The relationship between human motion and objects in the cognitive representation of visual events.

Margot D. Lasher
Assumption College

This paper concerns the cognitive representation of events which involve human motion.* I have elsewhere proposed a schematic structure for events which involve only the motion of the body, as in some athletics and dance (Lasher, 1981). In this paper I would like to describe the relationship between the motion of the body and objects which may be part of the event. The eventual goal is to develop schematic representations for all types of events of human motion.

The paper will look at two pairs of event units. The first event of each pair involves only the motion of the body; the second involves that same category of motion in relation to an object. Certain aspects of these paired events can be identified without further analysis. We are dealing with only one type of event unit, an event involving voluntary human motion. The AGENT of the event will always be the person in motion. The ACTION will consist of a description of that person's motions.

Event 1: A person leans, reaches out, and brings the arm back toward the body.

Event 2: A person leans, reaches out, picks up a glass and brings the glass back toward the body.

Figure 1 provides schematic structures for Events 1 and 2. Event 1 describes a motion unrelated to any object. It might be a movement in athletics or dance. The ACTION consists of a preparatory motion and a completing motion. There is experimental evidence for the psychological coherence of this preparatory-completing structure in the perception of events involving human motion (Lasher, 1981). A psychologically coherent event unit consists of a small, unfixed number of preparatory motions followed by a completing motion. Preparatory motions tend to be relatively stable motions, while completing motions are relatively unstable motions. The unstable completing motion is encoded as the intention of the entire event unit. The event unit is encoded as completed when this relatively unstable motion is finished and the person has returned to a position of stability in relation to the ground.

Each motion, whether preparatory or completing, can be described in terms of both the changes in the human body itself, labeled MOTION in Figure 1, and in terms of the variables which influence that motion, labeled MOTION RELATIONS. There are a potentially infinite number of motion relations. The actual representation of any event in a real mind will depend upon what variables are attended to at the time. I have assumed, for purposes of illustration, that we are attending to the direction of the lean and reach, which is forward, and to nothing else.

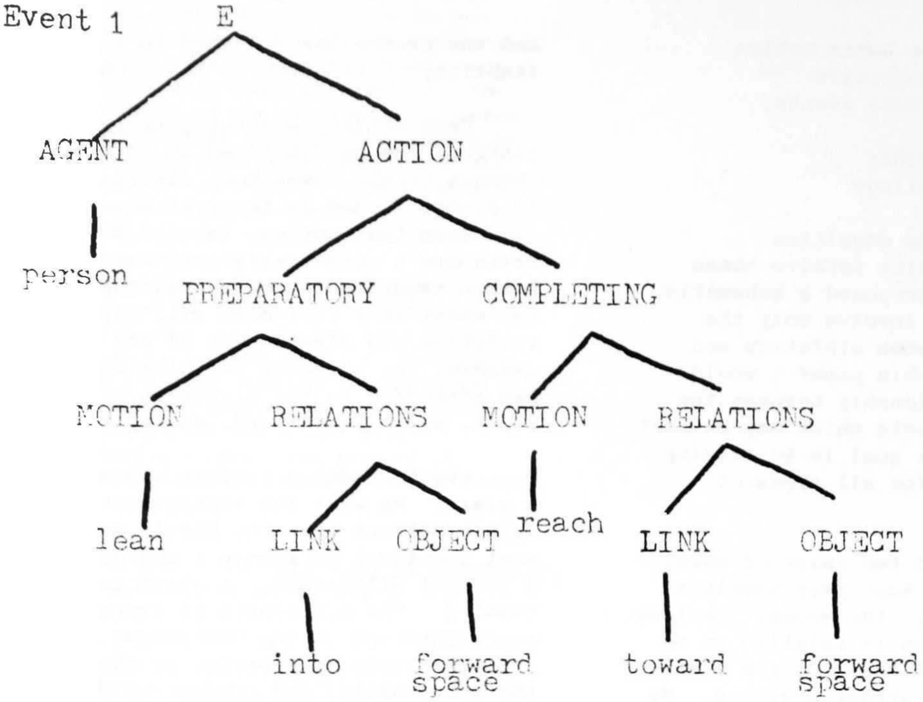
Event 2 adds a visible object to the scene, a glass. We want the representations to accurately reflect that the bodily motions are almost identical in Events 1 and 2, yet there is a crucial difference: a glass is picked up in Event 2. The difference is expressed in the representation of the COMPLETING MOTION, which is cognitively interpreted as the intention of the event unit. The completing motion is a motion which contains as one of its internal positions that of the fingers closing in upon and pressing against the glass. I have arbitrarily used the words reaches-picks up to represent this motion. The fact that in language these are two separate verbs is irrelevant to the representation of the visual event unit.

The MOTION RELATION which most influences the form of this completing motion is the OBJECT, the glass. The verbal nature of the representation is again not relevant. Holding might seem to be a separate motion because it is a verb rather than a preposition. From the viewpoint of the visual event, it is equivalent to a preposition: it is a link between the bodily motion and some other spatially represented variable which influences the form of the body's motion. The holding is not a separate motion. The body returns to a stable, balanced position, with the arm near the central axis of the torso, at the ends of both paired events. In Event 2, however, the body returns with a glass.

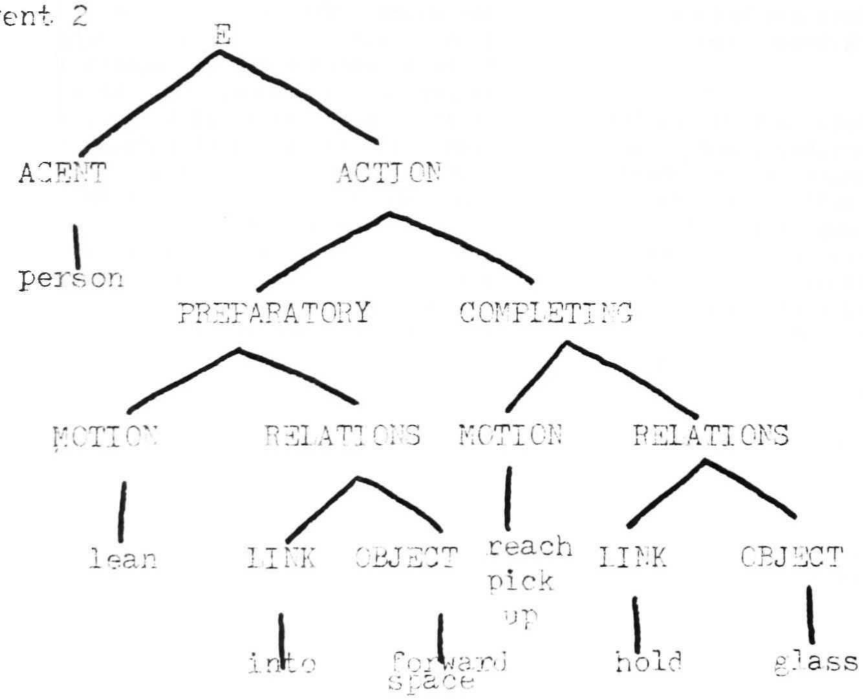
Events 1 and 2 were originally paired because they both belong to a general category of human motion events, reaching motions. Laban has spoken of such abstract categories of motion (Laban, 1966). We reach to pick something up, to hang something up, to put something down. The motion in bowling and the pitch in baseball are full reaching motions. When the reaching motion is extended to its fullest without the presence of an object, and one leg is extended

Figure 1

a. Event 1



b. Event 2



backwards in order to balance the reaching arm, we have an arabesque in ballet.

An interesting thing about human motion events is that any event involving voluntary human motion can be described in terms of AGENT-ACTION, without an OBJECT: any ACTION which can be performed in relation to an OBJECT can also be performed without the OBJECT. But in terms of a description of the visual structure, there is always a spatial direction involved. The person reaches upward or downward, toward the sky or toward the ground. In training for athletics or dance, aspects of the environment that are taken for granted in ordinary motion are often brought to conscious attention: the athlete is trained to notice the relationship between the body and the ground, and the body and the space around it. If the spatial aspects of the event are considered in relation to the motion itself, these spatial aspects act in the same ways as an object would act: as variables relevant to the intention of the motion.

Because of this similarity between visual directions and ordinary objects in a scene, I have represented them equivalently under the MOTION RELATIONS node. The variable is represented as the end-point of the motion under the OBJECT node. The LINK node is required to represent exactly how the OBJECT influences the MOTION. In Event 1, for example, the preparatory motion of leaning actually goes into, or bodily fills, the forward space; the completing motion of reaching goes toward, but not into, the forward space. Of course the forward space of the preparatory motion can be a different physical area from the forward space of the completing motion: space is forward or backward only in relation to the body.

The addition of an object in the second pair of events takes us into the interesting problem of representing causal relationships between event units.

Event 3: A dancer steps and leaps.

Event 4: A basketball player steps and takes a jump shot.

Figure 2 shows the visual relationships between the motion of the person, and the motions of the person plus the ball, in Events 3 and 4. Schematic representations of these events are given in Figure 3.

Until point 5 (Fig. 2) the visual structure of Events 3 and 4 overlap. At point 6, the person begins to descend to the ground (Events 3 and 4) and the ball continues to rise (point 7) until it falls at point 8 (Event 4 only). In both Events 3 and 4 the person must descend to the ground. But the ball has it own, independend motion in Event 4. We have in Event 4 two separate motion descriptions to deal with: the person in motion, whose structural description is almost identical to that of Event 3, and the object in motion, which spatially and temporally goes through its own motion.

We want our representations to mirror the similarity of the person's motions in Events 3 and 4. We also want the representation of Event 4 to correspond to the intuition that the basketball player's jump shot, and the basketball's descent through the hoop, are parts of the same event. We do not want to be in the position of saying that Event 4 is represented by two loosely joined schematic structures, one representing the players motion and the other the motions of the ball. These considerations lead to schematic structures which are very similar for Events 3 and 4, and in which the motions of the ball in Event 4 are part of the higher event of the person's motions.

Nevertheless, there is a real difference between the ball in Event 4 and the object in Event 2 (the glass which is picked up): the ball has a motion which is spatially and temporally non-co-extensive with the motion of the agent. In a basketball game people tend to watch the motion of the ball, and for many people, the event ends when the ball goes into or misses the basket, not when the player has descended again to the ground. Attaching the ball as a MOTION RELATION to the ACTION of taking a jump shot, as the glass is a MOTION RELATION to the ACTION of reaching/picking up, captures only part of the experience.

In order to represent these relationships I have described Event 4 as an event unit containing an embedded event unit (Figure 3b). As long as an object is carried, pushed, or pulled by a person, so that the object's motion retains a one-to-one relationship with those of the person, the object's motion is not represented as a separate unit (ie: the glass in Event 2). When the object's motion ceases to have a one-to-one correspondence with those of the person, the object's motion becomes represented as a separate event. However, the object's motion in this case is a consequence of the person's; the ball goes into the basket because the player steps, jumps, and releases the ball in a certain way. Players and coaches can often predict at the moment the ball leaves the player's hands whether the ball will hit the basket. This de-

Figure 2: Events 3 and 4

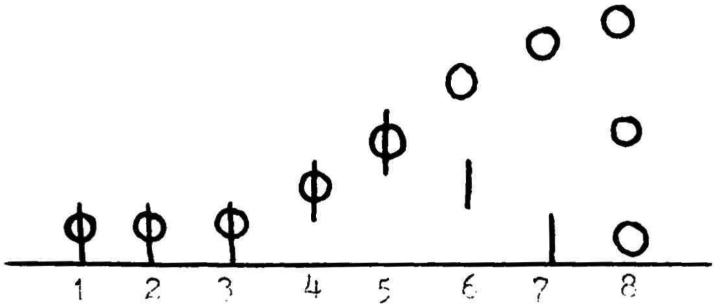


Figure 3
a. Event 3

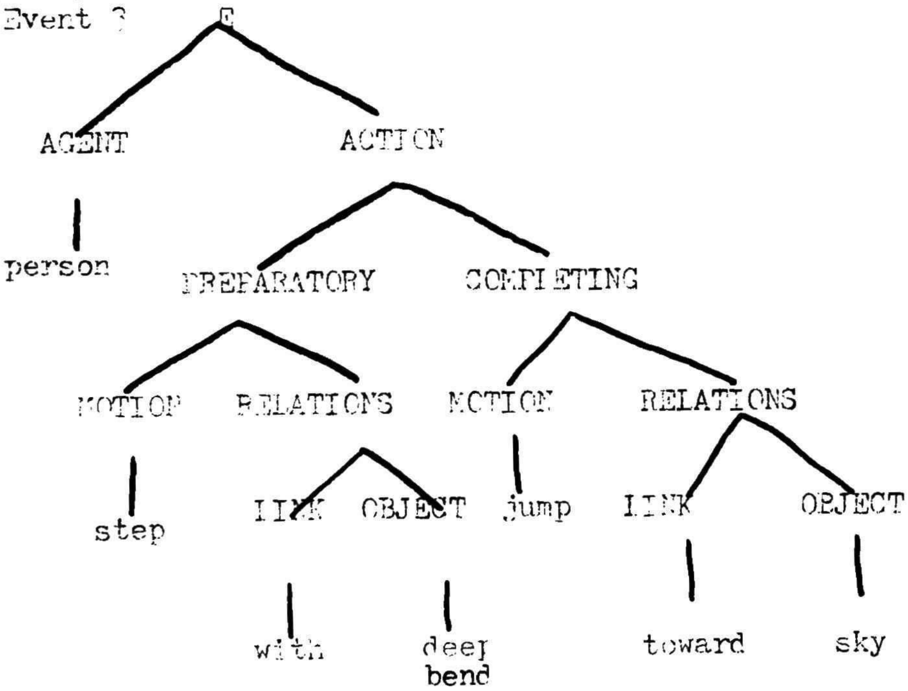
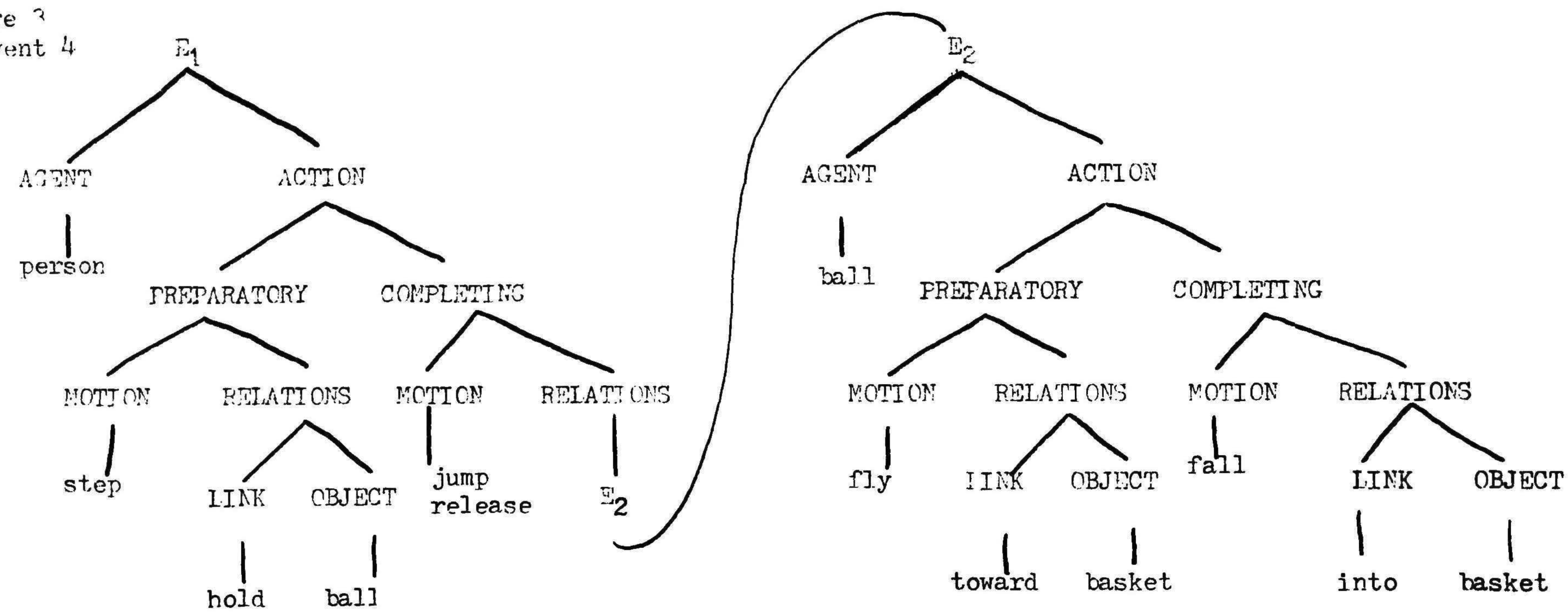


Figure 3
b. Event 4



pendence of the ball's motion upon the player's motion is captured by the embedded structure.

Events 3 and 4 fall into another abstract category of human motion events, run-leap. This category is especially interesting because the body defies gravity in a visually dramatic way at the center of the completing motion. When we are watching just a person in motion we normally watch the person descend to the ground at the end of the leap. The event is fully completed only when the person has returned to a position of stability on the ground.

But in the case of an embedded event in which the object has an independent motion of its own, the eye tends to follow the motion of the object (the ball, for example) once it leaves the person's hands. The follow-through aspect of the completing motion is not actually perceived. The dunk shot in basketball, however, is an interesting exception. In the dunk shot, the player's and the ball's descent occur together temporally and are very close together spatially. It is an interesting possibility that the excitement of the dunk-shot for spectators is partially due to the synchrony of endings of the embedded and embedding event units (Phillippe Poisson, personal communication).

Our excitement may arise from our ability, in this particular event, to actually see both the player and the ball return to the ground. This excitement would really be an aesthetic kind of excitement: the experience of a perfect fit between the schematic cognitive structure and the particular manifestation of that structure being observed.

References

Laban, Rudolf. The Language of Movement: Choreutics. London: Macdonald & Evans, 1966.

Lasher, Margot D. "The cognitive representation of an event involving human motion." Cognitive Psychology, 1981, 13, in press.

* I would like to thank John M. Carroll, Joseph M. O'Brien, and Philippe L. Poisson for helpful discussions of the ideas in this paper.