



# An Automated Structural Approach to Support Theatrical Performances by Introducing Gesture Recognition to a Cuing System

Kosuke Sasaki<sup>1</sup>, Anne-Catherine Luther<sup>2</sup>, Tomoo Inoue<sup>1</sup>(✉),  
and Wolfram Luther<sup>3</sup>

<sup>1</sup> University of Tsukuba, Tsukuba, Japan

{ksasaki, inoue}@slis.tsukuba.ac.jp

<sup>2</sup> RWTH Aachen, Aachen, Germany

anne.luther@rwth-aachen.de

<sup>3</sup> University of Duisburg-Essen, Duisburg, Germany

luther@inf.uni-due.de

**Abstract.** In this article, we explain how to process and interpret stage directions in order to support actors' performances. To this end, we specify the function of coverbal gesture and movement as an additional statement by an actor. We introduce a classification of communicative gestures by individuals and groups based on a description with typical body triangles and their sensor-detectable vertices moving in time. We claim that with the help of a speech and gesture recognition system, the occurrence of semantic isotopies to characterize a scene with speech and accompanying metaphoric gestures described in the stage directions can be detected and encoded in a feature vector. As a proof of concept, we highlight the ritual theater play *Orphée d'Afrique*, which was written as a second part of a homonymous chant novel and presents an implementation of a gesture- and voice-detection cuing system.

**Keywords:** Performance · Stage directions · Encoding standard · Body gesture and speech recognition system · Body gesture model · Cuing system

## 1 Introduction

A theater play is a cooperative task in which actors express themselves through text, gestures, sound and movements based on stage directions and a story, usually in the form of a script. They are assisted by prompters and technical staff responsible for actors' performances, set design, lighting, sound and music during rehearsals and performances. Some of these tasks have been taken over by technical installations and systems in research [13]. Takatsu et al. [15] proposed a system that detected speech and movement to support an actor's performance within an interacting group. The authors described the rehearsal cuing system architecture and developed a study to show that

the system helped actors by cuing the flow of events to each actor individually. An evaluation indicated the system might enhance actors' theatrical performance practice.

Ronfard [12] developed a complete workflow to publish recorded video data and the corresponding metadata from performances and related rehearsals online as Open Data. Recent improvements in ultrasound imaging enabling new opportunities for hand pose detection using wearable devices were reported in McIntosh [11]. Jessop [7] created and evaluated the Gestural Media Framework for gestural control of media in rehearsal and performance recognizing continuous gesture and translating Laban Effort Notation into the realm of technological gesture analysis. In [14], the authors propose "Digital-Script", a kind of stage directions that "focus on actor's performance factors, standing position, head direction and timing of actions." An implemented "theatrical practice support system detects user's standing position and head direction. In the virtual space, a virtual actor is shown as he performs along with Digital-Script information".

A crucial prerequisite for this is, in addition to suitable recognition hardware and software, to develop schemas for the standardized encoding of plays, orchestral scores, stage directions and scripts in order to be able to assess the participants' performance. For the encoding of drama and records of parliamentary debates, the element <stage> (for stage directions) was adopted in P5 Guidelines for Electronic Text Encoding and Interchange (TEI) [9, 16], and a formal transcription system for conversational gestures is described in [17].

According to Eco [2], *isotopy*, a concept first introduced by Greimas [5], has become "an umbrella term, a rather general notion that can allow for various more specific ones defining different textual phenomena"—even a gesture with a textual metaphorical interpretation. The recurrent occurrence of certain (clas)emes and subordinated lexemes in the text can be recognized and summarized as discursive isotopies together with their accompanying gestures described in the stage directions. Combining related discursive isotopies into a narrative strategy represents a reliable computer-based method for describing and quantifying literary constructs of interest.

To describe gestures, we adopt an approach followed by the authors in [8] providing a formal semantic analysis of coverbal iconic and deictic gestures with three important features based on the observation that speech and gesture together form a 'single thought'. First, the content of language and gesture are represented jointly in the same logical reasoning. Second, rhetorical relations connect the content of coverbal gesture and synchronous speech. Third, language and gesture are interpreted jointly within an integrated architecture for linking the form and meaning of utterances. Linguistic models can be applied to nonlinguistic objects in a metaphorical interpretation, but they supplement the nonlinguistic objects rather than replacing them, as the authors argue. Trippel [17] proposes the formal transcription system CoGesT, a machine processable computational model for annotating a subset of conversational gestures. Source and target location; trajectory described in terms of direction, type of movement, shape of the body part, and change in body shape during the movement; and modifiers for speed and size are composed to create a gesture feature vector; a BNF-grammar for the CoGesT system is offered in [4].

As examples of gesture types and their semantic content, the authors mention iconic, metaphoric, deictic and emblematic gestures. *Iconic*, word-like gestures clarify the spatial and dynamic features of a concrete concept/object; the notion *metaphoric*

refers to gestures that depict the content of an abstract idea through a physical action; *deictic* refers to pointing to a physically present or absent referent; *emblematic* gestures conform culturally and are universally understood by a community.

Section 2 deals with stage directions and proposes a classification of intended gestures and movements with emphasis on communicative gestures. Section 3 is devoted to gesture and body movement modeling based on relevant triangles, their trajectories and constraints combined with scene context in order to construct a feature state vector for motion description. In Sect. 4, M.-J. Hourantier's version of a ritual theater play *Orphée d'Afrique* [6] is presented as a use case. Section 5 describes a cuing system for an automated structural approach to its implementation and function. Finally, Sect. 6 offers conclusions and discusses further work.

## 2 Intentional Gestures and Movements Classification

The scope and frequency of stage or production directions vary according to the history of the theatre from quasi-nonexistence to taking an important place in the contemporary theatre [18]. Whereas in the dialogue the characters are speaking, in the didascalial the author himself directs, assigns positions to actors, has them speak or participate in a conversation and describes their gestures and actions. As Gallèpe [3] states, "Almost all of the didascalial indications concern bodily presence and its manifestations"; some describe the various locations. The didascalial mode is the most easily identifiable. First of all, it includes the stage directions formally inserted in multiple parts of the text and provides details about the body and its role in communication. The naming of the body in remarks made by the actors themselves is a second internal instruction mode [18]. The third mode is an implicit one: the body and its actions must be inferred from what the characters say or even from the tension existing between an intention and the content of the directions for obvious reasons of place and feasibility. Many didascalial are devoted to the indications concerning the personal façade of interacting characters. It seems preferable to focus exclusively on the didascalial dealing with notations of the gestures and positions of the bodies in the interactions: *Textual gestures* concern the phonation of the oral text of the interaction, *protaxial gestures* partially overlap pantomimic figurative or quasi-linguistic gestures, *coverbal gestures* constitute the nonverbal context of replicas, they correspond to the accompanying register, and *contextual gestures* denote activities that are not directly communicative but center on another action in relation to a concrete situation.

The method used by Gallèpe is as follows: he undertakes a methodical analysis of the didascalial of a number of German and French plays of various genres and periods constituting a basic corpus. For each of these 13 pieces, the first hundred replicas were considered. In 171 stage directions, the body is explicitly mentioned; in 77 it is only involved. The other didascalial relate to the various locations. They are classified according to four criteria: *unintentional gestures*; *gestures without intention of communication*; *intentional gestures of communication*; *undecidable*. The assignment of the gestures to a role is clearly described in the vast majority of cases. The corporal signifiers allowing translation within the gestures vary and are left open to interpretation. The different registers with occurrences found in the didascalial are as follows:

mimic (26), paraverbal utterance (23), gaze (19), static (7) and dynamic (29) posture, kinesics (20) (undifferentiated, hands, head, arm, torso, lower limbs), and proxemics (88). In this study, we decided to focus on the last four categories: static and dynamic posture, kinesics, and proxemics.

Greimas [5] differentiates between practical, mythical and communicative gestures. He mentions two lists covering behaviors such as (a) walking, running, sleeping, standing upright, etc. and (b) taking, giving, holding, pulling, pushing, etc. These two lists suggest the possibility of a very small inventory of simple and sufficiently general physical activities at the same time comparable to the very limited number of phonemes accounting for the totality of known articulations of natural languages. This sheds light on the parallelism between the gestural phenomena and the sememes covered by the verbs. In semantics, *sememe* is a bundle of minimal semantic strokes (called *semes*) whose formal correspondent is the lexeme.

According to an inventory of simple natural behaviors, it seems possible to reduce the gestural substance to figures in the visual plane and thus allow the division of this gestural text into minimal units whose combinatorial composition produces gestural statements and gestural speech itself. Unlike Greimas, Cosnier [1] emphasizes the absence of linear segmentation and the combinatorics of elementary units: gestures given their three-dimensional deployment can associate, combine, and condense. For this reason, we limit ourselves to a limited sample of gesture examined in Sect. 5.

### 3 Gesture and Body Movement Modeling

Recognizing the gestures and movements of a group of actors requires a few preliminary considerations and preparations, including their formal description and software. The process is as follows: Identify actors performing there, and record the positions. This makes it possible to judge whether each actor stays in a specific position correctly in each scene. Actors' body parts are also recorded so that their movements can be recognized by a machine-readable computational model that measures geometrical figures. It translates the stage directions into a machine-processable form (classifies the gestures in accompanying, replacing, explaining text), resulting in a script. With respect to a theater rehearsal, the script tags elements such as typical actors' constellations (i.e., how the scenes have been blocked) and their speech, gestures and movements based on the TEI P5. Formal description of body gestures is essential to recognizing actors' movements. The procedure is illustrated below.

- Place a depth camera in front of a stage to detect the vertices of the actors' body.
- A spatiotemporal description of an actor's movement as a sequence of body gestures is produced using involved triangles  $T_k$ . Also, a trajectory  $t$  can be constructed using an individual's ( $p$ ) positions and directions, start/source position  $s$  and endpoint/target position  $e$  on a surface, and timespan. Each position is described in an xyz coordinate system, each actor's endpoint position is checked to ensure he or she is in the correct position at the end of each scene.
- Spatial relations and constraints  $C$  rely on topological, directional, proximal or distance features and connectivity containment using contiguity predicates.

A depth camera is used to detect the vertices of each actor's head, neck, shoulders, elbows, hands, torso, femur balls, knees, and feet. Also, triangles constructed with vertices on the head, arms and legs, upper body, lower body, whole body, and three edges are measured to determine the normal vector pointing away from the body (cf. Table 1). Body gestures are modeled by moving the limbs and recording trajectories (cf. Table 2). For example, encounters between persons and communication can be described as distances of the corresponding body parts and directions of the triangles  $T_1, T_4, T_5, T_6, T_7$ , e.g., extending arms, a handshake or a hug. Using the triangles and their normal vectors, trajectories  $t_k(v_a), t_k(v_b), t_k(v_c)$ , distances  $d(v_{k1}, v_{k2})$  and gesture vectors  $f$  can be computed.

**Table 1.** Examples of body movements and their descriptions using 24 vertices  $v_k$ , 16 triangles  $T_k$ , and normal vectors  $nv_k$  pointing outside (os). In the current version,  $T_2, T_3, T_{13}$  and  $T_{14}$  cannot be acquired with the system, but they can be relevant to individual movements, depending on the camera positions.

T	Body part	Vertex a	Vertex b	Vertex c	$ab \times ac = nv$
1 <sub>fb</sub>	Front/Back (F/B) Body	Front/occiput	Right (R) forefoot/heel	Left (L) forefoot/heel	Forward/Backward
2	Head1	Chin	L cheekbone	R cheekbone	Forward
3	Head2	Occiput	Right ear	Left ear	Up
4	Head3	Nose	Right shoulder	Left shoulder	Forward
5	Upper body	Abdomen/navel	Left shoulder	Right shoulder	Forward
6	Right arm	Right shoulder	Right elbow	Right palm	To the right os
7	Left arm	Left shoulder	Left palm	Left elbow	To the left os
8	Right leg	Right femur b.	Right heel	Right patella	To the right os
9	Left leg	Left femur ball	Left patella	Left heel	To the left os
10	Lower body1	Abdomen/navel	Right forefoot	Left forefoot	Forward
11	Lower body2	Abdomen/navel	Right patella	Left patella	Forward/up
12	Lower body B	Pelvis	Right heel	Left heel	Backward
13	Left lower leg	Left patella	Left forefoot	Left heel	To the left os
14	R lower leg	Right patella	Right heel	Right forefoot	To the right os
15	Action hands	Nose	Right palm	Left palm	Forward

The advantage of using gesture modeling with triangles, vertices, and their normal vectors is that triangles always lie in a plane described by *Hesse normal form*  $nv_x(x - v_{ax}) + nv_y(y - v_{ay}) + nv_z(z - v_{az}) = 0$ . If one assumes the standard case of a plane stage, then it can be equipped with an orthogonal three-dimensional coordinate system  $(x, y, z)$ . The actors move on the floor  $z = 0$  following flat curves, and many constraints can be checked concerning distances or the position of a point with respect to a plane, circle or polygon by inserting the point coordinates into curve or surface equations.

**Table 2.** Examples of body movements and their descriptions: a non-exhaustive list.

Textual/graphic description in 3-dim. space	Triangle ${}_pT_k$ $v_a, v_b, v_c, nv$ Persons $p, q$	Trajectories $t$ ${}_pt (v.) s_v \rightarrow e_v,$ $d(p, q), d(v_i, v_j)$	Constraints $C$ $C_j({}_pt({}_pT_k), cond_j)$ $C_j({}_pt({}_pT_{k1}), {}_qt({}_qT_{k2}), cond_j)$
$\ddagger \rightarrow \leftarrow \ddagger^a$ 2, $2n$ persons	${}_pT_1 ({}_pT_5)$ ${}_qT_1 ({}_qT_5)$	Decreasing distances, opposite directions ${}_p, {}_q, nv$	Feet on the ground, triangles $\perp$ to the plane of motion
Handshake 2, $2n$ persons	${}_1T_6, {}_2T_6$	Hand distances decrease	Walking upright, feet on the ground
Go upright to a certain place $e$	$T_1$	Decreasing distance to the goal $d(v_b, e)$	Feet in the plane, triangles $\perp$ to the plane of motion
Hold an object in both hands	$T_6, T_7$	Same distance between both palms $v_b, v_c$	Hands are holding a musical instrument
Climb the seven steps reach attic, look to the sky	$T_1, T_2 (T_4)$	Repetitive distances between forefeet $v_b,$ $v_c$ ; direction of $NV(T_2) \nearrow$	Starting point $s$ , eight changing positions of left and right forefoot; end point $e$ on attic
Spinning in a $\bigcirc$ in the $xy$ -plane	$T_1$	Subsequent positions	Point in circle $nv_x/\sqrt{nv_x^2 + nv_y^2}$ passes through $[-1, 1]$
Lie/roll on the bed 1, 2 persons	$T_1, T_5,$ $(T_6, T_7); T_5$	Static horizontal position $e = s; nv$ in $yz$ -plane	Vertices of the triangles in/near the plane

<sup>a</sup>Two persons  $p, q$  walk upright, following trajectories  ${}_pt, {}_qt$  to meet each other, the normal vectors  ${}_pnv_1, {}_qnv_1$  of the front body triangles  ${}_pT_1, {}_qT_1$  are directed nearly parallel to the ground ( $z = 0$ ) and opposite, the distance between the triangles (smaller than  $\min_{1 \leq i, j \leq 3} d({}_pv_i, {}_qv_j)$ ) falls below a certain limit.

## 4 Gestures in Ritual Theater Play: A Use Case

The purpose of the second author’s dissertation [10], *Werewere Liking: Ritual and Writing: The Postcolonial Subject and Its Discursive and Narrative Strategies*, is to examine Liking’s five chant novels: *A la rencontre de...* (1980), *Orphée Dafric* (1981), *Elle sera de Jaspe et de Corail* (1983), *L’Amour-cent-vies* (1988) and *La mémoire amputée* (2004) from the angle of common research topics. In the novel *Orphée Dafric*, Orpheus goes through a series of exams and meditations—an individual initiation enriched with positive achievements—and including abstract images that he tries to render concrete and effective in accomplishing an archetypal rebirth.

The object of the research is to find coherence and differences between the five novels and to establish the paradigm of human development from a person who is alienated, ill, searching, and decadent to one who is responsible, engaged, and harmonious—a member of a new mankind, in five different variations: the binary, archetypal, prophesized, liminal and transgenerational accomplishment of human character development. Semantic isotopies and narrative strategies are made visible in the process of a character’s metamorphosis during the phases of denunciation, perspective development and synthesis. Systematic analyses of 10 preselected isotopy groups based on the semantic breakdown of the ambiguities in the corresponding 34 lexemes enable the various developments of the main characters to be described and

located at a higher level and distinguishable narrative strategies to be identified. (Clas) semes and subordinated lexemes are found by means of an encyclopedic definition; automatically searched related citations are listed in a table, then contextualized, actualized or virtualized in relation to both main and minor characters. A similarity measure was developed which determines differences in how the main characters evolve [10, p. 435]. After an initial exemplary analysis of the play *Orphée d'Afrique* [6], it turns out that in addition to the genre-related differences to the novels, content and structural peculiarities also occur that require an adaptation of the chosen method and selected isotopies. While, as in the novel the characters are in an initial state of unconsciousness, impotence, deprivation and isolation, other elements of initiation and development become decisive in the play. Formulated definitions, such as of the term initiation, need to be adapted, figure constellations and their dynamics must be rewritten.

As an extension, this approach analyzes the speech and movements of the actors in parallel using a recognition system to synchronize text isotopies and typical coverbal or mythical gestures: Stretch out, lie down, sit down, get up, take a few steps, climb stairs, stay motionless, leave or enter the scene, kneel down, bring something to the audience, take/hold something in hands, make several forms with the hands, perform ritual actions like baptism or passing within a circle or polygon (cf. Table 3).

**Table 3.** Actors, isotopies, accompanying movements, and gesture type.

LE NDINGA	(Chanting his sentence with the mvet*, goes to Orpheus.) * stringed musical instrument	Coverbal
ORPHEUS Engaged, responsible: active, positioned, fighting	With help from the tam-tam, O. looks for a rhythm that inspires him. A flute is born	(Pro)textual
ORPHEUS Searching (cherchant): deprivation, paralysis, wish, quest	Spinning in circles, looking for an object; An object circulated among the spectators will be the symbol of Orpheus's quest	Contextual Extra-communicatory
ORPHEUS Sick (malade): cause, symptom, cured	Lies down, turns on the bed	Contextual
ORPHEUS Decadent: decline, absurdity, perversion, restart	He climbs "seven steps" and reaches the attic, stops and is looking up at the sky	Contextual Mythical
ORPHEUS Initiated: transformed, firm, conscious, harmonious, healthy	Traverse several compartments of a circle; Should be assigned to the new isotopy "challenge"	Contextual Mythical

The system localizes people, one near or above the other, alone or in pairs, at the top or in the middle of a group and detects which of these patterns are in sync with text isotopies. Clear scripts and stated directions illustrating dialogues or concerning the gestures or movements of the actors, formally described and tagged, support the premise that this computerized approach could be successfully implemented.

## 5 System Design and Implementation

This section briefly explains how the cuing system has been designed, implemented and enhanced with gesture recognition.

### 5.1 Stage Directions

First, a standardized digital version of the play’s stage directions is made with the actors, with their accurately dimensioned body triangles in a basic position with respect to a standard rectangular flat stage. Then, actors’ gestures, movements, and actions are listed in a common timeline with accompanying text and instructions in the detailed standardized XML format.

### 5.2 Recognition of Actors’ Gestures

Actors’ gestures can be recognized as described in the following examples taken from Table 3.

For stair climbing, “Climb the seven steps to reach the attic, look at the sky”, the start and end points of the trajectory and the midpoint between the foot markers of triangle  $T_1$  occur along a straight line. The z-coordinate is controlled with regard to uniform increments from 0.5 to 6.5 times the height of the step. Finally, the direction of the normal vector of the head triangle  $T_2$  now points diagonally above.

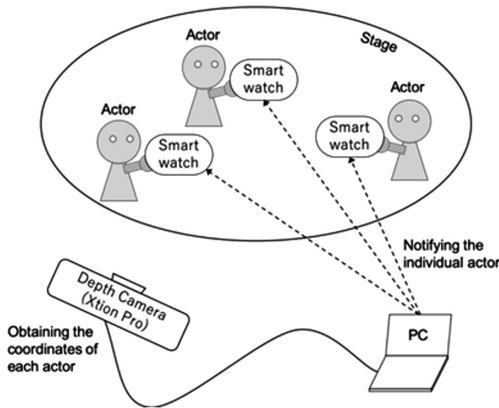


Fig. 1. The cuing system.

For “Spinning in a circle”, the trajectory of the feet vertices of  $T_1$  stays in the circle and  $nv_x/\sqrt{nv_x^2 + nv_y^2}$ , the cosine of the x- and y-coordinate of the normal vector  $nv$  in the plane, varies from 1 over  $-1$  back to 1. For “Lies down and turns on the bed”, take the upper body triangle  $T_5$  and the cosine of the y and z coordinates of the normal vector  $nv$  in the yz-plane. Normal vectors are calculated by the system.



### 5.3 Implementation

The system has been implemented according to the design. It uses Xtion PRO as a depth sensor and OpenNI to obtain the three-dimensional coordinates of each vertex. It calculates the normal vectors in the Sect. 3 to recognize actors' gestures. Each actor wears a smartwatch, and the server notifies the individual actor when he/she should start speaking or acting (Fig. 1).

```

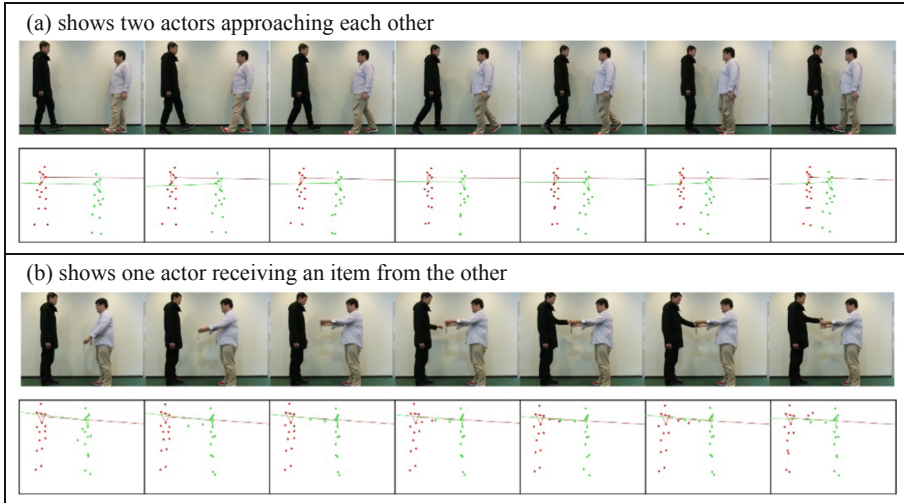
<?xml version="1.0" encoding="UTF-8"?>
<body>
  <head>Passing an Item</head>
  <sp n="0">
    <pos>-1287.4366,240.55136,3361.6633,1194.0463,266.115,3434.0425</pos>
    <speaker>Actor A</speaker>
    <move pref="stand" type="onStage"/>
    <p>I give you a present.</p>
  </sp>
  <sp n="1">
    <pos>-437.1841,228.42166,3336.9055,379.03928,232.34496,3465.1372</pos>
    <speaker>Actor B</speaker>
    <move pref="approach(A,B)" type="onStage"/>
  </sp>
  ...

```

Fig. 2. An example XML script.

An XML script file is stored in advance that includes encoded text and the position coordinates indicating where actors should stay in each scene (Fig. 2). In a play, it measures any positional gap between the coordinates in the script and the actors' actual positions. It also compares gestures with those in the script.

- (a) The system detects that actors' positions are converging by investigating the two normal vectors  $nv$  directed against each other (Table 2 “ $\uparrow \rightarrow \leftarrow \downarrow$ ”).
- (b) The system detects that the distance between the actors' right hands is decreasing (Fig. 3).



**Fig. 3.** In each set of images, upper images show the actors captured by RGB camera and lower images show the body vertices, triangles  $T_5$ , and normal vectors  $nv$ . The time axis runs from left to right.

The system also detects whether an actor is speaking or not to track the script automatically. Only the volume of their voices is used to avoid any influence from inaccuracy of speech recognition.

## 6 Conclusions and Further Work

With the help of a speech- and gesture-recognition system for cuing actors' utterances and gestures, we have shown that body gestures detailed in stage directions can be recognized and brought into a machine processable form. By combining semantic isotopies from an actor's speech and his body gestures, an individual story is generated that can largely be diversified by actors' spontaneous interactions.

However, a lot of practical work remains to be done to enhance the system. It should be extended to include full speech recognition, to detect and classify improvised actions and to deal with larger scale plays. With such a system, a database of collected descriptors of body gestures from a number of plays already recorded in the literature can be built, which is a prerequisite for investing in the implementation of a scene editor and player. Also, several open questions remain: How can the simultaneous actions of audiences be modeled? How should instructions concerning rhythm and facial expressions or behavioral gestures be described? What is the best way to describe complex actions like creating harmony or fighting effectively?

In addition, a research team of literacy, media and computer science experts could undertake an analysis of multimodal interactions in theater plays, including musical performances and dance, based on patterns from primitives of a multimodal gesture and sound language. There is a reasonable presumption that semantic isotopy analysis can

be most profitably extended to various kinds of theater with a mixture of language, intonation, gesture and dance for a better description and understanding of multimodal utterances.

## References

1. Cosnier, J., Vaysse, J.: Semiotics of communicative gestures. *New Semiot. Acts* **52**, 7–28 (1997)
2. Eco, U.: Two problems in textual interpretation. *Poet. Today* **2**(1a), 145–161 (1980). Roman Jakobson: Language and Poetry
3. Gallèpe, Th.: Joindre le corps à la parole: Quelle place pour le corps dans les textes de théâtre? Université Michel de Montaigne, Bordeaux, Cahier du C.I.E.L. Paris (1998–1999)
4. Gibbon, D., Gut, U., Hell, B., Looks, K., Thies, A., Trippel, Th.: A computational model of arm gestures in conversation. In: *Proceedings of Eurospeech*, Geneva, pp. 813–816 (2003)
5. Greimas, A.J.: *Sens, Semiotic Essays*. Seuil, Paris (1970)
6. Hourantier, M.-J.(Manuma ma Njock): *Orphée d’Afrique: théâtre-rituel d’initiation*, Paris, L’Harmattan (1981). Together with Liking W.: *Orphée Dafric, chant novel*
7. Jessop, E.A.: *A gestural media framework: tools for expressive gesture recognition and mapping in rehearsal and performance*. B.A., Amherst College (2008)
8. Lascarides, A., Stone, M.: A Formal semantic analysis of gesture. *J. Semant.* **26**(4), 393–449 (2009)
9. Lünen, H., Sperberg-McQueen, C.M.: A TEI P5 document grammar for the IDS text model. *J. Text Encoding Initiat.* **3** (2012). <http://journals.openedition.org/jtei/508>
10. Luther, A.C.: *Werewere Liking – Ritual und Schreiben: A Critical Debate in Africa and Latin America*. Peter Lang (2017). (in German)
11. McIntosh, J., Marzo, A., Fraser, M., Phillips, C.: EchoFlex: hand gesture recognition using ultrasound imaging. In: *Proceedings of CHI 2017, CHI Conference on Human Factors in Computing Systems*, pp. 1923–1934. ACM, New York (2017)
12. Ronfard, R., Encelle, B., Sauret, N., Champin, P.-A., Steiner, Th., et al.: Capturing and indexing rehearsals: the design and usage of a digital archive of performing arts. In: *Digital Heritage*, Grenade, Spain, September 2015, pp. 533–540. IEEE (2015)
13. Sasaki, K., Inoue, T.: Coordinating real-time serial cooperative work by cuing the order in the case of theatrical performance practice. *Mob. Inf. Syst.* **2019** (2019). Article ID 4545917, 10 pages
14. Shimada, M., Takano, T., Shigeno, H., Okada, K.: Supporting theatrical performance practice by collaborating real and virtual space. In: Yoshino, T., Chen, G.-D., Zurita, G., Yuizono, T., Inoue, T., Baloian, N. (eds.) *CollabTech 2016*. CCIS, vol. 647, pp. 17–30. Springer, Singapore (2016). [https://doi.org/10.1007/978-981-10-2618-8\\_2](https://doi.org/10.1007/978-981-10-2618-8_2)
15. Takatsu, R., Katayama, N., Inoue, T., Shigeno, H., Okada, K.: A wearable action cueing system for theatrical performance practice. In: Yoshino, T., Chen, G.-D., Zurita, G., Yuizono, T., Inoue, T., Baloian, N. (eds.) *CollabTech 2016*. CCIS, vol. 647, pp. 130–145. Springer, Singapore (2016). [https://doi.org/10.1007/978-981-10-2618-8\\_11](https://doi.org/10.1007/978-981-10-2618-8_11)
16. TEI P 5 Release, TEI P5: Guidelines for Electronic Text Encoding and Interchange. <http://www.tei-c.org/guidelines/p5/>
17. Trippel, Th., et al.: CoGesT: a formal transcription system for conversational gesture. In: *Proceedings of IRES*, pp. 2215–2218 (2004)
18. Ubersfeld, A.: *Lire le théâtre I*, Belin Lettres Sup. (1996). 240p. English version *Reading theatre*. Toronto Press (1999)