# Steps towards a formal semantics of dance

## 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 the nature of non-linguistic language-like systems. In this paper, we approach this goal by looking at narrative dance in the form of *Bharatanatyam*, which we treat as a type of visual narrative (developing the formalism pioneered in Abusch 2013 for printed comics without words). We first apply the formal linguistic method of eliciting minimal pairs to dance sequences that narrate short texts with one or more referents – the creation of minimal pairs in dance being a new endeavor at the level of experimental methodology; we then use the formalism of situation semantics and event semantics (see Kratzer 2020 for an overview article, and references therein) in order to capture the meanings that can be conveyed and/or inferred.

We argue that a formal semantics of dance can be modeled closely after a formal semantics of visual narrative (Abusch 2013) and incorporate features present in a formal semantics of music (Schlenker 2017a). One of our conclusions is that dance not only shares properties of these other fundamentally human means of expression, but that it also potentially incorporates rudimentary aspects of expressions that we find in sign languages (such as the marking of *referential loci* in the dancer's "signing space", cf. Lillo-Martin & Klima 1990, Liddell 1990). From the perspective of general human cognition, these conclusions further corroborate the idea that linguistic investigations beyond language (which we will call *Super Linguistics* (see Schlenker & Patel-Grosz 2018), using the word *super* in its original Latinate meaning, 'beyond') can yield insights into the very nature of the human mind and how it relates to the minds of other non-human animals.

*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'; Schlenker & Patel-Grosz 2018). Building on the pioneering work of Lerdahl & Jackendoff (1983), recent applications of linguistic methodology to music are instantiated by Katz (2017) and Schlenker (2017a), amongst others. 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 (including non-human primates).

The present investigation is part of the larger *Super Linguistics* program, and thus involves the application of linguistic methods to non-standard objects, i.e. objects of inquiry outside the realm of natural language. As such, we can outline our methodological assumptions as follows. Both the linguistic procedure of establishing

and eliciting minimal pairs[1] (which we apply in section 2) and the analytical framework of applying a formal semantic analysis (which we apply in section 3), qualify as part of linguistic methodology. The goal here is to show that these can be applied to narrative dance, which is a new object of study from a linguistics perspective, and to outline future applications to non-narrative dance. Finally, in section 4, we discuss commonalities between narrative dance and sign language, thus opening another new line of inquiry for future studies of body movement and the human mind, the eventual goal being to explore if mechanisms of communication in dance overlap with human language.

## 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 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, who zoom in on Western art music as instantiated by the works of Bach, Mozart, Saint-Saëns and Strauss; see also Charnavel 2016, who focuses on ballet and modern dance). In our study, we choose the second route, focusing on *Bharatanatyam*,[2] 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] While the construction of 'minimal pairs' may seem to have applications that are much broader than linguistic analysis (thus qualifying as 'basic science', so to say), pairwise comparison of expressions that only minimally differ is a central part of linguistic methodology, which goes 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'"). It 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).

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

(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 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 dance by untrained participants, which is relevant for future studies that build on our findings.[3]

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.[4] 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. We now proceed with describing the setup of our explanatory 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.

---

[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] 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."

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.[5]
    b. true iff $\exists x[x$ is a man & x is sitting on the ground
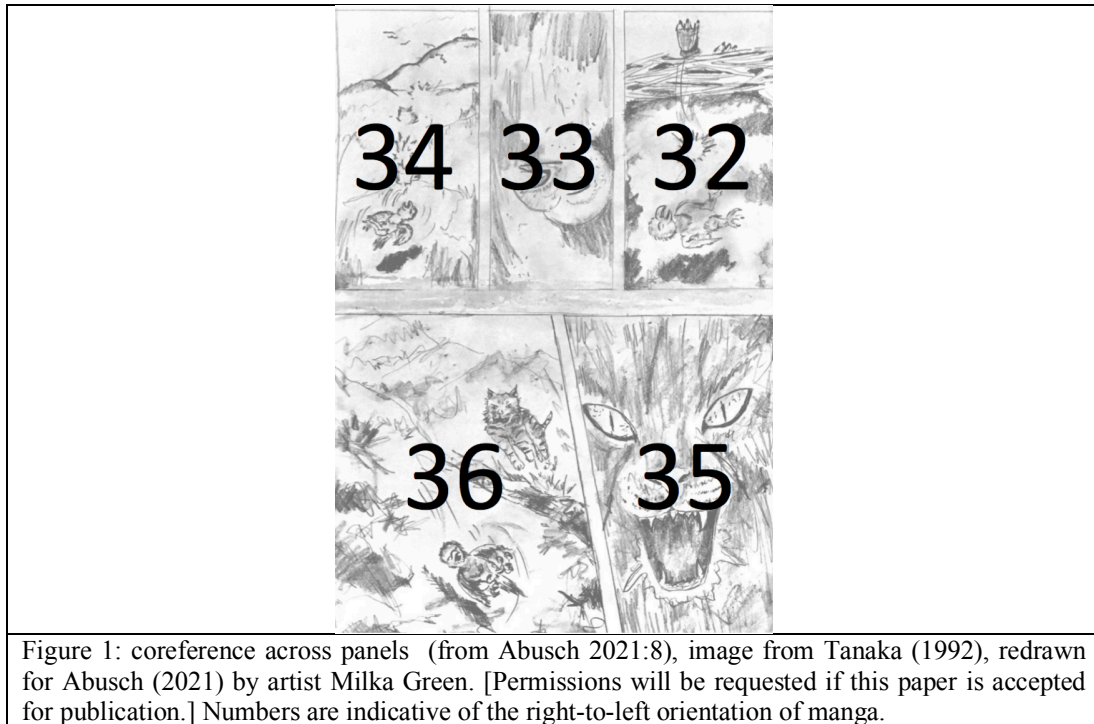        & $\exists y[y$ is a man & y is holding a spear & $y \neq x]]$

Abusch (2013) investigates comics without words (French *sourds*), i.e. purely visual narratives.[6] 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, 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).

---

[5] There is a non-trivial question of how the sentence in (3) relate 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.
[6] See also Cohn (2018, 2019, 2020) on the cognitive relevance of comics, and specifically on the syntax (hierarchical structure) of comics.

Figure 1: coreference across panels (from Abusch 2021:8), image from Tanaka (1992), redrawn for Abusch (2021) by artist Milka Green. [Permissions will be requested if this paper is accepted for publication.] Numbers are indicative of the right-to-left orientation of manga.

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, (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)*
   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"

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.] a situation semantics for comics without words, which becomes useful not only for analyzing individual moments in a dance sequence, but also for analyzing co-reference and disjoint reference in a visual narrative (Abusch 2013:13), 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 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) 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 chosen to be as natural as possible, with the aim of limiting artificial components in the narrative that are solely due to the experimental 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.

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. […]').[7] 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. 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 Music and Motion 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

---

[7] The intention behind fixing a perspectival center 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 *describe* the narrative by means of dance, rather than *acting out* the narrative. 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, (41).

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. 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.[8] 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).
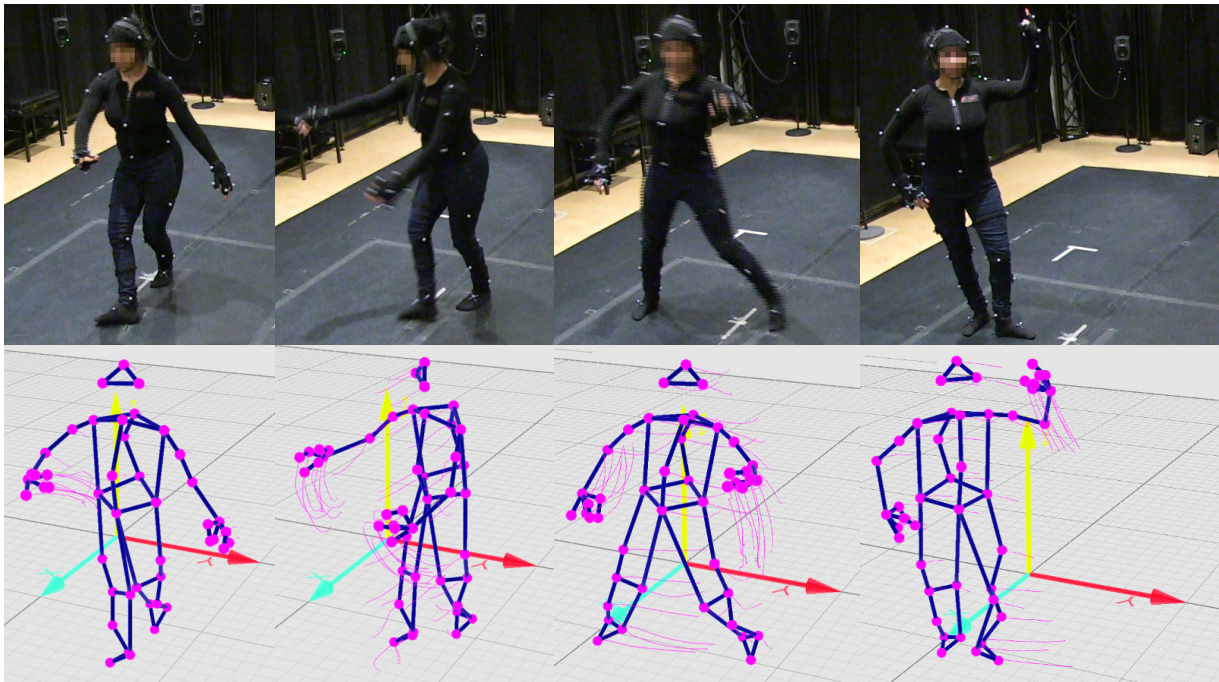


Figure 2: sequence of four dance positions (stills from the video recording and 3D motion capture rendering, with motion history trajectories)

---

[8] 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.

## 2.4    The production data

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 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].



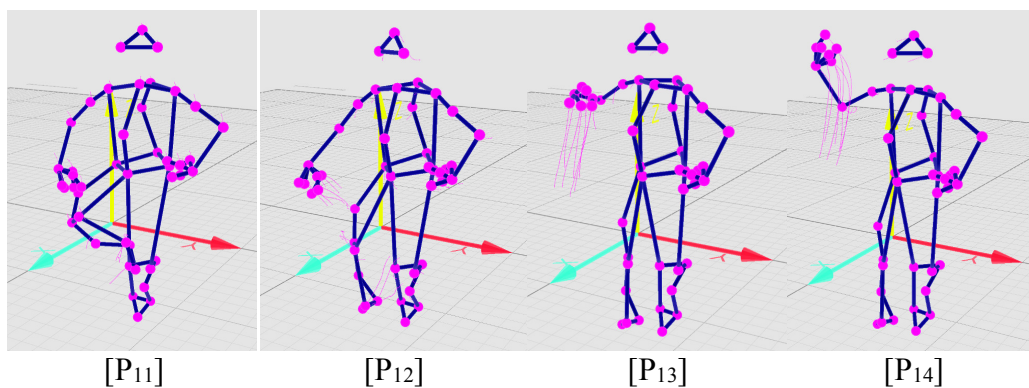[$P_{11}$]             [$P_{12}$]             [$P_{13}$]             [$P_{14}$]

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 *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 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].
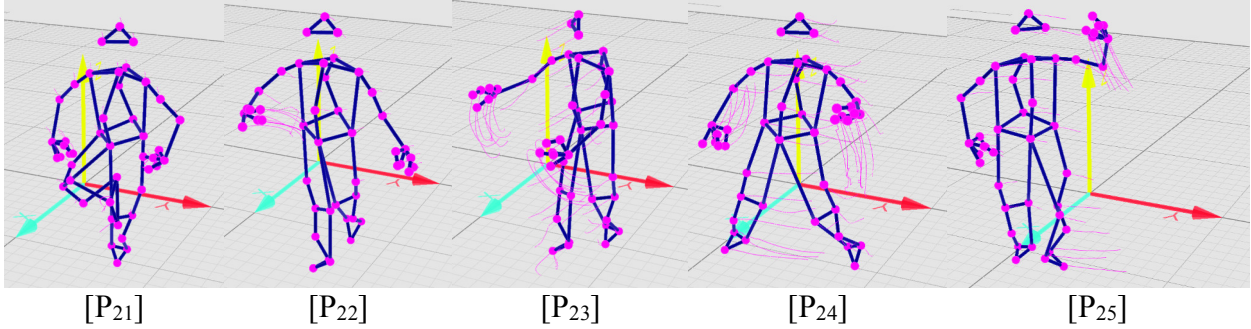
| [P₂₁] | [P₂₂] | [P₂₃] | [P₂₄] | [P₂₅] |

Figure 4: disjoint reference condition

Table 1 summarizes the presence and absence of these four cues ('another' in [P₂₂], pointing in [P₂₃], movement in [P₂₄], and mirroring in [P₂₅] *vs*. [P₁₄]) 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).[9] 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). 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 (14) and Figure 7, 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' mudra | – | – | – | – | – | – | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| pointing gesture | – | – | – | – | – | – | ✓ | – | ✓ | ✓ | ✓ | ✓ |
| move to new position | – | – | – | – | – | – | ✓ | ?[10] | ✓[11] | ✓ | ✓ | ✓ |
| mirroring the posture | – | – | – | – | – | – | ✓ | ✓ | – | – | – | – |
| 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.

---

[9] 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.

[10] 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 a definite description (*the clouds in the sky*) in the prompt, (7); see footnote 4 on Item 3.

[11] 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 moves in a similar way both in the coreference and in the disjoint reference condition, 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 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 4, when we discuss narratives with more than two referents.

(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.

Table 1 is 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. We expect that the shift in position from [$P_{21}$] to [$P_{25}$] in Figure 4 gives rise to a lower number on the y-axis (the axis in red, which goes from the center towards the bottom right), which we do not expect to find in [$P_{11}$]-[$P_{14}$] in Figure 3. We measured the dancer's position with regards to the X-axis and the Y-axis, and the results are summarized in Figure 5. In each of the 12 dance sequences, measurements of the 2 postures were taken that correspond to the predicates of the respective sentences (e.g. *sitting on the ground* and *holding a spear*). This figure should be interpreted as follows. The red (diamond-shaped) and pink (circle-shaped) dots are measurements from the coreferent conditions. While the origin of the graph was in the center of the motion-capture stills (e.g., in Figure 4), and the y-axis went rightward and downward while the x-axis went leftward and downward, the plot in Figure 5 has the origin in the bottom left corner, with the x-axis rightward and the y-axis upward (roughly turned 120° counter-clockwise). The fact that the red and pink dots roughly clutter together (in the top left corner of Figure 5 shows that the dancer barely moved out of her position in the coreferent conditions. By contrast, the orange (triangle-shaped) and black (square-shaped) dots, which represent the disjoint conditions, show a larger spread; while the orange dots largely (with one outlier) cluster with the red and pink dots (corresponding to the starting position of the dancer), the black dots are displaced to a lower Y-value and/or higher X-value. These displaced black dots correspond to the later position, associated with the new referent.
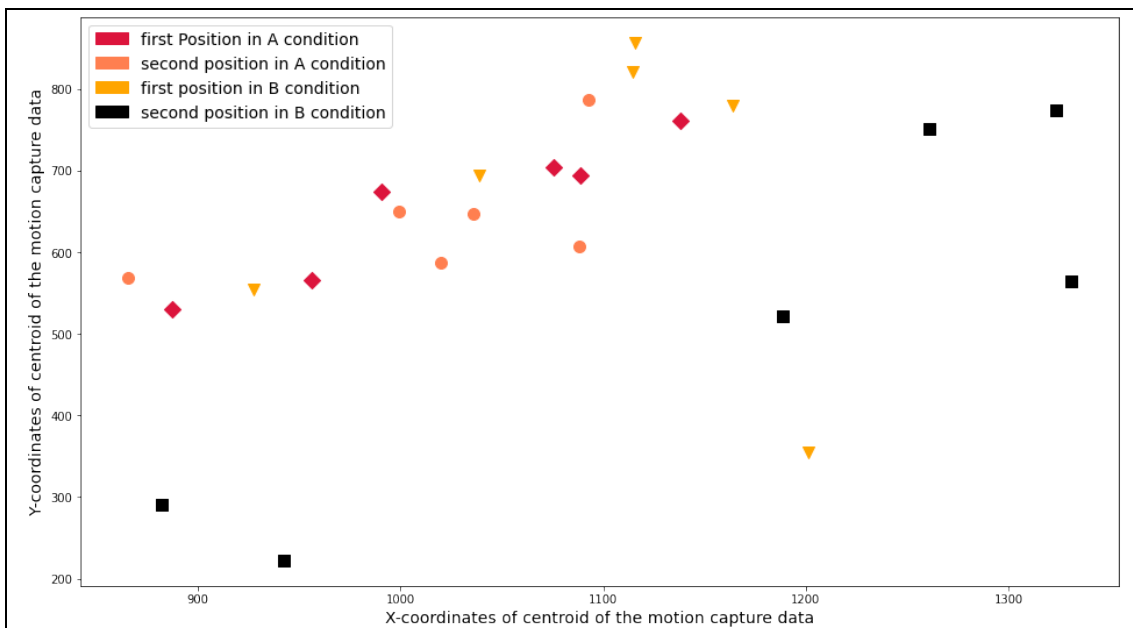


Figure 5: position of dancer on X-axis (horizontally) and Y-axis (vertically); the red (diamond-shaped)

and pink (circle-shaped) dots are from the coreferent condition; the orange (triangle-shaped) and black (square-shaped) dots are from the disjoint referent condition

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 starting position to the *same referent* position in the A condition *vs*. to the *different referent* position in the B condition. In the coreferent condition, this mean distance amounts to a 280 mm displacement in the x-y plane of the body centroid. By contrast, in the disjoint referent condition, it amounts to a 344.07 mm displacement in the body centroid in the x-y plane. This is summarized in the box plot below.
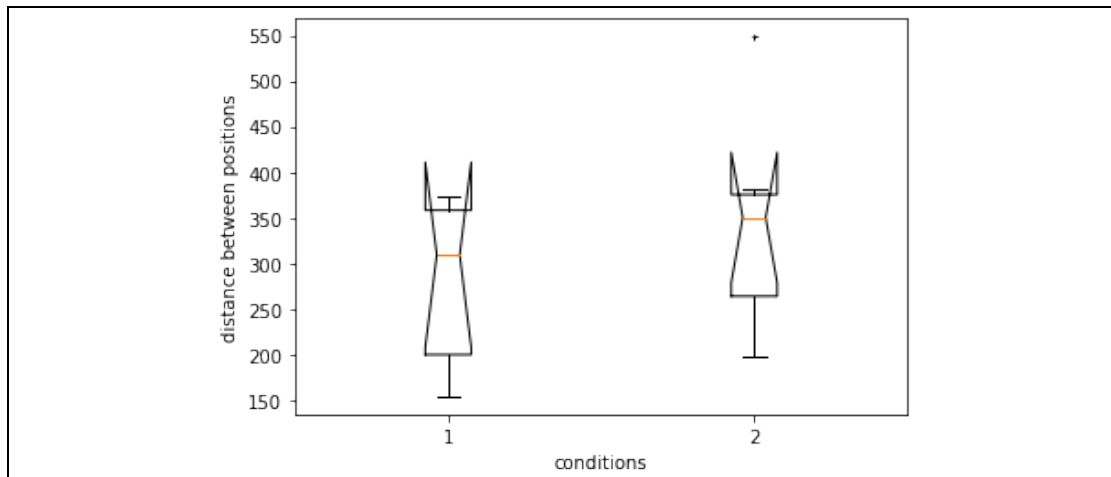


Figure 6: displacement in the x-y plane of the body centroid; in this graph, condition 1 is the coreferent condition (mean displacement = 280 mm), while condition 2 is the disjoint referent condition (mean displacement = 344.07 mm); the red line corresponds to the median displacements.

The marking of a new position on stage (and thus in the visual space), [P23] is a phenomenon that is reminiscent of referential loci present in sign language (cf. Lillo-Martin & Klima 1990, and see Schlenker 2017b for a recent survey article), opening new lines of inquiry for follow-up studies, which we address in section 4.[12] Assuming the new position also appears reminiscent of phenomena such as *Action Role Shift* in sign language, which is a phenomenon where a signer shifts her body, head position and/or eyegaze into the position of another character/referent, thereby indicating a shift in perspective towards that individual's perspective (e.g., Padden 1986, Lillo-Martin 1995, Quer 2005, Sandler & Lillo-Martin 2006, Herrmann & Steinbach 2009, 2012; see Davidson, 2015, for a recent discussion). For now, we will use terms such as *loci* in a purely descriptive way, as nothing in sections 2 and 3 depends on these positions being loci in a technical sense (and we may well be able to do away with *loci* altogether). Section 4 then picks up on the question of whether *loci* in a technical sense (and *Action Role Shift*) are relevant for the analysis of Bharatanatyam, and outline the possible hypothesis space.

---

[12] An anonymous reviewer raises the question of whether it is legitimate to apply an analytical notion from sign language (here: *loci*) to account for our finding on dance. Our initial position is to speak of 'positions in space' or 'positions on stage', which are integrated into sign language grammar as grammatical devices. 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 positions in *Bharatanatyam* should be treated as true equivalents of sign language *loci*.

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. The communicative challenge that arises in dance is fundamentally different from the challenge in comics. In comics, we find situations where 'copies' of the same individual occur across panels, but, technically, there is no continuity; the question is how a reader knows to identify the respective individual across panels. In dance, the problem is rather that an onlooker will always see the dancer in continuity; the question is rather how a dancer can convey that s/he is changing from one identity (*individual x*) to another (*individual y*). The disjoint reference case is thus the more revealing to study.

Before we continue with a brief discussion of perception, note that the hand gesture (mudra) that symbolizes 'another' is dispensable when it comes to introducing a new referent. To control for this, we ran a study where the descriptive differences between the two referents are sufficiently disambiguating (e.g., *man* vs. *woman*). A sample item is given in (14).

(14) a.  A man is sitting in the corner and a woman is standing in the middle of the room. The man is holding a spear. The woman is looking at him.
    b.  A man is sitting in the corner and a woman is standing in the middle of the room. The man is holding a spear. He is looking at the woman.

What we observed is that the dance sequences corresponding to such items did not include the 'another' mudra.[13,14] However, and most importantly, they still involved the introduction of a new position in signing space, followed by the dancer moving into the new position (location and orientation change) when describing an event that involves the respective individual. 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 definite descriptions in the text prompt. The need for introducing (and assuming) a new position indicates that every referent in our Bharatanatyam scenarios must be associated with its own position on the stage. The following figure exemplifies such a sequence.
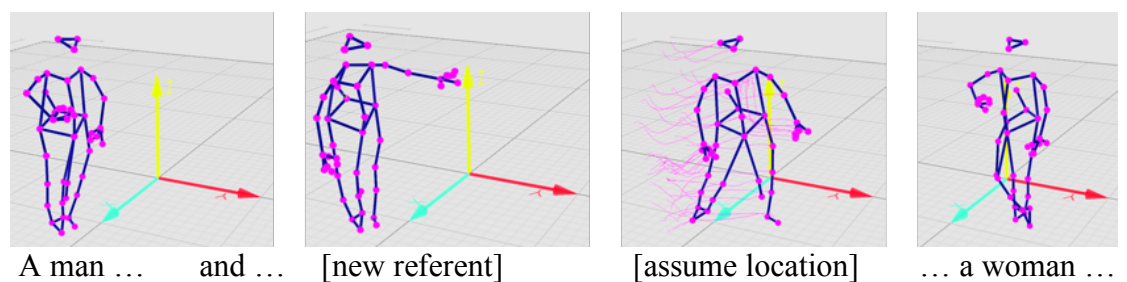


A man …     and …   [new referent]          [assume location]          … a woman …

Figure 7: disjoint reference without 'another'

---

[13] As pointed out in our discussion of Table 1, the 'another' mudra was robustly present in the twelve 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' mudra in the dance sequence.

[14] 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 7 express the meaning 'man' when the first referent (*a man*) is introduced, and the meaning 'woman' when the second referent (*a woman*) is introduced.

## 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*), plus 6 filler videos from a different production study (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 minimal pairs that were recorded and shown to the participants is given below.

| 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 (15). In other words, a higher rating would correlate with *disjointedness of reference*.

(15)  *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 ($t = 2.279$, $p < 0.05$)[15] 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; out of all participants, 19 grew up in Norway, 8 in Germany, 4 in Austria, and 1 in Sweden. 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 figures.

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.

---

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

Given that the dance sequences themselves exhibit some variation in how cues of disjointness are realized (see Table 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 / 56% in Item 3A to a total of 26 / 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 is the most crucial disjointness cue, as witnessed by the results of our pilot perception study.

In other words, looking at Figure 8 (repeated from Figure 7), the crucial ingredient for disjoint reference seems to be the introduction of a new location in space, followed by the dancer's movement into that new location. At the very least, this gives rise to an *iconic discontinuity inference*, i.e. there is a visually noticeable discontinuity in the dance sequence, which triggers the inference that there is also a discontinuity in the narrative that is being described. This is iconic (i.e. resemblance-based) in the sense that properties of the form (the physical dance) map onto properties of the meaning (the narrative). Section 3 shows how a formal semantic analysis can be applied to dance sequences of this type, focusing on the part where the dancer moves into the new location; in section 4, we revisit the data from the perspective of sign language grammar, asking the question of whether the introduction of a new location in space may be meaningful in itself (akin to so-called *loci* in sign languages).
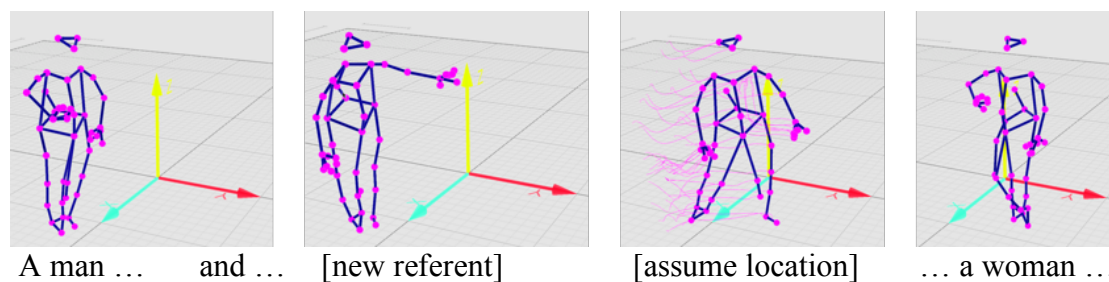


A man …      and …   [new referent]          [assume location]          … a woman …
Figure 8: disjoint reference without 'another'

## 2.6 Refining our mission statement

Prior to presenting our theoretical analysis in Section 3, we conclude Section 2 with an explicit statement about the scope of this paper, which culminates in a mission statement for future empirical investigations. As stated above, 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 (as illustrated by Figure 8) 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*)? 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 can precisify the hypothesis space 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 iconic discontinuity inferences are based on a 'break/disruption' in the form, which gives rise to a similar 'break/disruption' in the semantics.[16] 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 turned out to be 'yes' to all three, this would not invalidate the point that we are making in this paper, but rather strengthen it. It would yield 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.

Future experiments should thus probe into pantomime performed by un-trained participants (i.e. not by professional mimes) to test whether they spontaneously employ iconic discontinuity for the marking of disjointness (and possibly, as we discuss in Section 3, the structuring of events). We take a comparison of dance and other modes of expression to be highly important, and Section 4 outlines first steps, by comparing dance with sign language.

## 3. A super linguistic approach to narrative dance – formal theoretical analysis
### 3.1 Defining truth in visual narrative

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 (2021), building on Greenberg (2011, 2013), posits a generalized possible worlds model for informational entities, (16), based on the idea that any sentence, picture, etc., counts as an informational entity when it rules out some possibilities.

---

[16] 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**        'They fear the Ship might near **the Edge**
    **af Jorden**, hvor Havet nedstyrter sin Sluse,     **of the Earth**, [...]'

(16)  *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

We can start by illustrating Abusch's idea for the dance position [$P_{21}$] in Figure 7. 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 (16)), namely ones in which there is no person, or in which the person is not sitting. The statement in (17) is thus understood to provide new information about a given situation that we are describing.

(17)  There is a person who is sitting.

Crucially, Abusch argues that a picture achieves exactly the same result. In parallel to (17), the dance position in (18) can be understood to provide new information about a given situation (namely the current point in time in a narrative that is being told).[17] As Abusch observes, when it comes to the question of what a world or situation is like, (18) 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 (18) thus qualifies as an informational entity in line with (16). Abusch is careful to point out that pictures are often more informative than sentences; taken at face value, a naïve observer may infer from (18) 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 (17) can leave such information underspecified; there is no implication from (17) on how exactly the person is sitting.

(18)
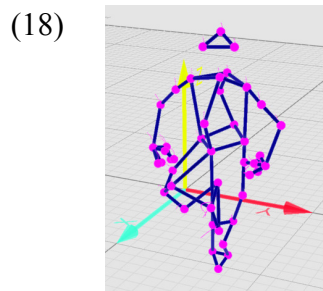


Refining the approach of Greenberg (2011, 2013), Abusch (2021) proceeds to identify the semantics of a picture with the set of possibilities that it admits. This means that we can define the semantics of a picture in terms of possible worlds, situations, or scenes. Treating any given dance position [$P_n$] as a picture, we can then posit
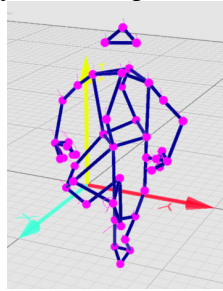
---

[17] Note that the dancer remains static in this 'sitting' position for an average of 112 frames / 560ms across the two conditions (specifically, 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), i.e., she is not constantly moving. 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.

satisfaction conditions as given in (19).[18] Truth in visual narrative is thus defined in terms of how well a dance position [$P_n$] maps to a situation/scene $\sigma_n$ in the narrative; i.e. the dance position in (19) counts as satisfied by a situation $\sigma$ (i.e. "true" in $\sigma$) if a sitting activity is taking place in $\sigma$. Note that Bharatanatyam encodes meaning in a highly iconic way; the meaning of the dance position in (19) is easily inferable, even to an onlooker unfamiliar with Bharatanatyam, based on resemblance between the dance posture and the sitting activity that is being described.[19]

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



a situation $\sigma$ satisfies          only if in $\sigma$ 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 Figure 8, 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    *Modeling disjoint reference in narrative dance on the basis of discontinuity*

#### 3.2.1    *Revisiting dance syntax (grouping) from the perspective of meaningful dance*

Let us take our example in Figure 8 as our point of departure, but simplify it, as given in Figure 9, by removing the pointing gesture that introduces a new location on the stage. (We return to the pointing gesture, and whether it is obligatory or optional, in Section 4.) One clear difference between [$P_{31}$] and [$P_{33}$] is the orientation of the dancer, due to the shift in [$P_{32}$].[20] Since this is the change that is, presumably, most evidence to an onlooker, we can ask whether this alone might be a trigger for inferences towards disjoint reference.

---

[18] 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.
[19] Greenberg (2020: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 (19) 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.   $[\![\mathbf{P_n}]\!]$ = {$s$ | there is an eventuality $e$ in $s$ such that $e$ *looks like* $\mathbf{P_n}$}
[20] 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.
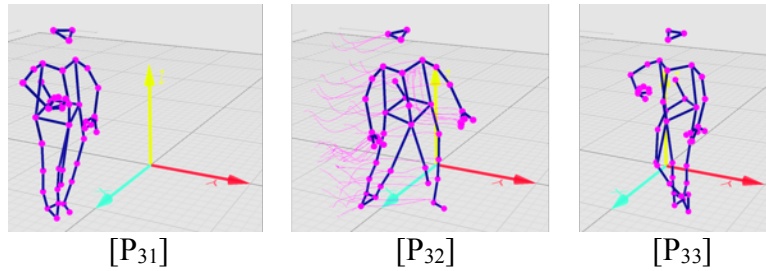
$[P_{31}]$        $[P_{32}]$        $[P_{33}]$

Figure 9: disjoint reference simplified

We can start by formalizing this insight on the basis of Charnavel's (2016, 2019)[21] 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 argue in the remainder of this section, culminating in (29)-(30) (in section 3.2.2), 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) 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 (20).

(20) *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 9, 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; we will come back to this point at the end of this section, in (26), and in section 4.

Let us now revisit our examples of coreference *vs.* disjoint reference. 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 10. In line with Abusch (2013:12, 2014:10), we posit the satisfaction conditions in (21) to (partially) describe the dance positions in Figure 10. 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 (21) is that dance positions $[P_n]$ are mapped to propositions $[\![P_n]\!]$.[22]

---

[21] 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.

[22] Note 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, 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 14.
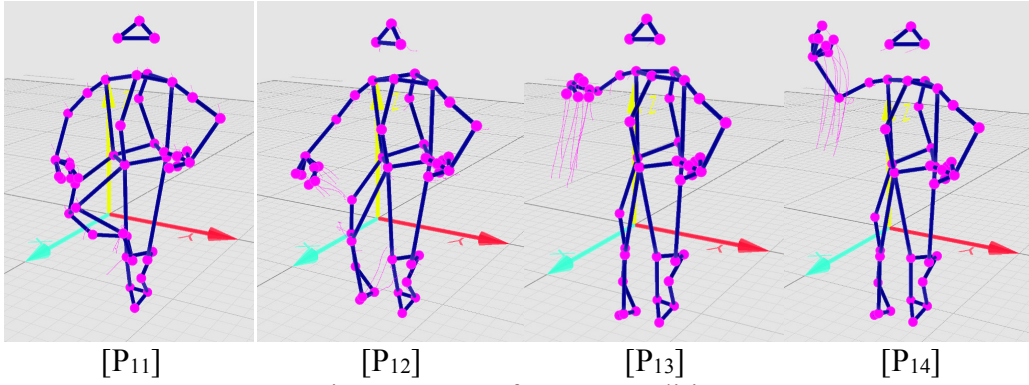
19

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

Figure 10: coreference condition

(21) a. A situation/scene $\sigma_{11}$ satisfies $[P_{11}]$ only if in $\sigma_{11}$ a person is sitting.

b. A situation/scene $\sigma_{14}$ satisfies $[P_{14}]$ only if in $\sigma_{14}$ a person is holding a spear.

The coreference condition, Figure 10, does not involve any change of orientation, i.e., we can assume that there is no grouping boundary that emerges between $[P_{11}]$ and $[P_{14}]$. Contrast this with the disjoint reference condition in Figure 11, 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.



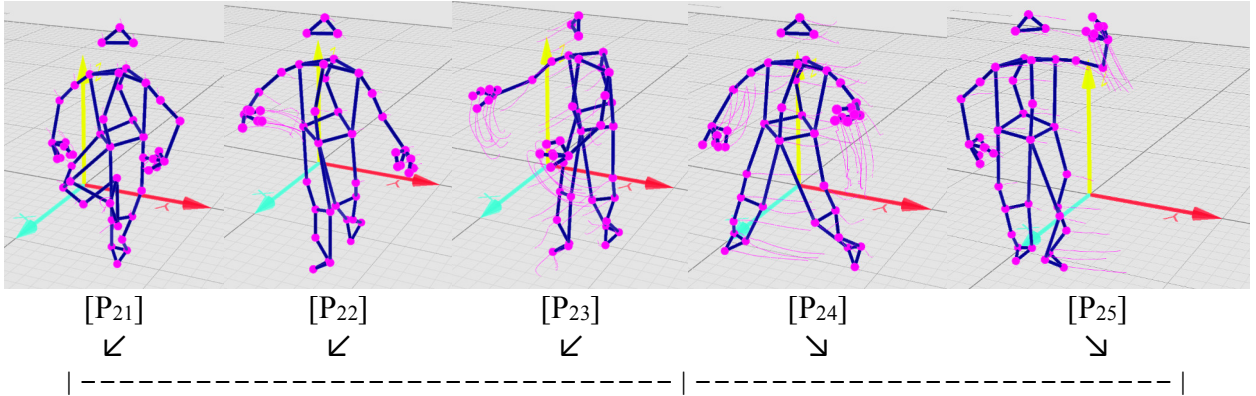$$[P_{21}] \qquad [P_{22}] \qquad [P_{23}] \qquad [P_{24}] \qquad [P_{25}]$$

Figure 11: grouping in the disjoint reference condition

The satisfaction conditions in (22a-b) are equivalent to those in (21a-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 $\alpha$ in $\sigma_{21}$ and the person $\beta$ in $\sigma_{25}$ are identical.

(22) a. $\sigma_{21}$ satisfies $[P_{21}]$ only if in $\sigma_{21}$ a person $\alpha$ is sitting.

b. $\sigma_{25}$ satisfies $[P_{25}]$ only if in $\sigma_{25}$ a person $\beta$ is holding a spear.

Crucially, 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 11, as given in (24) (using Charnavel's 2016, 2019 notation). For the purpose of illustration, (24) assumes that each of the positions in Figure 11 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).[23] 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 Charnavel's GPR10, (23), since such gestures are generally more intense than gestures that only involve individual body parts.

(23)    *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 (20), 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).

(24)    *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
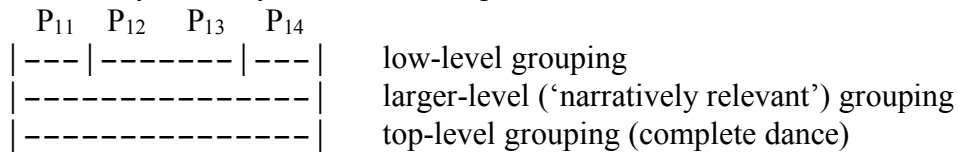         | ‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑ |        top-level grouping (complete dance)

By contrast, a dance sequence in the coreference condition given in Figure 10, would lack the larger-level grouping boundary, as schematically shown in (24) vs. (25). In fact, given the smoothness of the upward arm movement in Figure 10, 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.[24]

---

[23] For future research on how to sharpen the theory of boundary placement, a possible direction would be to compare dance sequences to sign language. 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. While Section 4 of our paper explores the striking similarity between positions on stage in Bharatanatyam and the *loci* (positions in signing space) of sign language, we limit our discussion of grouping boundaries to the precisification in (26), as an indepth exploration of low-level grouping is beyond the scope of this paper.

[24] 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.

(25) *structure of the coreference dance sequence*

$$P_{11} \quad P_{12} \quad P_{13} \quad P_{14}$$

```
|───|───────|───|    low-level grouping
|──────────────|    larger-level ('narratively relevant') grouping
|──────────────|    top-level grouping (complete dance)
```

One plausible approach to the contrast between disjoint reference, (24), and coreference, (25), 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.[25]

An open question with regards to the present discussion concerns the very nature of 'narratively relevant' grouping in (24) and (25), and how it relates to the underspecified notion of 'more pronounced' that Charnavel (2019:4) employs in (23). To reduce vagueness and make our hypothesis maximally concrete, we zoom in on the fact that the only 'narratively relevant' grouping in our data involves a change in position coupled to a change in orientation.[26] Based on this fact, we replace the notion of *larger-level ('narratively relevant') grouping* (in (24) and (25)) with a precisified notion of *event-structuring grouping* (in (26)).

(26) *Event-structuring grouping*
    In Bharatanatyam, an event-structuring 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 an event-structuring grouping boundary.

*3.2.2   Semantic grouping (or: semantic effects of grouping)*

We can now turn to the semantic effects of such grouping. A first approximation of the hypothesis that hierarchical grouping in narrative dance is mapped to situation structure is given in (29), a key idea and important departure from Lerdahl & Jackendoff (1983), which is inspired by Schlenker's (2017a:22-28) proposal that, in music semantics, grouping structure may reflect the organization of events in a narrative. Without going into the details of event composition and the relationship between events and situations that contain them (see Kratzer 2020 for an overview), we assume an approach where events can be decomposed into smaller events, while, on the other hand, some events are atomic (not having proper parts), see Schlenker (2017a:24), who builds on Varzi (2015) (see also Link 1983, Krifka 1989 and Landman 1991 for earlier relevant work on mereology). This is commonly represented as in (27) (compare Champollion 2014:9).

(27)      $e_1 \oplus e_2$

        $e_1$      $e_2$

---
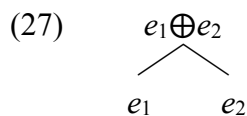
[25] We expect to find similar effects in other (non-narrative) dance forms.

[26] A more refined view 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). We are grateful to an anonymous reviewer for this suggestion.

For our purposes, the central issue can be illustrated for Item 1 of our production study, as repeated in (28); in line with (27), we can posit an event $e_1$, which is a sitting-on-the-ground event, and an event $e_2$, which is a spear-holding event. Each of these events has an agent, and the dancer's interpretative goal is thus to communicate whether the agent of $e_1$ and $e_2$ are identical or distinct. The dancer achieves this by placing an event-structuring grouping boundary, (26), in the case where the agents are distinct, (28b), and not in the case where the events $e_1$ and $e_2$ have an identical agent, (28a).

(28)  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)

We can use these insights to posit (29a), which, as its 'mirror image', yields (29b).

(29)  *grouping-based discourse structure*
   a.  If there is a group boundary between two dance positions $[P_n]$ and $[P_m]$, then a dance sequence that starts with $[P_n]$ and ends with $[P_m]$ is interpreted as *discontinuous*, i.e., it maps to two distinct atomic events, $e_1$ and $e_2$.
   b.  In the absence of a group boundary, a dance sequence that starts with $[P_n]$ and ends with $[P_m]$ is interpreted as *continuous*, i.e., as corresponding to a plurality of events $e_1 \oplus e_2$ (or, in some cases, to a single atomic event $e_1$).

Crucially, we propose that grouping boundaries mark discontinuity, as a type of visual iconicity; in other words, grouping-induced discontinuity is treated as an iconic effect (which departs from the purely syntactic notion of grouping in Lerdahl & Jackendoff 1983 and Charnavel 2016, 2019). The grouping boundaries themselves may thus be a by-product of visual iconicity (the marking of a situational break by virtue of a syntactic break) rather than an explanation of perceived discontinuity in their own right.

   We propose that grouping is used to indirectly convey disjointness (e.g. disjoint reference when two characters are introduced into a narrative), by virtue of an event-based reference determination rule, given in (30), building on Abusch (2013:13).[27]

(30)  *grouping-based reference determination*
   a.  If a narrative dance sequence corresponds to a plurality of events $e_1 \oplus e_2$ and contains two similar entities α and β, coreference (i.e. α=β) arises by default when there is no indication that parts of $e_1 \oplus e_2$ contain more than one entity of this type.
   b.  If a narrative dance sequence corresponds to two distinct atomic events $e_1$ and $e_2$, which contain two similar entities α (in $e_1$) and β (in $e_2$), respectively, disjoint reference (i.e. α≠β) arises by default.

---

[27] As Abusch (2013) points out, the identification of entities in a single situation (or, in our case, in a single event), α = β, may well reflect low-level processes of indexing in vision, see Pylyshyn (2003); this would be much in line with the discussion in section 2.6. 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.

While we now proceed to apply (30) to our two paradigm examples (in (32) and (33)), it is worth pointing out that a more simplistic rendering could be stated as in (31).

(31) *grouping-based reference determination (simplified)*
    If a narrative dance sequence contains two similar entities $\alpha$ and $\beta$, which are associated with action descriptions $e_1(\alpha)$ and $e_2(\beta)$, then
    a.  coreference (i.e. $\alpha=\beta$) arises by default if $e_1(\alpha)$ and $e_2(\beta)$ are **not** separated by an event-structuring grouping boundary, (26), and
    b.  disjoint reference (i.e. $\alpha\neq\beta$) arises by default if $e_1(\alpha)$ and $e_2(\beta)$ **are** separated by an event-structuring grouping boundary, (26).

We can now proceed with the coreference sequence (in Figure 10) and render its semantics as given in (32). We have already established the two satisfaction conditions in (32a) and (32b). By grouping-based coreference, we now infer that the narrative sequence corresponds to an event $e_{11}\oplus e_{14}$ in (32c) (loosely based on Abusch, 2013); this is an event that has a subpart $e_{11}$ and a subpart $e_{14}$, which each involve existential quantification over a person ($\alpha$ and $\beta$, respectively). Since both are part of the described event $e_{11}\oplus e_{14}$, we can, by (30b), identify $\alpha$ and $\beta$.

(32) *Semantic analysis of the coreference condition* [Figure 10]
    a.  $\sigma_{11}$ satisfies $[P_{11}]$ only if in $\sigma_{11}$ a person $\alpha$ is the agent of a sitting event $e_{11}$.
    b.  $\sigma_{14}$ satisfies $[P_{14}]$ only if in $\sigma_{14}$ a person $\beta$ is the agent of a spear-holding event $e_{14}$.
    c.  *by grouping-based discourse structure* [via (29a)]
        since there is no grouping boundary between $[P_{11}]$ and $[P_{14}]$, the narrative dance sequence correspond to the plurality of events $e_{11}\oplus e_{14}$.
    d.  *by grouping-based reference determination* [via (30a)]
        since the narrative dance sequence that encompasses $[P_{11}]$ and $[P_{14}]$ corresponds to the plurality of events $e_{11}\oplus e_{14}$, coreference (i.e. $\alpha=\beta$) arises by default.

Conversely, grouping alone may be sufficient to block coreference (i.e. referent identification) in a simple narrative like the disjoint reference narrative that we discussed (in Figure 11), blocking identification of the agent of the atomic event $e_{21}$ and the agent of the atomic event $e_{25}$. This is shown in (33).

(33) *Semantic analysis of the disjoint reference condition* [Figure 11]
    a.  $\sigma_{21}$ satisfies $[P_{21}]$ only if in $\sigma_{21}$ a person $\alpha$ is the agent of a sitting event $e_{21}$.
    b.  $\sigma_{25}$ satisfies $[P_{25}]$ only if in $\sigma_{25}$ a person $\beta$ is the agent of a spear-holding event $e_{25}$.
    c.  *by grouping-based discourse structure* [via (29b)]
        since there is a grouping boundary between $[P_{21}]$ and $[P_{25}]$, the narrative dance sequence corresponds to a description of the atomic events $e_{21}$ and $e_{25}$ (and not to a description of the plurality $e_{21}\oplus e_{25}$).
    d.  *by grouping-based reference determination* [via (30b)]
        since the narrative dance sequence that encompasses $[P_{21}]$ and $[P_{25}]$ corresponds to a description of two distinct atomic events $e_{21}$ and $e_{25}$, disjoint reference (i.e. $\alpha\neq\beta$) arises by default.

In brief, we propose that grouping in dance serves as a way to organize (sub-)events. Specifically, the introduction of larger-level group ('narratively relevant') boundaries

(i.e. *event-structuring grouping*) 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 (and discontinuity inferences may give rise to all kinds of effects unrelated to reference); yet, it is quite plausible that every change of character requires a grouping boundary to be placed.[28]

We wish to conclude this section with a comment on the methodology. An anonymous reviewer raises the question of whether it is legitimate to combine analytical elements from approaches that are based on different domains; specifically, while the method for establishing grouping boundaries was indeed designed for dance in general (Charnavel 2016, 2019), the situation semantics that we employ was originally conceived for comics (Abusch 2013, 2014, 2021), and the connection between grouping and event-organization was specifically designed for music (Schlenker 2017a). Our take on the compatibility of these different tools is that grouping, situations, and events are all modality-independent, so these are exactly the kinds of tools that we expect to be freely applicable across different means of expression, particularly when we investigate semantics (rather than, say, syntax, which may be more closely tied to a given modality).

### 3.3   *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) 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* is a likely candidate for such a meaning atom, which could mark discontinuities of protagonists/characters, locations, situations, and so forth, all of which would amount to a shift from one virtual source to another. 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 (34), 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) that

---

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

the meanings that can be conveyed, while consistently being iconic (resemblance-based), are not limited to 'literal' meanings. For instance, while the movement in (34) 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'. The actual meaning that is expressed will naturally be much more abstract, as indicated.

(34)
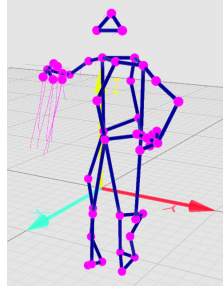
a situation σ satisfies  only if the virtual source in σ is involved in a (partial or total) upward movement.

A skeptical reader may question the appropriateness of analyzing an example like (34), 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 (2019). Here, a ballet choreography conveys a dialogue between two virtual entities (sources), not independently present in the music.

We now proceed, in section 4, with an exploration of parallels between meaningful mechanisms that can be used in dance, on the one hand, and sign language, on the other hand. The aim of doing so is to explore the extent to which dance may, in fact, incorporate a rudimentary sign language. 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 fully fledged (signed) languages. Specifically, section 4 explores whether *loci* of the type that we find in sign language should be added to our analysis of disjoint reference in Bharatanatyam, which, if answered in the affirmative, allows for two possible scenarios: *loci* as a complement to the grouping-based analysis (i.e. both are necessary), or *loci* as an alternative to the grouping-based analysis (i.e. only *loci* play a role in disjoint reference).

## 4.  *Loci* as a commonality of dance gestures and sign language

In section 3 above, we explored the option that a change in position and orientation may give rise to a grouping boundary, which can serve as a semantically meaningful way of encoding discontinuity inferences, e.g., in marking disjoint reference. A different question that we can now ask is whether the different positions on the stage that we saw could have a function that goes beyond being a mere grouping indicator.

The key theoretical question amounts to whether the change of position on part of the dancer is purely a marker of grouping (as explored in section 3), or also introduces loci, as found in sign language. We state three competing hypotheses in (35).

(35)  a.  $H_1$ = visual discontinuity inferences (a type of visual iconicity, which we captured in terms of grouping-induced disjointness) are sufficient to account for the change of position on part of the dancer when managing distinct discourse referents (and no additional mechanisms, such as *loci*, are needed)

   b.  $H_2$ = designated positions in space (*loci* of the type that we find in sign languages) are needed to account for the change of position on part of the dancer when managing distinct discourse referents, in addition to visual discontinuity marking

   c.  $H_3$ = designated positions in space (*loci* of the type that we find in sign languages) are the only mechanism needed to account for the change of position on part of the dancer when managing distinct discourse referents, i.e. visual discontinuity marking by virtue of grouping is entirely uninformative with regards to coreference.

$H_3$ is a hypothesis that essentially promotes a *loci*-based approach to being *the* approach for explaining coreference *vs.* disjoint reference in narrative dance. Notably, as discussed in section 2.6, there are good reasons to assume that visual discontinuity inferences occur quite broadly, so $H_3$ would not deny the existence of such inferences, but it would maintain that such inferences have no bearing on coreference *vs.* disjoint reference. For now, rather than aiming to decide between the hypotheses in (35), we pursue the modest goal of motivating $H_2$ and $H_3$ by showing that *loci* may, in fact, be relevant for theorizing about dance semantics.

To clarify, consider Figure 11, repeated from above, and analyzed in terms of grouping alone in (33).
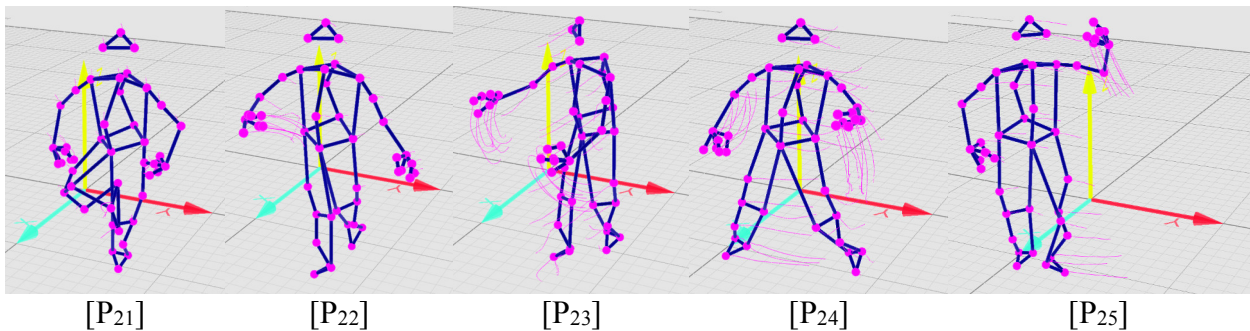


| [P$_{21}$] | [P$_{22}$] | [P$_{23}$] | [P$_{24}$] | [P$_{25}$] |

Figure 12: grouping in the disjoint reference condition

If we now focus on P$_{23}$ and P$_{24}$, we may incorporate sign-language like *loci* and posit satisfaction conditions such as (36), from the perspective of Abusch's (2013, 2014, 2021) picture semantics. Here, $i$ is a virtual locus, and $g_c(i)$ is an individual that is associated with the locus $i$ in context $c$ by virtue of an assignment function $g_c$. This is an over-simplification of how loci work in sign language, but it will suffice for present purposes. (See Kuhn 2016 for a recent discussion, which compares possible formal semantic analyses of loci.)

(36)  a.  For any assignment function $g_c$ in context $c$ and virtual locus $i$,
         $\sigma_{23}$ satisfies [P$_{23}$] only if in $\sigma_{23}$ there is an individual $x$ in the narrative such that $x = g_c(i)$.

   b.  $\sigma_{24}$ satisfies [P$_{24}$] only if in $\sigma_{24}$ the perspectival center of the narrative is *at* $g_c(i)$.

The question with regards to (36) can be stated as follows: what, if any, is the shared cognitive underpinning of virtual loci in narrative dance on the one hand, and the referential loci of Sign Languages on the other hand (see Lillo-Martin & Klima 1990; cf. Schlenker 2017b for an overview)? Specifically, we may wonder if narrative dance incorporates a rudimentary form of sign language. In sign language, a so-called *locus* (plural *loci*) is a position in signing space (e.g., Schlenker 2017b:151 for a recent overview) that is associated with a given referent; we can then ask if narrative dance also establishes such loci. Schlenker (2017b:149) describes the relevance of loci for reference as follows: "In simple cases, an antecedent is associated with a position or 'locus' in signing space, and an anaphoric link is obtained by pointing toward that locus to recover it semantic value". 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, 2017b: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 [P24].

We carried out two follow-up production studies to test for possible loci, and the findings indicate that loci might indeed play a role in narrative dance. In Study 2, we built on the *disjoint reference* condition of Study 1, investigating what would happen if one or the other referent is subsequently picked up in the story. The aim was to see if positions on stage (i.e. *loci*) would be picked up again, in the same way in which this happens with sign language loci. A sample item is given in (37). Both conditions share the first two sentences; they differ in the third sentence. In this item, we first introduce a child that is eating a mango, and then a child that is holding a spear. The subsequent transitive sentence picks up both of these referents; in (37a), the subject of the transitive clause corresponds to the referent that was introduced first, and the object to the referent that was introduced second; example (37b) has the inverse relationship.

(37) a.  The artist sees a child eating a mango outside the temple.
        Then she sees another child holding a spear.
        *The eating child* watches *the child with the spear*.

     b.  The artist sees a child eating a mango outside the temple.
        Then she sees another child holding a spear.
        *The child with the spear* watches *the eating child*.

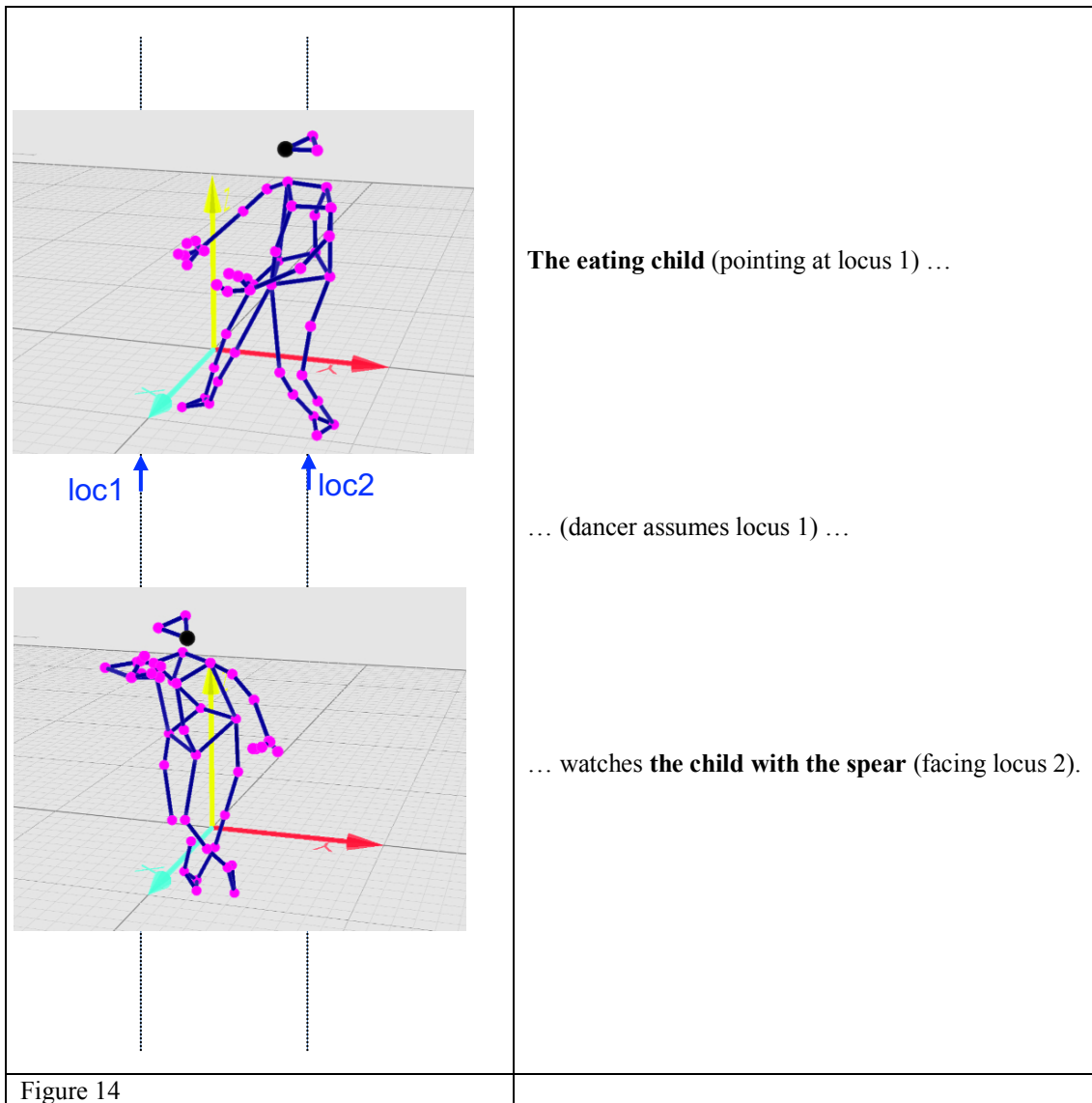What we find is that the dancer establishes two separate positions for the two children, as shown in the following figure. Reading these pictures from top to bottom, we find that the dancer starts in a neutral (central) position and then introduces a position for the first referent (labeled *loc1*), followed by a position for the second referent (labeled *loc2*); the locus positions are indicated with a blue arrow.

| | |
|---|---|
|  | **The artist** (neutral locus) sees … |
|  | … **a child** (locus 1) eating a mango outside the temple. … |
|  | … Then she sees **another child** (locus 2) holding a spear. |
| Figure 13 | |
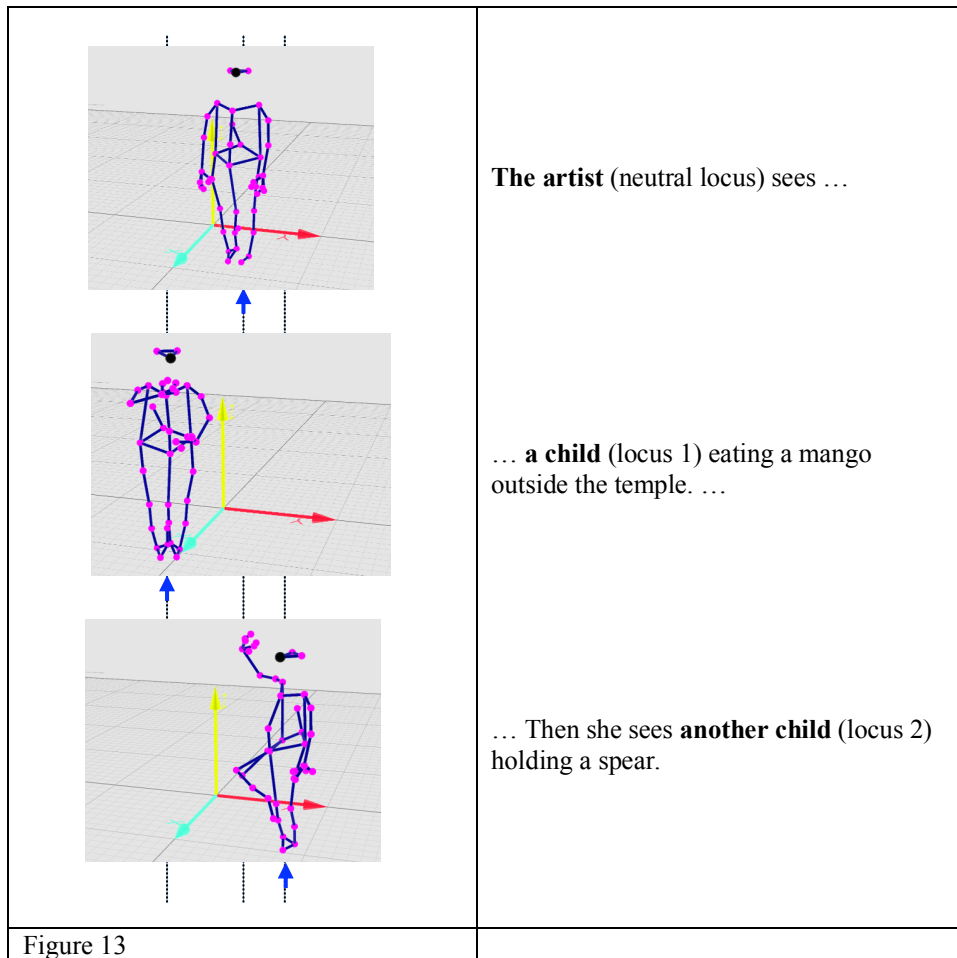
Once we zoom in on the third sentence, where the two conditions differ from each other, we find that anaphoric dependencies can indeed be established by virtue of these *dance loci*. Let us start with (38), repeated from (37a).
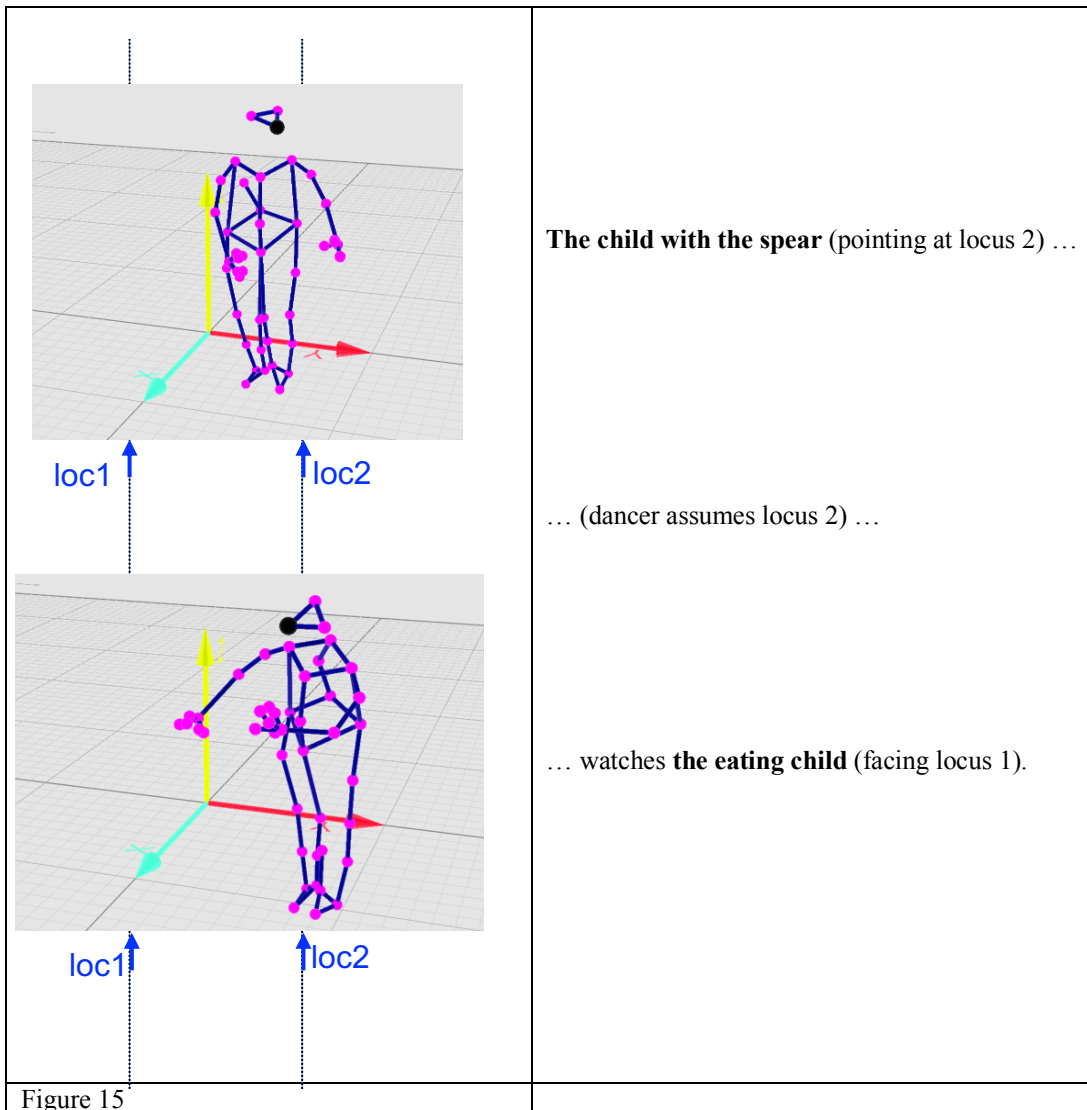
(38)   The artist [neutral] sees a child [loc1] eating a mango outside the temple.
Then she sees another child [loc2] holding a spear.
*The eating child* watches *the child with the spear*.

Figure 14 corresponds to the third sentence in (38); here, the dancer first points at locus 1 (to reactivate the reference for *the eating child*) and then physically moves into the locus 1 position, orienting herself towards locus 2 (i.e., towards *the child with the spear*). In addition to the fact that pointing at locus 1 is strongly reminiscent of sign language (where anaphoric links can be established by pointing towards a locus), we observe that the movement into locus 1 (where the dancer takes up locus 1 as a new viewpoint) is reminiscent of Action Role Shift in sign language (e.g., Padden 1986, Lillo-Martin 1995, Quer 2005, Sandler & Lillo-Martin 2006, Herrmann & Steinbach 2009, 2012), where a signer assumes the perspective of another individual through a body, head position and/or eyegaze shift; moreover, the orientation towards locus 2 is reminiscent of so-called *agreeing verbs* in sign language, which incorporate directionality (e.g., Padden 1988).

| | |
|---|---|
| | **The eating child** (pointing at locus 1) … |
| loc1     loc2 | … (dancer assumes locus 1) … |
| | … watches **the child with the spear** (facing locus 2). |

Figure 14

The opposite example, (39), repeated from (37b), comes out in parallel. What we see in Figure 15 is the mirror image of what we saw in Figure 14. Here, the dancer first points at locus 2, and then moves into locus 2, orienting herself towards locus 1, assuming this new viewpoint/perspective. Once again, this is strongly reminiscent of loci, Action Role Shift, and agreeing verbs in sign languages.

(39)  The artist (neutral) sees a child (loc1) eating a mango outside the temple.
Then she sees another child (loc2) holding a spear.
*The child with the spear* watches *the eating child*.

| | |
|---|---|
|  <br> loc1  loc2 | **The child with the spear** (pointing at locus 2) … |
| | … (dancer assumes locus 2) … |
|  <br> loc1  loc2 | … watches **the eating child** (facing locus 1). |
| Figure 15 | |

Since *moving into the locus* and *pointing at a locus* seem to have the same function (namely identifying a referent in the narrative), we expect that either one of them is sufficient and they do not always have to co-occur. In examples (37)-(39), we have already seen cases where the dancer orients towards a locus without moving into the locus, namely in the final part (reminiscent of agreeing verbs in sign language), where one child (in loc1/2) watches the other child (in loc2/1). We can now ask whether there are also cases where the dancer moves into a locus without first pointing at the locus. As a matter of fact, this is attested when the *first* referent is introduced into the narrative. In (40), we have added the symbol '☞' to every pointing gesture that we observed in the dance sequence; while the dancer points at *loc1* (in bold type) in one dance sequence, (40b), this pointing gesture is missing in the other dance sequence, (40a). 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 (40b) may be an artifact of the dancer's choice to introduce the temple before introducing the first child in (40b), but not in (40a); if *loc1* is introduced after introducing the temple referent, this would indicate that *loc1* is, strictly speaking, not the first referent.)

31

(40) a.   The artist[neutral] sees a child[loc1] eating a mango outside the temple.
Then she sees another ☞ child[loc2] holding a spear.
*The eating ☞ child[loc1]* watches *the ☞ child with the spear*[loc2].

b.   The artist[neutral] sees a **☞ child[loc1]** eating a mango outside the temple.
Then she sees another ☞ child[loc2] holding a spear.
*The ☞ child with the spear*[loc2] watches *the eating ☞ child*[loc1].

We carried out a third study, to test what would happen with more than two possible referents. A sample item for Study 3 is given in (41); our findings explicitly replicate the findings from Study 2. The dancer introduces three separate positions on the stage, which we can, once again, think of as rudimentary loci (*locus 1* for the woman, *locus 2* for the man, and *locus 3* for the child; in terms of position on the stage, these loci roughly correspond to the three positions given above: *locus 1* corresponds to *neutral*, *locus 2* to *locus 1*, and *locus 3* to *locus 2*). In the final (transitive) clause of the various conditions in (41a-f), the dancer moves into the locus of the respective subject (i.e., *the man* in (41a), *the child* in (41b), and so forth) and faces the locus of the respective object, in parallel to what we saw above for Study 2.

(41) *Shared by all conditions:* A woman [loc1] is standing outside the temple, a man [loc2] is sitting on the ground, and a child [loc3] is playing.
a.   The woman[loc1] is holding a book. The man[loc2] is looking at the child[loc3].
b.   The woman[loc1] is holding a book. The child[loc3] is looking at the man[loc2].
c.   The man[loc2] is holding a book. The woman[loc1] is looking at the child[loc3].
d.   The man[loc2] is holding a book. The child[loc3] is looking at the woman[loc1].
e.   The child[loc3] is holding a book. The woman[loc1] is looking at the man[loc2].
f.   The child[loc3] is holding a book. The man[loc2] is looking at the woman[loc1].

An open question, which at this point cannot be addressed due to the limitations of our initial perception experiments (as reported in section 2.5), but should be addressed in future research, is whether these positions in space exhibit other properties of sign-language *loci*. For instance, sign-language *loci* do not need to mark to the actual positions of object in physical space, but can define arbitrary positions. The question is thus whether audience members who pick up on the *loci* identify them with the actual (literal) positions of entities or not. Arbitrariness in the position of *loci* is a central property of grammatical *loci*, which would allow us to tease them apart from purely iconic uses (in which the dancer iconically incorporates the relevant character/referent). The relevant hypotheses are stated in (42).

(42) 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 grammatical *loci*.

Regardless of the eventual decision between (42a) and (42b), the very fact that rudimentary loci may be employed in narrative dance opens avenues of future investigation for the semantics of dance.

## 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 (as is instantiated by the situation and event semantics described in Kratzer 2020) can be fruitfully applied to meaning in narrative dance.

Using the formal linguistic method of eliciting minimal pairs, we carried out a production study (sections 2.3 and 2.4) that focused on the topic of coreference (*the same man*) vs. disjoint reference (*another man*). We determined that a dancer can use a change of orientation in order to signal discontinuity/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 rate the disjointness of reference higher in the disjoint reference video than in the coreference video. While it is unclear what cues they used to carry out their ratings, this indicates that the discontinuity can also be processed as an actual indicator of disjointness.

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), 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). We explored (in section 3.2) the idea that change of orientation gives rise to grouping boundaries (i.e., boundaries between sections/segments of a dance sequence) which are interpreted as organizing a narrative into events/sub-events (inspired by Schlenker 2017a). We treat grouping-induced discontinuity as an iconic effect (thus instantiating visual iconicity).

In section 4, we raise the question of whether narrative dance can also recruit positions in dance as a type of rudimentary *loci*, in line with the loci of sign language (e.g., Lillo-Martin & Klima 1990, Liddell 1990). We outlined the open question which is whether visual iconicity that recruits grouping structure (such as a change of orientation) is sufficient to mark disjointness via discontinuity, or whether *loci* are also needed to account for the way meaning is conveyed in the dance sequences that we investigated.

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 shared possibility to not only intentionally communicate meaning to an audience, but to use grouping (constituency) for communicating meaning on event organization in a narrative. 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 (34), 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.

**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. (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, 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 (2020). The Iconic-Symbolic Spectrum. Slides for a presentation at Institut-Jean Nicod, 9th October 2020.

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.

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.

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]

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., 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.

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.

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.

Philippe Schlenker, 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.

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 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.

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. (2019) What is Super Semantics? *Philosophical Perspectives* 32, 365-453.

Schlenker, P. and P. Patel-Grosz (2018). What is Super Linguistics? Presentation at the Workshop "Super Linguistics – an introduction", University of Oslo, 10 December 2018.

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.