Bekkering H, Wohlschläger A, Gattis M (2000) Imitation of gestures indicates that, similarly to music perception, music production too top of the acquisition of motor flexibility. More generally, this finding planning based on musical syntax as a slowly acquired skill built on syntax-based motor control with expertise might hint at the action specific, immediate movement selection. Moreover, the increase of coherently with the context and forehead those ones underlying musical structure, motor plans for distal musical goal are generated movements' specifications, especially in more expert pianists being more affected by the priming effect of the contextual syntactic structure.

Taken together, these findings indicate that, given a contextual musical structure, motor plans for distal musical goal are generated coherently with the context and forehead those ones underlying specific, immediate movement selection. Moreover, the increase of syntax-based motor control with expertise might hint at the action planning based on musical syntax as a slowly acquired skill built on top of the acquisition of motor flexibility. More generally, this finding indicates that, similarly to music perception, music production too relies on generative syntactic rules.

References

Motor learning in dance using different modalities: visual vs. verbal models
Bettina Bläsing1, Jenny Coogan2, José Biondi2, Liane Simmel2, Thomas Schack2
1 Neurocognition and Action Research Group & Center of Excellence Cognitive Interaction Technology (CITEC), Bielefeld University, Germany; 2 Palucca Hochschule für Tanz Dresden, Germany; 3 named Tanzmedizin Deutschland e.V., Fit for Dance Praxis and Institut für Tanzmedizin, München, Germany

Keywords
Motor learning, Observation, Visual model, Verbal instruction, Dance

Introduction
Observational learning is viewed as the major mode of motor learning (Hodges et al. 2007). Empirical evidence shows that observational learning primarily takes place in an implicit way, by activating shared neural correlates of movement execution, observation and simulation (Jennefer 2004; Cross et al. 2006, 2009). It has been shown that the use of language (in terms of verbal cues) can facilitate or enhance motor learning by guiding attention towards relevant features of the movement and making these aspects explicit (see Wulf and Prinz 2001). In dance training (and other movement disciplines), observational learning from a visual model is most commonly applied, and is often supported by verbal cue-giving. Evidence from practice suggests that explicit verbal instructions and movement descriptions play a major role in movement learning by supporting the understanding, internalizing and simulating of movement phrases. In modern and contemporary dance, however, choreographers often do not expect the dancers to simply reproduce movement phrases in adequate form, but to develop movement material on their own, in accordance with a given idea, description or instruction, aiming at a more personal expression and higher artistic quality of the developed movement material.

In this study, we investigate dancers’ learning of movement phrases based on the exclusive and complementary use of visual model observation and verbal instruction (movement description).
Dance students learned comparable movement material via two different modes: via observation of a model and via listening to a verbal movement description (as example, a part of a model sequence is displayed in Fig. 1). In a second step, the complementary mode was added. After both learning steps, the students’ performance of the learned movement phrases was recorded and rated by independent experts. A retention test was applied to evaluate long-term effects of the learning processes. We expected the dance students to learn successfully from the visual model, their most commonly practiced mode of movement learning. From the verbal instruction, we expected that performed movement phrases would vary more strongly, but could possibly be performed with more artistic quality. We also expected performance after the second learning step to be improved compared to the first learning step in both conditions.

**Method**

**Learning task:** Eighteen students (age: 18.4 ± 1.0 years, 11 female) from the BA Dance study program at the Palucca Hochschule für Tanz Dresden learned two dance phrases of similar length (approx. 30 s) and complexity, one via visual observation of a demonstration video, the other one via a recorded verbal description (see Fig. 1). In a first learning step (Step 1), one of the dance phrases was presented five times either visually (video) or verbally (audio), and the participant was instructed to learn it by watching or listening, and by marking movements as required. After a short practice, the participant performed the learned dance phrase while being recorded on video. In a second learning step (Step 2), the participant was twice presented the same dance phrase in the complementary presentation mode (i.e., video for the verbally learned phrase and vice versa), and the performance was recorded again. The other dance phrase was then learned and performed using the same procedure, but was presented in the remaining learning mode (verbal or visual) in the Step 1, complemented by the other mode in the Step 2. The order of the dance phrases (Phrase 1, Phrase 2) and of the initial learning modes (visual, verbal) was balanced between the participants (the experimental design of the study is illustrated in Table 1). The experimental procedure took place in a biomechanics laboratory and lasted approximately one hour for each participant. Additional to the evaluation of the recorded performances, questionnaires and psychometric tests were applied to investigate the students’ learning success and their personal impressions of the different learning processes.

**Expert ratings of the reproduced material:** Two independent experts rated the recorded performance trials from the recorded and cut video clips, one of each demonstration condition (visual, visual + verbal, verbal, verbal + visual). The experts rated each of the recorded performances by filling out a questionnaire consisting of six-point Likert-scale type questions assigned to two categories, accordance with the model (AM: 10 questions) and artistic performance quality (PQ, 5 questions). For each category of questions, ratings of the questions were averaged to achieve general measures for the main criteria AM and PQ. Each expert independently watched the recordings from the students’ performances and marked one answer for each question, without knowing about the learning condition of the recorded performance. Non-parametric tests (Wilcoxon signed-rank, Mann–Whitney U) were used to compare the averaged ratings of the two experts for the different conditions (visual, visual + verbal, verbal, verbal + visual) within each criterion (AM, PQ) and for the two criteria within each demonstration condition.

**Retention test:** Thirteen of the dance students (8 female) participated in a retention test that was carried out 10–13 days after the experimental learning task. The retention test included the video-recorded performance of the remembered movement material, psychometric tests and questionnaires. In the performance part of the test, each student was asked to perform both dance phrases as completely as possible. Students were allowed to practice for several minutes before being recorded, but were not given any assistance in reproducing the phrases. Each student was recorded individually and on his/her own in a separate dance studio. The video recordings of the students’ performance in the retention test were annotated for the completeness of the phrases by two annotators. Each phrase was segmented into eleven partial phrase, or elements, of similar content (note that the phrases had been choreographed to resemble each other in complexity, duration and structure). The annotators independently watched the recordings and marked the completeness of each of the eleven elements as value between 0 and 1 (0: the element was not danced at all, or was not recognizable; 1: the element was clearly recognizable and was performed without error); ratings of the two annotators were then averaged. Each student thereby received for each of the two phrases a value between 0 (no partial phrase was reproduced at all) and 11 (all partial phrases were reproduced perfectly). Non-parametric tests (Wilcoxon signed-rank, Mann–Whitney U) were used to compare averaged completeness scores between dance phrases (Phrase 1, Phrase 2) and learning modes (visual first, verbal first).

**Results**

**Expert ratings:** Ratings of the two experts were positively correlated for both criteria, AM (r = 0.528; p < .001) and PQ (r = 0.513; p < .001). After Step 1, ratings of PQ were significantly better than ratings for AM (visual: 3.82, 3.33; Z = −2.987, p < .003; verbal: 3.73, 2.69; Z = −3.529, p < .001), whereas ratings did not differ after Step 2. AM ratings after learning only from verbal description was lower (2.69) than after all other conditions (verbal + visual: 3.48, Z = −3.724, p < .001; visual: 3.33, Z = −3.624, p < .001; verbal + visual: 3.65, Z = −3.682, p < .001), and AM ratings after visual + verbal learning were higher than after visual learning (Z = −2.573, p = .01). PQ ratings did not differ for any of the learning conditions.

---

**Fig. 1** Images illustrating approximately two-thirds of Phrase 1, choreographed by Jenny Coogan and performed by Robin Jung. The phrase was presented as video of 26 s and as audio recording of a verbal description (speaker: Alex Simkins). Phrase 2, choreographed by José Biondi, was of similar length and complexity and contained similar movement elements as Phrase 1, and was performed and spoken by the same dancer and speaker in the video and audio recording, respectively. The verbal description of the dance sequence shown in the pictures reads as follows: “Stand facing the front left diagonal of the room in first position. At the same time extend your left leg forward and your two arms sideways to the horizontal. Allow your right hand to continue moving until it arrives to a high diagonal. Gradually let the shape melt back into its beginning position as you shift your weight into the right hip, bending both knees, sinking your head to the left to make a big C-curve. Continue into falling, then catch the weight with a step of the left leg crossing to the right. Follow with two steps sideward, in the same direction while throwing both arms in front of your shoulders. Keeping your arms close to you, spiral to the right diagonal, then, kick your right leg, left arm and head forward as you throw your right arm behind you. Bring the energy back into you quickly bending both elbows and the right knee close to the body, spine vertical. Drop your arms and take a step back onto your right leg turning fully around while dragging your left leg behind you. Finish with the weight low, left leg behind, spine rounded forward, arms wrapped around the body, right arm front, left arm back. Stretch your legs and gradually lengthen your spine horizontally. Allow your arms to follow the succession of your spine, right front, left back.”
Retention test: Completeness scores given by the two annotators were highly correlated for both sequences (Phrase 1: $r = 0.942$, $p < .001$; Phrase 2: $r = 0.930$, $p < .001$). No differences were found between the groups (Group 1: Phrase 1 verbal first, $N = 5$; Group 2: Phrase 1 visual first, $N = 8$) in general, and no differences were found between the two sequences (Phrase 1: 7.64; Phrase 2: 6.90). Scores were better for the first visually learned phrase (8.32) than for the phrase first learned from verbal description (6.23) ($Z = -1.992$, $p = .046$). When the sequences were regarded separately, groups differed for Phrase 2 (Group 1: 9.17; Group 2: 5.48), but not for Phrase 1 (Group 1: 7.42; Group 2: 7.78), with Group 1 performing better than Group 2 ($Z = -2.196$, $p = .028$) (see Fig. 2). When comparing ratings for the individual elements (1 to 11), primary effects were found for both dance phrases, in terms of higher scores for the first 3 and 2 parts in Phrase 1 and Phrase 2, respectively (Phrase 1: element 1 differed from 6, 7, 8, 9 and 11; 2 differed from 5, 6, 7, 8, 3 differed from 4, 5, 6, 7, 8, 9 and 10; Phrase 2: 1 differed from 3, 4, 5, 6, 7, 8, 9, 10 and 11; 2 differed from 4, 7, 9 and 10; all $p < .05$).

**Discussion**

Interdisciplinary projects linking dance and neurocognitive research have recently come to increasing awareness in artistic and scientific communities (see Bläsing et al. 2012; Sevdalis, Keller 2011). The presented project on observational (implicit) and verbal (explicit) movement learning in dance has been developed within an interdisciplinary network (Dance engaging Science; The Forsythe Company | Motion Bank), motivated by scientific, artistic and (dance-) pedagogical questions. We compared expert ratings for the recorded performance of two different movement phrases in 18 dance students who had learned one phrase initially via verbal description and the other one via observation of a video model. After dancing the phrase and being recorded, students received the complementary modality to learn from, and were recorded performing again. Ratings for performance quality were better than rating for model reproduction after the first learning step (one modality), but not after the second learning step (two modalities). After learning from only one modality, ratings for accordance with the model were better if the first learning modality was visual than verbal, whereas ratings for performance quality did not differ for visual vs. verbal learning. When the students had to reproduce the learned movement material in a retention test, the (initially) visually learned material was reproduced more completely than the verbally learned material, however, when the dance phrases were regarded separately, this result was only significant for one of the phrases. The results corroborate findings regarding observational learning of movements in dance and other disciplines or tasks, but also suggest dissociation between the exact execution of a model phrase and the artistic quality of dance, even in the learning phase. As expected, accordance with the model phrases was stronger after visual learning and after two compared to one modalities (which might as well have been influenced by the additional practice, as this was always the second learning step.) Regarding artistic quality of performance, the students danced the newly learned material after learning from verbal description as well as after learning from visual observation, but not better, as we had expected. Questionnaires and psychometric tests are currently being analyzed to complement the reported findings of this study. We expect the outcomes to contribute to our understanding of explicit and implicit motor learning on the basis of different modalities, and also to yield potential implications for teaching and training in dance-related disciplines. While explicit learning (via verbal instruction) and implicit learning (via observation and practice) have been found to work synergistically in skilled motor action (Taylor and Iyvi 2013), the situation might be different for dance and potentially for dance-like movement in general (see Schachner and Carey 2013), in which skilful movement execution largely depends on kinesthetic awareness; further research is needed at this point. Further implications could be derived for

![Fig. 2](image-url)
learning in general, specifically regarding the potential benefit of combining different modes (or modalities) for conveying information in order to shape and optimize learning success.

References
Schachner A, Carey S (2013) Reasoning about ‘irrational’ actions—when intentional movements cannot be explained, the movements themselves are seen as the goal. Cognition 129:309–327
A frontotemporoparietal network common to initiating and responding to joint attention bids
Nathan Caruana, Jon Brock, Alexandra Woolgar
ARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University, Sydney, Australia

Joint attention is the ability to interactively coordinate attention with another person to objects of mutual interest, and is a fundamental component of daily interpersonal relationships and communication. According to the Parallel Distributed Processing model (PDPM; Mundy, Newell 2007), responding to joint attention bids (RJA) is supported by posterior-parietal cortical regions, while initiating joint attention (IJA) involves frontal regions. Although the model emphasizes their functional and developmental divergence, it also suggests that the integration of frontal and posterior-parietal networks is crucial for the emergence of complex joint attention behavior, allowing individuals to represent their own attentional perspective as well as the attentional focus of their social partner in parallel. However, little is known about the neural basis of these parallel joint attention processes, due to a lack of ecologically valid paradigms.

In the present study, we used functional magnetic resonance imaging to directly test the claims of the PDPM. Thirteen subjects (9 male, \[ M_{\text{age}} = 24.85, SD = 5.65 \]) were scanned as they engaged with an avatar whom they believed was operated by another person outside the scanner, but was in fact controlled by a gaze-contingent computer algorithm. The task involved catching a burglar who was hiding inside one of six houses displayed on the screen. Each trial began with a ‘search phase’, during which there was a division of labor between the subject and their virtual partner. Subjects were required to search a row of three houses located at either the top or bottom of the screen, whilst the avatar searched the other row. When the subject fixated one of their designated houses, the door opened to reveal an empty house or the burglar (see Fig.1a). The location of the subject’s designated houses was counterbalanced across acquisition runs. Subjects were instructed that whoever found the burglar on each trial had to guide their partner to that location by first establishing mutual gaze and then looking at the appropriate house.

On RJA trials, subjects searched their designated houses, each of which would be empty. The avatar would then complete his search and guide the subject to the burglar’s location. The avatar responded by gazing at the location fixated by the subject, regardless of whether it was correct or not. Again, positive feedback was provided when joint attention was achieved at the burglar’s location. Negative feedback was also provided if the subject failed to make a responsive eye movement within three seconds, or if they responded or initiated by fixating an incorrect location.

During the search phase, the avatar’s gaze behavior was controlled so that he only completed his search after the subject completed their search and fixated back on the avatar. This meant that subjects were required to monitor the avatar’s attention during their interaction, before responding to, or initiating a joint attention bid. This paradigm—as in ecological interactions—establishing mutual gaze was therefore essential in determining whether the avatar was ready to guide the subject, or respond to the subject’s initiation of joint attention. The onset latencies of the avatar’s gaze behavior (i.e. alternating between search houses, establishing mutual gaze, and executing responding or initiating saccades) were also jittered with a uniform distribution between 500 and 1,000 ms. This served to enhance the avatar’s ecological appearance.

The subject’s social role as a ‘responder’ or ‘initiator’ only became apparent throughout the course of each trial. Our paradigm thereby created a social context that (1) elicited intentional, goal-driven joint attention (2) naturally informed subjects of their social role without overt instruction, and (3) required subjects to engage in social attention monitoring.

In order to account for the effect of non-social task features, the neural correlates of RJA and IJA were investigated relative to non-social control conditions that were matched on attentional demands, number of eye movements elicited and task complexity. During these trials, the avatar remained on the screen with his eyes closed, and subjects were told that both partners were completing the task independently. In the IJA control condition (IJA), subjects found the burglar, looked back to a central fixation point and, when this turned green, saccaded towards the burglar location. In the RJA control condition (RJA), the fixation point became an arrow directing them to the burglar location (see Fig. 1b).

A synchronization pulse was used at the beginning of each acquisition run to allow for the BOLD and eye tracking data to be acquired.