Coreference and disjoint reference in the semantics of narrative dance[1]

Pritty PATEL-GROSZ — *University of Oslo*
Patrick Georg GROSZ — *University of Oslo*
Tejaswinee KELKAR — *University of Oslo*
Alexander Refsum JENSENIUS — *University of Oslo*

Abstract. This paper presents an exploratory production study of Bharatanatyam, a figurative (narrative) dance. We investigate the encoding of coreference vs. disjoint reference in this dance and argue that a formal semantics of narrative dance can be modeled in line with Abusch's (2013, 2014, 2015) semantics of visual narrative (drawing also on Schlenker's, 2017a, approach to music semantics). A main finding of our investigation is that larger-level group-boundaries (Charnavel, 2016) can be seen as triggers for discontinuity inferences (possibly involving the dynamic shift from one salient entity to another).

Keywords: co-reference, disjoint reference, dance semantics, iconic semantics, picture semantics.

1. Background and motivation

In this paper, we aim to contribute to new lines of research that look at different cognitive systems (in the cognitive science/neuroscience sense; cf. Rebuschat et al., 2011) and how they relate to each other. Our research builds on recent pioneering investigations that explicitly connect language and linguistics to other fields such as music cognition (Schlenker, 2017a) and dance (Charnavel, 2016). A well-known predecessor of such approaches is the work of Lerdahl and Jackendoff (1983), who proposed a generative theory of music. In this paper, the question of interest is whether dance has something akin to a (compositional or non-compositional) "semantics", which we can describe by means of linguistic tools. We thus adopt the methodological approach of Lerdahl and Jackendoff, applying linguistic methodology to other cognitive systems in order to investigate underlying commonalities.

In exploring the semantics of dance, we directly build on Charnavel's (2016) question of whether dance has hierarchical structure (which derives from *grouping*; see also Lerdahl and Jackendoff, 1983). On the semantics side, we proceed to ask what kind of meaning may be encoded in dance. This type of research being foundational, we start by carrying out a production study to establish the range of possible factors that we can investigate. A long-term goal of this investigation is to establish the common semantic properties of (non-) linguistic cognitive systems.

1.1. Co-reference and disjoint reference in dance?

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 approaches. They can either try to make generalizations across genres (e.g., Charnavel, 2016; Schlenker, 2017a) or focus on a case study (see Katz and Pesetsky, 2011, who zoom in on Western art music as instantiated by the works of Mozart and Bach). In our study, we choose the second route, focusing on *Bharatanatyam*,[2] a narrative dance form (outlined in section 1.2 of this paper). This allows us to draw on the insights of Abusch (2013, 2014, 2015) with regards to visual narrative. 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) for music; these may be manifested in dance through different types of spatiotemporal movement descriptors, e.g. the *quality* of a given movement may be described as "smooth" vs. "jagged" (see for example Guest, 2004, and Napoli and Kraus, 2015, for overviews on the parameters of dance and movement).

Our strategy to approach dance semantics was to single out a phenomenon that we could investigate by means of a production study, namely coreference vs. disjoint reference. The encoding of coreference and disjoint reference between noun phrases is illustrated (very coarsely) in (1) and (2), 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 (1) and (2), and we take (1a) to roughly have the truth conditions in (1b), whereas (2a) roughly has the truth conditions in (2b). The difference between (1) and (2) that is at the center of our exploration is that (1) introduces a single discourse referent whereas (2) introduces two separate discourse referents (see also Kamp and Reyle, 1993).

(1)   *coreference*
    a.   *A man* came into the room and *that man* closed the window.
    b.   true iff $\exists x[x$ is a man & x came into the room & x closed the window]

(2)   *disjoint reference*
    a.   *A man* came into the room and *another man* closed the window.
    b.   true iff $\exists x[x$ is a man & x came into the room
                  & $\exists y[y$ is a man & y closed the window & $y \neq x]]$

As linguists, we are interested both in the meanings of natural language expressions, such as the sentences in (1a) and (2a), and in how they are compositionally derived from their parts. Another relevant question in formal semantics concerns the difference between types of content (i.e., using the terminology of Potts, 2015, the difference between truth-conditional *at-issue* content and *non-at-issue* content, which encompasses presuppositions, conventional implicatures, and conversational implicatures). We will return to this second question later.

Focusing on the coreference/disjoint reference distinction, Abusch (2013) investigates comics without words (French *sourds*), i.e. purely visual narratives. She focuses on mangas such as

---

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

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 we see an eaglet depicted in panels 32, 33, 34, and 36, 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. In line with Discourse Representation Theory (Kamp and Reyle, 1993), she proposes that the referents in comic panels are existentially quantified, (3a-c), and coreference arises from post-semantic identification of discourse referents in the pragmatics (which is a type of pragmatic enrichment), (3d). Such existential quantification is plausible in visual narratives, as there are no definite descriptions comparable to *the eaglet* in natural language.

(3)　　*coreference in comics without words (Abusch 2013)*
　　　a.　　panel 34: "[*an eaglet*]$_1$ bounced down a cliff face"
　　　b.　　panel 35: "[*a bobcat*]$_2$ looked and opened its mouth"
　　　c.　　panel 36: "[*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"

Crucially, the questions and insights that Abusch (2013) addresses for comics without words should 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.

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. Coreference vs. disjoint reference is a very basic and fundamental distinction in natural language semantics. As a consequence, if we find that it can be encoded in dance, this can be taken to lend initial support to a view that a semantics of dance may be conceivable.

1.2. Enter Bharatanatyam

Bharatanatyam is a classical South Indian dance that originates in Tamil Nadu (see Puri, 1986, 2004; Williams, 2003; Ramesh, 2013, 2014); it 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 or spoken word (e.g., singing of the narrative), it is not necessarily accompanied by music, and we recorded our stimuli (described in section 2) without music.

Traditionally, Bharatanatyam is used to tell 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 around 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 danceform. 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 following set (Ikegami, 1971:373):

(4)   *possible meanings associated with the patāka ('flag') mudra* (Ikegami, 1971:373)
       '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 focused 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 spontaneous dance, which is relevant for future studies that build on our findings.[3]

To move away from low-level symbols such as hand gestures (which may simply have a sign-based semantics), our strategy was to look at more abstract and global types of meaning such as the coreference/disjoint reference distinction. We now proceed with describing the setup of our explanatory production study.

2. Experimental design

In our investigation of Bharatanatyam, we are working with Kavitha Laxmi, who is the artistic director of the Kala Saadhana dance institute in Oslo and a professional Bharatanatyam dancer.[4] For our exploratory production study, we recorded dance sequences based on a set of items that we constructed in order to probe for coreference vs. disjoint reference. We designed our stimuli as short narrative texts. The items were designed in a way that aims to utilize conventional meanings such as the ones associated with hand gestures, illustrated in (4) above (including objects such as 'palmyra leaf', cf. (7)). The context for all items is given in (5); this context (an artist having designed a statue for a temple) was based on recent dance productions at the Kala Saadhana dance institute with the aim of limiting artificial components in the narrative that are solely due to the experimental design. What is

---

[3] 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.
[4] For illustration, a dance sequence can be found at https://www.youtube.com/watch?v=O-LpIysAKE4

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.

We recorded 6 mini-narratives in 2 conditions (coreference vs. disjoint reference), i.e. 12 dance sequences in total. Two sample narratives are given in (6) and (7). This setup allows us to elicit minimal pairs in our production study. 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 studies, we included unembedded variants (e.g. 'A woman is sitting on the ground. […]'). 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, but it can also be danced without them. We recorded our stimuli without music.

The dance sequences were recorded in the Music and Motion Lab of the Department of Musicology, University of Oslo. The 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 advantages of such a production study is that we can compare minimal pairs 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. 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 1. 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).

While we limit the discussion in this paper to a qualitative analysis, the methodology (motion capture) lends itself to quantitative follow-up analyses using numerical methods (see Kelkar and Jensenius, 2018, for an example).
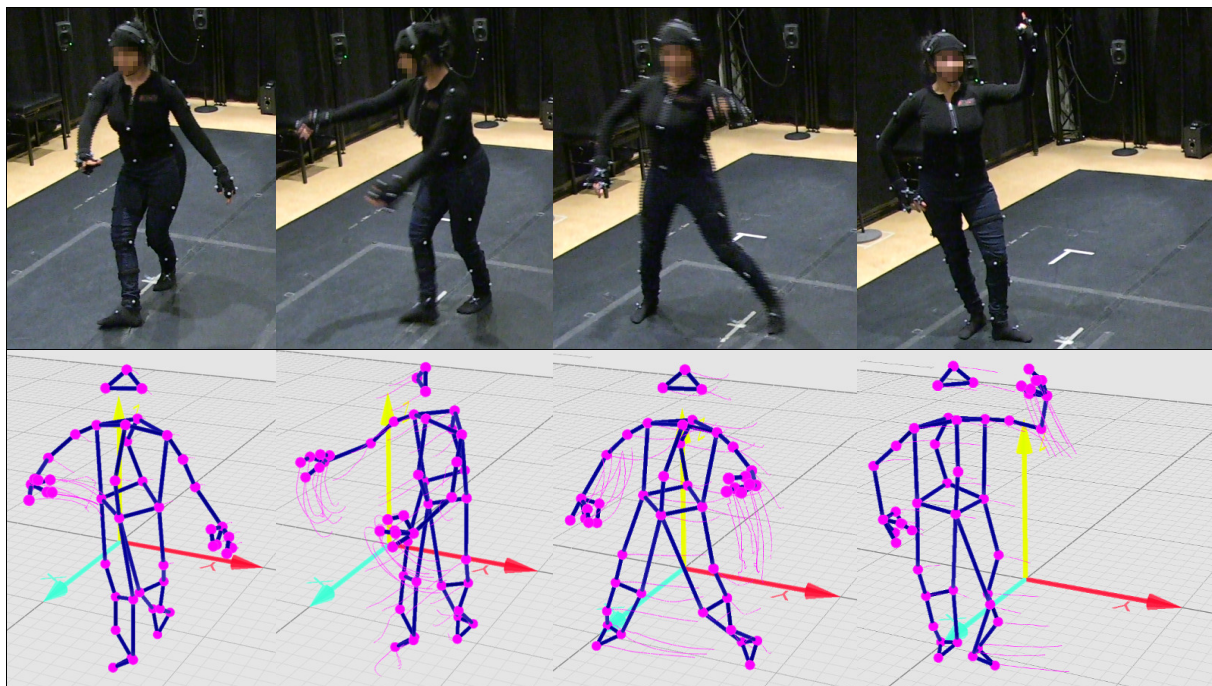

Figure 1: sequence of four dance positions (stills from the video recording and 3D motion capture rendering, with motion history trajectories)

In the next section, we proceed with a qualitative analysis of the results.

3. Qualitative analysis of the results

We start by analyzing the coreference sequence, (6a), adapted in (8); as shown in Figure 2, we can zoom in on the movement and study different parts. In Figure 2, each label $[P_n]$ represents a dance position; these positions are stipulated at arbitrary cut-off points, 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].

Figure 2: coreference condition
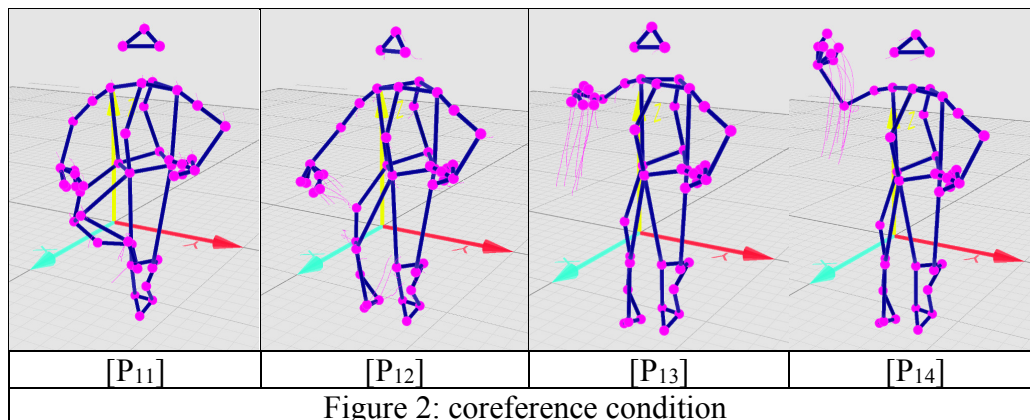
What Figure 2 shows is that the coreference condition simply involves a fluid motion from displaying a sitting position to displaying a spear-holding position. It does not seem to be necessary (in the given context) to separately mark coreference between the "sitter" and the "spear holder". By contrast, the disjoint reference condition, repeated in (9) from (6b), has additional complexity, as illustrated in Figure 3. 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 2 and [P_{21}] in Figure 3, 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 *mudra* (here: 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 2), 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].

Figure 3: disjoint reference condition
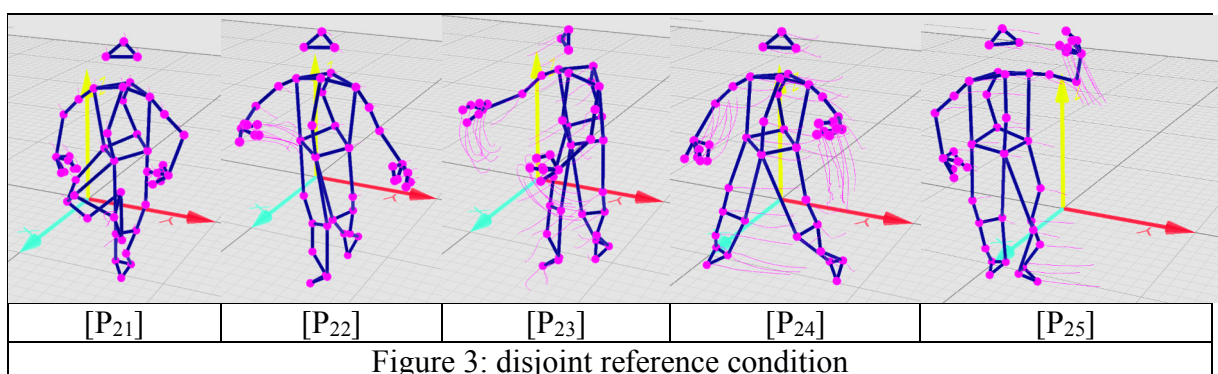
The marking of a new position on stage (and thus in the visual space), [P_{23}] is a phenomenon that is reminiscent, from a linguist's perspective, of the loci in sign language (see Schlenker,

2017c, for a recent survey article), opening new lines of inquiry for follow-up studies. Assuming the new position also appears reminiscent of phenomena such as *Action Role Shift* in sign language (see Davidson, 2015, for a recent discussion).

For present purposes, we take the sequence in [P22]-[P24] to be crucial for an understanding of how disjoint reference, in particular, can be encoded in dance. While we have not yet carried out perception studies on the basis of these data, we carried out an informal pilot study in which two of our research assistants (who had not yet seen the dance sequences) watched the dance sequences in random order with the task of determining for each sequence whether it described a story about a single individual or two separate individuals. Even after being informed about the 'another'-symbolizing mudra, they found this mudra difficult to track. Instead, they reported that the introduction of a new position on stage was a major cue for introducing a second individual, while the absence of such a new position implied the lack of such a second individual. We thus expect this to be a feature of the dance that may carry over into other narrative dance forms, and possibly also into non-narrative dance forms as long as the dancer pursues an intention of narrating a story. Section 4 shows how a semantics of dance could be formally implemented (drawing on Abusch, 2015) and which role *grouping* may play in a dance semantics (drawing on Charnavel, 2016). Section 5 briefly returns to the mirroring that we observed in comparing [P25] in Figure 3 to [P14] in Figure 2.

4. Theoretical interpretation of results

In theoretically interpreting the findings, one factor that is clearly relevant (though it has not previously been connected to semantic interpretation) is the notion of *grouping* (see Lerdahl and Jackendoff, 1983). Charnavel (2016) argues that dance shares hierarchical grouping with language and music, which may be interpreted as giving rise to a syntactic structure of dance.

Building on Lerdahl and Jackendoff (1983), Charnavel (2016:13) posits five *grouping well-formedness rules* three of which are quoted in (10).

(10)   *grouping well-formedness rules (Charnavel, 2016:13)*
    a.   *GWFR2:* A dance constitutes a group.
    b.   *GWFR3:* A group may contain smaller groups.
    c.   *GWFR5:* If a group G1 contains a smaller group G2, then G1 must be exhaustively partitioned into smaller groups.

By virtue of the grouping well-formedness rules in (10a-c), grouping serves to create hierarchical structure (in the sense of an exhaustive partitioning of a dance sequence into sub-sequences [groups], which may, in turn, be partitioned into further sub-sequences [groups]). What becomes central at this point are the *grouping preference rules* that determine the sub-sequences (or constituents) of a dance sequence. Charnavel (2016) proposes fifteen such grouping preference rules, of which the most relevant (for our purposes) are given in (11).

(11)   *grouping preference rules (Charnavel, 2016:18,19,24)*
    a.   *GPR1* (change of direction): Consider a sequence of positions p_1, p_2, p_3, p_4, p_5, p_6. The transition p_3-p_4 may be seen as a group boundary if the path formed by p_1-p_2-p_3 does not have the same direction as the path formed by p_4-p_5-p_6.

    b.   *GPR2* (change of orientation): Consider a sequence of positions p_1, p_2, p_3, p_4. The transition p_2-p_3 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_3-p_4.

    c.   *GPR10* (intensification): When the effects picked out by the local rules of change (GPR1-GPR8) are relatively more pronounced, a larger-level group boundary may be placed.

In section 6, we lay out the hypothesis that hierarchical grouping in narrative dance can be mapped to situation structure. We can start by asking how grouping may be used to convey disjointness (e.g. disjoint reference when two characters are introduced into a narrative), as specifically instantiated by Bharatanatyam. We propose the informal rule in (12).

(12) *grouping-based coreference* (first approximation)
    a.   In the absence of a group boundary, a dance sequence $[P_n]$-$[P_{n+1}]$ is interpreted as continuous (e.g. describing a narrative about a single individual).
    b.   If there is a group boundary between two dance positions $[P_n]$ and $[P_{n+1}]$, then a dance sequence $[P_n]$-$[P_{n+1}]$ is interpreted as discontinuous (e.g. describing a narrative about two separate individuals).

The workings of grouping-based coreference is illustrated for the disjoint reference condition in Figure 4, where an orientation shift occurs between positions $[P_{23}]$ and $[P_{24}]$. A reader may wish to verify that such an orientation shift does not occur in the coreferent condition, given in Figure 2 above. In terms of Charnavel's grouping preference rules, (11), it is not completely clear whether the rule at work is GPR1 (change of direction) or GPR2 (change of orientation), as a change of direction seems to be combined with a change of orientation in this sequence; however, it is clear that the change from $[P_{23}]$ to $[P_{24}]$ is quite pronounced, in line with GPR10 (intensification).

[P₂₁]	[P₂₂]	[P₂₃]	[P₂₄]	[P₂₅]
↙	↙	↙	↘	↘

→ *time*

Figure 4: orientation shift in the disjoint reference condition
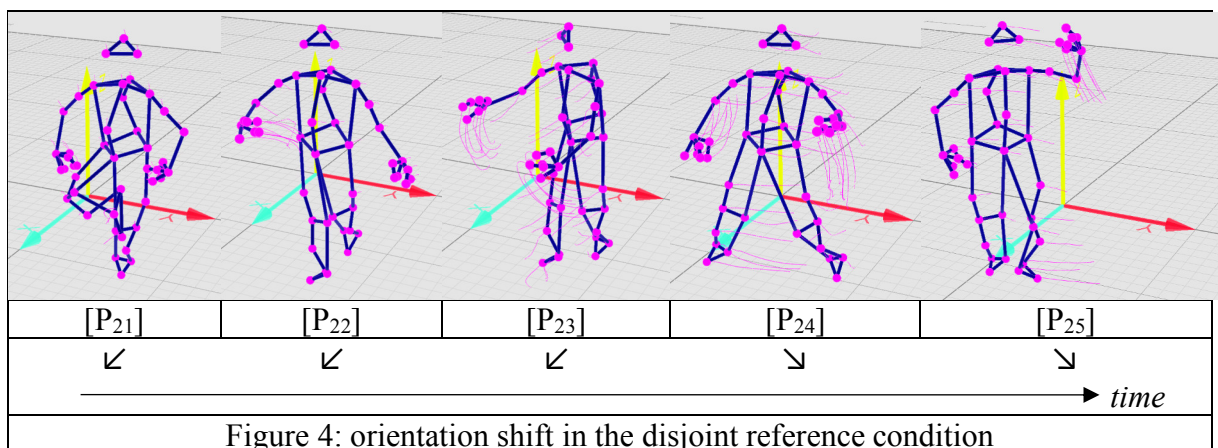
Crucially, if we factor in smaller changes in the dance sequence as group-inducing (at a lower level), then we can posit at least a three-level hierarchical structure, as given in (13), using Charnavel's notation. For the purpose of illustration, we assume that each of the positions in Figure 4 is associated with a low-level group boundary, given that the orientation direction of body parts constantly changes (hand-and-arm in $[P_{21}$-$P_{22}]$, upper body and arms in $[P_{22}$-$P_{23}]$, and so forth). The role of *global (whole-body) gestures* comes into play in connection with

Charnavel's GPR10, since such gestures are generally more intense than gestures that only involve individual body parts. In line with GPR1 and GPR2, as stated in (11), we 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).

(13)   *structure of the disjoint reference dance sequence*
          P$_{21}$ P$_{22}$ P$_{23}$ P$_{24}$ P$_{25}$
          |---|---|---|---| low-level grouping
          |---------|-----| larger-level grouping
          |---------------| top-level grouping (complete dance)

The core idea here is that group boundaries themselves appear to be meaningful in narrative dance in that they signal *discontinuity*; we expect to find similar effects in other (non-narrative) dance forms.

5. Towards a formal semantic analysis of narrative dance positions

We now take steps towards a formal semantic rendering of the generalizations in section 4. An important first step consists in defining how exactly we should approach the semantics of pictures, i.e. how we could define truth in a visual narrative. In order to answer this question, we build on Abusch (2015), who posits a generalized possible worlds model for informational entities; her idea is that any sentence, picture, etc., counts as an informational entity when it rules out some possibilities, based on the definition in (14).

(14)   *possible worlds model of information content* (Abusch 2015:2)
   any informational entity such as a sentence or picture rules out some possibilities
          [= possible worlds, situations, or scenes] and admits others

Let us illustrate Abusch's idea for the dance position [P$_{21}$] in Figure 4. (This example is closely modeled after Abusch's own example that involves two octahedrons.) 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 (14)), namely ones in which there is no person, or in which the person is not sitting. The statement in (15) is thus understood to provide new information about a given situation that we are describing.

(15)   There is a person who is sitting.

Crucially, Abusch argues that a picture achieves exactly the same result. In parallel to (15), the dance position in (16) can be understood to provide new information about a given situation (namely the current point in time in a narrative that is being told). As Abusch observes, when it comes to the question of what a world or situation is like, (16) 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 (16) thus qualifies as an informational entity in line with (14). Abusch is careful to point out that pictures are often more informative than sentences; taken at face value, a naïve observer may infer from (16) 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 (15) can leave such information underspecified; there is no implication from (15) on how exactly the person is sitting.

(16)
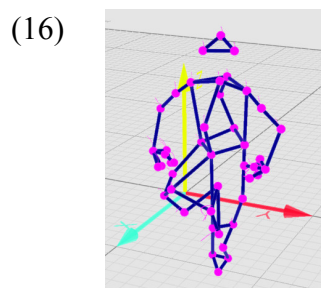
Abusch (2015) proceeds to identify the semantics of a picture with the set of possibilities that it admits. This means that we can define the semantics of a picture in terms of possible worlds, situations, or scenes. Treating any given dance position $[P_n]$ as a picture, we can then posit satisfaction conditions as given in (17).[5] 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 (17) counts as satisfied by a fictional situation σ (i.e. "true" in σ) if a sitting activity is taking place in σ.

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

a situation σ satisfies   only if in σ a person is sitting.
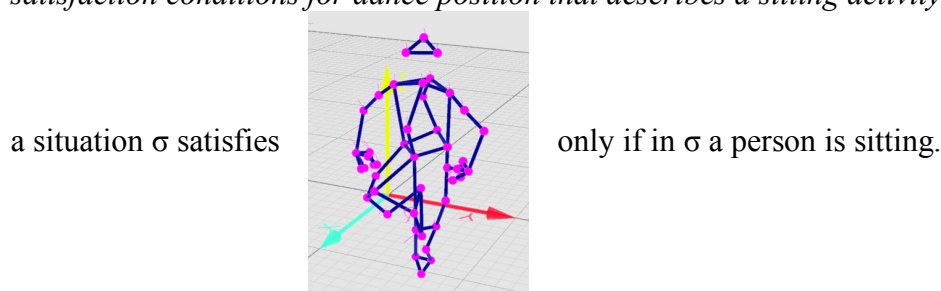
For now, the rendering in (17) is connected to narrative—or figurative—dance, which encodes a visual narrative. However, as long as we allow for more abstract, iconic atoms of meaning, it should be clear how this view carries over to all dance forms, including non-narrative dance forms. Schlenker (2017a) identifies meaning in music with inferences that can be drawn on a (fictional) virtual source (e.g. an increasing volume may symbolize that a source is gaining in size, and/or moving closer). Combining Schlenker's *source-based semantics* with Abusch's picture semantics, we could thus posit more abstract satisfaction conditions such as (18), which corresponds to $[P_{13}]$ in Figure 2. (One interesting future direction to explore in this respect touches on correlations between pitch and vertical motion in the bodily responses of (untrained) listeners to music, as discussed by Kelkar and Jensenius, 2018.)

---

[5] Note that this is glossing over the viewpoint-dependence of pictures. As discussed by Abusch (2013, 2015), pictures are generally related to the objects that they depict by means of projection lines that are oriented towards a given viewpoint.

(18)   *satisfaction conditions for abstract upward movement of a source*
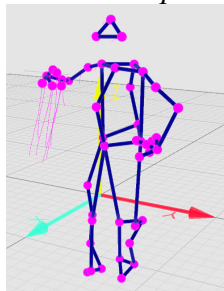
a situation σ satisfies   only if the virtual source in σ is involved
                                                     in a (partial or total) upward movement.

Once we look beyond Bharatanatyam, including non-figurative dance, we may thus need a more iconic semantics in line with Schlenker (2017a). An example of future venues for exploration is given by Kelkar and Jensenius (2018), who outline six ways in which (untrained, i.e., for our purposes, "naïve") listeners move their two hands to accompany music that they are hearing. From the perspective of Schlenker (2017a), it is plausible that the two hands separately (or jointly) represent virtual sources, which may give rise to meaning inferences on part of an onlooker.

6. Towards a formalization of grouping-based coreference

Having established an approach to "truth" in narrative dance (in line with Abusch 2015), we can now proceed to reviewing the rule of grouping-based coreference that we introduced in section 4. To that end, let us reconsider the coreferent dance sequence from Figure 2, repeated in Figure 5. In line with Abusch (2013:12, 2014:10), we posit the satisfaction conditions in (19) to (partially) describe the dance positions in Figure 5. We will henceforth 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 (19) is that dance positions $[P_n]$ are mapped to propositions $[\![P_n]\!]$.[6]

$[P_{11}]$	$[P_{12}]$	$[P_{13}]$	$[P_{14}]$
Figure 5: coreference condition			
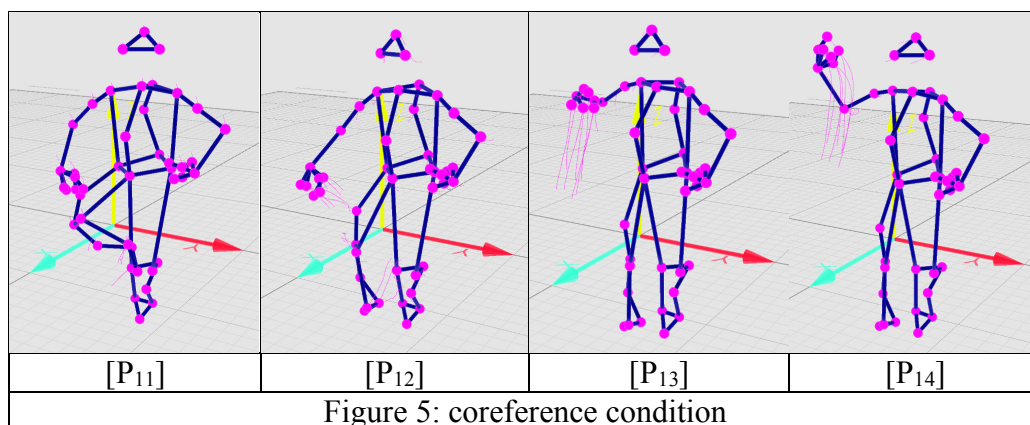
(19)   a.   A situation/scene σ_{11} satisfies $[P_{11}]$ only if in σ_{11} a person is sitting.
        b.   A situation/scene σ_{14} satisfies $[P_{14}]$ only if in σ_{14} a person is holding a spear.

---

[6] Note that, since dance is continuous, discrete positions such as $[P_{11}]$ and $[P_{12}]$ must be stipulated. For now, we keep treating dance positions as static images, but one open question concerns the continuity (movement) between them.

Based on our findings with regards to coreference vs. disjoint reference, we formalize the grouping-based coreference rule (or grouping-based coreference principle) as given in (20), building on Abusch (2013:13). We illustrate this rule below, though it is worth guiding the reader's attention to the phrase 'narratively relevant' in (20a) and (20b); as of yet, this is a notion that we leave undefined, to be further explored at a later stage. The underlying intuition is that we are only concerned with situations that are delimited and separated by larger-level grouping boundaries in the sense of (13). This is crucial since a complete narrative dance will of course always describe larger situations that contain *all* dance sequences [P_n] that it contains, i.e. if we were to eliminate 'narratively relevant' from (20), the rule would become void.

(20) *grouping-based coreference* (second approximation)
   a.   In the absence of a group boundary, a dance sequence [P_n]-[P_{n+1}] is interpreted as corresponding to a larger *narratively relevant* situation σ_{top}.
   b.   If a narrative dance sequence corresponds to a single *narratively relevant* situation σ_{top} and contains two similar entities α and β, coreference (i.e. $\alpha = \beta$) arises by default when there is no indication that parts of σ_{top} contain more than one entity of this type.

As Abusch (2013) points out, the identification of entities in a single situation, $\alpha = \beta$, may well reflect low-level processes of indexing in vision, see Pylyshyn (2003); as pointed out by Abusch, such extra-linguistic (or pre-linguistic) processes are not in contradiction with the formal semantic approach that we (and Abusch) pursue.

We can now proceed with the coreference sequence in Figure 5 and render (19) as given in (21). We have already established the two satisfaction conditions in (21a) and (21b). By grouping-based coreference, we now derive a larger narratively relevant situation σ_{top} in (21c) (loosely based on Abusch, 2013); this is a situation that has a subpart σ_{11} and a subpart σ_{14}, which each involve existential quantification over a person (α and β, respectively). Since both are part of the same overarching narratively relevant situation, we can, by (20b), identify α and β.

(21) a. σ_{11} satisfies [P_{11}] only if in σ_{11} a person α is sitting.
   b. σ_{14} satisfies [P_{14}] only if in σ_{14} a person β is holding a spear.
   c.   *by grouping-based coreference*
        a narratively relevant situation σ_{top} satisfies [P_{11}-P_{14}] only if σ_{top} has a part σ_{11} such that in σ_{11} a person α is sitting, and σ_{top} has a part σ_{14} such that in σ_{14} a person β is holding a spear [via (20a)] and $\alpha = \beta$ [via (20b)].

The important part here is that [P_{11}] and [P_{14}] in Figure 5 are not separated by a larger-level grouping boundary. Contrast this with the disjoint reference condition in Figure 6, adapted from Figure 4. Here, a larger-level group boundary is introduced between [P_{23}] and [P_{24}] due to a change in orientation and direction.

[P21]	[P22]	[P23]	[P24]	[P25]
↙	↙	↙	↘	↘

| – | – |

Figure 6: grouping in the disjoint reference condition
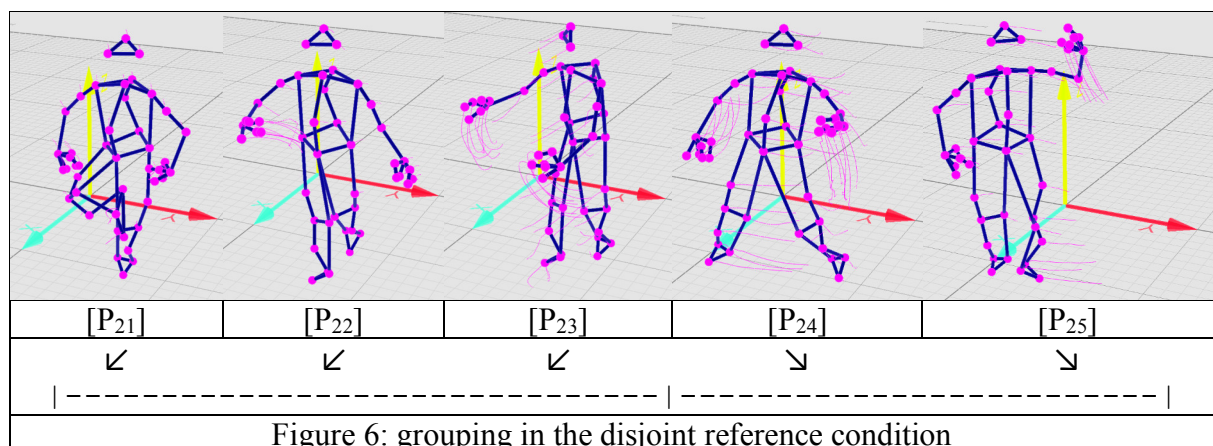
The satisfaction conditions in (22a-b) are equivalent to those in (21a-b). (We return to the mirroring of the spear holding in [P25] vs. [P14] below.) The crucial difference is that the group boundary between [P23] and [P24] (which, by transitivity, counts as a group boundary between [P21] and [P25]) blocks the inference to a larger narratively relevant situation σ_{top}.

(22) a. σ_{21} satisfies [P21] only if in σ_{21} a person α is sitting.
       b. σ_{25} satisfies [P25] only if in σ_{25} a person β is holding a spear.

Grouping alone may thus be sufficient to block coreference (i.e. referent identification) in a simple narrative like this one, i.e. identification of the agent in the two situations σ_{21} and σ_{25}. An open question (at this point) concerns the exact nature of narratively relevant situations. Since situations are recursively embedded in larger situations, any visual narrative of the type in Figure 6 will contain one larger (non-narratively-relevant) situation that contains σ_{21} and σ_{25}. For present purposes, we exclude such top-level situations, but eventually we aim to determine more precisely which levels matter.

Note that, much in line with Schlenker's (2017a) ideas for the syntax/semantics mapping in music, we propose that grouping in dance serves as a way to organize (sub-)events. Specifically, the introduction of larger-level group boundaries serves to signal discontinuity. Such a signal can have different functions; in other words, it is not necessarily the case that every single grouping boundary indicates a change of character; yet, it is quite plausible that every change of character requires a grouping boundary to be placed.[7]

We can now conclude the discussion of grouping-based coreference and disjoint reference, and move on to a separate question, asking about the types of meaning that are encoded in such dance sequences.

7. Mirroring and the question of *at-issueness in dance*

If we review Figure 6, we observe that disjoint reference is encoded at several levels (going beyond inferences based on grouping). First, the dancer uses a designated *mudra* (hand-and-arm gesture) that symbolizes 'another, a different', as visible in [P22]. She then explicitly introduces a new locus on stage, as visible in [P23], which she then assumes, in [P24]. From

---

[7] We are grateful to an anonymous reviewer for flagging this point.

the perspective of Abusch's (2013, 2014, 2015) picture semantics, as applied to visual narrative in dance, we could posit satisfaction conditions such as (23).

(23) a. σ_{22} satisfies [P$_{22}$] only if in σ_{22}, there is an individual y such that y is distinct from the most salient individual x.
  b. σ_{23} satisfies [P$_{23}$] only if in σ_{23} there is a virtual locus i.
  c. σ_{24} satisfies [P$_{24}$] only if in σ_{24} the narrative is *at* the virtual locus i.

In terms of narrative progression, each of (23a-c) seems to redundantly encode disjoint reference (in addition to what is already achieved by grouping). Note that this does not void grouping-based coreference, (20), which was intended as a more general rule (or principle) of narrative dance that also applies beyond conventional aspects of bharathanatyam. However, there are open questions with regards to, in particular, (23b-c): what, if any, is the shared cognitive underpinning of virtual loci in narrative dance on the one hand, and the loci of Sign Languages on the other hand (see Schlenker, 2017c, for an overview)? A particularly promising idea in this regard may be the hypothesis that even sign language loci may at times be "iconic depictions of their denotations" (Schlenker, 2017c:174, building on research such as Liddell, 2003, and the work by Judy Kegl, as in Neidle et al., 2000), in parallel to the dancer's *virtually assuming* of the position associated with the new locus in [P$_{24}$].

More intriguingly, for present purposes, is the role of "mirroring" in [P$_{25}$], which is illustrated in direct comparison in Figure 7. An initial hypothesis could be that this is a trivial reflection of the orientation change. However, in a post-experimental debriefing with the dancer, this mirroring forms an additional part of ensuring that an audience can follow the narrative, i.e. it is a designated means of marking disjoint reference (in addition to (23a-c)).

[P$_{14}$] *coreference*	[P$_{25}$] *disjoint reference*
Figure 7: mirroring in Item 1	
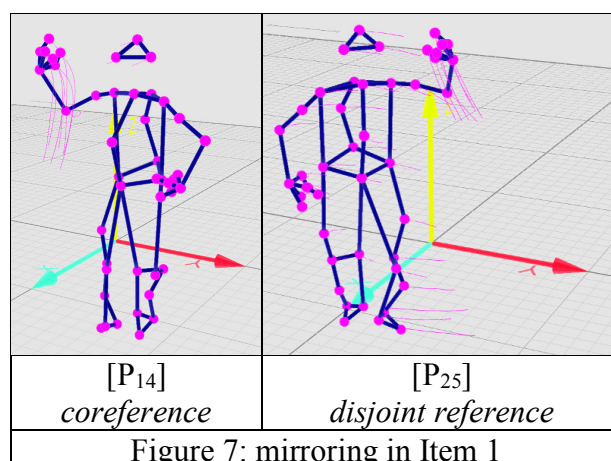
This naturally raises the question of where in the semantics mirroring could be represented. The satisfaction conditions for [P$_{14}$] and [P$_{25}$] are restated in (24a-b), repeated from (21b) and (22b). Crucially, what mirroring may contribute is a non-at-issue inference, as given in (24c) for [P$_{25}$].

(24) a. σ_{14} satisfies [P$_{14}$] only if in σ_{14} a person β is holding a spear.
  b. σ_{25} satisfies [P$_{25}$] only if in σ_{25} a person β is holding a spear.

c.   *non-at-issue inference (modeled as a definedness condition)*
        σ_{25} is defined for [P_{25}] iff the agent of the activity in σ_{25} is distinct from the most salient individual in the current narrative.

Of course, (24c) looks at dance from a linguistic angle, and it is difficult to argue that visual narratives contain something akin to presuppositional content. (For instance, it is rather difficult to conceive how tests for projective content could adequately be carried out.)[8] Nevertheless, the relationship between [$P_{22\text{-}24}$] and [P_{25}] could be likened to that of S_1 and S_2 in (25). It is a standard assumption that S_2 presupposes the same information that S_1 asserts, due to the definite description (with possessive pronoun) *his sister*. Similarly, we conjecture that [P_{25}] may presuppose the same information that is conveyed 'at issue' in [$P_{22\text{-}24}$].

(25)   [$_{S1}$ Bill has a sister.] [$_{S2}$ And *his sister* lives in Tromsø.]

In the linguistics literature, there are precedents for non-at-issue content being conveyed outside of speech. For instance, in the realm of speech-accompanying gestures, co-speech gestures (which accompany spoken words) have been argued to encode non-at-issue meaning (see Schlenker, 2017b; Tieu et al., 2017), and Schlenker (2017d) argues that pro-speech gestures (which replace spoken words) can trigger presuppositional inferences (amongst other types of inferences that they can trigger). Moreover, since sign language loci and gestural loci can be linked to presuppositional content (e.g., Schlenker, 2017c:170-171), even the sequence in [P_{22}-P_{24}] may be analyzed at the level of non-at-issue meaning.

Before concluding, it is also worth commenting more on the exact rendering of the inference in (24c). An anonymous reviewer points out that grouping breaks (which we discussed in section 6) can be seen as having "the discourse semantic function [of] introducing a new center" (corresponding to the management of a stack of entities in a dynamic semantics). S/he inquires what the type of these centers should be ("a protagonist, a location, a situation, or a combination of them"). For the purposes of the Bharatanatyam narrative that we have been working with, the center seems to be a protagonist/character in the narrative (rather than a location or situation). However, in a broader view (moving beyond Bharatanatyam) it is plausible that centers are more abstract corresponding to virtual sources in the style of Schlenker (2017a); a larger-level grouping boundary would then indicate a shift from one virtual source (on a stack of contextually given entities) to another virtual source.

8. Conclusion

In this paper, we presented an exploratory production study to investigate the encoding of coreference vs. disjoint reference in Bharatanatyam, a figurative (narrative) dance that serves to tell a story. We maintained that a formal semantics of narrative dance is possible in line with Abusch's (2013, 2014, 2015) approach to the semantics of visual narrative. While our

---

[8] One test for (non-)at-issueness that may be fruitful involves embedding under negation (or under a negative predicate such as *deny*; we are grateful to Philippe Schlenker for pointing this out). For Item 1, as illustrated in (6b), a relevant item may thus be designed as in (i) below, where the prediction is that the inference in (24c) is not canceled in such a context. We plan to explore such items in future follow-up studies.
i.   The witness describes that a strong man was sitting on the ground.
        Then she denies that another man was holding a spear.

analysis was closely tied to the nature of Bharatanatyam, we maintain that at least two aspects of the analysis would carry over to other (non-narrative) dance forms:

[i.] dance positions [P_n] (as arbitrary discrete moments in dance sequences) can be described by means of satisfaction conditions that involve a fictional virtual source in the spirit of Schlenker (2017a), as illustrated in (18).

[ii.] a larger-level group-boundary between two dance positions [P_n] and [P_{n+1}] (as triggered in line with grouping preference rules as posited by Charnavel, 2016) triggers discontinuity inferences, which may, for instance indicate non-identity of two virtual sources, see Figure 6.

Open questions to be explored in future studies concern the question of what happens if more than two characters are introduced into a Bharatanatyam narrative; specifically, a question that has arisen from the discussion so far is the extent to which loci in Bharatanatyam can be likened to loci in sign languages. Another goal is to move on to non-figurative dancing and test whether the insights from this study carry over (and to what extent). Here it could also be interesting to investigate to what extent expressive qualities (cf. Krumhansl and Schenck, 1997), as seen in the spatiotemporal features of the gestures (e.g. "jaggedness" in the quality of movement), complement or contradict some of the linguistic meanings. Future studies also need to move on from production to perception, to investigate how observers (both trained and untrained) interpret a given dance sequence, e.g. if and how observers can recognize whether a narrative dance sequence involves one or more protagonists. Moving beyond a sophisticated dance form such as Bharatanatyam, such perception studies would benefit from a shift towards simplified dances/gestures (e.g. using the medium of a simple dance form such as "finger dance" to construct stimuli; see Charnavel, 2016:fn.11).

References

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. (2015). Possible worlds semantics for pictures. Manuscript, December 2015, Cornell University. [http://hdl.handle.net/1813/44654]

Charnavel, I. (2016). Steps towards a Generative Theory of Dance Cognition. Manuscript, June 2016, Harvard University. [http://ling.auf.net/lingbuzz/003137]

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

Guest, A. H. (2004). *Labanotation*. New York: Routledge.

Heim, I. (1982). The Semantics of Definite and Indefinite Noun Phrases. PhD dissertation, UMass Amherst.

Ikegami, Y. (1971). A stratificational analysis of the hand gestures in Indian classical dancing. *Semiotica* 4, 365–391.

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]

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]

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.

Lerdahl, F. and R. Jackendoff (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.

Napoli, D. J. and L. Kraus (2015). 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.

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.

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 theater. 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.

Schlenker, P. (2017a). Outline of Music Semantics. *Music Perception* 35, 3–37.

Schlenker, P. (2017b). Gestural Grammar. Manuscript, June 2017, Institut Jean-Nicod / ENS / NYU. [http://ling.auf.net/lingbuzz/003497]

Schlenker, P. (2017c). Sign Language and the Foundations of Anaphora. *Annual Review of Linguistics* 3, 149–177.

Schlenker, P. (2017d). Gestural Semantics. Manuscript, October 2017, Institut Jean-Nicod / ENS / NYU. [http://ling.auf.net/lingbuzz/003485]

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.

Williams, D. (2003). In the Shadow of Hollywood Orientalism: Authentic East-Indian Dancing. *Journal for the Anthropological Study of Human Movement* 12, 78–99.