without them.[11] For reasons of simplicity, we recorded our stimuli without music, as this reduced any potential influence from the music (e.g. from its beat) onto the dance sequence.

The dance sequences were recorded in the fourMs Lab of the Department of Musicology, University of Oslo. A professional Bharatanatyam dancer was recorded by one video camera and eight motion capture cameras, using an infrared, marker-based Qualisys motion capture system with eight wall-mounted Oqus 300 cameras, capturing at 200 Hz. A total of 45 reflective markers ("dots" to be tracked by the cameras) were placed on the body of the dancer. The advantage of such a production study is that we can compare controlled pairs of narratives (such as (6a) vs. (6b)) and see how intended meanings can be encoded. *After* recording the 12 dance sequences without any accompaniment, we recorded the same 12 dance sequences while slowly reading out the text; this allowed us to map the recorded movements (and related gestures) to intended meanings in case of uncertainty, while at the same time minimizing a potential source language effect. An open question, which goes beyond the scope of this paper, is whether we expect to find differences in how a dancer conveys meaning in planned / choreographed dance moves *vs.* spontaneous dance moves.[12] The dancer did not choreograph the dance sequences in advance, but read the dance sequences before beginning the dance sequence. While the production thus involves a certain amount of planning (and is not fully spontaneous), it still retains a certain amount of spontaneity.

For the analysis, the recordings were post-processed in the Qualisys Track Manager software (QTM 2.16). This software generates a 3-dimensional (3D) rendering based on the multi-camera recording of the reflective markers, as illustrated for four dance positions in Figure 2. In the remainder of this paper, we use the 3D renderings in order to focus on the "global" (full-body) gesture aspects of the dance sequence that are relevant for us (glossing over details that may be present in the live video recording yet lost in the 3D rendering).[13]

[11] Puri (2004:52-53) observes that traditional Bharatanatyam recitals in India are always accompanied by South Indian *Karnatic* music. This involves a dance master (*naṭṭuvanār*), who beats a rhythm and rhythmically chants syllables, in addition to a drummer, at least one singer and at least two instrumentalists. Modern Bharatanatyam performances are frequently more experimental, especially on a global stage, and depart from this traditional setup. More experimental approaches are exemplified by the work of Subathra Subramaniam and Mayuri Boonham, who founded the British dance company Angika (1997-2009), whose performance *Triple Hymn* combined Bharatanatyam dance with Western art music, as discussed in O'Shea (2003:180) (see also Kedhar 2020). A recent *TEDxBrighton* presentation by Subathra Subramaniam, which discusses her approach to Bharatanatyam, can be found here: https://youtu.be/nqvki2hSDzE – This presentation contains four Bharatanatyam sequences that are performed in silence, entirely without musical accompaniment.

[12] There is no reason to prioritize either planned or spontaneous expressions for super linguistic analysis, which has been productively applied to both comics (e.g. Abusch 2013, 2021), which are planned, and speech-accompanying gestures (e.g. Tieu et al. 2017), which are typically spontaneous.

[13] The twelve video sequences that we analyze in section 2.4, and which were used as stimuli for the perception study in section 2.5, can be viewed at this link:
https://www.dropbox.com/sh/fj5j3ryjkyz6ro0/AAA98e4Dlkt8zJUXoLX1ySJQa?dl=0
The naming format is such that "Set1_2A.avi" is the second item in the coreference condition, whereas "Set1_2B.avi" is the second item in the disjoint reference condition, with the "Set1_" prefix marking that this was the first set of 12 recordings, see Appendix.

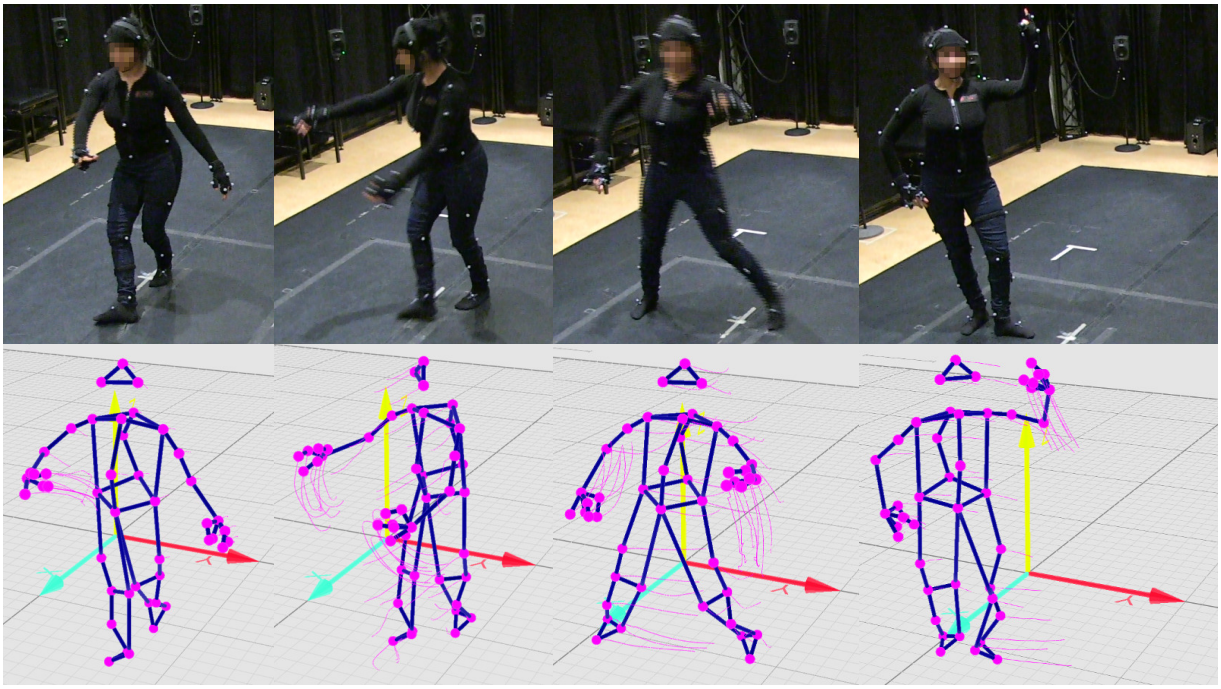


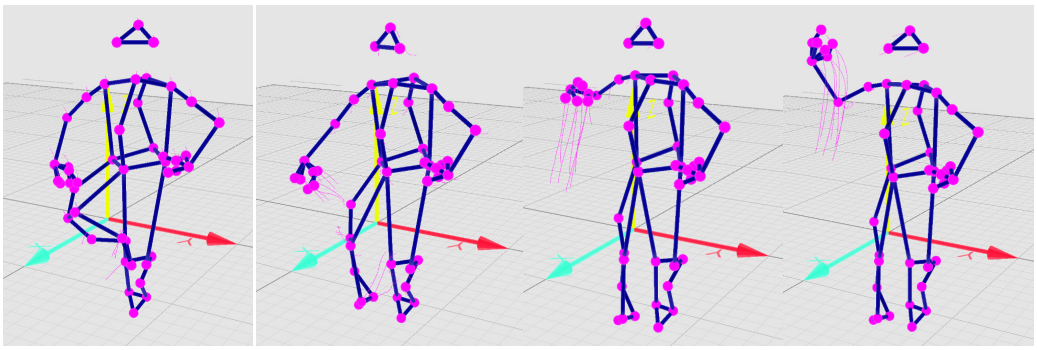
Figure 2: sequence of four dance positions (stills from the video recording and 3D motion capture rendering, with motion history trajectories)

2.4 The production data

2.4.1 Qualitative analysis

We start by analyzing the coreference sequence, (6a), adapted in (8); as shown in Figure 3, we can zoom in on the movement and study different parts. In Figure 3, each label [P_n] represents a dance position; these positions are stipulated at semi-arbitrary cut-off points (chosen on the basis of our qualitative analysis), since a dance performance is by its very nature non-discrete. As indicated in (8), we can identify the dance position [P_{11}] with an activity of *sitting on the ground*, whereas the dance position [P_{14}] represents an activity of *holding a spear*. Intermediate stages (such as [P_{12}] and [P_{13}]) cannot be as easily connected to parts of the written narrative.

(8) The artist sees a strong man [$_{P11}$ sitting on the ground].
Then she sees that *the same man* [$_{P14}$ is holding a spear].

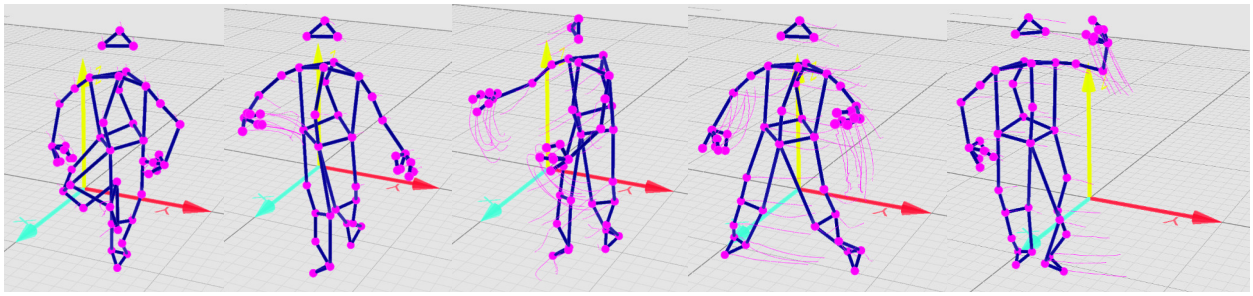


[P_{11}] [P_{12}] [P_{13}] [P_{14}]

Figure 3: coreference condition

Figure 3 illustrates that the coreference condition involves a single, simple fluid motion, from displaying a sitting position to displaying a spear-holding position. It does not seem to be necessary (in the given context) to separately mark coreference between the "sitter" and the "spear holder". By contrast, the disjoint reference condition, repeated in (9) from (6b), has additional complexity, as illustrated in Figure 4. Once again, we can identify a dance position that symbolizes a *sitting on the ground* activity, [P_{21}]; an attentive reader will notice a remarkable consistency between [P_{11}] in Figure 3 and [P_{21}] in Figure 4, which are taken from two separate recordings. We can also identify a dance position that symbolizes a *spear holding* activity, [P_{25}]. Most interestingly, for our purposes, the marking of disjoint reference can be broken down into three different dance positions that are assumed between [P_{21}] and [P_{25}]. Step by step, we notice that after giving up the sitting position [P_{21}], the dancer first uses a hand-and-arm gesture that symbolizes "another/different", in [P_{22}] (roughly: a round movement of the right hand and arm from the left to the right). She then marks a new position in the visual space, [P_{23}], and she then assumes the new position, [P_{24}]. Eventually, she assumes the spear-holding position in [P_{25}], but does so in a way that mirrors the spear-holding position in the coreferent condition ([P_{14}] in Figure 3), i.e. it is now the left arm that is raised (as opposed to the right arm) and the dancer faces towards the left (as opposed to the right).

(9) The artist sees a strong man [$_{P21}$ sitting on the ground].
Then she sees that [$_{P22+P23+P24}$ *another man*] [$_{P25}$ is holding a spear].



[P_{21}] [P_{22}] [P_{23}] [P_{24}] [P_{25}]

Figure 4: disjoint reference condition

Table 1 summarizes the presence and absence of these four cues ('another' in [P_{22}], pointing in [P_{23}], movement and change of orientation in [P_{24}], and mirroring in [P_{25}] *vs*. [P_{14}]) across all twelve dance sequences. While Items 1 and 2 have already been given in (6) and (7), the remaining four items are given in (10)-(13).[14] An empirical data point that becomes clear from Table 1 is that the four cues do not seem to be rigid parts of a conventionalized sequence; while Item 1 exhibits all four, the posture mirroring is only present in Items 1 and 2 – this is trivially due to the fact that some postures are symmetric, using both hands and arms the same way, so mirroring would be vacuous (at least in Items 3, 4, and 5, which is why the relevant cells are marked as *n/a*). From Table 1, it may appear as if the 'another' gesture were the most robust component of the dance sequences in this study, but, as discussed in Figure 8 in

[14] A reader may notice that we varied simple definites (*the child*) with demonstratives (*that woman*) and DPs that contain *same* (*the same man*) in the coreference conditions. This being an exploratory study, the goal was to see if this would make any difference whatsoever; our analysis of the resulting dance sequences show that all three types of definite descriptions were danced in the same way.

section 2.6.2, this seems to be an artifact of the particular prompts, which contained the word 'another'.

description of cue	coreference condition						disjoint reference condition					
	1A	2A	3A	4A	5A	6A	1B	2B	3B	4B	5B	6B
'another' gesture	–	–	–	–	–	–	✓	✓	✓	✓	✓	✓
pointing gesture	–	–	–	–	–	–	✓	–	✓	✓	✓	✓
move to new position	–	–	–	–	–	–	✓	?[15]	✓[16]	✓	✓	✓
mirroring the posture	–	–	n/a	n/a	n/a	–	✓	✓	n/a	n/a	n/a	–
Table 1: presence of cues across the 12 dance sequences (✓ = present)												

(10) *Item 3:* The artist watches a child eating a mango outside the temple.
Then {*the child / another child*} is entering the temple.

(11) *Item 4:* The artist watches a man holding a book.
Then she sees {*the same man* / *another man*} looking at a water lily.

(12) *Item 5:* The artist sees a woman praying in silence.
Then {*that woman / another woman*} walks to a basket of fruits.

(13) *Item 6:* The artist watches a girl dancing in the sunlight.
Then {*the girl / another girl*} trips over a stone.

2.4.2 Quantitative analysis

Table 1 in the previous section was based on the authors' qualitative analysis of the dance sequences, based on viewing of the sequences. The motion-capture technology allows us to corroborate this qualitative analysis with quantitative data, focusing in particular on the move to a new position, as was illustrated by [P_{24}] in Figure 4. To measure the systematic presence of such a move in the disjoint reference condition, we proceeded as follows. Out of the 45 reflective markers that had been placed on the body of the dancer, we calculated the *centroid* of the foot markers (also referred to as the *body centroid*), i.e., the point that lies between the right-foot marker and the left-foot marker. To do so, we first segmented and annotated the motion-capture recordings by determining, e.g., when the dancer had finished assuming position [P_{21}] in Figure 4, and when the dancer started moving again. We thus determined the

[15] In Item 2, the dancer's position and orientation is different in the coreference *vs.* disjoint reference condition, in line with our overall generalization, but the change is less pronounced than in the other items, and it is not preceded by an explicit pointing towards that position. A potential confound is the presence of the predicate *point* (in *pointing at the clouds in the sky*) in the prompt, (7), as flagged by an anonymous reviewer. The presence of a definite description (*the clouds in the sky*) in the prompt may independently be a confound, as discussed in footnote 16 with regards to Item 3.

[16] Item 3 is given in (10). In this dance sequence, the artist points at a new position and briefly moves into that position in order to mark *another child* (as opposed to *the child*); however, she then assumes the same orientation on stage, both in the coreference and in the disjoint reference condition, and moves in a similar way, ending up in nearly the same location towards the end. Crucially, the confounding factor in this example is the recurrence of the definite description *the temple*, which appears to be associated with the location at which both narratives converge. What corroborates this assumption is that the dancer points back at the location of the second child in the disjoint condition before pointing towards location of the temple; by contrast, she points at the location of the first child in the coreference condition before pointing at the temple. This is parallel to what we discuss in section 2.6.1, when we discuss narratives with more than two referents.

beginning and the end of [P_{21}]. In this specific example, position [P_{21}] in Figure 4 lasted from frame 1870 until frame 1949 in the recording, whereas position [P_{25}] lasted from frame 2750 until frame 3317. We tracked the body centroid throughout these two time-intervals and then compared the averages. We expect that the shift in position from [P_{21}] to [P_{25}] in Figure 4 gives rise to a larger distance between the [P_{21}]-centroid and the [P_{25}]-centroid, than what we expect to find in Figure 3, between the [P_{11}]-centroid and the [P_{14}]-centroid. In each of the 12 dance sequences, the two centroids were calculated that correspond to the predicates of the respective sentences (e.g. *sitting on the ground* and *holding a spear*).

The results of the quantitative analysis are summarized in Figure 5. This figure should be interpreted as follows. Each of the six shapes (disk, diamond, square, pentagon, cross and star) corresponds to one of the six items, i.e. a pair of two narratives that differ in coreference vs. disjoint reference. The color and pattern combinations mark the first and second position within each condition of each item, i.e., what we discussed above for item 1 by using the labels [P_{11}]/[P_{21}] (first position) and [P_{14}]/[P_{25}] (second position). Two black dotted arrows have been added to the Figure as an aid for the reader, indicating the distance between the centroids associated with [P_{11}] (yellow disk with dots) and [P_{14}] (red disk with grid pattern), and the distance between the centroids associated with [P_{21}] (green disk with horizontal lines) and [P_{25}] (blue disk with diagonal lines). What the reader should pay attention to is the observation that the yellow (1st position coreferent), red (2nd position coreferent) and green dots (1st position disjoint referent) roughly cluster together on the plot, whereas the blue dots (2nd position disjoint referent) are further removed from that cluster. The red triangle, which corresponds to the 2nd position in the coreference condition of item 2, is an outlier in this respect; however, the triangles are not an exception to the generalization that the second position in the disjoint condition is associated with a different location in space than the second position in the coreferent condition; while the starting positions (yellow and green triangle) are close together, the later positions (red and blue triangle) are visibly distinct.

Note that, across all items, the 2nd position and the 1st position in the coreferent condition are typically distinct (e.g. the yellow cross is quite far removed from the red cross), with item 1 being the exception rather than the rule; this is simply due to the fact that the dancer generally moves as part of the performance and does not stay static in the same location. The important difference is the relative distance from the 1st to the 2nd position across the two conditions, i.e. coreference vs. disjoint reference. (Even in the case of the crosses, the distance from the green to the blue cross is larger than the distance from the yellow to the red cross.)

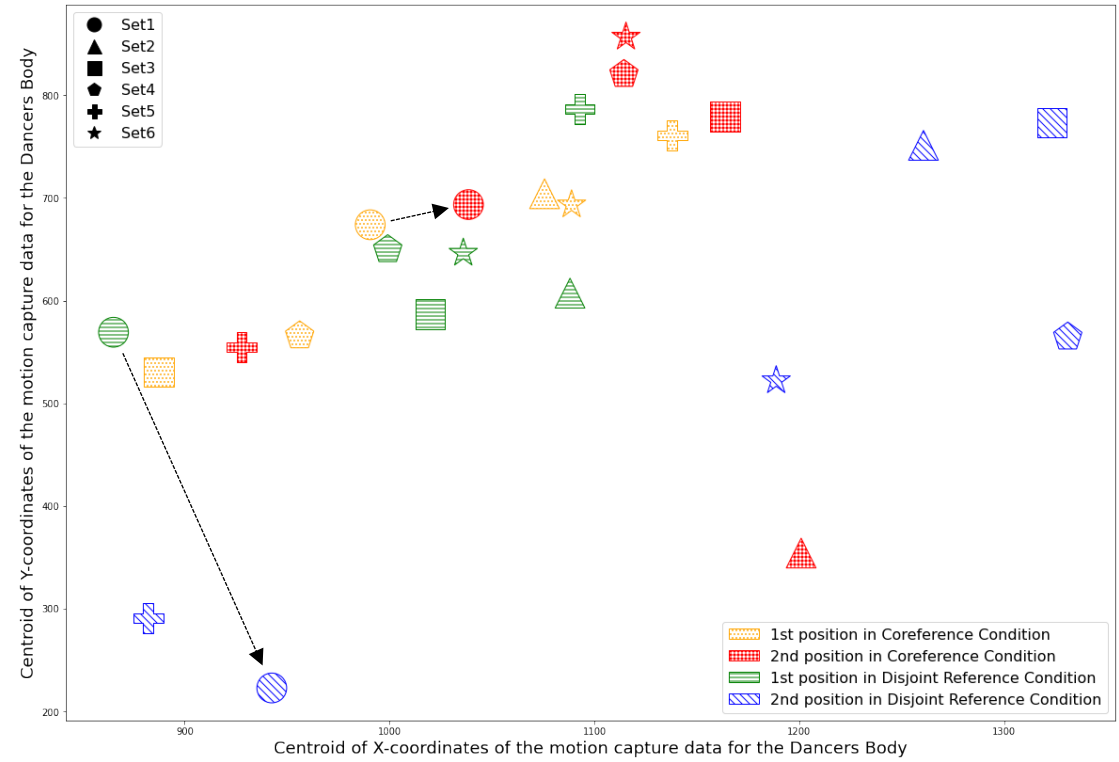


Figure 5: position of dancer on X-axis (horizontally) and Y-axis (vertically);[17] the six shapes correspond to the six items (pairs of narratives); the yellow and red dots are from the coreferent condition; the green and blue dots are from the disjoint referent condition; dotted arrows have been added to item 1 for illustration purposes only, to show which two dots belong to the same condition (coreference vs. disjoint reference); the beginning of the arrows marks the first position, and the end marks the second position.

As an additional descriptive statistic, we can calculate the mean displacement in the x-y plane of the body centroid, corresponding to the distance traveled from the first position to the second position in the coreference condition *vs.* in the disjoint reference condition. In the coreference condition, this mean distance amounts to a 280 mm displacement in the x-y plane of the body centroid. By contrast, in the disjoint reference condition, it amounts to a 344.07 mm displacement in the body centroid in the x-y plane. The difference is summarized in the box plot in Figure 6.

[17] In the motion-capture stills (e.g., in Figure 4), the origin of the graph was in the center of the pictures, and the y-axis went rightward and downward while the x-axis went leftward and downward. By contrast, the plot in Figure 5 has the origin in the bottom left corner, with the (horizontal) x-axis rightward and the (vertical) y-axis upward.

Figure 6: displacement in the x-y plane of the body centroid; this graph, compares the coreference condition (mean displacement = 280 mm) to the disjoint reference condition (mean displacement = 344.07 mm); the red horizontal line corresponds to the median displacements.

The quantitative analysis thus corroborates the generalization that the disjoint reference conditions systematically exhibit a change in position that is missing in the coreference conditions. We will return to this change in position in section 3.

So far, we have only focused on the production side of encoding coreference and disjoint reference in dance; in section 2.5, we present a perception study aimed at determining whether onlookers can actually draw the intended meaning references from watching the video sequences.

2.5 Verifying the production data

In order to gain additional insight on the meaningful components of the recorded production sequence, we carried out a pilot perception study in which participants who did not have prior experience with Bharatanatyam watched videos with stick-figure exports (using the same software that produced the stills above). Each participants saw 6 videos in both conditions (*coreference* and *disjoint reference*),[18] plus 6 filler videos from a different set of recordings (probing for reflexives such as *he calls himself a hero*). A total of 18 videos was shown to 32 participants in a pseudo-randomized order. The complete list of items that were recorded and shown to the participants is given in Table 2.

[18] The video stimuli can be found at the following DropBox link:
https://www.dropbox.com/sh/fj5j3ryjkyz6ro0/AAA98e4Dlkt8zJUXoLX1ySJQa?dl=0

Item Order	Video Name	Condition	Narrative
1	Set2_6B	filler (non-reflexive)	The artist watches a boy with pen and paper. The boy is drawing another boy.
2	Set2_5A	filler (reflexive)	The artist watches a man walk up to the statue. Then the man calls himself a hero.
3	Set1_5A	coref	The artist sees a woman praying in silence. Then that woman walks to a basket of fruits.
4	Set1_4B	disjoint	The artist watches a man holding a book. Then she sees another man looking at a water lily.
5	Set1_3A	coref	The artist watches a child eating a mango outside the temple. Then the child is entering the temple.
6	Set2_3A	filler (reflexive)	The artist sees a boy running around the temple. Afterwards the boy sees himself in the mirror.
7	Set1_1A	coref	The artist sees a strong man sitting on the ground. Then she sees that the same man is holding a spear.
8	Set1_6B	disjoint	The artist watches a girl dancing in the sunlight. Then another girl trips over a stone.
9	Set1_2B	disjoint	The artist sees a woman waving a palmyra leaf in the sunlight. Afterwards another woman is pointing at the clouds in the sky.
10	Set2_3B	filler (non-reflexive)	The artist sees a boy running around the temple. Afterwards the boy sees another boy in the mirror.
11	Set1_6A	coref	The artist watches a girl dancing in the sunlight. Then the girl trips over a stone.
12	Set1_1B	disjoint	The artist sees a strong man sitting on the ground. Then she sees that another man is holding a spear.
13	Set1_5B	disjoint	The artist sees a woman praying in silence. Then another woman walks to a basket of fruits.
14	Set2_6A	filler (reflexive)	The artist watches a boy with pen an paper. The boy is drawing himself.
15	Set1_2A	coref	The artist sees a woman waving a palmyra leaf in the sunlight. Afterwards that woman is pointing at the clouds in the sky.
16	Set2_5B	filler (non-reflexive)	The artist watches a man walk up to the statue. Then the man calls another man a hero.
17	Set1_4A	coref	The artist watches a man holding a book. Then she sees the same man looking at a water lily.
18	Set1_3B	disjoint	The artist watches a child eating a mango outside the temple. Then another child is entering the temple.

Table 2: original textual narratives of videos that were seen by participants in the perception study

Participants (n=32) watched the videos on a laptop, aided by a research assistant. They then filled out a questionnaire with the following instructions.

> You will watch 18 short videos (less than ½ minute each). Each video contains a stick figure animation in which a professional dancer is performing a short story by means of dance.
>
> Some of the videos contain a story about 1 person. Others contain a story about 2 people. We are interested in the following question: Can you identify the number of people in the story?
>
> For each of the videos, please place an X into the box that best reflects your impression. Do you think that this dance sequence describes a story that involves 1 person or 2 people?

Participants were then asked to give a rating on the following scale.

definitely **one** person	quite likely **one** person	unsure **(one or two)**	quite likely **two** people	definitely **two** people

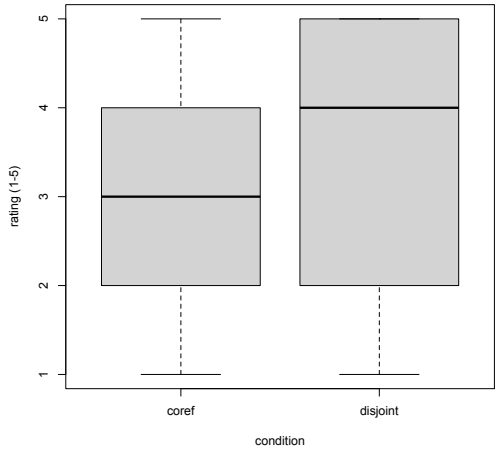
For the analysis, the scale was converted to a numerical scale, assigning the values in (14). In other words, a higher rating would correlate with *disjointedness of reference*.

(14) *conversion of ratings to a numerical scale*
5 = definitely **two** people
4 = quite likely **two** people
3 = unsure (**one or two**)
2 = quite likely **one** person
1 = definitely **one** person

The pilot study yielded an average rating of 2.92 for the coreference condition, compared to an average rating of 3.27 for the disjoint condition. A linear mixed effects regression yielded a significant main effect ($t = 2.279$, $p < 0.05$)[19] of disjointedness. We interpret the statistical significance as follows: The data reject the null hypothesis (H_0) that the mean ratings in the coreference condition are identical to the mean ratings in the disjoint condition. The 32 subjects for this study were not professional dancers, and they did not have prior exposure to Bharatanatyam. In a brief follow-up survey, participants were asked about the country they grew up in; all 32 participants grew up in Northern Europe. This first pilot study thus suggests that participants who are not professional dancers can infer intended meanings from dance sequences – even though they only see stick figure renderings.

It is worth making two further remarks in connection with these results: First, it is evidently not clear how participants drew these inferences. While they were allowed to provide optional comments, none of them commented on the actual items; participants only occasionally used the option to comment in connection with the fillers, for reasons that are unclear. Second, we ran the same questionnaire study with 5 professional Bharatanatyam dancers; their ratings were not included in the summary and analysis above. However, the results were consistent with those from the untrained participants, in that their mean ratings were 2.8 for coreference and 3.33 for disjoint reference (compared to 2.92 for coreference and 3.27 for disjoint reference in untrained participants). This suggests that experience and prior exposure to the dance form may only play a negligible role in perception. The results are summarized by the box plots in Figure 7.

[19] While we report the results for the raw scores, the analysis with z-scores also yielded a significant main effect ($t = 2.812$, $p < 0.01$).



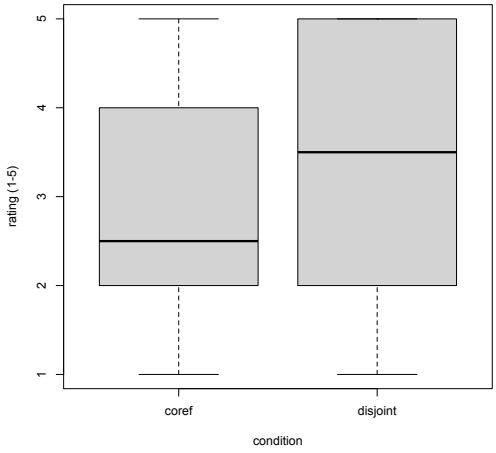


Figure 7: disjointness ratings from perception study (1 = definitely one person, 5 = definitely two people); the left box plot shows the ratings from untrained participants (*n*=32), while the right box plot shows the ratings from the professional Bharatanatyam dancers (*n*=5); the thick line marks the median.

Given that the dance sequences themselves exhibit some variation in how cues of disjointness are realized (see Table 1 in section 2.4.1), we follow up with a *post hoc* analysis of the individual selections, as made by the participants in the perception study, which we list in Table 3. Recall that the *A* variants (1A, 2A, …) are the coreference conditions, while the *B* variants (1B, 2B, …) are the disjoint reference conditions.

rating	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B	6A	6B
def. **two**	8	7	6	3	9	15	4	8	1	8	11	9
likely **two**	8	10	3	3	9	11	6	14	12	13	7	11
unsure	1	3	1	2	5	1	4	1	1	0	2	1
likely **one**	10	4	13	14	2	3	10	5	9	3	6	5
def. **one**	5	8	9	10	7	2	8	4	9	8	6	6

Table 3: raw counts of answers per condition (*n=32*)

We can briefly zoom in on the results in Table 3; in Items 1, 3, 4, 5, and 6, the change from coreferent (*A*) condition to disjoint referent (*B*) condition systematically increases the total number of participant who selected either *definitely two people* or *quite likely two people* (e.g. from a total of 18 counts / 56% in Item 3A to a total of 26 counts / 81% in Item 3B). Crucially, the only item that does not comply with this tendency is Item 2, where the disjoint referent condition (2B) had less selections in the *definitely/quite likely **two** people* category (6 counts / 19%) than the coreference condition (2A) (9 counts / 28%). This, in fact, is entirely expected, since Item 2 was the only dance sequence (see Table 1) that did not involve an obvious instantiation of pointing at a new location and then assuming that new location. We tentatively conclude that pointing and moving into a new location (accompanied by a change of orientation) is the most crucial disjointness cue, as witnessed by the results of our pilot perception study.

2.6 Further empirical explorations

2.6.1 Design of second exploratory study (Study 2)

The second exploratory study (Study 2, Dec. 2017) used the same approach as the first exploratory study (Study 1, Nov. 2016), as was described in sections 2.3-2.5. Study 2 aimed to answer questions that arose from the qualitative analysis of Study 1. As a reminder, Item 1 from Study 1 is repeated from (6) in (15); both conditions share the initial sentence in (15a), whereas condition 1 (coreference) continues with the sentence in (15b) and condition 2 (disjoint reference) continues with the sentence in (15c).

(15) *Study 1 – Item 1*
- a. The artist sees a strong man sitting on the ground.
- b. *Cond. 1 (coref):* Then she sees that *the same man* is holding a spear.
- c. *Cond. 2 (disjoint):* Then she sees that *another man* is holding a spear.

We now proceed by discussing the three relevant sets of recordings from Study 2 (Set 1, Set 2, Set 4) that build on the findings from Study 1.[20] Each of these three sets builds on a disjoint reference case, either with 'another' (Set 1 of Study 2) or with definite descriptions that are sufficiently distinct (e.g., *man, woman, child*) to potentially do away with 'another' (Set 2 and Set 4). In addition, each of these three sets is designed such that a referent that is introduced early in the narrative is picked up again later in the narrative (see, e.g., (16a-c) for illustration, where (16a) introduces *a child eating a mango*, and (16c) picks up the same referent with the definite description *the eating child*).

In Set 1 of Study 2 (henceforth *Study 2.1*), we recorded 3 pairs of dance sequences (6 dance sequences in total), one of which is illustrated in (16) (see Appendix for the whole set); both conditions share the initial sentences in (16a) and (16b), whereas condition 1 (parallel subject-object reference) continues with the sentence in (16c) and condition 2 (switched subject-object reference) continues with the sentence in (16d). (16a-b) reproduces the disjoint condition of Study 1, whereas (16c-d) probe what happens if the same referents are picked up again later in the narrative – this is what we aimed to investigate in Study 2.1.

(16) *Study 2.1 – Item 1*
- a. The artist sees a child eating a mango outside the temple.
- b. Then she sees another child holding a spear.
- c. *Cond. 1 (parallel):* The eating child watches the child with the spear.
- d. *Cond. 2 (switched):* The child with the spear watches the eating child.

We carried out a qualitative analysis of the results and found that (16a-b) reproduce the findings from Study 1, i.e., the dancer uses the same cues in order to mark disjoint reference between *a child* and *another child*. While replication was not the aim of Study 2.1, it is worth emphasizing that Study 2.1 (Dec. 2017) replicated the findings of Study 1 (Nov. 2016) after a 13-month interval, i.e., the strategies employed by the

[20] We do not include a discussion of Set 3 of Study 2, where we probed for the encoding of temporal reference (i.e. *earlier, now, before that, later, after that*) in Bharatanatyam narratives. (See Appendix for the items.) This was elicited to see if pointing would play a role in temporal reference as well, but it is orthogonal to the discussion in this paper.

dancer seem to be consistent. We lay out the qualitative analysis of (16c-d) in section 2.6.2.

In Set 2 of Study 2 (henceforth *Study 2.2*), we recorded the 6 dance sequences given in (17a-g). The aim of Study 2.2 was to see how the dancer would represent a narrative with three characters (rather than two characters). All 6 dance sequences started with the initial sentence in (17a). Each of them continued with a sequence of two sentences, as given in (17b-g). The aim of this study was two-fold: first, in (17a), we aimed to check how the dancer would introduce three referents; crucially, these referents are descriptively distinct (*woman/man/child*) and the expectation was that they would not require the 'another' gesture discussed in section 2.4.1 (something that we also tested in Set 4 of Study 2). Conditions 1-6 (in (17b-g)) were designed to explore which strategies the dancer would employ in order to retrieve referents that have been introduced at an earlier point (in (17a)).

(17) *Study 2.2*
- a. A woman is standing outside the temple, a man is sitting on the ground, and a child is playing.
- b. *Cond. 1:* The woman is holding a book. The man is looking at the child.
- c. *Cond. 2:* The woman is holding a book. The child is looking at the man.
- d. *Cond. 3:* The man is holding a book. The woman is looking at the child.
- e. *Cond. 4:* The man is holding a book. The child is looking at the woman.
- f. *Cond. 5:* The child is holding a book. The woman is looking at the man.
- g. *Cond. 6:* The child is holding a book. The man is looking at the woman.

In Set 4 of Study 2 (henceforth *Study 2.4*), we recorded 2 pairs of dance sequences (4 dance sequences in total), one of which is illustrated in (18) (see Appendix for the whole set); both conditions share the initial sentences in (18a-b), whereas condition 1 (topic shift) continues with the sentence in (18c) and condition 2 (continued topic) continues with the sentence in (18d). (18a) reproduces the disjoint condition of Study 1, crucially without the word *another* in the English-language narrative. (18c-d) probe what happens if the same referents are picked up again later in the narrative – this is the same question that we pursued in Study 2.1 and Study 2.2.

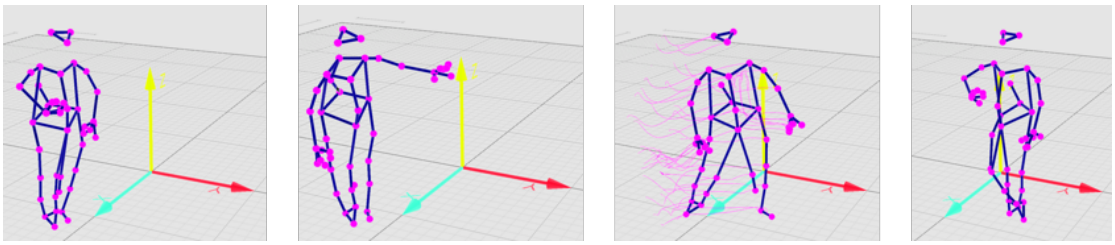
(18) *Study 2.4 – Item 1*
- a. A woman is sitting outside and a man is standing in the middle of the room.
- b. The woman is holding a book.
- c. *Cond. 1 (topic shift):* The man is looking at her.
- d. *Cond. 2 (continued topic):* She is looking at the man.

2.6.2 Qualitative analysis of results from Study 2

The findings from Study 2 (beyond the replication of aspects of Study 1 that we briefly discussed in section 2.6.1) are two-fold. First, Study 2.2 and Study 2.4 are set up in a way where the introduction of disjoint referents does not require 'another' in the English-language prompt, as the descriptive differences between the two referents are sufficiently disambiguating (e.g., *man* vs. *woman*). We repeat the relevant parts of the narratives in (19a) (from (17a)) and in (19b) (from (18a)).

(19) a. *Introduction of disjoint referents in Study 2.2*
A woman is standing outside the temple, **a man** is sitting on the ground, and **a child** is playing.
b. *Introduction of disjoint referents in Study 2.4*
A woman is sitting outside and **a man** is standing in the middle of the room.

What we found in the realization of (19a) and (19b) is that the dancer utilized the pointing gesture to introduce a new referent at a new position on the stage, followed by a move to that referent's position and change of orientation, in line with our summary of the Study 1 findings in Table 1. Figure 8 (from Item 2 of Study 2.4; see Appendix) exemplifies such a sequence. However, in sharp contrast, the dancer did not use the 'another' gesture. We take this to indicate that the 'another' gesture was either an artifact of two undistinguishable descriptions (*a man ... a(nother) man*) or a source language effect due to the English-language prompt, rather than a crucial component of introducing a new referent.[21,22] We can thus conclude, that – in dance, as in any natural language – the 'another' gesture is not a necessary part of introducing a new referent; it is simply an artifact of having two indistinguishable descriptions in the text prompt. By contrast, the introduction (and assumption) of a new position indicates that every referent in our Bharatanatyam scenarios is associated with its own position on the stage.



A man … and … [new referent] [assume location] … a woman …

Figure 8: disjoint reference without 'another'

In addition to probing the introduction of a new referent without 'another', all three sets in Study 2 include a subsequent part in the narrative where a referent is picked up again that has been introduced earlier on. To see this, reconsider item 1 from Study 2.1, which is repeated in (20).

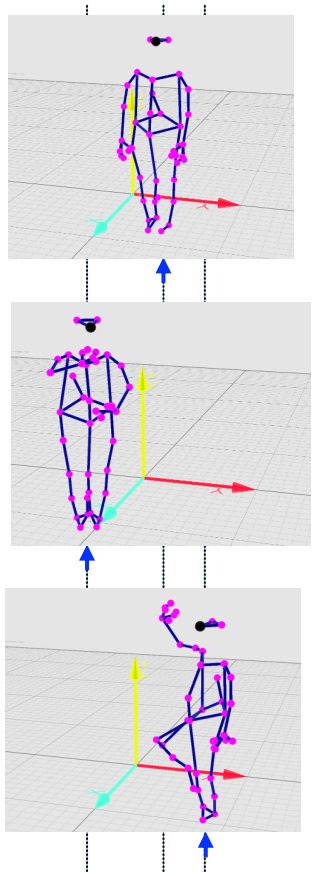
(20) *Study 2.1 – Item 1*
a. The artist sees a child eating a mango outside the temple.
b. Then she sees another child holding a spear.
c. *Cond. 1 (parallel):* The eating child watches the child with the spear.

[21] As pointed out in our discussion of Table 1, the 'another' gesture was robustly present in the six disjoint-reference dance sequences of our first study, but all of the prompts for the dancer included the word 'another', which may have influenced the way the narratives were performed. Crucially, in follow-up studies that did not contain the word 'another' in the prompt, we never found the 'another' gesture in the dance sequence.

[22] Bharatanatyam has means of encoding concepts such as 'man', 'woman' and 'child', through the use of hand gestures (*mudras*) and body postures; we discussed the recordings with the dancer afterwards, and she confirmed that her gestures and posture in the sequence in Figure 8 express the meaning 'man' when the first referent (*a man*) is introduced, and the meaning 'woman' when the second referent (*a woman*) is introduced.

d. *Cond. 2 (switched):* The child with the spear watches the eating child.

What we found in all three sets of Study 2 is the following. First, the dancer introduces two positions on the stage, using the mechanism outlined in Figure 8. To see this, consider Figure 9. Reading these pictures from top to bottom, we find that the dancer starts in a neutral (central) position and then introduces a position for the first referent, followed by a position for the second referent; the positions are indicated with a blue arrow.



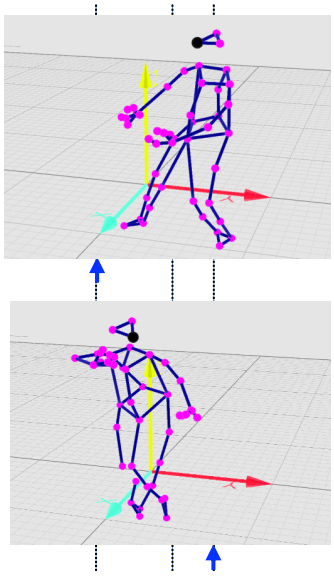
The artist (neutral position) sees …

… **a child** (position 1) eating a mango outside the temple. …

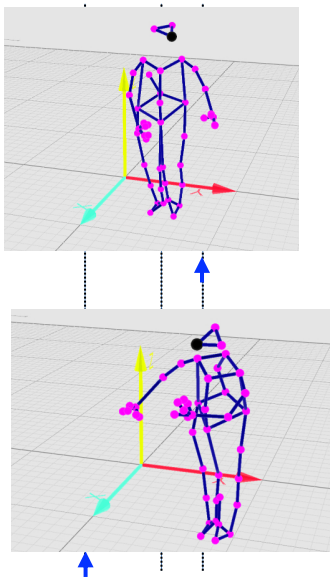
… Then she sees **another child** (position 2) holding a spear.

Figure 9

Zooming in on the third sentence of (20), where the two conditions differ in (20c) vs. (20d), we find that anaphoric dependencies can be established by virtue of re-using the positions on the stage. Let us start with the sequence based on (20c). Figure 10 shows that the dancer first points at position 1 (reactivating the *eating-child* referent) and then moves into position 1, orienting herself towards position 2 (i.e., towards *the child with the spear*).

	**The eating child** (pointing at position 1) … … (dancer assumes position 1) … … watches **the child with the spear** (facing position 2).
Figure 10	

The sequence based on (20d) comes out in parallel, as shown in Figure 11. What we see in Figure 11 is the mirror image of what we saw in Figure 10. Here, the dancer first points at position 2 (after briefly reverting to the neutral position), and then moves into position 2, apparently in order to (re-)assume this viewpoint/perspective, orienting herself towards position 1.

	**The child with the spear** (pointing at position 2) … … (dancer assumes position 2) … … watches **the eating child** (facing position 1).
Figure 11	

The results from Study 2.2 and Study 2.4 were equivalent to those reported in Figure 9, Figure 10, and Figure 11, reproducing the same pattern, so we will not discuss them in detail here; however, we come back to them in section 3.2. Specifically, the mechanism in Figure 10 and Figure 11 is also utilized in the examples with three referents (Study 2.2).

2.6.3 Additional data points

To summarize the discussion in section 2.6.2, one crucial ingredient for disjoint reference seems to be the introduction of a new position/location in space, followed by the dancer's movement into that new location, as was illustrated in Figure 8.

The change of position that we observe in these dance sequences is plausibly connected to the way in which non-signers who narrate a conversation may use body posture in combination with direct quotation, a parallel that we will elaborate on in sections 3.2.1-3.2.3. A particularly striking example at the intersection of Bharatnatyam dance and speech-accompanying gestures can be found in a *TEDxOakParkWomen* presentation by Aishwarya Ravindran (henceforth: AR), which is described in (21). Here, AR is orally narrating an event while accompanying it with movements from Bharatanatyam (the topic of her presentation).

(21) *[AR faces towards the right of the stage]* "Mum", I would say, "can I join, too?"
[AR points towards the position of her sister, previously introduced earlier in the narrative]
[AR moves into the mother's position and changes orientation to face towards the left of the stage] My mother would patiently sit me down and say: "Your legs are too tiny. And your arms too weak. Wait. As you grow older and stronger, then you, too, can join your sister."
Video: https://youtu.be/EJ1G_tvk59Q?t=107

As the reader can verify in this short sequence, AR changes her position and orientation as sketched in Figure 12. While this is part of a speech-accompanying dance sequence, rather than a silent dance sequence, it mirrors exactly what we found in our exploratory study.

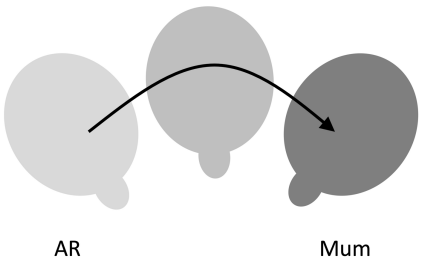


Figure 12: change of position and orientation in connection with a new referent on the stage https://youtu.be/EJ1G_tvk59Q?t=107

A natural question that emerges at this point is whether the change in position (here: from left to right on the stage) is more important, or the change in orientation (here: from facing leftward to facing rightward, from the dancer's perspective). A change of orientation without a change in position is schematically given in Figure 13.

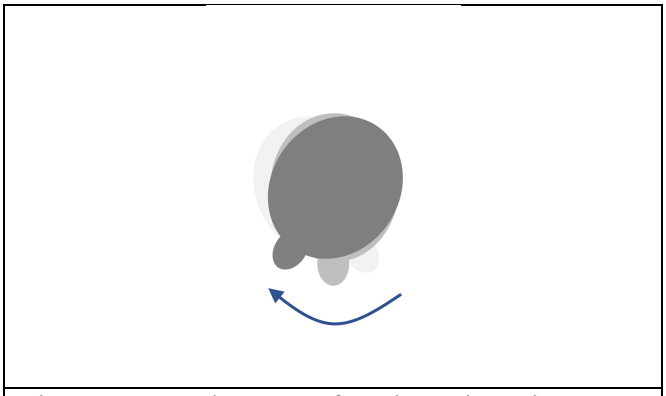


Figure 13: change of orientation in same position

There are good reasons to assume that the change in position is more important than the change in orientation. For one, changes of orientation occur frequently in our dance recordings in order to track the direction that one and the same character is facing (as in the examples with orientation towards another character that we discussed in section 2.6.2). Moreover, change in position can trigger disjoint reference even in the absence of a change in orientation. This becomes clear when AR concludes the above-mentioned presentation with a dance narrating the story of the goddess Devi (Durga) and the demon Mahisha (Mahishasura). To distinguish between Devi (at 10:05 of the video) and Mahisha (at 10:12), AR moves from a position in the center of the stage to one at the left of the stage: https://youtu.be/EJ1G_tvk59Q?t=605 She then returns to the original position when she picks up the Devi persona again (at 10:33). Crucially, while a change in position can mark a switch to a different discourse referent, it does not entail a switch to a different discourse referent. In the role of Mahisha, AR moves across the stage (10:20-10:24) and *through* the position that was previously associated with Devi – and which is subsequently once again associated with Devi. This does not trigger a referent switch. Notably, in this dance sequence, AR only changes position and does not change orientation, facing the audience in both locations. This is schematically given in Figure 14.

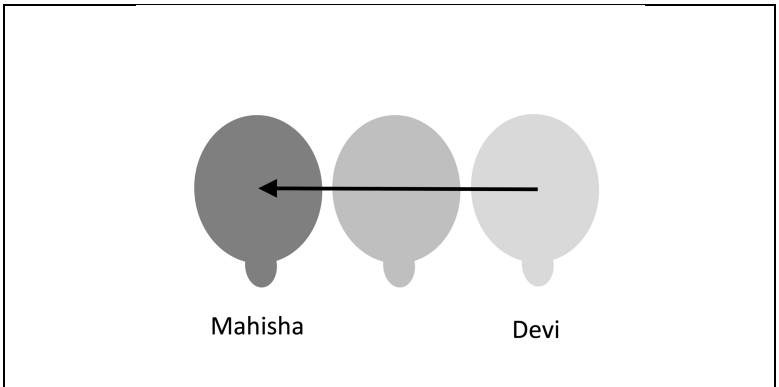


Figure 14: change of position without change of orientation in connection with a new referent on the stage https://youtu.be/EJ1G_tvk59Q?t=605

It should be pointed out that AR's hand gestures sufficiently disambiguate between the characters: She symbolizes Devi by holding her hands in front of the torso, palms facing the audience, fingers pointing upward on the right hand, downward on the left

hand. By contrast, Mahisha is shown by holding both hands up to the side of AR's head, to form horns with the index fingers. The change of position in Figure 14 is thus an accompaniment that is not strictly necessary, yet redundantly marks the two different referents.

We thus maintain that change of position (optionally combined with a change in orientation) is the operative cue for establishing disjoint reference in these cases. Moreover, both of these examples show that the decisive change is a change of position and not a change of orientation *per se*.

Another relevant data point can be found in a *TedxLondonBusinessSchool* presentation by Pancham Gajjar (henceforth PG), who dances 'the story of the tortoise and the hare' in Bharatanatyam. When introducing the two characters (03:51-04:16), PG does not introduce separate positions for them; the two characters are disambiguated only by virtue of the accompanying oral narrative (she describes the story while dancing it) and their mode of movement (fast vs. slow): https://youtu.be/Duzr01VZfLc?t=231 However, when PG describes a conversation between the two characters, we encounter, once again, the change in position and orientation seen in Figure 12 (06:46-07:07): https://youtu.be/Duzr01VZfLc?t=406

3. A super linguistic approach to narrative dance – formal theoretical analysis

3.1 Defining truth in visual narrative

One of the core tenets of Super Linguistics (see Patel-Grosz et al. forthcoming; see also Schlenker 2019a) holds that the formalisms of theoretical linguistics enable us to posit a precise analysis of meaning outside of natural language. This is the designated aim of section 3 of this paper, which will focus, in particular, on modelling the tracking of individuals (coreference *vs.* disjoint reference) and perspective taking in Bharatanatyam, based on the findings in section 2. To arrive at a formal semantic rendering of narrative dance sequences, an important first step consists in defining how we should approach the semantics of pictures, i.e., how we define truth in a visual narrative. Abusch (2013, 2014, 2021), building on Greenberg (2011, 2013), posits a generalized possible worlds model for informational entities, (22), based on the idea that any sentence, picture, etc., counts as an informational entity when it rules out some possibilities.

(22) *possible worlds model of information content* (Abusch 2021:2)
any informational entity, such as a sentence or picture, rules out some possibilities [= possible worlds, situations, or scenes] and admits others

Similar definitions have been given in other areas of super linguistic inquiry, such as the study of animal vocalizations, where Schlenker et al. (2017) provide the definition of meaning that we cite in (23).

(23) *a lean notion of meaning* (Schlenker et al. 2017)
"[T]he meaning of an expression [is assimilated] to the bi-partition it establishes among situations in which it is true vs. false."

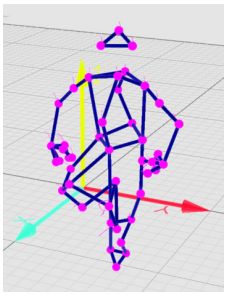
We can start by illustrating Abusch's idea for the dance position [P_{21}] in Figure 8, repeated in (25) below. Assume, for our purposes, that the world is populated by finitely many undistinguishable persons and nothing else. In such a scenario, if I say

"There is a person who is sitting.", I rule out a range of possible scenarios (in line with (22)), namely ones in which there is no person, or in which there is a person who is doing something that does not look like sitting. The statement in (24) is thus understood to provide new information about a given situation that we are describing.

(24) There is a person who is sitting.

Crucially, Abusch argues that a picture achieves exactly the same result. In parallel to (24), the dance position in (25) can be understood to provide new information about a given situation (namely the current point in time in a narrative that is being told).[23] As Abusch observes, when it comes to the question of what a world or situation is like, (25) rules out possibilities in which no sitting activity takes place, while ruling in possibilities in which a sitting activity takes place. The dance position in (25) thus qualifies as an informational entity in line with (22). Abusch is careful to point out that pictures are often more informative than sentences; taken at face value, a naïve observer may infer from (25) that (in addition to being in a sitting position) the person in the narrative has one leg straight and one leg at an angle. (Of course, this may simply be part of a conventionalized gesture for 'sitting'.) Sentences like (24) can leave such information underspecified; there is no implication from (24) on how exactly the person is sitting.

(25)



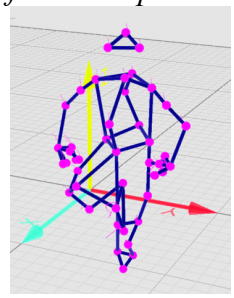
Refining the approach of Greenberg (2011, 2013), Abusch (2021) proceeds to identify the semantics of a picture with the set of possibilities that it admits. This means that we can define the semantics of a picture in terms of possible worlds, situations, or scenes. Treating any given dance position [P_n] as a picture, we can then posit satisfaction conditions as given in (26).[24] Truth in visual narrative is thus defined in terms of how well a dance position [P_n] maps to a situation/scene σ_n in the narrative; i.e. the dance position in (26) counts as satisfied by a situation σ (i.e. "true" in σ) if a sitting activity is taking place in σ. Note that Bharatanatyam can encode meaning in a highly iconic way; the meaning of the dance position in (26) is easily inferable, even

[23] Note that the dancer remains static in this 'sitting' position for an average of 112 frames / 560ms across the two conditions, i.e., she is not constantly moving. (Specifically, she remains in this position for 145 frames / 725ms in the coreferent condition, and 79 frames / 395ms in the disjoint referent condition, but this seems to be a coincidental difference, as the conditions have not yet diverged at this point.) This is relevant for our discussion, in connection with our methodological choice of 'transforming' dance sequences into stills that more closely resemble comics; given that predicates ('sitting on the ground', 'holding a spear', etc.) in the Bharatanatyam dance sequences that we recorded are typically expressed by virtue of the dancer assuming static positions, it is justifiable to base the analysis of the dance sequences on such static postures rather than on the transitional movements between the static postures.

[24] This is glossing over the fact (as discussed by Abusch and Greenberg) that pictures are generally related to the depicted objects by means of projection lines that are oriented towards a given viewpoint.

to an onlooker unfamiliar with Bharatanatyam, based on resemblance between the dance posture and the sitting activity that is being described.[25]

(26) *satisfaction conditions for dance position that describes a sitting activity*

a situation σ satisfies  only if in σ a person is sitting.

For present purposes, we simplify in two respects: by analyzing a dance sequence in the form of still shots, as in (25), we abstract away from both the continuity of dance and the three-dimensionality, essentially transforming the dance sequence into a two-dimensional cartoon. This simplification is warranted as it allows us to directly apply the approach of Abusch and Greenberg without first incorporating continuous movement and a third dimension.

3.2 Analyzing Dance Semantics with Tools from Gesture Semantics

While Bharatanatyam dance can be analyzed with the tools devised for the analysis of visual narrative, it is a special type of visual narrative: by using the visual-gestural modality, it has commonalities with speech accompanying gestures, and also with sign languages.[26] Since dance and gesture share the property of being distinct from natural language, as opposed to sign languages, which are full-fledged natural languages, we maintain that gestures provide a better reference point for narrative dance than sign languages. At the same time, of course, all three (sign language, gestures and dance) plausibly build on the same cognitive underpinnings with regards to the three phenomena that we discuss in sections 3.2.1-3.2.3.

In what follows, we proceed to argue that tools and terminology inspired by sign-language research can enlighten our understanding of meaning making in Bharatanatyam. Specifically, three phenomena from sign language, which have received much coverage in recent years, are relevant for our discussion: (i.) *loci* (see, e.g., Lillo-Martin & Klima 1990, Liddell 1990, Sandler & Lillo-Martin 2006, Schlenker 2017b), (ii.) *agreeing verbs*, also known as *agreement verbs* or *directional verbs* (see, e.g., Padden 1988, Kegl 2004, Liddell 2000, 2003, Schlenker & Chemla 2018), and (iii.) *action role shift*, also known as *action report* or *constructed action*

[25] Greenberg (2021:slide 48) defines an *iconic semantics* (which he contrasts with *symbolic semantics*) as one where lexical entries are "rule-like" and "sign-dependent", i.e. stated in a way where the expression in denotation brackets (the *sign*) also occurs in the denotation (to the right of the equals sign); in this vein, we could approximate the iconic meaning of a dance posture as in (i.), of which (26) would then be an instantiation. The sign-dependence and rule-like nature of (i.) is captured by the fact that the sign P_n occurs both on the left and on the right of the equals sign.

i. $[\![\mathbf{P_n}]\!] = \{s \mid$ there is an eventuality e in s such that e *looks like* $\mathbf{P_n}\}$

[26] Of course, we do not aim to imply that sign languages are anything less than full-fledged languages, or that dance is close to a full-fledged language; the question rather relates to the very atoms of meaning that are shared by humans, possibly innate, and which can be encoded by means of body movements that are recruited both in dance and in full-fledged (signed) languages.

(see, e.g., Padden 1986, Lillo-Martin 1995, Quer 2005, Sandler & Lillo-Martin 2006, Herrmann & Steinbach 2009, 2012, Davidson 2015).

To be entirely clear, we do *not* argue that such aspects of sign language (*loci*, *agreeing verbs*, *role shift*) are found in dance. We propose that dance utilizes mechanisms provided by general cognition for the tracking of individuals in a narrative, as well as perspective taking with regards to such individuals. The connection to sign language are two-fold: (i) formalisms that have been developed for the analysis of the linguistic mechanisms in sign language can also be applied to the non-linguistic mechanisms in dance; (ii) the linguistic mechanisms at work in sign language and the non-linguistic mechanisms in dance may share a common cognitive foundation that humans activate when using the body to communicate in visual space.

In order to command a maximally neutral terminology, we introduce the following two concepts.

For the positions in space that dancers can utilize in order to track individuals in a narrative, as discussed in sections 2.4-2.6, we introduce the term *indexical base* (which we shorten to *base*); the notion of 'base' is mnemonic and captures the visual appearance (e.g., in Figure 9) that the dancer moves from one 'base' to another when portraying different characters in the narrative. Our notion of *base* is inspired by the sign-language notion of *locus*, but refers to a non-linguistic mechanism for reference tracking. We will also discuss dance sequences in which the dancer points from one base to another points, thus establishing a *base-to-base linking* (see Figure 10 and Figure 11); our formal analysis of *base-to-base linking* is inspired by the analysis of so-called *agreeing verbs* in sign languages, once again without arguing in any way for an identification of the two mechanisms.

For the perspective-taking mechanism where a dancer moves into a position (base) associated with a given referent in order to perform that referent's actions, we introduce the transparent term *action-performance*. Here, too, our notion of *action-performance* is inspired by the sign-language notion of *role shift*, but refers to a non-linguistic mechanism for demonstrating the actions of an individual.

In sections 3.2.1-3.2.3, we review our empirical findings from section 2 in light *bases* and *action-performances*. As we do so, we adapt formal notation that was originally introduced in sign language research, and which has already been productively applied to the gestures of non-signers, e.g., by Schlenker & Chemla (2018) and Schlenker (2020), among others. The benefits of providing an analysis with formal tools is that it introduces precision to the analysis of how coreference and disjoint reference are communicated in narrative dance, which future research can build on.

3.2.1 Reference tracking through indexical bases

In a performance of narrative dance (such as Bharatanatyam), *indexical bases* are positions in space that a dancer can use to track referents by virtue of body movement in visual space. In previous research (Schlenker & Chemla 2018, Schlenker 2020), similar mechanisms were found in the speech-accompanying gestures of non-signers, for which the term *gestural loci* was coined. (We remain agnostic as to whether *gestural loci* are a linguistic phenomenon, similar to sign language loci, or a non-linguistic phenomenon, similar to our indexical bases.) In what follows, we adapt the formal analysis of gestural loci for our formal analysis of indexical bases.

A representative example of gestural loci is cited in (27). Here, a co-speech gestures like IX-hand-a involves an open hand, palm up, associated with a position

(a/b/c) in gesturing space, accompanying the bracketed constituent in bold type (e.g. *John*). The speaker thus introduces three positions in signing space, which are associated with the three referents that are mentioned. These three positions are classified as *gestural loci*. In (27a-e), the speaker uses a pro-speech (speech-replacing) gesture, pointing at one of these three loci (or at the speaker/hearer in the case of IX-1/ IX-2) in order to retrieve the respective discourse referent.

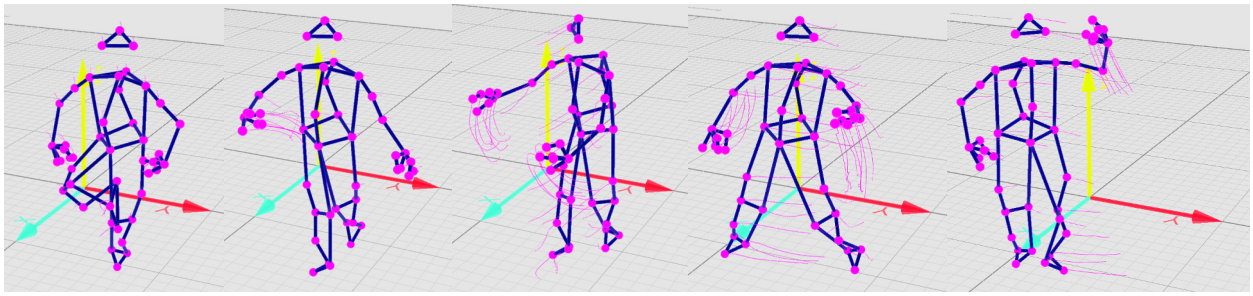
(27) Yesterday, I had a long conversation with IX-hand-a **[John]**,
Then with IX-hand-b **[Mary]**, then with IX-hand-c **[Sam]**.
You know who the company's gonna promote?
a. IX-a. b. IX-b. c. IX-c. d. IX-1. e. IX-2.
= John = Mary = Sam = me = you
(Schlenker 2020:897, Video 3845 https://youtu.be/KLpow-YBNRs)

We observed the exact same mechanism in our Bharatanatyam dance sequences. Our qualitative analysis in Figure 9–Figure 11 of section 2.6.2 is summarized in (28), by virtue of the addition of the *indexical bases* BAS-a and BAS-b; we wish to emphasize that (28) is an informal representation of the dance sequence, since all of (28a-d) is narrated by virtue of dance (i.e. body movement) without any accompanying speech or sign. The bold type and bracketing, building on (27), indicates which referent is associated with a given base. In (28a-b), the dancer establishes bases for the two children (BAS-a and BAS-b), which she then retrieves in (28c-d).[27] This analysis is still incomplete in that we have not yet incorporated the directionality of the predicate (*watch*); we elaborate on this point in section 3.2.2.

(28) *Study 2.1 – Item 1 (preliminary analysis 1, to be revised)*
a. The artist sees **BAS-a [a child]** eating a mango outside the temple.
b. Then she sees **BAS-b [another child]** holding a spear.
c. *Cond. 1 (parallel):*
BAS-a [The eating child] watches the child with the spear.
d. *Cond. 2 (switched):*
BAS-b [The child with the spear] watches the eating child.

We can now revisit our initial example of disjoint reference in Figure 15, repeated from Figure 4. If we now focus on [P_{23}], we may incorporate *indexical bases* and posit satisfaction conditions such as (29), from the perspective of Abusch's (2013, 2014, 2021) picture semantics. (We return to [P_{24}] and [P_{25}] in section 3.2.3.)

[27] An anonymous reviewer raises the question of whether it is legitimate to apply an analytical notion from sign language (here: a formalism originally introduced to handle *loci* and *gestural loci*) to account for our finding on dance. In the pre-theoretic part, section 2.6, we speak of 'positions in space' or 'positions on stage', and we now proceed to argue that such 'indexical bases' reflect the use of general cognitive strategies for tracking referents in a narrative. While we do not wish to argue that *Bharatanatyam* is a language, or even language-*like*, the interesting question that emerges is whether the grammatical *loci* in sign language and the positions on stage found in dance share cognitive underpinnings, the exploration of which would further our understanding of human cognition more generally. We do not, at any point, intend to argue that the *indexical bases* of *Bharatanatyam* should be treated as equivalents of *sign language loci*, or even *gestural loci*.



[P_{21}] [P_{22}] [P_{23}] [P_{24}] [P_{25}]

Figure 15: disjoint reference condition

In (29), *a* is an indexical base, and $g_c(a)$ is an individual that is associated with the base *a* in context *c* by virtue of an assignment function g_c.

(29) For any assignment function g_c in context *c* and indexical base *a*,
σ_{23} satisfies [P_{23}] only if in σ_{23} there is an individual *x* in the narrative such that $x = g_c(a)$.

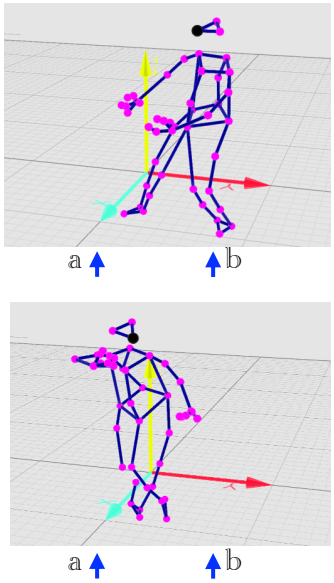
One may well ask whether the *indexical bases* of Bharatanatyam dance are phenomenologically distinct from the *gestural loci* that were proposed in the literature. It is not clear that such a question is motivated to begin with; it seems rather plausible that reference tracking by virtue of positions in space is a communicative tool that is broadly made available by human cognition as soon as the visual-gestural modality is involved, and can then be recruited for diverse communicative modes, including gesture, dance, silent comics, and sign languages.

3.2.2 Base-to-base linking in dance

Having introduced the notion of indexical bases, we now introduce *base-to-base linking*. This is a phenomenon where a dancer moves into the base associated with an individual *a*, and orients herself towards the base associated with another individual *b*, in order to convey that *a* performs an action towards *b* (e.g., *a watches b*). Once again, a similar phenomenon has been documented in the speech-accompanying gestures of non-signers (Schlenker & Chemla 2018, Schlenker 2020). In (30), PUNCH-a is a pro-speech (speech-replacing) gesture, which consists of a punching movement that is directed towards the gestural locus (a) associated with the addressee's brother. By contrast, SLAP-b is a pro-speech slapping gesture directed towards the gestural locus (b) associated with the addressee's sister.

(30) When I was a kid, I often got into fights with IX-hand-a **[your brother]**,
But also with IX-hand-b **[your sister]**. One morning,
Your brother, I tried to PUNCH-a, and then your sister, I tried to SLAP-b.
(Schlenker 2020:900, Video 3905 https://youtu.be/BVSuyFsuj4o)

If we consider Figure 10, repeated in Figure 16 with different annotation in line with (30), we notice that the indexical bases of Bharatanatyam dance can be employed to the same communicative effect. Here, directionality towards the base associated with the object referent is marked by facing in that direction.

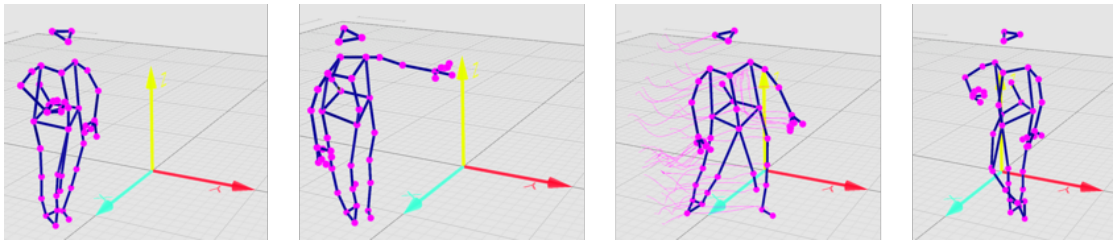
	**BAS-a [The eating child]** (pointing at position a) … … (dancer assumes position a) … … watches-**b** the child with the spear (facing position b).
Figure 16	

We can thus revise our analysis in (28) as given in (31), where *watches* has been replaced by *watches{-b/-a}* to indicate the directionality of the dancer's gaze and orientation.

(31) *Study 2.1 – Item 1 (preliminary analysis 2, to be revised)*
- a. The artist sees **BAS-a [a child]** eating a mango] outside the temple.
- b. Then she sees **BAS-b [another child]** holding a spear].
- c. *Cond. 1 (parallel):*
  BAS-a [The eating child] watches-**b** the child with the spear.
- d. *Cond. 2 (switched):*
  BAS-b [The child with the spear] watches-**a** the eating child.

3.2.3 Performance within performance (in dance)

A final generalization that cuts across all of the dance sequences that we collected involves the dancer moving into the position (base) of a referent after introducing that position. This was illustrated by the sequence in Figure 8, repeated as Figure 17 below. Here, in line with the discussion in section 3.2.1, a base is introduced by virtue of pointing in the second picture; the movement into this base is shown in the third and fourth picture. It is natural to ask whether something similar is found in gesture – and sign language – and we can answer in the affirmative.



A man … and … [new referent] [assume location] … a woman …

Figure 17: disjoint reference without 'another'

Schlenker (2020:926) demonstrates that a highly similar phenomenon is found in gestures; adopting terminology from sign languages, (32b) is analyzed as an instance of gestural *action role shift* (or gestural action report), whereas (32a) is a case of gestural *attitude role shift* – a reportative/quotative device for conveying (in this example) what Robin and Francis said or thought. In sign languages, the phenomenon of action role shift involves markers such as a change of eye gaze, upper body posture, or head orientation, associated with the agent of a narrative, whom the signer temporarily embodies in order to vividly describe their actions. The phenomenon has been compared to non-signers who imitate the pitch and intonation of someone that they are quoting (see, e.g., Davidson 2015).

(32) (*Notation:* RS_i indicates that the speaker shifts his body to adopt the position of a fictional character found in gestural locus i (here we will have $i = a$ or $i = b$). The gesture that follows RS_i is realized from this shifted position.)

I was standing next to IX-hand-a **[little Robin]** and IX-hand-b **[little Francis]**, and I was holding a really yoummy chocolate bar. And I asked: Who wants it?

a. And so of course, IX-hand-a **[little Robin]** goes: RS_a IX-1.
   And IX-hand-b **[little Francis]** goes: RS_b IX-1.
b. Next thing I know, IX-hand-a **[little Robin]** turns to IX-hand-b **[Francis]** and RS_a SLAP-b. And so IX-hand-b **[Francis]** RS_b SLAP-a.

(Schlenker 2020:925-926, Video 4053 https://youtu.be/r0dhqgQk2k0)

The action role shift of sign languages is a grammatical mechanism, and there is no reason to assume that such a mechanism exists in narrative dance. Yet, the phenomenon that we observed in Figure 17, and, even more vividly, Figure 16, while non-linguistic in nature, is strikingly similar, allowing us to pursue a parallel formal analysis. We introduce the neutral descriptive term *action-performance*, and provide the final analysis in (33) for Figure 16; here, we have added PER_a and PER_b to symbolize the part of the dance sequence where the dancer assumes the position of the referent in order to *perform* the referent's actions. Note that the dancer also assumes the relevant positions in (33a) and (33b), not only in (33c-d).

(33) *Study 2.1 – Item 1 (final analysis)*

a. The artist sees **BAS-a [a child]** $\mathbf{PER_a}$ eating a mango outside the temple.
b. Then she sees **BAS-b [another child]** $\mathbf{PER_b}$ holding a spear.
c. *Cond. 1 (parallel):*
   BAS-a [The eating child] $\mathbf{PER_a}$ watches-**b** the child with the spear.
d. *Cond. 2 (switched):*
   BAS-b [The child with the spear] $\mathbf{PER_b}$ watches-**a** the eating child.

To conclude, we have found that Bharatantyam dance can incorporate mechanisms of reference tracking and perspective taking that have parallels in both gesture and sign language, and conceivably also in other visual modes of human expression. We maintain that these mechanisms are most plausibly based on general human cognition, which is why they occur both as non-linguistic mechanisms (in dance) and as linguistic mechanisms (in sign language).

Study 2.2, discussed in section 2.6.1 above, tested what would happen with more than two possible referents, namely with the three referents *man, woman* and *child*.

The findings are summarized in (34), and explicitly replicate the findings from Study 2.1, as given in (33). The dancer introduces three separate positions on the stage, which we can, once again, think of as *indexical bases*: *a* for the woman, *b* for the man, and *c* for the child. In terms of position on the stage, these bases roughly correspond to the three positions discussed in Figure 9: base *a* in (34) corresponds to the neutral position in Figure 9, base *b* to position 1, and base *c* to position 2. In the final (transitive) clause of the various conditions in (34a-f), the dancer moves into the base of the respective subject (i.e., *the man* in (34a), *the child* in (34b), and so forth) and faces the base of the respective object, in parallel to what we saw above for Study 2.1. Once again, we can model this as a combination of *action-performance* (e.g. ***PER$_b$***) with *base-to-base linking* (e.g. *looking-**c***).

(34) *Study 2.2*

Shared by all conditions:
BAS-a [A woman] PER$_a$ is standing outside the temple, **BAS-b [a man] PER$_b$** is sitting on the ground, and **BAS-c [a child] PER$_c$** is playing.

a. **BAS-a [The woman] PER$_a$** is holding a book. **BAS-b [The man] PER$_b$** is looking-**c** at the child.

b. **BAS-a [The woman] PER$_a$** is holding a book. **BAS-c [The child] PER$_c$** is looking-**b** at the man.

c. **BAS-b [The man] PER$_b$** is holding a book. **BAS-a [The woman] PER$_a$** is looking-**c** at the child.

d. **BAS-b [The man] PER$_b$** is holding a book. **BAS-c [The child] PER$_c$** is looking-**a** at the woman.

e. **BAS-c [The child] PER$_c$** is holding a book. **BAS-a [The woman] PER$_a$** is looking-**b** at the man.

f. **BAS-c [The child] PER$_c$** is holding a book. **BAS-b [The man] PER$_b$** is looking-**a** at the woman.

Since *moving into a base in order to perform a referent's actions* (PER$_i$) and *pointing at a base in order to first establish a referent* (BAS-i) seem to have the same function in sequences such as (33)-(34) (namely identifying a referent in the narrative), we expect that either one of them may be sufficient for marking a referent in narrative dance, and that they do not always have to co-occur. Specifically, we can ask whether there are cases where the dancer moves into a base without first pointing at the base. As a matter of fact, this is attested when the *first* referent is introduced into the narrative. In (33a-b), we collapsed the two conditions at the beginning, and did not analyze them separately. In (35a) and (35b), we revisit them separately; we have separated the three positions (neutral position, position 1 and position 2), and added the symbol '☞' to every pointing gesture that we observed in the dance sequence; while the dancer points at position 1 (in bold type) in dance sequence (35b), this pointing gesture is missing in dance sequence (35a). Since the dance sequences have not yet diverged at this point of the narrative, we take this initial pointing to be optional and in free variation. (In fact, the presence of the first pointing gesture in (35b) may be an artifact of the dancer's choice to introduce the temple before introducing the first child in (35b), but not in (35a); if position 1 is introduced after

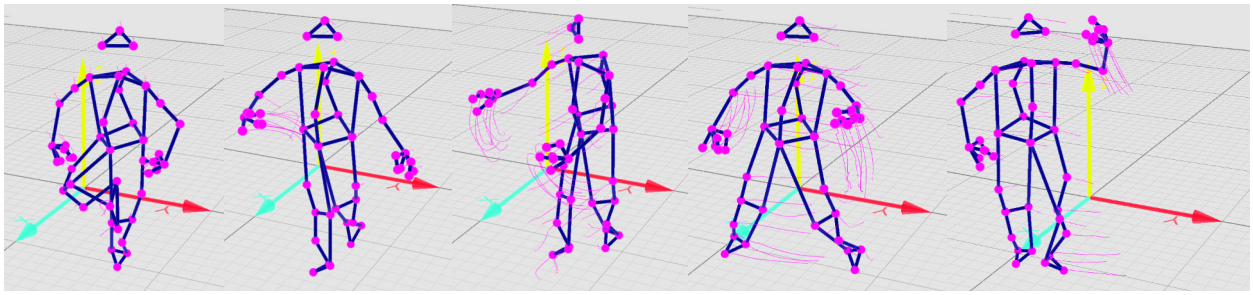
introducing the temple referent, this would indicate that position 1 is, strictly speaking, not the first referent to be introduced.)

(35) a. The artist$_{[neutral]}$ sees a child$_{[pos\ 1]}$ eating a mango outside the temple.
Then she sees another ☞ child$_{[pos\ 2]}$ holding a spear.
The eating ☞ child$_{[pos\ 1]}$ watches *the ☞ child with the spear*$_{[pos\ 2]}$.

b. The artist$_{[neutral]}$ sees a ☛ **child$_{[pos\ 1]}$** eating a mango outside the temple.
Then she sees another ☞ child$_{[pos\ 2]}$ holding a spear.
The ☞ child with the spear$_{[pos\ 2]}$ watches *the eating ☞ child*$_{[pos\ 1]}$.

It is also worth pointing out that the dance videos from *Tedx* talks that we discussed in section 2.6.3 utilize explicit pointing much less than the dancer in our Study 1 and Study 2, relying more heavily on the movement into a referent's position, which we have now classified as *action-performance*.

We can conclude this section by briefly revisiting our discussion of Figure 15, which we repeat in Figure 18. Previously, we posited an analysis of [P$_{23}$] in (29), which is repeated in (36a). We can now ask about the nature of [P$_{24}$] and [P$_{25}$].



[P$_{21}$] [P$_{22}$] [P$_{23}$] [P$_{24}$] [P$_{25}$]

Figure 18: disjoint reference condition

We have analyzed [P$_{24}$] as an instantiation of *dance action role shift*. One somewhat simplistic way of modeling this in a formal semantics is given in (36b). If we add the analysis of [P$_{25}$] in (36c), there is only one missing step, namely the identification of the individual x introduced in (36a) with the individual y that is the agent of the eventuality in (36c).

(36) a. For any assignment function g_c in context c and indexical base a,
σ_{23} satisfies [P$_{23}$] only if in σ_{23} there is an individual x in the narrative such that $x = g_c(a)$.
b. σ_{24} satisfies [P$_{24}$] only if in σ_{24} the perspectival center of the narrative is *at* $g_c(a)$.
c. σ_{25} satisfies [P$_{25}$] only if in σ_{25} an individual y is holding a spear.

To capture this identification, we can pose the tentative generalization in (37), based on which we infer that $y = g_c(a) = x$ from the combination of (36a), (36b) and (36c).

(37) *Pragmatics of agent identification in Bharatanatyam*
The agent of eventualities in Bharatanatyam narratives is identified with the perspectival center.

Note that the analysis in (36) (and also (29)), which draws on concepts such as assignment functions, and connects indexical bases to individual variables (here: *a*), may seem to be a *linguistic* analysis of a *non-linguistic* phenomenon. We do not consider this a problem, since earlier work in Super Semantics has argued that variables may be the one analytical tool that has applications across a range of non-linguistic phenomena, such as silent comics (Abusch 2013, 2014, 2021) and even music (Schlenker 2019a, 2021). The analytical use of variables does not entail he presence of complex linguistic structures or of linguistic phenomena such as, say, quantifier binding, neither of which is likely to be found in dance, silent comics, or music.

To conclude our discussion of action-performance, it is worth exploring an alternative approach, in section 3.2.4, which may qualify as more general than the approach in (36), and thus potentially an even more suitable candidate for the analysis of non-linguistic phenomena.

3.2.4 Modeling action-performance in dance as demonstration

Developing the idea of *action-performance* (section 3.2.3) further, an alternative to (36) and (37) could be based on Davidson (2015). Building on Clark & Gerrig (1990), Davidson argues that spoken language quotation involves an iconic *demonstration* (or *performance*), an analysis that she applies to action role shift in sign language. Examples from spoken language are given in (38), where the quotation includes a particular intonation pattern (38a), or an upset facial expression (38b), or where the quotation only consists of a speech-replacing gesture without accompanying words (38c). Davidson's idea is that the speaker of (38a-c) demonstrates/performs Bob's and Mary's behavior during the original event, rather than just verbatim repeating the content of their utterances.

(38) a. Bob saw the spider and was like "ahh! [in a scared voice]."

b. **:-/**_____
I saw Mary studying for finals and she was all "I'll never be prepared"

c. Bob was like [gobbling gesture].

(quoted from Davidson 2015:485,489)

To formally model demonstration/performance, Davidson introduces a demonstration type *d*, for which she gives the definition in (39).

(39) Definition: a demonstration *d* is a demonstration of [an event] *e*
(i.e. demonstration(*d, e*) holds) if *d* reproduces properties of *e* and those properties are relevant in the context of speech.
(Davidson 2015:487)

Glossing over the compositional steps (e.g. the semantics of *(be) like*), which the reader can find in Davidson (2015), she proposes the analysis in (40b-c) for the 'plain' quotation example in (40a). In words, there is an event of which John is/was the agent, and which the speaker performs/demonstrates by virtue of performing the demonstration d_1, which amounts to a demonstration of "I'm happy" that can be enriched by intonation, gestures, and other extralinguistic features.

(40) a. John was like "I'm happy"
b. ⟦"I'm happy"⟧ = d_1 (a particular demonstration involving two words and perhaps other intonation, gestures, etc.)
c. ⟦John was like "I'm happy"⟧ = ∃*e*. [agent(*e,* John) ∧ demonstration(d_1, *e*)] (Davidson 2015:487, slightly adapted)

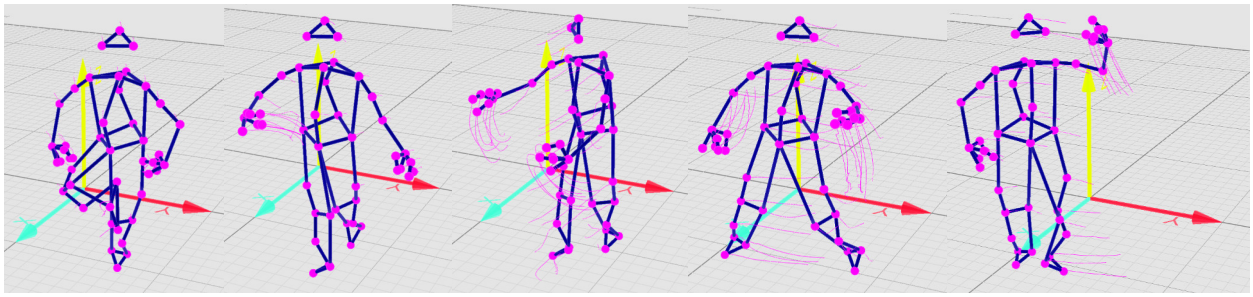
Davidson's analysis of the example in (41a) is transparently entirely parallel, as shown by (41b-c), i.e., the only difference is the nature of d_1, with all else being equal.

(41) a. John was like [gobbling gesture].
b. ⟦gobbling gesture⟧ = d_1 (a particular demonstration involving a gobbling gesture)
c. ⟦John was like [gobbling gesture]⟧
= ∃*e*. [agent(*e,* John) ∧ demonstration(d_1, *e*)]
(Davidson 2015:488, shortened and slightly adapted)

Davidson (2015) proceeds by applying such an analysis to action reports (i.e. action role shift) in sign language, an important aspect of which she captures by the following quote (p. 508): "in reported action [i.e. action role shift], one becomes the other actor because one wants to show *aspects of the behavior of the actor in the speaking event*".

Crucially, what we have found in our dance sequences is that such iconic demonstration also occurs in Bharatanatyam, where we labeled it *action-performance*. This is striking for the reason that the meaning of the *entire* dance sequence operates on an iconic semantics, i.e. we are dealing with a narrative already driven by iconic semantics, which in turn embeds iconic demonstrations whenever the dancer assumes one character or the other.

An analysis of Figure 15, repeated in Figure 19 can thus also be fleshed out in the spirit of Davidson's demonstration-based approach.



[P_{21}] [P_{22}] [P_{23}] [P_{24}] [P_{25}]

Figure 19: disjoint reference condition

A Davidson-style analysis of the action-performance in Figure 19 is given in (43)-(44). To explain this analysis, we can start with the spoken language example (42), which differs from (41) in the addition of the predicate *eating*. Here, the phrase *like [gobbling gesture]* is a case of an iconic demonstrational event modification, in that it modifies the eating event denoted by *John was eating*. The analysis (42c) differs from (41c) in that (42c) includes an eating event, enriched by virtue of a demonstration.

(42) a. John was eating like [gobbling gesture].

b. ⟦gobbling gesture⟧ = d_1 (a particular demonstration involving a gobbling gesture)

c. ⟦John was eating like [gobbling gesture]⟧
= ∃*e*. [agent(*e,* John) ∧ eating(*e*) ∧ demonstration(d_1, *e*)]

(Davidson 2015:488, shortened and slightly adapted)

Reconsider the original narrative that Figure 19 is based on, repeated in (43) (from (9)). A demonstration-based approach to the change in position and orientation that we encounter in [P_{24}] and [P_{25}] is sketched in (44c-d). In full parallel to (42c), we posit that (44a-b) includes a sitting event enriched by a demonstration of the first (sitting) man's actions, while (44c-d) includes a spear-holding event, enriched by a demonstration of the second (spear-holding) man's actions. The role of d_2 in (44c-d) is to essentially capture the demonstration/performance-aspect connected to the change in position and orientation that we observe (i.e., the dancer "becoming" the second man for the duration of the action-performance). The pervasiveness of demonstration/performance in narrative dance is captured by the idea that we move from one instance of action-performance, (44a-b), into another, (44c-d), separated by parts of the narrative that appear to be unshifted in [P_{22}] and [P_{23}].

(43) The artist sees a strong man [$_{P21}$ sitting on the ground].
Then she sees that [$_{P22+P23+P24}$ *another man*] [$_{P25}$ is holding a spear].

(44) a. ⟦appropriate position, orientation and movement⟧ = d_1 (a particular demonstration that reflects an imitation/reenactment of the **first** man's actions)

b. ⟦P_{21} ⟧ = ∃ *e*. [agent(*e,* man) ∧ sitting(*e*) ∧ demonstration(d_1, *e*)]

c. ⟦appropriate position, orientation and movement⟧ = d_2 (a particular demonstration that reflects an imitation/reenactment of the **second** man's actions)

d. ⟦P_{24}+P_{25}⟧ = ∃ *e*. [agent(*e,* man) ∧ theme(*e*, spear) ∧ holding(*e*) ∧ demonstration(d_2, *e*)]

Note that a full-fledged analysis of (43), in line with the discussion in section 3.2.3, would also reflect the bases (BAS) and action-performances (PER), as given in (45). If we were to follow Davidson to the letter, $PER_{a/b}$ in (45) could be analyzed as a lexical item that has the same meaning as *like* in the spoken English example (42). We do not pursue such an analysis here, as is would introduce linguistic components (namely a conventionalized *like*-meaning) into the analysis of dance that are, as of now, unwarranted.

(45) The artist sees **BAS-a [a strong man] PER_a** sitting on the ground.
Then she sees that **BAS-b [another man] PER_b** is holding a spear.

An interesting question that we leave open for future research concerns the relationship of specific individuated demonstrations, (44a) and (44c), and the entire dance sequence (e.g., Figure 19), which is a performance/demonstration in its own right. If we were to consider the entire dance sequence to be of Davidson's type *d*,

then this would require demonstrations to have internal complexity, and to be able to recursively contain smaller demonstrations.

3.2.5 How arbitrary are indexical bases?

An open question, which at this point cannot be addressed due to the limitations of our exploratory studies (as reported in section 2), but should be addressed in future research, is whether the positions in space that we have identified as *indexical bases* share other properties of the reference-tracking positions in visual space that have been observed in sign language and in the gestures of non-signers. For instance, sign-language *loci* do not need to mark the actual positions of objects in physical space, but can define arbitrary positions. The question is thus whether audience members who pick up on the *indexical bases* of a dance performance identify them with the actual (literal) positions of entities or not. Arbitrariness in the position of *bases* would be a property that they share with the grammatical *loci* of sign language; an attestation of arbitrariness would thus allow us to tease apart abstract reference-tracking uses of indexical bases from purely iconic uses in which the dancer iconically incorporates (a 'moving picture' of) the relevant character/ referent. The relevant hypotheses are stated in (46).

(46) a. H_1 = relative stage positions in a narrative dance are isomorphic to the relative positions of respective characters/referents/individuals in the space of a described situation. The positions are thus non-arbitrary and iconically represent the positions of characters in the narrative.

b. H_2 = positions on the stage in a narrative dance are arbitrary, without an implication that they correspond to the relative positions of characters in the described situation. The position thus function as abstract reference-tracking devices.

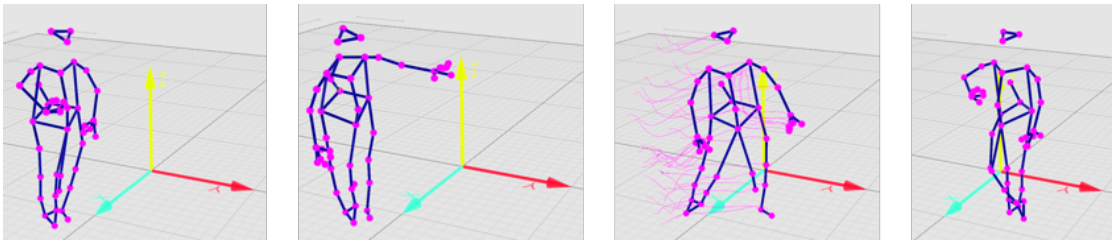
Initial evidence for H_2 stems from the dance performance of Aishwarya Ravindran (AR), which we discussed in section 2.6.3. As mentioned above, we can identify a base *a* for the goddess Devi (at 10:05 of the video) and a base *b* for the demon Mahisha (at 10:12). (See https://youtu.be/EJ1G_tvk59Q?t=605) While representing Mahisha, AR moves across the stage (10:20-10:24) and passes through the Devi-associated base position – which is subsequently once again associated with Devi. Such an overlap in the movement range would seem to be incompatible with H_1, as one referent (Mahisha) should not be able to move *through* the position of the other referent (Devi) if their bases were to iconically represent their positions in the narrative.

Regardless of the eventual decision between (46a) and (46b), the very fact that indexical bases are employed in narrative dance opens avenues of future investigation for the semantics of dance. We can conclude our analysis section by highlighting that our investigation of disjoint reference *vs*. coreference marking in Bharatanatyam dance has shed new light on the question of how reference tracking and perspective taking can be communicated by the human body in visual space.

4. An alternative analysis on the basis of iconic discontinuity inferences

In section 3, we argued that narrative dance can employ communicative mechanisms such as indexical bases, base-to-base linking, or action-performance. A reader may

wonder if the introduction of such concepts and notions to the analysis of dance is entirely justified given the empirical basis that we are working with. One question that arises for the most basic example, repeated in Figure 20 from Figure 17, is the following: Couldn't this just be accounted for in terms of visual iconicity?[28] The relevant logic would be as follows: the change of position and orientation of the dancer triggers a strong iconic discontinuity inference, which is interpreted as a 'break' in the narrative that signals disjoint reference, i.e., the introduction of a separate referent.



A man … and … [new referent] [assume location] … a woman …

Figure 20: disjoint reference without 'another'

It is worth exploring such an analysis for the examples with two referents, discussed in sections 2.4 and 2.5. However, already the possible encoding of three referents, which we discussed in section 3.2.3 (example (34)) does not lend itself to an analysis in terms of mere visual discontinuity.

To lay out the hypothesis space, the key theoretical question for the cases with two referents amounts to whether the change of position on part of the dancer is purely an iconic marker of discontinuity, or also introduces indexical bases, as fleshed out in section 3.2. We state three competing hypotheses in (47); our analysis in section 3.2 is compatible with both H_2 (*bases* are necessary for the management of discourse referents in dance) and H_3 (*bases* are sufficient for the management of discourse referents in dance). The newly added hypothesis H_1 aims to do away with bases altogether, and rely on iconic discontinuity alone. This is the hypothesis we explore in the remainder of section 4. We will conclude that H_1 is not a viable option, and that *bases* are, in fact, relevant for theorizing about dance semantics, in line with H_2 and H_3. We leave open which of H_2 and H_3 is correct, as the presence of bases entails the presence of visual discontinuity, which makes it difficult to tease the two apart on the basis of the data that we have collected. Note also that H_3 would not deny the existence of visual discontinuity inferences, but it would maintain that such inferences have no bearing on coreference *vs*. disjoint reference.

(47) To account for the change of position on part of the dancer when managing distinct discourse referents …

a. H_1 = … visual discontinuity inferences (a type of visual iconicity) are sufficient, and no additional mechanisms, such as *indexical bases*, are needed.

[28] Natural language examples of iconic discontinuity are employed in poetry, in the form of line breaks. In the following Danish poem (*Columbus* by Johannes V. Jensen, 1906), the line break between 'the Edge' and 'of the Earth' has been previously analyzed as a marker of the break/disruption that comes with reaching an edge (cited from Brandt 2013:563, who attributes it to Kjørup 2003).

De frygter at Skibet skal nærme sig **Kanten af Jorden**, hvor Havet nedstyrter sin Sluse,

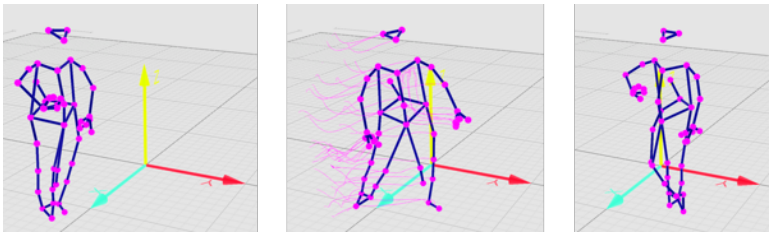
'They fear the Ship might near **the Edge of the Earth**, [...]'

b. H_2 = … designated positions in space (*indexical bases*) are needed in addition to visual discontinuity marking.

c. H_3 = … designated positions in space (*indexical bases*) are sufficient, and visual discontinuity marking is entirely uninformative with regards to coreference.

4.1 Revisiting meaningful dance from the perspective of dance syntax (grouping)

Let us take Figure 21 as our point of departure. This is a simplified version of Figure 20, which we construct by removing the pointing gesture that introduces a new location on the stage. One clear difference between [P_{31}] and [P_{33}] is the position and orientation of the dancer, due to the shift in [P_{32}].[29] Since this is the change that is, presumably, most evident to an onlooker, we can ask whether this alone might be an iconic trigger for inferences towards disjoint reference. From the perspective of a formal semantic analysis, the question then arises how to model or derive such a visual discontinuity inference.



[P_{31}] [P_{32}] [P_{33}]

Figure 21: disjoint reference simplified

We can attempt to formalize the idea of a visual discontinuity inference on the basis of Charnavel's (2016, 2019)[30] work, who applies Lerdahl and Lackendoff's (1983) notion of *grouping* to dance. *Grouping* is defined as a hierarchical segmentation into smaller groups/sections (Lerdahl & Jackendoff 1983:8-9); for Charnavel, a dance sequence that contains a jump might thus be segmented into at least three groups: the section before the jump, the jump, and the section after the jump. Lerdahl & Jackendoff introduce grouping as a syntactic notion, but we will explore the idea that it can play a semantic (event-structuring) role, as foreshadowed in Schlenker (2017a:4). Specifically, we explore the option that grouping alone may be used (as an instantiation of a visual iconicity inference from discontinuity) to convey disjointness in the sense that grouping boundaries indicate two disjoint (non-identical) narratively relevant events, which potentially contain different characters of a narrative (building on Abusch 2013:13).

Charnavel proposes that a complete dance constitutes an overarching group, which can be exhaustively partitioned into smaller groups, determined by *grouping preference rules*, one of which – the central one for our case study – is given in (48).

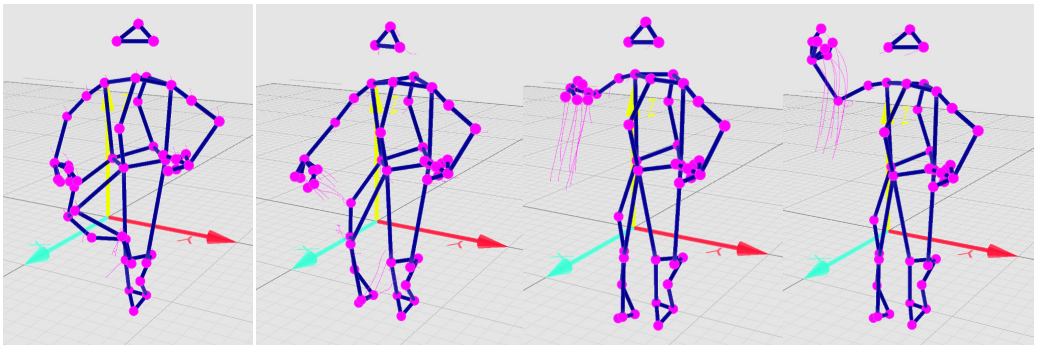
[29] There is a non-trivial question of whether the *direction* in which the dancer is moving changes as well (compare Charnavel's 2019:4 Grouping Preference Rule 4 [GPR4] – *change of direction*). In our dance sequences, the extent to which the dancer's movement involves directionality is limited, and direction thus reduces to orientation (i.e. which direction the dancer is facing) for all relevant purposes.

[30] We include both the reference to the published work (Charnavel 2019) and the earlier manuscript (Charnavel 2016), since each of the two texts contains material that is not included in the other text.

(48) *Grouping Preference Rule 2 (GPR2):* change of orientation
Position p_2 may be seen as a group boundary if the orientation of the body (part) in p_1-p_2 is different from the orientation of the body (part) in p_2-p_3.
(Charnavel 2019:4, see also Charnavel 2016:18-19)

Applying GPR2 to our dummy example in Figure 21, we infer that there must be a grouping boundary between P_{31} and P_{33}, since the orientation of the dancer has changed drastically. Crucially, note that the dancer's position has also changed, and change of position may actually be the operational cue, as discussed throughout sections 2 and 3. However, the relevant changes that Charnavel (2019:4) takes to induce grouping boundaries involve a change of (i) moving entity, (ii) orientation, (iii) contact point with the floor/weight shift, (iv) direction, (v) speed, and (vi) dynamics/quality. None of these transparently maps towards a change in position.

We can now revisit our examples of coreference *vs*. disjoint reference from the perspective of grouping. Having established an approach to "truth" in narrative dance in section 3.1 (in line with Abusch 2021), let us reconsider the coreferent dance sequence from Figure 3, repeated in Figure 22. In line with Abusch (2013:12, 2014:10), we can posit the satisfaction conditions in (49) to (partially) describe the dance positions in Figure 22. Recall that we use the dance position label, $[P_n]$ to stand in for the actual dance position. This notation is parallel to the way in which Abusch (2013, 2014) labels the panels in a comic. What becomes explicit from (49) is that dance positions $[P_n]$ are mapped to propositions $[\![P_n]\!]$.[31]



$[P_{11}]$ $[P_{12}]$ $[P_{13}]$ $[P_{14}]$

Figure 22: coreference condition

(49) a. A situation/scene σ_{11} satisfies $[P_{11}]$ only if in σ_{11} a person α is sitting.
b. A situation/scene σ_{14} satisfies $[P_{14}]$ only if in σ_{14} a person β is holding a spear.

The coreference condition, Figure 22, does not involve any change of orientation, i.e., we can assume that there is no relevant grouping boundary that emerges between $[P_{11}]$ and $[P_{14}]$. We also observe that the discourse referent α in (49a) is identified with the discourse referent β in (49b) ($\alpha = \beta$), which is transparently based on the absence of any cue towards disjoint reference.

[31] Recall that, since dance is continuous, discrete positions such as $[P_{11}]$ and $[P_{12}]$ must be stipulated in order to apply Abusch's analysis, which was designed for comics. For now, as previously discussed, we keep treating dance positions as static images, but one open question concerns the continuity (movement) between them, and whether a sign-language based semantics would be more adequate, see footnote 32.

Contrast this with the disjoint reference condition in Figure 23, adapted from Figure 4. Here, a grouping boundary is introduced between [P_{21}] and [P_{25}] (roughly between [P_{23}] and [P_{24}]) due to a change in orientation and position.

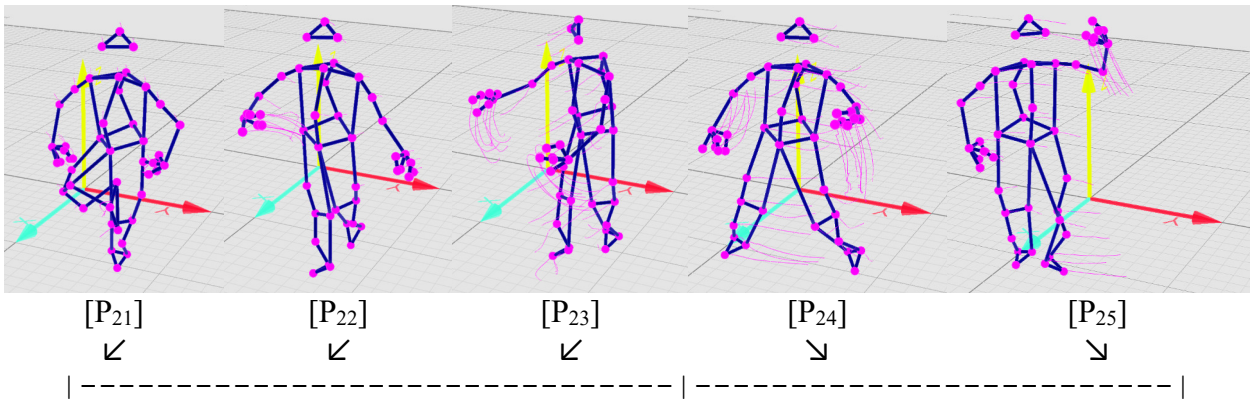


Figure 23: grouping in the disjoint reference condition

The satisfaction conditions in (50a-b) are equivalent to those in (49a-b). The crucial difference seems to be that the group boundary between [P_{23}] and [P_{24}] (which, by transitivity, counts as a group boundary between [P_{21}] and [P_{25}]) somehow blocks an inference to the end that the person α in σ_{21} and the person β in σ_{25} are identical, i.e., we draw a disjoint reference inference ($\alpha \neq \beta$).

(50) a. A situation/scene σ_{21} satisfies [P_{21}] only if in σ_{21} a person α is sitting.
b. A situation/scene σ_{25} satisfies [P_{25}] only if in σ_{25} a person β is holding a spear.

Importantly, if we factor in smaller changes in the dance sequence (e.g., changes in the position of the right hand and arm) as group-inducing (at a lower level), then we can posit at least a three-level hierarchical structure for Figure 23, as given in (52) (using Charnavel's 2016, 2019 notation). For the purpose of illustration, (52) assumes that each of the positions in Figure 23 is separated from the preceding/following position by a low-level group boundary, given that the orientation of body parts constantly changes (in [P_{21}-P_{22}], the right leg changes orientation; in [P_{22}-P_{23}], the upper body changes its orientation while the movements of the hands change direction, and so forth).[32] The role of *global (whole-body) gestures* (the larger change of orientation in the transition from P_{23} to P_{24}) comes into play in connection with

[32] For future research on how to sharpen the theory of boundary placement, a possible direction would be to compare dance sequences to sign language also with regards to grouping. We are grateful to an anonymous reviewer for suggesting such an approach. Sign language, like dance, is also a modality that employs continuous body movement; moreover, since sign languages are natural languages, they uncontroversially involve the processing of such body movements into discrete linguistic units that are hierarchically organized. In the designated sign language literature, Wilbur (2003) and Malaia & Wilbur (2012) analyze the representation of event boundaries in verbs, building on Zacks et al.'s (2007, 2009) work on event boundaries outside of natural language. Liddell & Johnson (1989) and Johnson & Liddell (2010, 2011) also provide important insights on continuous body movement, by studying differences between the movement of a sign in a sign language and the transition movement between signs. One caveat in this respect concerns the fact that dance is *unlike* sign language in that even a dance form that is (potentially) highly conventionalized, such as Bharatanatyam, cannot be considered to be a language.

Charnavel's GPR10, (51), since such gestures are generally more intense than gestures that only involve individual body parts.

(51) *Grouping Preference Rule 10 (GPR10): intensification*
When the effects picked out by the local rules of change (GPR1-GPR6) are relatively more pronounced, a larger-level group boundary may be placed.
(Charnavel 2019:17, see also Charnavel 2016:24)

In line with GPR2, as stated in (48), we can position the larger-level group boundary *between* [P_{23}] and [P_{24}], i.e., in the transition between them (as opposed to identifying it with one of these dance positions); these larger-level group boundaries seem to be relevant for the narrative (i.e. 'narratively relevant') in that they can trigger disjoint reference inferences.

(52) *structure of the disjoint reference dance sequence*

P_{21}	P_{22}	P_{23}	P_{24}	P_{25}	
| – – – |	– – – |	– – – |	– – – |	– – – |	low-level grouping
| – – – – – – – – – – |			– – – – – – – |		larger-level ('narratively relevant') grouping
| – – – – – – – – – – – – – – – – – – |					top-level grouping (complete dance)

By contrast, a dance sequence in the coreference condition given in Figure 22, would lack the larger-level grouping boundary, as schematically shown in (52) vs. (53). In fact, given the smoothness of the upward arm movement in Figure 22, it is not even possible to posit a low-level group boundary between P_{12} and P_{13}. Such a low-level group boundary can arguably be placed between P_{11} and P_{12}, where the arm *starts* to move, and between P_{13} and P_{14}, where the arm *stops* moving.[33]

(53) *structure of the coreference dance sequence*

P_{11}	P_{12}	P_{13}	P_{14}	
| – – – |	– – – – – – – |	– – – |		low-level grouping
| – – – – – – – – – – – – – – – |				larger-level ('narratively relevant') grouping
| – – – – – – – – – – – – – – – |				top-level grouping (complete dance)

One plausible approach to the contrast between disjoint reference, (52), and coreference, (53), is that group boundaries themselves are meaningful in narrative dance in that they signal/encode *discontinuity*, which can be used for structuring a narrative into events and sub-events.[34] We could capture this intuition by virtue of the statement in (54) (building on Abusch 2013:13); we will inspect this statement further in section 4.2. To foreshadow the outcome, we will conclude that it is not possible to determine what counts as 'narratively relevant' grouping and how to differentiate it from grouping that is not 'narratively relevant'.

[33] In Charnavel's (2019:4) system, there are several grouping preference rules (GPRs) that could be employed to place a grouping boundary between a moment where a body part is still and a moment where it is moving: her GPR1 is based on a change in which entity is moving; her GPR5 is based on a change of speed, which clearly occurs when the speed changes from 0 to *larger-than-0* and back to 0; in addition, her GPR6 is based on a change of the quality of movement, and 'no movement' has arguably a different quality of movement from 'some movement'. It is orthogonal for the present discussion, which of these GPRs is, in fact, employed to place these two grouping boundaries.
[34] We expect to find similar effects in other (non-narrative) dance forms.

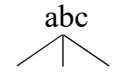
(54) *grouping-based reference determination*
If a narrative dance sequence contains two similar entities α and β, which are associated with action descriptions $e_1(\alpha)$ and $e_2(\beta)$, then
a. coreference (i.e. α=β) arises by default if $e_1(\alpha)$ and $e_2(\beta)$ **are not** separated by a narratively relevant grouping boundary, and
b. disjoint reference (i.e. α≠β) arises by default if $e_1(\alpha)$ and $e_2(\beta)$ **are** separated by a narratively relevant grouping boundary

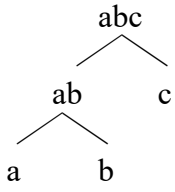
4.2 Semantic grouping (or: semantic effects of grouping)

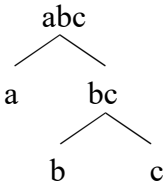
After having mapped out how coreference and disjoint reference could be linked to differences in syntactic grouping, we can now turn to the semantic side of things, i.e. to the question of which semantic effects arise from syntactic grouping in a dance sequence. A first approximation of the potential semantic effects of hierarchical grouping in narrative dance can be rendered as follows. In an important departure from Lerdahl & Jackendoff (1983), Schlenker (2017a:22-28) proposes that grouping structure in *music* semantics may reflect the organization of events in a narrative (see also Link 1983, Krifka 1989 and Landman 1991 for earlier relevant work on mereology). Building on Varzi (2015), Schlenker (2019b:76-79) fleshes out an idea of event structure in music semantics, which builds on the premise that "events are naturally perceived with a part-of structure" (Schlenker 2019b:79). His illustration of well-formed part-of structures is given in (55) (omitting cases of potentially ill-formed 'discontinuous constituents' or 'multi dominance').

(55) *Possible decompositions of abc – simplified notation*
where "abc" is an event with three atomic sub-events "a", "b", "c",
a. abc → a, b, c
b. abc → ab, c
c. abc → a, bc
(adapted from Schlenker 2019b:77)

The event decomposition in (55) gives rise to the part-of structures in (56) (from Schlenker 2019:77). As a comparable phenomenon outside of music, Schlenker highlights Cohn et al.'s (2014) work, which proposes a hierarchical (event-based) structure for comics as a type of visual narrative.

(56) a.


b.


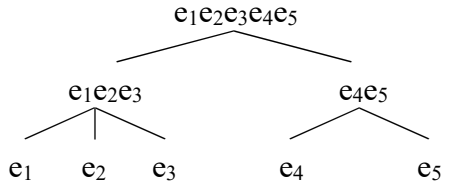
c.


Once we established groupings as given in (52) and (53) in section 4.1, we could propose that they are directly mapped onto eventive part-of structures, as sketched in (57b) and (58b). Crucially, in the disjoint reference case, we observe that the events that correspond to a given referent (α vs. β) are on different 'branches' of the part-of structure in (57b). The α referent is associated with the $e_1e_2e_3$ branch whereas the β referent is associated with the e_4e_5 branch.

(57) a. *syntactic grouping: structure of the disjoint reference dance sequence*

P_{21}	P_{22}	P_{23}	P_{24}	P_{25}	
|–––|	–––|	–––|	–––|	–––|	low-level grouping
|––––––––––|		–––––––|			larger-level ('narratively relevant') grouping
|––––––––––––––––––|					top-level grouping

b. *semantic grouping: part-of structure of the disjoint reference event*



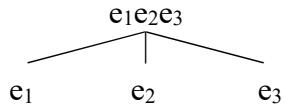
(where $P_{21} \rightsquigarrow e_1$, $P_{22} \rightsquigarrow e_2$, $P_{23} \rightsquigarrow e_3$, $P_{24} \rightsquigarrow e_4$, $P_{25} \rightsquigarrow e_5$)

For (58b), we observe that $[P_{12}]$ and $[P_{13}]$ are plausibly mapped to a single event (e_2) as they lack a low-level grouping boundary; more crucially, due to the absence of a larger-level grouping boundary, the part-of structure that emerges in (58b) is a flat structure.

(58) a. *syntactic grouping: structure of the coreference dance sequence*

P_{11}	P_{12}	P_{13}	P_{14}	
|–––|	–––––––|		–––|	low-level grouping
|––––––––––––––|				larger-level 'narratively relevant' grouping
|––––––––––––––|				top-level grouping (complete dance)

b. *semantic grouping: part-of structure of the disjoint reference event*



(where $P_{11} \rightsquigarrow e_1$, $P_{12}P_{13} \rightsquigarrow e_2$, $P_{14} \rightsquigarrow e_3$)

The relevant step in deriving disjoint reference from syntactico-semantic grouping alone would be the intermediate level in (57b), which divides $e_1e_2e_3$ (and thus the α referent) from e_4e_5 (and thus the β referent). However, there is no inherently privileged status about this level in the part-of structure, i.e., it is unclear how a mechanism could be conceived that would derive disjoint reference in (57b), and wouldn't equally derive disjoint reference between e_1 and e_2 (or between e_2 and e_3) in (58b). One possibility is to make reference to the fact that the intermediate stages in (57b) are non-atomic in that $e_1e_2e_3$ and e_4e_5 are still composed of atomic sub-events. Such an approach is spelled out in (59).

(59) *grouping-based reference determination (revised)*
If a narrative dance sequence contains two similar entities α and β, which are associated with action descriptions $e_1(\alpha)$ and $e_2(\beta)$, then
a. disjoint reference (i.e. $\alpha \neq \beta$) arises by default if $e_1(\alpha)$ and $e_2(\beta)$ **are** on different branches of a part-of event structure that have non-atomic subevents, and
b. coreference (i.e. $\alpha = \beta$) arises by default if $e_1(\alpha)$ and $e_2(\beta)$ **are not** on different branches of a part-of event structure that have non-atomic subevents.

However, while this may allow us to differentiate between a (57b) scenario and a (58b) scenario, it is more problematic that the operative mechanism in creating and differentiating these two branches is intensification. Charnavel's rendering in (60) (repeated from (51)) presupposes a notion of *relatively more pronounced* changes, which is in itself underspecified. [35]

(60) *Grouping Preference Rule 10 (GPR10): intensification*
When the effects picked out by the local rules of change (GPR1-GPR6) are relatively more pronounced, a larger-level group boundary may be placed.
(Charnavel 2019:17, see also Charnavel 2016:24)

The bottom line seems to be that disjoint reference is triggered by a change in position and change in orientation, which would seem to be the *only* grouping boundary that matters as 'narratively relevant' for (57)-(58). Other potential changes in line with Charnavel's (2019) grouping preference rules, such as a change of speed, or a change of dynamics/quality, do not give rise to a similar disjoint reference effect.

We are thus forced to conclude that 'narratively relevant' grouping is captured by the statement in (61), i.e., it involves a change in position coupled to a change in orientation.

(61) *Narratively relevant grouping*
In Bharatanatyam, a 'narratively relevant' grouping boundary is established when the dancer [i.] assumes a new position and [ii.] changes the orientation of her entire body. Other dance moves do not create a 'narratively relevant' grouping boundary.

An attentive reader will already have observed that (61) more or less amounts to proposing that a 'narratively relevant' grouping boundary is established when the dancer performs action role shift in dance as outlined in 3.2.3-3.2.4.

This is crucial, as an approach that makes *direct* reference to mechanisms such as indexical bases (associated with change of position) and action-performance (associated with change of orientation) is more parsimonious than an approach that builds on a stipulated notion of 'narratively relevant' grouping, which ends up being parasitic on indexical bases and action-performance in all but name; we can thus do away with 'narratively relevant' grouping altogether. This conclusion is strengthened by the fact that something along the lines of indexical bases and action-performance is needed in any case to account for the narrative sequences with three separate referents.

To conclude this section, we should emphasize that we here outlined an analysis that implements *iconic discontinuity effects* by virtue of syntactic grouping, from which grouping-based part-of structures of events are derived in the semantics. There may well be other ways to model iconic discontinuity effects that don't run into the same problems that the event-structuring analysis encounters. However, the burden of

[35] A more informed approach to the question of how 'more pronounced' movements could be distinguished from 'less pronounced' movements may be based on the discussions of sonority in sign languages, where, for instance, arm movements are classified as 'more sonorant' than hand movements, and thus more pronounced/prominent (see Brentari 1998, Schlenker et al. 2016). While a further exploration of this issue is beyond the scope of this paper, we are grateful to an anonymous reviewer for this suggestion.

the proof is clearly on scholars who argue that iconic discontinuity is sufficient to derive disjoint reference, given that indexical bases and action-performance provide a parsimonious way to derive disjoint reference; the existence of indexical bases and action-performance in dance is further supported by the previously established status of similar mechanisms that humans can employ to visually encode meaning, both in sign language, and in the speech-accompanying gestures of non-signers.

5. Implications and conclusions

This paper provides an initial study of meaning in dance sequences from a linguistic perspective, thus further enriching and expanding the empirical domain of linguistic analysis beyond language. We carried out the following steps to show that a formal semantic methodology can be fruitfully applied to meaning in narrative dance (as pioneered for meaning in pictures, Greenberg 2011, 2013, pictorial narrative, Abusch 2013, 2014, 2021, and music, Schlenker 2017a, 2019b).

Using the linguistic method of controlled elicitation, we carried out two production studies (sections 2.3, 2.4 and 2.6) that focused, in particular, on the topic of coreference (*the same man*) vs. disjoint reference (*another man*). We determined that a dancer can use a change of position and orientation in order to signal disjointness in the disjoint reference condition. A perception study (section 2.5) that used stick figures based on motion capture recording showed that untrained audience members inferred a disjointness of reference more often in the disjoint reference video than in the coreference video. While it is unclear which cues they used to draw their inferences, this indicates that the change of position and orientation can also be processed as an indicator of disjoint reference – or, as we argue, a mechanism for referent tracking.

Much in line with previous applications of formal linguistic methodology to meaning in pictures (Greenberg 2011, 2013), pictorial narrative (Abusch 2013, 2014, 2021) and music (Schlenker 2017a, 2019b), we proposed a formal semantic analysis that treats dance positions as informational entities with information content (section 3). The meaning of a given dance position can thus be modeled in a possible worlds semantics (Abusch 2021:2) (section 3.1).

Inspired by Schlenker's (2020) approach to the semantics of speech-accompanying gestures of non-signers, we argued that narrative dance can recruit communicative mechanisms that may be availed by human cognition more generally; we argued that narrative dance makes use of: (i) *indexical bases*, positions on the stage that are associated with a given referent, section 3.2.1; (ii) *base-to-base linking*, where a dancer is positioned in one base and carries out an action directed towards the base associated with the object of that action, section 3.2.2; and (iii) *action-performance*, where a dancer assumes the perspective of a character in the narrative in order to demonstrate/perform their actions, sections 3.2.3 and 3.2.4.

Acknowledging that a reader may be skeptical of positing such abstract mechanisms for dance, we explored the option that disjoint reference may be simply due to visual discontinuity, in section 4. We showed that such an option is not viable, as it requires a stipulation of 'narratively relevant' discontinuity and 'non-narratively relevant' discontinuity, which essentially boils down to positing indexical bases and action-performance after all. The only type of discontinuity that seems to be 'narratively relevant' for reference tracking is discontinuity based on the assumption of a new position on stage, accompanied by a change in orientation.

We can now revisit our initial aims of expanding linguistic investigation of non-standard objects, and draw two tentative conclusions on the questions of what unifies dance with language and other communicative competencies, and what distinguishes them from one another: On the one hand, what unifies dance and natural language, in particular, are the presence of hierarchical grouping structure in both (which dance shares with music and other pictorial narratives such as comics), and the access to shared communicative tools for reference tracking and perspective taking in visual communication (e.g., the indexical bases in dance, which descriptively resemble the loci of sign languages and gestural communication). On the other hand, what differentiates dance from language (and maybe from language-*like* systems, more generally) is that dance, like music, operates predominantly on iconic resemblance between the dance sequences and the intended meaning, but does so at a highly underspecified level (see the discussion around (62) in section 5.2 below, where one and the same dance move could alternatively be understood to mean 'someone is raising the right arm' or 'the sun is rising'). This sets dance and music apart from more concrete pictorial representations like comics, and from non-iconic symbolic representation, as is central to natural language.

5.1 Extensions 1: Refining our mission statement

We conclude our paper with an explicit statement about its scope, which culminates in a mission statement for future empirical investigations. This will be followed by a brief discussion of further directions in section 5.2. In this paper, we have presented and analyzed dance sequences performed in an exploratory setup by a single professional dancer. This gives rise to two empirical questions: [i.] Is the strategy employed in our data a general strategy that is employed by Bharatanatyam dancers (*generalizability*), or is it an individual choice? [ii.] Is this strategy specific to dance or does it draw on general gestural resources (*specificity*)? For the first question, we have minimally observed that the same dancer was consistent across the space of one year, employing the same strategies in our first study (November 2016) and in our second study (December 2017). For the second question, we argued in section 3 that the main tools (indexical bases, base-to-base linking, action-performance) are crucially not specific to dance, but rather reflect a broader inventory of tools for meaning-making that human cognition seems to avail. While both questions should be explored further in designated experimental papers (going beyond the scope of this paper, which has an exploratory theoretical focus rather than an experimental one), we briefly elaborate further, in line with our current goals.

As far as generalizability is concerned, we consider the question to be a non-issue of whether this strategy is employed by *all* Bharatanatyam dancer, since the inferences that we found were based on mechanisms that are more broadly employed, in sign languages and the speech-accompanying gestures of non-signers. The question is not whether all Bharatanatyam dancer do, in fact, use such a strategy; rather, our conclusion is that the strategy is available to dancers (in Bharatanatyam and beyond), whether they avail of it or not.

As for specificity, a natural question is whether we would expect the same strategy to be utilized in other dance forms, in pantomime, and in (silent) gestural communication more generally. If the answer turns out to be 'yes' to all three, this will not invalidate the point that we are making in this paper, but rather strengthen it. It would corroborate the conclusion that there are universal resources available in body-based communication (subsuming dance, pantomime, and gesture) that can be

applied for a semantic effect, much in line with our finding that meaningful mechanisms from gestural communication resemble mechanisms that we find in dance.

Future experiments may probe specifically into the pantomime performed by untrained participants (i.e., not by professional mimes) to test whether they spontaneously employ similar strategies to our dancer, including a type of indexical base, base-to-base linking and action-performances. More in-depth comparison of dance and other modes of expression is a highly important step that needs to be taken in future follow-up research.

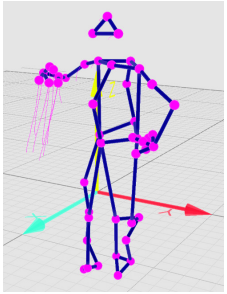
5.2 Extensions 2: Abstract iconic meanings in – and beyond – narrative dance

Naturally, a long-term goal of exploring the semantics of dance should include an in-depth investigation of abstract iconic meaning atoms as posited by Schlenker (2017a, 2019b) for music. Abstract meaning atoms in music have been argued to license inferences on properties or actions of a *virtual source*: when listening to a piece of music, listeners may imagine that the music conveys information about some object, person, animal, landscape or other entity. This entity would be a virtual source of the music; for instance, when listening to a low-pitched melody, we may imagine a big animal represented by the melody, or a vast landscape. By contrast, when listening to a high-pitched melody, we may imagine a small animal. Such virtual sources are underspecified by the abstract meaning that music can convey; e.g., the music alone will rarely (if ever) disambiguate whether the virtual source is a big animal or a vast landscape – it will simply be compatible with both. Abstract iconic meaning atoms can then be used to encode certain properties of virtual sources. *Discontinuity*, as discussed in section 4, may well be a likely candidate for such a meaning atom, which could accompany the marking of discontinuities in protagonists/characters, locations, situations, and so forth, all of which would amount to a shift from one virtual source to another. This is independent from our conclusion in section 4, that discontinuity cannot be operative in the distinction between coreference and disjoint reference in our data. Other iconic meaning atoms may be manifested in dance through the inferences arising from different types of spatiotemporal movement descriptors; e.g., the *quality* of a given movement may be described as "smooth" vs. "jagged" (see for example Laban 1975, Guest 2005, and Napoli and Kraus 2017, for overviews on the parameters of dance and movement).

Once we expand our semantic formalism to more abstract, iconic atoms of meaning, the approach clearly carries over to all dance forms, including non-narrative dance forms. In line with Schlenker's (2017a) approach to music, we can assume that dance movements allow for inferences on a virtual source, i.e., a center of attention in the narrative, such as a main character or a central event. We capture this by positing more abstract satisfaction conditions such as (62), which is a dance movement taken out of its original context. Outside of a dedicated narrative context, inferences will typically be more abstract; by abstract, we mean (in the spirit of Schlenker 2017a, 2019b) that the meanings that can be conveyed, while consistently being iconic (resemblance-based), are not limited to 'literal' meanings. For instance, while the movement in (62) may well serve to convey 'literal' meanings such as 'someone raises the right arm', it could also serve to convey 'non-literal' meanings such as 'the sun is rising'. In other words, (62) could easily accompany the sunrise in Strauss's

Zarathustra, which Schlenker (2017a, 2019b) discusses.[36] The actual meaning that is expressed will naturally be much more abstract, as indicated.

(62)

a situation σ satisfies  only if the virtual source in σ is involved in a (partial or total) upward movement.

A skeptical reader may question the appropriateness of analyzing an example like (62), which is taken out of its original context; however, this example is purely included for illustration purposes. For a concrete example of meaning in abstract dance, the reader may wish to consult the discussion of Balanchine's ballet Symphony in C in Appendix IV of Schlenker (2019a). Here, a ballet choreography conveys a dialogue between two virtual entities (sources), not independently present in the music.

Appendix

Stimuli November 2016

Set 1 (coreference vs. disjoint reference):
Context for all items: "An artist has designed a statue for a temple. She is at the temple, watching how people interact with the statue; the room is full of people."

Item 1
a. The artist sees a strong man sitting on the ground.
b. *Cond. 1 (coref):* Then she sees that *the same man* is holding a spear.
c. *Cond. 2 (disj):* Then she sees that *another man* is holding a spear.

Item 2
a. The artist sees a woman waving a palmyra leaf in the sunlight.
b. *Cond. 1 (coref):* Afterwards *that woman* is pointing at the clouds in the sky.
c. *Cond. 2 (disj):* Afterwards *another woman* is pointing at the clouds in the sky.

Item 3
a. The artist watches a child eating a mango outside the temple.
b. *Cond. 1 (coref):* Then *the child* is entering the temple.
c. *Cond. 2 (disj):* Then *another child* is entering the temple.

Item 4
a. The artist watches a man holding a book.
b. *Cond. 1 (coref):* Then she sees *the same man* looking at a water lily.
c. *Cond. 2 (disj):* Then she sees *another man* looking at a water lily.

Item 5
a. The artist sees a woman praying in silence.
b. *Cond. 1 (coref):* Then *that woman* walks to a basket of fruits.

[36] The relevant video can be found at https://youtu.be/a1pqRbQypqM

c. *Cond. 2 (disj):* Then *another woman* walks to a basket of fruits.

Item 6
a. The artist watches a girl dancing in the sunlight.
b. *Cond. 1 (coref):* Then *the girl* trips over a stone.
c. *Cond. 2 (disj):* Then *another girl* trips over a stone.

Set 2 (reflexive vs. non-reflexive): ***not used/discussed in this paper***

Item 1
a. The artist watches a man grasping something.
b. *Cond. 1 (refl):* Then the man cuts *himself* with a sword.
c. *Cond. 2 (non-refl):* Then the man cuts *another man* with a sword.

Item 2
a. The artist watches a woman cutting an apple for her son.
b. *Cond. 1 (refl):* Then the woman puts a coat on *herself*.
c. *Cond. 2 (non-refl):* Then the woman puts a coat on *him*.

Item 3
a. The artist sees a boy running around the temple.
b. *Cond. 1 (refl):* Afterwards the boy sees *himself* in the mirror.
c. *Cond. 2 (non-refl):* Afterwards the boy sees *another boy* in the mirror.

Item 4
a. The artist watches a girl reading.
b. *Cond. 1 (refl):* Then the girl gets up to feed *herself*.
c. *Cond. 2 (non-refl):* Then the girl gets up to feed *her sister*.

Item 5
a. The artist watches a man walk up to the statue.
b. *Cond. 1 (refl):* Then the man calls *himself* a hero.
c. *Cond. 2 (non-refl):* Then the man calls *another man* a hero.

Item 6
a. The artist watches a boy with pen and paper.
b. *Cond. 1 (refl):* The boy is drawing *himself*.
c. *Cond. 2 (non-refl):* The boy is drawing *another boy*.

Set 3 (de se vs. de re): ***not used/discussed in this paper***

Item 1
a. The artist watches a woman walking around without realizing that it is his wife.
b. *Cond. 1 (de se):*
He thinks *himself* to be a lucky man, because he sees a beautiful girl.
c. *Cond. 2 (de re):*
He thinks *her husband* to be a lucky man, because he sees a beautiful girl.

Item 2
a. The artist watches his own shadow on the ground without recognizing himself.
b. *Cond. 1 (de se):*
Then he notices (understands) that *he himself* is casting a shadow on the statue.
c. *Cond. 2 (de re):*

Then he notices that *the person whose shadow he is watching* is casting a shadow on the statue.

Item 3

a. The artist sees a girl assisting an old woman without noticing that the girl is his daughter.
b. *Cond. 1 (de se):*
   He believes *himself* to be a good man.
c. *Cond. 2 (de re):*
   He believes *the girl's father* to be a good man.

Stimuli December 2017

Set 1 (parallel vs. switched reference with potential agreeing verbs):

Item 1

a. The artist sees a child eating a mango outside the temple.
b. Then she sees another child holding a spear.
c. *Cond. 1 (parallel):* The eating child watches the child with the spear.
d. *Cond. 2 (switched):* The child with the spear watches the eating child.

Item 2

a. The artist watches a man holding a book.
b. Then she sees that another man is sitting on the ground.
c. *Cond. 1 (parallel):* The man holding a book looks at the sitting man.
d. *Cond. 2 (switched):* The sitting man looks at the man holding a book.

Item 3

a. The artist sees a woman entering the temple.
b. Then she sees another woman eating fruit.
c. *Cond. 1 (parallel):* The woman entering waves at the woman eating fruit.
d. *Cond. 2 (switched):* The woman eating fruit waves at the woman entering.

Set 2 (three referents in all combinations):

Item

a. A woman is standing outside the temple, a man is sitting on the ground, and a child is playing.
b. *Cond. 1:* The woman is holding a book. The man is looking at the child.
c. *Cond. 2:* The woman is holding a book. The child is looking at the man.
d. *Cond. 3:* The man is holding a book. The woman is looking at the child.
e. *Cond. 4:* The man is holding a book. The child is looking at the woman.
f. *Cond. 5:* The child is holding a book. The woman is looking at the man.
g. *Cond. 6:* The child is holding a book. The man is looking at the woman.

Set 3 (temporal reference): ***not used/discussed in this paper***

Item 1

a. A woman enters the temple.
b. *Cond. 1 (present perfect & present):*
   (earlier) She has performed a dance.
   She prays to the gods for peace (now).
c. *Cond. 2 (present perfect & past perfect):*

(earlier) She has performed a dance.
(before that) She had prayed to the gods for peace.

d. *Cond. 3 (present perfect & future):*
(earlier) She has performed a dance.
She will pray to the gods for peace (later).

e. *Cond. 4 (past perfect & past):*
(earlier) She had performed a dance.
Then she prayed to the gods for peace (after that).

Item 2

a. A woman sits on the ground.
b. *Cond. 1 (future & present):*
She will read a book (later).
She eats a mango (now).
c. *Cond. 2 (future & present perfect):*
She will read a book (later).
(earlier) She has eaten a mango.
d. *Cond. 3 (future & future perfect):*
She will read a book (later).
(before that) She will have eaten a mango.
e. *Cond. 4 (future & future):*
She will read a book (later).
Then she will eat a mango (after that).

Set 4 (disjoint reference without 'another' with potential agreeing verbs):

Item 1

a. A woman is sitting outside and a man is standing in the middle of the room.
b. The woman is holding a book.
c. *Cond. 1 (topic shift):* The man is looking at her.
d. *Cond. 2 (continued topic):* She is looking at the man.

Item 2

a. A man is sitting in the corner and a woman is standing in the middle of the room.
b. The man is holding a spear.
c. *Cond. 1 (topic shift):* The woman is looking at him.
d. *Cond. 2 (continued topic):* He is looking at the woman.

References

Abner, Natasha, Kensy Cooperrider, and Susan Goldin-Meadow (2015). Gesture for Linguists: A Handy Primer, *Language and Linguistics Compass* 9, 437–449.

Abusch, D. (2013). Applying discourse semantics and pragmatics to co-reference in picture sequences. *Proceedings of Sinn und Bedeutung 17*, ENS Paris, 9–25.

Abusch, D. (2014). Temporal succession and aspectual type in visual narrative. In L. Crnič and U. Sauerland (Eds.): *The Art and Craft of Semantics: A Festschrift for Irene Heim, vol. 1*, MITWPL 70, pp. 9–20. Cambridge, MA: MIT Working Papers in Linguistics.

Abusch, D. (2021). Possible-Worlds Semantics for Pictures. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann

(eds.): *The Wiley Blackwell Companion to Semantics*: Oxford: John Wiley & Sons. [DOI: 10.1002/9781118788516.sem003]
Brandt, Line (2013). *The Communicative Mind: A Linguistic Exploration of Conceptual Integration and Meaning Construction*. Newcastle: Cambridge Scholars Publishing.
Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.
Champollion, L. (2014). Integrating Montague Semantics and Event Semantics. Lecture notes for the NASSLLI 2014 course, University of Maryland, College Park, 23-27 June 2014.
Charnavel, I. (2016). Steps towards a Generative Theory of Dance Cognition. Manuscript, June 2016, Harvard University. [http://ling.auf.net/lingbuzz/003137]
Charnavel, I. (2019). Steps Toward a Universal Grammar of Dance: Local Grouping Structure in Basic Human Movement Perception. *Frontiers in Psychology* 10, 1364.
Cohn, N., R. Jackendoff, P.J. Holcomb, and G.R. Kuperberg (2014. The grammar of visual narrative: neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia* 64, 63–70.
Cohn, Neil (2018). In defense of a "grammar" in the visual language of comics. *Journal of Pragmatics* 127, 1-19.
Cohn, Neil (2020). Your Brain on Comics: A Cognitive Model of Visual Narrative Comprehension. *Topics in Cognitive Science* 12, 352-386.
Cohn, Neil (2019). Being explicit about the implicit: inference generating techniques in visual narrative. *Language and Cognition* 11, 66-97.
Davidson, K. (2015). Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38, 477–520.
Deo, Ashwini (2015). Diachronic Semantics. *The Annual Review of Linguistics* 1, 179-197.
Greenberg, Gabriel (2011). The semiotic spectrum. PhD dissertation, Rutgers University.
Greenberg, Gabriel (2013). Beyond Resemblance. *Philosophical Review* 122, 215-287.
Greenberg, Gabriel (2021). The Iconic-Symbolic Spectrum. Manuscript, UCLA. Available at https://ling.auf.net/lingbuzz/005787
Grønn, Atle, and Kjell Johan Sæbø (2011). *A, The, Another*: A Game of Same and Different, *Journal of Logic, Language and Information* 21, 75-95.
Guest, A. H. (2005). *Labanotation*. New York: Routledge.
Heim, I. (1982). The Semantics of Definite and Indefinite Noun Phrases. PhD dissertation, Umass Amherst.
Herrmann, Annika, and Markus Steinbach. 2009. Visible context shifts: Quotation in sign languages. Talk presented at the 42nd Annual Meeting of the Societas Linguistica Europaea (SLE), Universidade de Lisboa, September 2009.
Herrmann, Annika, and Markus Steinbach. 2012. Quotation in sign language. A visible context shift. In Ingrid van Alphen and Isabelle Buchstaller (eds.): *Quotatives: Cross-linguistic and cross-disciplinary perspectives*. Amsterdam: John Benjamins, 203-228.
Fodor, J. A., M. F. Garrett, E. C. T. Walker, and C. H. Parkes. (1980). Against definitions. *Cognition* 8, 263-267.
Hockett, Charles F. (1942). A System of Descriptive Phonology. *Language* 18, 3-21.

Ikegami, Y. (1971). A stratificational analysis of the hand gestures in Indian classical dancing. *Semiotica* 4, 365–391.
Jensenius, Alexander R. (2017). 2014: To Gesture or Not? An Analysis of Terminology in NIME Proceedings 2001–2013. In Alexander R. Jensenius, and Michael J. Lyons (eds.): *A NIME Reader. Current Research in Systematic Musicology*, vol 3. Cham: Springer, 451-464. https://doi.org/10.1007/978-3-319-47214-0_29
Jensenius, Alexander R., Marcelo M. Wanderley, Rolf I. Godøy, and Marc Leman (2010). Musical gestures: Concepts and methods in research. In Rolf I. Godøy & Marc Leman (eds.): *Musical gestures: Sound, movement, and meaning*. New York: Routledge, 12–35.
Johnson, Robert E., and Scott K. Liddell (2010). Toward a Phonetic Representation of Signs: Sequentiality and Contrast. *Sign Language Studies* 11, 241-274.
Johnson, Robert E., and Scott K. Liddell (2011). Segmental Framework for Representing Signs Phonetically. *Sign Language Studies* 11, 408-463.
Kamp, H. and U. Reyle (1993). *From discourse to logic*. Dordrecht: Kluwer.
Katz, J. and D. Pesetsky (2011). The Identity Thesis for Language and Music. Manuscript, January 2011, Institut Jean Nicod / MIT. [http://ling.auf.net/lingbuzz/000959]
Katz, Jonah (2017). "Harmonic syntax of the 12-bar blues: a corpus study." Music Perception 35, 165-192.
Kegl, Judy. (2004). ASL syntax: Research in progress and proposed research. *Sign Language and Linguistics* 7, 173-206. [Reprint of 1977 ms., MIT.]
Kelkar, T., and Jensenius, A. R. (2018). Analyzing Free-Hand Sound-Tracings of Melodic Phrases. *Applied Sciences* 8(1), 135. [https://doi.org/10.3390/app8010135]
Kedhar, Anusha. (2020). *Flexible Bodies: British South Asian Dancers in an Age of Neoliberalism*. Oxford: Oxford University Press.
Kendon, Adam. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
Kjørup, Frank (2003). *Sprog Versus Sprog: Mod en versets poetic.* Copenhagen: Museum Tuscalanums Forlag, Copenhagen University.
Kratzer, Angelika (2020). Situations in Natural Language Semantics. In Edward N. Zalta (ed.): *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), URL: https://plato.stanford.edu/entries/situations-semantics/
Krifka, Manfred (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (eds.): *Semantics and contextual expression*. Dordrecht, Netherlands: Foris, 75-115.
Krumhansl, C. L. and D. L. Schenck (1997). Can dance reflect the structural and expressive qualities of music? A perceptual experiment on Balanchine's choreography of Mozart's Divertimento No. 15. *Musicae Scientiae* 1, 63–85.
Kuhn, Jeremy (2016). ASL Loci: Variables or Features? *Journal of Semantics* 33, 449–491.
Laban, Rudolf. (1975). *Laban's Principles of Dance and Movement Notation*. 2nd edition edited and annotated by Roderyk Lange. London: MacDonald and Evans. [First published 1956.]
Landman, F. (1991). *Structures for Semantics*. Dordrecht: Kluwer.
Lerdahl, F. and R. Jackendoff (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Liddell, S. 1990. Four Functions of Locus: Reexamining the Structure of Space in ASL. In C. Lucas (ed.): *Sign Language Research: Theoretical Issues*. Washington, DC: Gallaudet University Press, 176-98.
Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.
Liddell, Scott K. (2000). Indicating verbs and pronouns: Pointing away from agreement. In Karen Emmorey and Harlan Lane (eds.): *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*. Mahwah: Lawrence Erlbaum Associates, 303-320.
Liddell, Scott K., and Robert E. Johnson (1989). American Sign Language: The Phonological Base. *Sign Language Studies* 1064, 195-277.
Lillo-Martin, Diane. 1995. The Point of View Predicate in American Sign Language. In K. Emmorey and J.S. Reilly (eds.): *Language, Gesture & Space*. Hillsdalle, NJ: Lawrence Erlbaum Associates, 155-170.
Lillo-Martin, Diane, and Edward Klima. 1990. Pointing out difference: ASL pronouns in syntactic theory. In S. Fischer and P. Siple (eds.): *Theoretical Issues in Sign Language Research 1: Linguistics*. Chicago: University of Chicago Press, 191-210.
Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (eds.): *Meaning, use and interpretation of language*. Berlin, Germany: de Gruyter, 303-323.
Malaia, E., and Wilbur, R. (2012). Telicity expression in the visual modality. In V. Demonte and L. McNally (eds.): *Telicity, change, and state: A cross-categorial view of event structure*. Oxford University Press, Oxford, UK.
McNeill, David (1992). *Hand and mind: what gestures reveal about thought*. Chicago, IL: University of Chicago Press.
McNeill, David (2012). *How Language Began, Gesture and Speech in Human Evolution*. Cambridge: Cambridge University Press.
Napoli, D. J. and L. Kraus (2017). Suggestions for a Parametric Typology of Dance. *Leonardo* 50, 468–476.
Neidle C., J. Kegl, D. McLaughlin, B. Bahan and R. G. Lee (2000). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: MIT Press.
O'Shea, Janet. 2003. At Home in the World?: The Bharatanatyam Dancer as Transnational Interpreter. *The Drama Review* 47, 176-186. [Reprinted in Jens Richard Giersdorf and Yutian Wong (eds.) (2019): *The Routledge Dance Studies Reader*, 3rd edition, New York: Routledge, 66-76.]
Padden, Carol A. 1986. Verbs and role-shifting in American Sign Language. In Carol A. Padden (ed.): *Proceedings of the fourth national symposium on sign language research and teaching*. Silver Spring, MD: National Association of the Deaf.
Padden, Carol. 1988. *Interaction of morphology and syntax in American Sign Language*. New York, NY: Garland.
Patel-Grosz, Pritty, Salvador Mascarenhas, Emmanuel Chemla, and Philippe Schlenker. Forthcoming. Super Linguistics: an Introduction. https://ling.auf.net/lingbuzz/005242
Potts, C. (2015). Presupposition and Implicature. In S. Lappin, and C. Fox (Eds.), *The Handbook of Contemporary Semantic Theory*, pp. 168–203. Oxford: Wiley.
Puri, R. (1986). Elementary Units of an Action Sign System: The Hasta or Hand Positions of Indian Classical Dance. *Semiotica* 62, 247–277.

Puri, R. (2004). Bharatanatyam Performed: A Typical Recital. *Visual Anthropology* 17, 45–68.
Pylyshyn, Z. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.
Quer, Josep. 2005. Context shift and indexical variables in sign languages. In *SALT XV*. Ithaka, NY: CLC Publications.
Ramesh, R. (2013). Indian traditions: a grammar of gestures in dance, theatre and ritual. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf (Eds.), *Body – Language – Communication, Vol. 1*. Handbooks of Linguistics and Communication Science (HSK), pp. 306–320. Berlin: de Gruyter.
Ramesh, R. (2014). Gestures in Southwest India: Dance 57heatre. In A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf (Eds.): *Body – Language – Communication, Vol. 2*. Handbooks of Linguistics and Communication Science (HSK), pp. 1226–1233. Berlin: de Gruyter.
Rebuschat, P., M. Rohrmeier, J. A. Hawkins and I. Cross (Eds.). (2011). *Language and Music as Cognitive Systems*. Oxford: Oxford University Press.
Sandler, Wendy & Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press.
Schlenker, P. (2017a). Outline of Music Semantics. *Music Perception* 35, 3–37.
Schlenker, P. (2017b). Sign Language and the Foundations of Anaphora. *Annual Review of Linguistics* 3, 149–177.
Schlenker, P. (2017c). Super monsters I: Attitude and Action Role Shift in sign language. *Semantics and Pragmatics* 10. http://dx.doi.org/10. 3765/sp.10.9
Schlenker, P., E. Chemla, and K. Zuberbühler (2017). Semantics and Pragmatics of Monkey Communication. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.220
Schlenker, Philippe, Valentina Aristodemo, Ludovic Ducasse, Jonathan Lamberton, and Mirko Santoro. (2016). The Unity of Focus: Evidence from Sign Language (ASL and LSF). *Linguistic Inquiry* 47, 363-381.
Schlenker, P., and E. Chemla (2018). Gestural agreement. *Natural Language and Linguistic Theory* 36, 587-625.
Schlenker, P. (2019a). What is Super Semantics? *Philosophical Perspectives* 32, 365-453.
Schlenker, P. (2019b). Prolegomena to Music Semantics. *Review of Philosophy and Psychology* 10, 35-111. https://doi.org/10.1007/s13164-018-0384-5
Schlenker, P. (2020). Gestural grammar. *Natural Language and Linguistic Theory* 38, 887-936.
Schlenker, P. (2021). Musical Meaning within Super Semantics. *Linguistics and Philosophy*. https://doi.org/10.1007/s10988-021-09329-8
Supalla, T. (1982). Structure and acquisition of verbs of motion and location in American Sign Language. Ph.D. Thesis, University of California, San Diego.
Tanaka, M. (1992–1994). *Gon 1*. Episodes 1-4. Tokyo: Kodansha Limited. Current edition Kodansha Comics (2011).
Tieu, L., R. Pasternak, P. Schlenker and E. Chemla. (2017). Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: a journal of general linguistics* 2(1): 102, 1–27.
Truswell, Robert (2017). Dendrophobia in bonobo comprehension of spoken English. *Mind and Language* 32, 395–415.
Varzi, A. C. (2015). Mereology. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. URL: http://plato.stanford.edu/archives/win2015/entries/mereology/

Wilbur, R. (2003). Representations of telicity in ASL. In *Chicago Linguistic Society 39*, 354-368.
Williams, D. (2003). In the Shadow of Hollywood Orientalism: Authentic East-Indian Dancing. *Journal for the Anthropological Study of Human Movement* 12, 78–99.
Zacks, J. M., Kumar, S., Abrams, R. A., and Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition* 112, 201-216.
Zacks, J., Speer, N., Swallow, K., Braver, T., and Reynolds, J. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin* 133, 273-293.