PARSING MOTION FOR MEANING:
INFANTS' INDIVIDUATION OF ACTIONS FROM CONTINUOUS MOTION

by

Tanya Lee Sharon

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1999

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Tanya Lee Sharon

entitled

Parsing Motion for Meaning:
Infants' Individuation of Actions
from Continuous Motion

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Karen Wynn	5-27-99 Date
Paul Bloom	5-27-99 Date
Lynn Nadel	5/27/99 Date
Mary Peterson	5-27-99 Date
Felice Bedford	5/27/99 Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director Karen Wynn	5-27-99 Date

STATEMENT BY AUTHOR

SIGNED: ______________

ACKNOWLEDGMENTS

First thanks must go to the friends who made my education: Wen-Chi Chiang, Lori Markson, and Gil Diesendruck. Thanks to Gil for his helpful conversations and for his contributions to our discussion group. Thanks to Lori for her friendship and support, for excellent career advice, and most especially for her camaraderie during our first year. Finally, special thanks to Wen-Chi Chiang for being the best buddy and fellow lab-mate anyone could ever hope to have, and a fine and considerate friend with whom anything can be shared. You made my experience memorable, wonderful, and worthwhile.

I would also like to thank:

Karen Wynn for sharing her knowledge, providing first-rate training, and forbearing through difficult times.

Mary Peterson for her help and suggestions throughout the course of this work.

Paul Bloom for sharing thought-provoking insights and opinions in many conversations.

Betty Glisky for helpful conversations at a number of critical moments.

The Department of Psychology at the University of Arizona for providing a stimulating, truly exciting environment for learning psychology and cognitive science.

The Children's Research Lab at the University of Texas for providing an office while I finished this dissertation.

Thanks to Margaret, for everything.

TABLE OF CONTENTS

ABSTRACT

Almost nothing is known regarding infants' abilities for parsing the on-going activity in their surroundings into distinct and meaningful parts. However, the individuation of actions is a fundamental ability, as explicated in a four-part introduction. Based on a review of general principles of individuation across multiple ontological domains, three possible mechanisms for action individuation in infants are identified and tested. The results of a series of studies show some important limitations in infants' abilities to parse actions from continuous motion. Although infants can perceptually discriminate different types of actions (such as jumps and falls) performed by a puppet, and can individuate and enumerate sequences of such actions when the acts are separated by brief motionless pauses, their ability to individuate actions embedded within a continuous stream of motion is limited: Neither repeating cycles in the action sequences nor marked differences in extent of motion are sufficient cues. The results instead suggest that tangent discontinuities in the path of motion are an important cue to infants' ability to parse actions from on-going motion. Implications for infants' conceptual structure for actions, and additional potential mechanisms of action individuation, are also discussed.

Chapter 1: Introduction

The human infant begins life nearly helpless physically, with only limited vision and a severely restricted behavioral repertoire. It is perhaps unsurprising then that the newborn has at times been viewed as intellectually at a loss as well. The British empiricists developed this view the furthest: on their account, the world of the infant was one great, undifferentiated "blooming, buzzing confusion" from which even the most basic parts had yet to be resolved. It was only through the accumulation of sufficient experience over the first few years of life that the infant gradually developed the capacity to perceive discrete, meaningful units in continuous experience. However, a wealth of research in the past fifteen years has amply demonstrated that infants are nowhere as limited as the empiricists thought. For example, newborns can already use prosodic cues to detect word boundaries (Christophe, Dupoux, Bertoncini, & Mehler, 1994). In addition, there is considerable evidence that very young infants parse their surroundings into distinct objects, and expect these objects to endure over time and behave in predictable ways (Spelke, 1988; Spelke, 1991). The human mind seems designed to carve distinct parts from continuous experience.

There is however a significant gap in our knowledge: namely, infants' individuation of actions. Infants no more live in a static world than adults: they are constantly surrounded by other people interacting with them, each other, and the environment. Yet in contrast to the established traditions investigating infants individuation of objects and sounds, inquiry into infants' ability to parse the on-going activity of the world around them into distinct and meaningful parts is almost non-existent. But careful consideration reveals that this ability is just as essential as the perception of objects and phonemes, for four different reasons.

First, action individuation is required to grasp the fundamental phenomenon of causality. The most elementary cases of simple mechanical causality (e.g., a basic launching event where one object hits a second, stationary object, causing it to move) will not be perceived as causal if the two actions are not parsed separately. Individuating objects is insufficient to give causality; for example, it cannot support the perception of causality in a sequence where the causal and resulting actions fail to map in a one-to-one fashion to objects (i.e., when one billiard ball impacts two other balls, causing both to move). This is not to say that individuating actions is sufficient for perceiving mechanical causality, but it is certainly necessary. Without such individuation, distinct but contiguous actions, as with a launching event, will necessarily be perceived as one undifferentiated portion of motion.

Second, individuating actions is a necessary precursor to producing them efficiently and productively. A child who wants to break a window in the house of the neighborhood bully wants to make one throw (and a quick getaway) not "some ball-throwingness". And a child who wants to throw a ball but has not yet realized that pulling back one's arm is a necessary component of the throw, and jumping up and down in excitement afterwards is not, will likely have an ineffective throw. Individuating distinct actions from the behavior stream provides the mental representations that guide motor production—having individuated a 'throw' properly, a child has a representation of what kind of thing a 'throw' is and can invoke it whenever that particular type of action with that particular effect is needed or desired. Further, once one has individuated actions, one can both perform them separately and recombine them in new and productive ways. Thus a child who has decomposed the act of tying a shoelace into the distinct actions of tying a knot and making a bow can use those same actions on something else (wrapping a gift with a bow), or use them individually as needed (for example, tying a knot in a strap so it doesn't slip). Unlike species whose young are born already hard-wired with most of the behaviors they will ever need, human infants are born with a minimal number of preprogrammed behavior sequences. Thus, a significant part of childhood learning necessarily involves discovering both new ways to act in the world, and new ways to combine familiar actions to achieve the results one wants. The origins of this ability begin with the ability to parse streams of action into distinct parts. Of course, by adulthood many acts are so thoroughly familiar they are overlearned, and there is much less call for creating novel action representations. However, adults may still face this need sometimes, as when learning an entirely novel sport. One of the basic tasks for students of tai chi, for example, is simply to learn to recognize the different basic moves contained in a continuous, twenty-minute routine. For young children the demands for creating novel actions representations will be much more constant.

A third major motivation for parsing behavior streams into discrete actions lies in the need to interpret and explain the behaviors of others—in short, to understand social causality. Common sense folk psychology is founded on the attribution of desires and beliefs (Fodor, 1981; Searle, 1983); we talk constantly in terms such as "she opened the window because she wanted fresh air", and "he swept the counter to hide the evidence". Yet beliefs and desires are mental phenomena; an observer must infer these on the basis of manifest behavior. Parsing behavior provides distinct behavioral units on which useful and reasonably reliable inferences can be made. Thus, in the above example, segmenting the action "opening" from the other behaviors of the agent supports the attribution of a 'desire' mental state to that agent (specifically, "wanting to cool down"). Similarly, parsing "sweeping" from the behavior stream of the second agent could contribute to a complex chain of inferences and attributions regarding motivations, beliefs and desires. Imagine in more detail a scene where a boy comes into a kitchen, walks up to a counter and takes three cookies off a plate. He eats them, looks down, notices he's left crumbs everywhere, and sweeps them off the counter. He then shifts the cookies on the plate so there's no noticeable gap and quickly leaves the kitchen. From this sequence of different acts, an observer would easily infer that the boy wanted the cookies, wasn't supposed to eat them, and believed he could hide evidence of his consumption. All of this would be impossible however if the different acts were not parsed separately; if the boy's actions from entrance to exit were all parsed as one continuous stream, it would be impossible to point to specific parts and the inferences they support (for example, it is the act of shifting the cookies on the plate that provides the giveaway clue that the cookies were not supposed to be eaten.) These examples thus point again to the importance of action individuation for the comprehension of causal structures, in the social realm just as much as the physical.

A fourth and final reason why we must individuate actions is in order to communicate effectively with others: For language to be comprehensible, two people must have the same bounded portion of motion in mind when they refer to a 'jump', a 'hug', or a 'hit'. A speaker who uses 'kick' to refer to any and all portions of a soccer match, or 'pour' to denote all aspects of cooking a meal, will not be understood. Of course, this requirement logically leaves open the possibility that it is language itself which provides the criteria for individuation, as argued by theorists such as Quine (1981). However, experimental work makes this proposal, at least it its strongest form, untenable. Two groups of researchers explicitly tested Quine's thesis by investigating children's generalizations of novel labels to novel objects and substances (Imai & Gentner, 1994; Soja, Carey, & Spelke, 1991). Both groups of researchers found that children did not depend on linguistic structure to begin to divide their world into either these different fundamental kinds nor into individual objects: the children already possessed these concepts[1]. There is no reason to think that the case would be different with actions. In fact, evidence will be presented below that prelinguistic children are capable of individuating actions under at least some conditions. The purpose of the studies reported in this dissertation is to explore what cues might support this early ability for individuating actions.

Chapter 2: General Principles of Individuation

This section provides an overview of mechanisms for individuating entities in a number of domains, with an eye towards uncovering common, possibly general principles of individuation. Infants' abilities for segmenting objects and sounds will be reviewed first; then, since infants' action individuation is virtually unexplored, we will summarize the research with adults in this area. This survey suggests three main points. First, humans possess multiple mechanisms for individuating entities in any given domain. Second, discontinuities[2] in a relevant dimension (whether brightness, amplitude, speed, etc.) appear to be a primary cue for individuating entities in any domain. Third and finally, people detect not just discontinuities but consistent patterns: humans possess a strong sensitivity to regularities in their environment, a sensitivity which contributes to the ability to find meaningful structure in their surroundings. These points will be developed below.

2.1 Objects

Infants' ability to individuate objects from their surroundings has been a topic of investigation since Piaget. He held that the perception of objects is dependent on the ability to first conceive of external entities which persist over time and space. Since, in his view, very young infants are unable to differentiate between themselves and their world, they experience an object simply as a passing pattern of sensation, an unrelated series of pictures. It is only with extensive experience over the first two years of life that infants learn to segment their surroundings into external, enduring entities (Piaget, 1954).

More recent research has pointed to a different picture, revealing that even young infants possess at least some ability to parse their visual surroundings into objects. Parents have long noted that infants will begin reaching for objects as soon as they develop sufficient motor control, around 4.5 months. Experimental tests undertaken with 6-month-olds proved that infants accurately guide their reaches to grasp the edges of interesting objects of various sizes, thus showing the ability to segment objects from their surroundings (Clifton, Rochat, Litovsky, & Perris, 1991). What cues might support this ability?

Evidence from a number of labs strongly supports the importance of motion information. In fact, in research done by Kellman and colleagues (Kellman & Spelke, 1983; Kellman, Spelke, & Short, 1986) patterns of relative motion were the primary cue used by infants. In a typical experiment, young infants were habituated to a display of a rod partially occluded by a box. At test, infants looked longer at displays of a broken rod than at a continuous rod (indicating that they found the former novel or unexpected) only in the cases where, in habituation, the two rod ends moved in common behind the box. After habituation to stationary displays, or displays in which the visible portions of the rod lacked common and distinctive motion (e.g., the box moved in front of a stationary rod, or the box and rod all moved together), infants failed to discriminate between the test displays, suggesting that detecting patterns of motion is a key component of infants' early object individuation. The importance of motion information has also been demonstrated by Yonas and colleagues, who likewise propose that infants perceive object boundaries by detecting the relative motion patterns of surfaces and edges (Arterberry, Craton, & Yonas, 1993). Thus, work by multiple researchers converges to show the importance of continuities and discontinuities in motion information to infants' ability to individuate objects.

Early research by Kellman and colleagues also showed some limitations in infants' ability to individuate objects. Specifically, infants at 4 to 6 months fail to perceive objects in accordance with many of the Gestalt principles that influence adult perception. For example, infants perceived two adjacent, stationary objects as one unit even if the objects differed in color, texture, and shape, and contain misaligned edges (Spelke, Breilinger, Jacobsen, & Phillips, 1993). Conversely, infants who were habituated to a display in which a rod moved laterally above a box, and an irregular polygon of a different color and texture moved in concert below the box, perceived one unified object, despite the lack of similarity and good continuation—in other words, despite the discontinuities in color and form (Kellman & Spelke, 1983). At the same time, young infants did use spatial separation (including separation in depth) to individuate objects (Kestenbaum, Termine, & Spelke, 1987; von Hofsten & Spelke, 1985). This overall pattern of results—a reliance on common motion and spatial separation, but not other Gestalt cues such as similarity and good continuation, in the perception of object unity in infants under 7 months—has been replicated many times (e.g., Kellman, Spelke, & Short, 1986; Kestenbaum, Termine, &

Spelke, 1987; Termine, Hrynick, Kestenbaum, Gleitman, & Spelke, 1987). Spelke thus proposes that young infants perceive objects by analyzing three-dimensional surface arrangements and their motions, grouping together all surfaces that are adjacent and undergo no relative motion (Spelke, 1990). Spelke suggests that infants' early reliance on discontinuities in motion and depth over those in luminance and shape may be adaptive, in that the former have greater ecological validity than the latter as cues to object boundaries, a point made also by Kellman (Kellman, 1996a; von Hofsten & Spelke, 1985).

Kellman offers a slightly different formulation of the results (Kellman, 1996b; Kellman & Shipley, 1991). He posits two processes of object perception: a 'primitive process' reliant on common motion but insensitive to the relationship between edges, and a 'rich process' which depends on the relationship between oriented edges, specifically, whether or not they are relatable. The concept of relatability formalizes the Gestalt principle of good continuation in mathematical terms; informally, it specifies that edges are relatable when they can be connected by a smooth, monotonic curve whose endpoints match the two edge tangents (for details, see Kellman & Shipley, 1991). The rich, edge-sensitive process is posited to be unavailable until the second half of the first year of life. For example, it is around 7 months that infants begin to show sensitivity to the alignability of edges in static displays (Spelke, 1990) and to subjective contours induced by static and kinematic illusory contour displays (Bertenthal, Campos, & Haith, 1980).

Kellman's theory has been tested by Johnson and colleagues, leading to a more complete and complex picture of infant object perception. In one series of studies

(Johnson & Aslin, 1996), 4-month-olds were habituated to a video display of a rod moving laterally behind a box. In one condition the rod edges were relatable, in a second they were not. Infants appeared to rely on relatability as an indication of whether the rod edges were connected or not: when tested with displays of broken and complete rods, infants in the 'nonrelatable' condition tended to look longer at the connected rod display (although the difference was only marginally significant), suggesting that they were surprised that the nonrelatable edges were connected, despite their common motion in the habituation. Johnson's results thus support the importance of relatability, albeit possibly at an earlier age then posited by Kellman.

A separate series of studies explored the importance of depth cues to infants' perception of object unity (Johnson & Nanez, 1995). Four-month-olds were habituated to a 2-dimensional rod-and-box occlusion display (which included a textured background); results showed they still preferred the broken rod at test, revealing that 3-dimensional depth cues were not necessary to support their object perception. Further tests however established that without the textured background, infants failed to perceive the rod as a continuous entity. Johnson argues that the textured background appears critical to infants' correctly perceiving the depth relations in the 2-dimensional display; without the depth cue provided by the accretion and deletion of texture elements, infants were unable to resolve the relationship among the rod parts.

The recovery of proper depth relations between portions of the visual display also plays an important role in the theories of Spelke and Kellman. Under Spelke's theory, infant perception of objects is based on the analysis of three-dimensional surface arrangements, and thus obviously requires first recovering information about depth relations. Kellman's model assumes that depth information is available in both the 'primitive' and 'rich processes', in the latter specifically in a form similar to Marr's 2.5-D sketch (Kellman & Shipley, 1992). However, infants' use of differences in depth is itself an impressive accomplishment, given that depth information must be recovered from the 2-dimensional retinal array. Evidence suggests that young infants use multiple cues to recover depth information. Not surprisingly, given their general sensitivity to motion, kinematic cues such as optical expansion and texture flow are among the first to emerge (Banks & Salapatek, 1983). Around 4 months of age infants begin to acquire stereopsis (the perception of depth based on binocular disparity), and by 6 months all infants have it (Held, 1983). The final depth cues to emerge are static monocular cues. Evidence collected by Yonas and colleagues shows that a host of 2-dimensional, pictorial cues to depth become available at roughly the same time, around 7 months of age. Among these are relative size (Yonas, Granrud, & Petterson, 1985), familiar size (Granrud, Haake, & Yonas, 1985), linear perspective (Yonas, Cleaves, & Petterson, 1978), and interposition (Granrud & Yonas, 1984).

Sensitivity to regularities in the visual tableau is another possible mechanism by which to find structure in the environment. In the past ten years, computational approaches to vision research have showed how various non-accidental regularities (such as collinearity, parallelism and certain kinds of vertices such as T-junctions and forks) can be used in the process of individuating and recognizing objects (Hummel & Biederman, 1992; Lowe, 1985). These features get their name from the fact that they are unlikely to arise by chance, and tend to be robust across different alignments between a viewer and an object. For example, the perception of parallel lines tends to reliably indicate the presence of an object. Unfortunately, the use of non-accidental properties has not been specifically tested in infants. It is possible however to make some conjectures based on infants responses to static displays, such as those discussed above. Based on the work of Spelke and colleagues, it would appear that infants do not use regularities such as parallel lines (e.g., the parallel edges of the rod) and T-junctions (as when the rod passed behind the occluder) until the second half of the first year of life. More recent work by Needham suggests that that age-limit may need to be revised downward. She has argued that infants show sensitivity to configural features (such as color, shape, and patterns) as early as 4 months under simplified conditions (Needham, Baillargeon, & Kaufman, 1998). Although discussed in terms of configural features, many of the displays used in her research also offer parallel lines and T-junctures as cues for segregation, leaving open the possibility that infants at an early age may make use of these cues under at least some conditions. Nevertheless, the consensus remains that motion cues are the primary source of information in early infancy.

In sum, the evidence to date shows that infants as well as adults use multiple cues to individuate objects. Among these, discontinuities in patterns of relative motion and depth are fundamental.

2.2 Sounds

The single most important sound source around infants are the conversations of other people. Normal speech however consists of bursts of continuous sound; the gaps we perceive between words are the product of extensive experience, as can readily be appreciated any time one attempts to determine the units of a completely novel language. Learning language and comprehending speech requires first segmenting the speech stream into distinct and valid parts. This task is made even more complicated by the fact that speech sounds are both highly variable and non-linear. That is, the same phoneme can possess different acoustic attributes in different contexts, and different phonemes can overlap in time due to coarticulation. Clearly parsing speech is a highly complex task, yet all normal infants begin the process well before a year of age. Research suggests they employ a number of mechanisms.

First, as in other domains, spatio-temporal discontinuities form one basic cue for parsing speech sounds. The temporal resolution of 6-month-olds infants is nearly adultlike, capable of detecting gaps as small as 12 milliseconds, (Trehub, Schneider, & Henderson, 1995) and thus enabling them to detect stop-consonant closures and other pauses. Further, they are also sensitive to equally small changes in duration, changes which are an important part of segmental and prosodic cues (Morrongiello & Trehaub, 1987).

Second, from early infancy babies are sensitive to differences between phonemes (Werker & LaLonde, 1988). This precocious ability is based on a sensitivity to distinctive features of different phonemes (such as being voiced or voiceless, a glide or a fricative), although the exact acoustic invariants defining phonemes are still unknown[3]. In turn, sensitivity to phonemes, when combined with sensitivity to the distributional characteristics of their native language, allows infants to find word boundaries in a stream of speech based on phonotactic cues such as the presence of legal and illegal clusters at the beginning of words (Friederici & Wessels, 1993; Jusczyk, Frederici, Wessels, Svenkerud, & Jusczyk, 1993). An example for an English-speaking adult would be parsing the novel sequence "zybski" as "zyb-ski" because the phonotactics of English dictate that 'bs' cannot mark the start of a syllable.

A third very important cue to infants' parsing of sounds is the presence of certain prosodic markers such as changes in stress and accent. Infants show a remarkably early and strong sensitivity to prosody, perhaps because this type of acoustic information is transmitted particularly well to the developing fetus. In one study, French infants only four days old detected changes in stimuli based on syllables (the basic rhythmic unit of French) but were unable to detect analogous changes in Japanese stimuli, where the rhythmic unit is the mora (Mehler, Dupoux, Nazzi, & Dehaene-Lambertz, 1996). Similarly, French newborns used prosodic cues to discriminate between pairs of stimuli that either contained or did not contain a word boundary (Christophe, Dupoux, Bertoncini, & Mehler, 1994). Prosody is also used to indicate saliency, especially through changes in stress. A number of investigators have accumulated evidence suggesting that stress plays a critical role in infants' early speech parsing, either by providing an initial entry into word-level units (Gleitman, Gleitman, Landau, & Wanner, 1988) or by contributing to the metrical rhythm of speech (Cutler & Norris, 1988; Jusczyk, Cutler, & Redanz, 1993), or both (Echols, 1996). By 9 months of age infants have developed a strong strategy for syllable segmentation based on stress (Morgan, 1996; Morgan & Saffran, 1995). This is consistent with the general importance of discontinuities: Stress patterns can be thought of as informative and easily accessible discontinuities in amplitude.

Notably, 'motherese' or infant-directed speech possesses many of the attributes just discussed: specifically, it uses a slow tempo, high pitch, and most important, highly exaggerated intonation and stress. This pattern of speech turns out to be very helpful to infants in detecting different syllables (Karzon, 1982) and segmenting speech into different clauses (Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989). In other words, infant-directed speech aids infants by exaggerating the discontinuities in pitch contour.

A final cue, as mentioned above, is infants sensitivity to the distributional characteristics of sounds in their native language. This sensitivity is, remarkably, sufficient on its own to support parsing of a speech stream: 8-month-olds infants who were exposed to two minutes of a continuous string of 3 different repeating, 3-syllable nonsense words were able to segment the stream based purely on the statistical relationships between neighboring speech sounds (see also Goodsit, Morse, Ver Hoove, & Cowan, 1984; Saffran, Aslin, & Newport, 1996). Specifically, infants' response was based on the transitional probability between syllables: within the 2-minute sequence, the transitional probabilities within words was higher (1.0 in all cases) than that between words (0.3 in all cases). Thus, infants appear to possess a powerful ability for detecting statistical regularities in the speech stream.

In summary, evidence indicates that infants begin to successfully segment speech streams before they are a year old using a variety of cues (see also Morgan & Saffran, 1995; Myers, et al., 1996). In addition to their sensitivity to phonemic contrasts and the prosody of speech, infants show a sophisticated ability for utilizing regularities in the stimulus input. In their use of multiple cues, infants are similar to adults. Further, infants use these various cues for parsing units at various levels, from syllables to words to clauses (see Aslin, Jusczyk, & Pisoni, 1998 for an overview).

Research into infants' processing of non-speech sounds has not been as comprehensive as study of their speech processing. A number of researchers have made connections between the two, suggesting in particular that many of the same principles may govern infants' processing of other complex sounds such as music (Aslin, Jusczyk, & Pisoni, 1998; Trehub & Trainor, 1990). For example, Krumhansl and Jusczyk (1990) found that five-month-olds detected phase boundaries in music using some of the same cues of changes in pitch and duration used to detect phrase boundaries in speech.

Exploration of infants' general auditory processing has been shaped by the highly influential work of Albert Bregman (Bregman, 1990). Bregman has studied the principles underlying adult auditory processing for decades, and has developed a detailed, comprehensive theory. His focus is on discovering the dimensions the auditory system uses to separate and discriminate the mixed acoustic energy received by the ear. To this end, he studies auditory stream segregation, a phenomenon whereby a complex auditory signal (rapidly repeated sequences of sounds) is perceived as either one stream of sound or two depending on similarities and dissimilarities in a variety of factors[4]. Bregman identifies two basic classes of primitive processes for auditory parsing: temporal or sequential integration, and spectral integration. Both processes break down the incoming array of energy into separate temporal and frequency regions and then analyze and regroup the regions according to various heuristics. Sequential grouping is influenced by factors such as fundamental frequency, temporal proximity, intensity, and spatial information. For example, temporally continuous sound energy that is judged to arise from the same location is usually grouped together, but spatially distinct sounds are not. Discontinuities in intensity also contribute to the formation of units. In spectral integration, the challenge for the auditory system is to assign specific frequency components to distinct sounds despite a temporal overlap. Bregman argues that the auditory system solves this problem by looking for correlations between elements that are unlikely to have occurred by chance. Thus, similarities and dissimilarities in sound frequency are important, as are common changes in the frequencies of different partials or in their amplitudes—an auditory analogue of the Gestalt principle of common fate. (Note however that other researchers have failed to find an effect of common frequency modulation—see Carolyn, 1991; Summerfield and Culling, 1992.) In addition, differences in location (i.e., spatial separation) again play a role.

Studies of auditory stream segregation in infancy show that many of these processes of auditory parsing are operative early in infancy. Infants within the first half-year of life can parse auditory input into distinct streams based on frequency proximity, similarity of intensity, and spectral similarity (Demany, 1982; Fassbender, 1993). In additional work, newborns only 3 and 4 days old used complex properties of timbre and spatial location to segregate sounds. The ability to individuate sounds based on their location in space appears to be innate, although actual performance shows a U-shaped function between birth and 3-5 months as control shifts from subcortical to cortical structures (Muir, Clifton, & Clarkson, 1989). In summary, work investigating infants' processing of both speech and non-speech signals converge to show that infants from birth possess a highly developed capacity for processing auditory input into distinct entities.

2.3 Actions

As discussed in the introduction, the question of how infants might parse the action in their surroundings into distinct actions has been essentially ignored until quite recently. Thus, for a useful overview of this domain we must turn to adults, where the issue has not been quite so neglected. Darren Newtson in particular has conducted a number of studies exploring the processes by which adults make sense of behavior. By his account, individual actions are picked out by the selection of successive "points of definition" (or 'breakpoints') in the behavior stream *according to a criterion of relative change* in the stimulus (Newtson, 1977; Newtson, Rindner, Miller, & LaCross, 1978).

For example, when drinking a cup of coffee, setting the cup down both defines the occurrence of 'putting the cup down' and divides it from the next action, perhaps 'reaching for a donut'. Breakpoints thus are those points which contain the most distinctive, and hence instructive, information. In one study (Newtson & Engquist, 1976) subjects who viewed triads of slides consisting of three consecutive breakpoints were significantly more accurate in describing the depicted action, rated the sequence as more intelligible, and were more accurate in judging the correct slide ordering, than subjects who viewed slides of three non-breakpoints. Similarly, subjects who viewed short films later showed significantly greater recognition accuracy for slides of breakpoints than for non-breakpoints. Breakpoints thus capture meaningful changes in the agent's behavior.

It is important to note two important aspects of breakpoints: first, they are defined by relative change, and second, they can simultaneously support more than one level of segmentation. These facts give the mechanism for parsing actions an adaptive degree of flexibility. For example, additional studies by Newtson established that subjects viewing the same videotaped behavior sequence were able to segment it into units of varying sizes, but the subjects instructed to segment into the smallest possible natural units (rather than the largest) developed more elaborate theories about the actor and were more confident in their impressions (Newtson, 1973). This implies that adults can vary the level of parsing in response to situations where greater accuracy or confidence is needed. This prediction was supported by results from a second study, in which two groups of subjects viewed a videotape of problem-solving behavior (a student assembling a model of a molecule). For one group, an unexpected behavior (taking off one shoe and sock) was inserted into the model-making; subjects in this group subsequently employed more units of perception (made more divisions), suggesting that they increased the 'resolution' of their parsing in a search for more information that might explain the unexpected behavior (Newtson, 1973).

In summary, research has shown that adults are able to parse the behavior of others into meaningful units, and can adjust their individuation processes adaptively to different situations. Whether infants possess any comparable abilities is the question motivating the studies conducted here. There are a number of prima facie reasons for thinking they might.

First, infants are highly attentive to motion from birth (Fantz & Nevis, 1967) – a fact frequently utilized by researchers who want to ensure infant attention to their stimuli. In fact, despite the relative weakness of infant vision compared to hearing, looking is the major means for extracting information from the environment for young infants (von Hofsten, 1983).

Second, research shows that young infants are able to analyze motion for a variety of information. In addition to individuating objects on the basis of motion, as discussed above, infants extract information relevant to social causality. For example, Rochat and colleagues (Rochat, Morgan, & Carpenter, 1997) showed 6-month-olds two displays of moving colored balls. The displays were almost identical in all their movement parameters except that in one, the movement between 2 balls was contingent—one ball appeared to 'chase' the other, which would 'flee' when the first got too close. Infants discriminated between the displays, thus showing an early sensitivity to different types of movement information, information that to adults specifies social causality or its absence. Similarly, research shows that infants can analyze patterns of movement for goal-directedness (Gergely, Nadasdy, Csibra, & Biro, 1995; Woodward, 1998), and discriminate between point-light displays of biomechanical versus 'scrambled' motion (Bertenthal, Profitt, & Kramer, 1987). Although these findings do not inform us directly regarding infants' ability for individuating actions (since they require only the discrimination of types of actions, not the individuation of distinct acts), they do suggest that young infants possess the basic capacity for attentive analysis of motion which would be required to parse streams of motion into parts.

Evidence more germane to the question of parsing behavior sequences, although still indirect, comes from a line of research in children's language acquisition concerned with syntactic bootstrapping. Under the syntactic bootstrapping hypothesis (Gleitman, 1990), children are aided in inferring the meaning of a new word by the syntactic frames in which they hear it presented. For example, children could make different inferences about the meaning of a novel verb when hearing it in a transitive frame (which requires a direct object) compared to an intransitive frame (where no direct object is required). In tests of this hypothesis, children were shown videos of two characters engaged in multiple forms of activity (for example, a rabbit who repeatedly hops up only to be forced back down to a squat by one hand of a rabbit, while both characters continually whirl an arm in circles in the air). While viewing this video, children heard either a transitive sentence ("The rabbit is gorping the duck") or an intransitive sentence (The rabbit and the duck are gorping"). At test, children heard "Where's gorping now?" while they watched two new videos side by side. In one, the duck and rabbit both wheeled their arms, without squatting; in the other, the rabbit repeatedly forced the duck to squat, without arm-wheeling. Children's looking times to the two test videos was systematically influenced by the syntactic frame they had originally heard, suggesting that they had parsed the complex scene of motion in the original video into different components.

Finally, recent research provides unequivocal evidence of a basic ability to individuate actions in infancy. In two studies, Wynn (1996) explored infants' abilities to individuate and enumerate physical actions, and specifically the role of temporal discontinuity in their ability. As discussed above, research suggests that spatio-temporal discontinuities are a basic component of individuating entities across multiple ontological domains: surfaces separated in space are perceived as belonging to two separate entities, and portions of acoustic energy which are separated in time by silence, or originate from different locations in space, are judged as different acoustic events. These facts suggest that infants may use spatiotemporal discontinuities to segment scenes of motion into distinct units: Portions of motion separated by gaps in space and/or time may mark different actions. These considerations motivated the two studies by Wynn.

In the first study, one group of 6-month-old infants was habituated to a puppet jumping 2 times on a stage, pausing briefly between each jump. A second group of infants was habituated to the puppet jumping 3 times, again pausing between each jump. Thus, the actions in the sequences were always temporally bounded by portions of motionlessness. After habituation, both groups of infants were presented with 2- and 3-jump trials. Infants looked reliably longer at trials containing the new number of jumps, showing they had enumerated the jumps. Both the duration and the tempo of all sequences were controlled, thus precluding the possibility that infants could respond based on rhythm. As enumeration requires that the counted entities first be recognized as separate entities, their discrimination of the test sequences also demonstrates that infants individuated the actions. In a second study the motion/absence of motion boundaries were removed by having the puppet gently wag back and forth between jumps and briefly following the last jump in a sequence. Infants again showed a preference for the new number of actions, showing that a temporal discontinuity in motion *per se* was not required for infants to detect action boundaries. However, infants in the first experiment discriminated the new number from the old more strongly than did the infants in the second experiment, showing that while a temporal discontinuity in motion is a very useful cue for individuating actions, it is not necessary.

Chapter 3: Cues for Action Individuation in Infants

The studies reported here explore what other cues young infants use, in addition to temporal discontinuities, to parse action streams into discrete acts. This is especially important since much of the daily activity with which infants are confronted is continuous, moving without pause from one action to the next. How might infants parse such on-going activity into discrete parts? Based on Wynn's two studies and the literature reviewed in Chapter 2, three hypotheses are considered.

The simplest alternative is that infants merely pick out more salient portions of motion from less salient portions. Research has shown that infants do not attend to all aspects of their surroundings equally: some stimuli have a greater capacity to draw infants' interest. For example, as noted above, movement information is very salient to young children. Likewise, the cue of stress in spoken language is both highly informative and perceptually salient.

But note that salience is an inherently relational construct: a given portion of visual or acoustic energy might be perceived as more or less salient depending on other information present in the environment, and even given the very same display different aspects are often preferred at different ages. For example, 4 month olds find the eyes more salient than the mouth, whereas 5-month-olds find these equally salient (Caron, Caron, Caldwell, & Weiss, 1973). What is more, there are limits: infants attend selectively to louder noises or brighter lights, but not to the point of discomfort. Nevertheless, a few generalizations are possible. Research has generally found that infants prefer stimuli which are relatively bigger (Fagan & Fantz, 1975; Lawson & Ruff, 1984; Linn, Reznick, Kagan, & Hans, 1982), moving rather than stationary (Kaufmann & Kaufmann, 1980; McKenzie & Day, 1977), moving more quickly versus slowly (Nelson & Horowitz, 1983), and more 'complex' (Berlyne, 1958; Columbo, 1985; Columbo, O'Brien, Mitchell, & Horowitz, 1986)[5]. The frequent finding in infant research of a baseline preference for two items over one is another example: Infants seem to operate under a general rule that 'more is better'.

In Experiment 2 of Wynn (1996) (the 'jump-wagging' study), the wagging was smaller and intuitively much less striking than the jumps. The above findings thus suggest that the wagging may have been less noticeable and less interesting for infants, forming a 'ground' against which the more salient jumps stood out as 'figure'. More generally, under a salience hypothesis infants would parse a stream of actions into units using the relative salience of the motion, in effect using a 'discontinuity in salience' to identify action boundaries. The limitation with this account is the relative nature of salience; if initial studies support this hypothesis, additional studies to confirm the differences in perceived salience would be needed.

A second hypothesis is that infants require certain kinds of perceptual contrasts to detect action boundaries. Specifically, they may depend on tangent discontinuities in the path of motion to parse an action stream. A number of researchers (Kellman & Shipley, 1991; Shipley & Kellman, 1990) accord similar discontinuities a central role in object perception. The importance of such discontinuities arises from a basic fact of projective geometry: when one object intersects with or partly occludes another, there will almost always be an abrupt change in the orientations of the projected edges. Such changes are called tangent discontinuities because the tangent lines at points along the projected edges (lines which capture the orientation of a line at a particular point), will change orientation abruptly rather than continuously (for a formal definition and examples, see Figs. 1A-1C).

Insert Figures 1A-1C about here

Kellman and Shipley (Kellman & Shipley, 1991; Shipley & Kellman, 1990) incorporate tangent discontinuities as a critical step in their model of object perception. Once edges defined by discontinuities in luminance are located and assigned depth information, the visual system searches for tangent discontinuities and then uses these to interpolate boundaries. Together with the locally given edges, these are then used to derive an enclosed area which is perceived as a unit—an object or figure. This theory has been used to explain a host of perceptual phenomena, including the perception of occluded and illusory contours in 2 and 3 dimensions, and the perception of units across time (Kellman & Shipley, 1992). Considerable empirical work supports their model (Kellman & Cohen, 1984; Kellman & Shipley, 1992; Shipley & Kellman, 1992).

Applied to the individuation of actions, this account thus suggests that tangent discontinuities in the path of motion could be an important cue to action boundaries for infants. For example, infants in the 'jump-wagging' study may have individuated jumps and wagging based on the sharp contrast (vertical versus horizontal) in their paths of motion: since their paths were transverse in the vertical (xy) plane, the switch between them created a tangent discontinuity (see Fig. 2).

Insert Figure 2 about here

A final and somewhat different possibility is that infants may be sensitive to repeating patterns of motion in the perceptual input. Infants well under a year are capable of both detecting and utilizing regularities in their environment. Although some of the most impressive examples come from the domain of language (see Saffran, Aslin, & Newport, 1996, discussed above), infants show this ability in other areas as well. Haith and colleagues have demonstrated that young infants are adept at extracting information about the spatio-temporal pattern of a sequence of alternating pictures. For example, in one study 2- and 3-month-olds who were presented with pictures in a regular series (twice on the left then once on the right) quickly learned to anticipate the next picture location, as evidenced by both anticipatory visual fixations and decreased reaction times (Canfield & Haith, 1991). Five-month-olds detected a similar 4-element pattern (Canfield & Smith, 1996).

It is possible then that infants at this age may detect and use regularities in the motion around them in determining action boundaries. In the second study above, infants could specifically have used the repeating perceptual pattern of "up-down-side-to-side" in the motion to detect unit boundaries, segmenting the motion at the point it completed one cycle and began another.

Note that there are important differences in the units picked out under the different hypotheses. According to the repeating patterns hypothesis, a unit is defined by the structure of the pattern; thus in the study above, the sequences would be parsed into 'jump-wagging' units. In contrast, the tangent discontinuity hypothesis implies that infants would parse these sequences into separate 'jump' and 'wagging' segments, based on changes in the path of motion between motion along the vertical axis and motion along the horizontal axis. Finally, the salience hypothesis implies that infants would individuate only jumps and treat the wagging as a background. These differences have ramifications for the numerical task that the infants confronted: If infants parse using either repeating patterns or salient differences in extent, a sequence of 2 jumps each followed by wagging would be enumerated as "two" (as 2 'jump-wagging's or 2 'jump's, respectively). However, if infants strictly use tangent discontinuities in the path of motion, such a sequence will likely be enumerated as "four" (2 'jump's and 2 'wagging's), and infants therefore would need to make a discrimination between 4 and 6 rather than 2 and 3[6]. This casts some doubt on the tangent discontinuity hypothesis: While young infants' ability to discriminate up to 3 items is fairly well established (Antell & Keating, 1983; Starkey, Spelke, & Gelman, 1990; Strauss & Curtis, 1981), evidence to date regarding their ability to distinguish between slightly higher numbers is equivocal. A few studies have shown some ability to distinguish between 4 and 5 objects (van Loosbroek & Smitsman, 1990) or 4 and 6 (Tan & Bryant, 1996), but others have produced negative results (Starkey, Spelke, & Gelman, 1990). Infants only appear able to discriminate higher numbers of simultaneously presented objects when the discrepancy is on the order of two-to-one (e.g., between eight and sixteen dots—see Xu and Spelke, 1998). Infants' ability to discriminate sequentially presented items has not been tested at all above the number 4 (Canfield & Smith, 1996). This makes the tangent discontinuity hypothesis a less likely explanation for the results of Wynn (1996), and thus a somewhat less compelling candidate hypothesis, although obviously still possible.

A series of experiments was conducted to test among these three hypotheses, beginning with the cue of repeating cycles. First, a control study (Experiment 1) was run to determine whether infants could discriminate different kinds of actions. This set up a test of the repeating cycles hypothesis in Experiment 2: infants were presented with continuous, heterogeneous sequences of actions (e.g., jump-fall-jump) whose structure afforded no repeating pattern of motion. Experiments 3 and 4 each tested alternative explanations for infants' performance in Experiment 2.

EXPERIMENT 1

The first study tested infants' ability to perceptually discriminate two kinds of actions, the jumps and falls of a puppet. This experiment provided basic information on the extent of infants' perceptual capacities. If infants failed to discriminate between the two types of actions[7], it would suggest that infants this age are insensitive to the different directions of motion implicated in both the tangent discontinuity and repeating cycles hypotheses; by default, the salience hypothesis would thus gain support. Conversely, if infants did discriminate the actions, it would become possible to evaluate the role of repeating patterns of motion by creating sequences whose heterogeneous composition (e.g., jump-fall-jump) lack such a pattern.

Method

Subjects

Twenty-two full-term 6-month-old infants participated (mean age 6 months, 9 days, range 5 months 20 days to 6 months 16 days), with 6 girls and 5 boys in each familiarization condition. Four additional infants were dropped after failing to complete at least 4 test trials due to fussiness (3 infants) or experimenter error (1 infant).

Apparatus

All events took place on a yellow display stage (55 cm tall, 82 cm wide, and 40 cm deep), against which the black and orange Daffy Duck puppet contrasted strongly. A black curtain was lowered between trials to hide the stage, and additional curtains to the sides kept infants focused on display events. Infants sat in an infant seat 90 cm from the stage; as a result, the 30 cm high puppet subtended a visual angle of 18 degrees. Two video cameras recorded events for post-session review. Recording of infants' looking time was done live by an observer who could see the infant but not the display stage.

Design

Infants were shown 6 familiarization trials in which the puppet performed 2 actions, with a 0.5 second motionless pause in between the actions[8]. Half the infants saw "jumps" and half saw "falls", making 2 familiarization conditions. On each "jump", the puppet rose approximately 6 cm and returned to its point of origin. On each "fall", the puppet rotated approximately 75 degrees to the right in the fronto-parallel plane, from a vertical to an almost horizontal position, and then returned to its original orientation. Each action took 1 second to complete. Infants were then presented with 6 test trials of 2 actions apiece. Test trials alternated between the novel action kind and the familiarized, 'old' action kind. Order of presentation (old-new and new-old) was counterbalanced across subjects.

Procedure

Prior to testing, informed consent was obtained from subjects' parents, and each infant was shown the puppet and allowed to touch it. Classical music played softly throughout the experiment. Testing began with a brief introduction to the stage and its operations: a gloved hand appeared from above and patted the bottom and sides of the stage, and the curtain was raised and lowered.

For all trials, the curtain was raised to reveal a stationary puppet. Approximately 2 seconds later the experimenter began performing the specified actions (either jump-jump or fall-fall) to the beat of a metronome (audible only to the presenter, who wore earphones). Timing of infants' looking began 0.5 seconds after the completion of the action sequence and continued until the infant either looked away from the display for more than 2 s after a minimum look of 0.5 seconds, or reached a cumulative looking time of 30 seconds. The curtain was then lowered for approximately 2 seconds before being raised to begin the next trial. There was a break of approximately 20 seconds in between the familiarization and the test trials, during which the infant was turned away from the stage and allowed to interact with his or her parent. All timing in this and the other experiments was done by one of two highly experienced observers.

Results and Discussion

A 2 (Familiarization condition—Jumps or Falls) by 2 (Sex) by 2 (Trial Kind- novel or familiarized action) repeated-measures ANOVA on infants' mean looking times showed only a significant effect of Trial Kind; no other main effects or interactions were significant. Infants looked on average for 4.8 seconds on trials containing the familiar action and 6.6 seconds on trials containing the new action, $F(1,18)=4.9$, $p<.05$ (see Fig. 3). Thus, infants perceptually discriminated these two kinds of actions.

Insert Figure 3 about here

EXPERIMENT 2

The above result shows that infants at 6 months are able to discriminate between jumps and falls. Without directly supporting any of the three hypotheses, this result is consistent with the possibility that infants are sensitive to different forms of motion, as required by the tangent discontinuity and repeating patterns hypotheses. Experiment 2 is a direct test of whether the latter cue is necessary for infants to successfully parse action sequences into discrete units. In this experiment, infants were familiarized with heterogeneous sequences of actions (e.g, jump-fall-jump) which afforded no repeating cycle. Further, to provide as clear a test as possible, the possibility that infants could simply separate the jumps and falls based on the cue of temporal discontinuity between actions was removed by having the puppet move continually in between the actions and briefly following the last action.

Method

Subjects

Thirty full-term 6-month-olds participated (mean age 6 months 1 day, range 5 months 7 days to 6 months 16 days), with approximately equal numbers of boys and girls in each familiarization condition. An additional 17 infants were excluded after failing to complete at least 4 test trials due to fussiness (9), disinterest (5), sleepiness (1), and experimenter error (2).

Design

Subjects were randomly divided into 2 groups, one of which was familiarized with sequences of 2 heterogeneous actions (jump-fall and fall-jump) and the other with sequences of 3 heterogeneous actions (jump-fall-jump and fall-jump-fall). There were 6 familiarization trials in total, 3 of each composition, alternating across trials in an ABAABB order. There were also 6 test trials, alternating between new and old number of actions. Order of test trial presentation (new number first or old number first) was counterbalanced between subjects.

To achieve novelty of test sequences, test trials necessarily consisted of homogeneous sequences composed of all jumps or all falls (there exist only two possible heterogeneous 2-action sequences, both of which were already presented in the 2-action familiarization trials). To control for the possibility of infants responding in the test trials on the basis of the tempo or duration of the sequences, rather than number, half of each familiarization group received test trials in which the old- and new-number sequences were matched for overall duration, with the tempo of the new-number and old-number test sequences differing equally (one faster, one slower) from that of the familiarized sequences. For the other half of each familiarization group, test sequences were similarly matched for tempo and controlled for duration (see Table 1). Tempo was defined as the time interval between actions (all actions took 1 second to complete). As the design of the experiment required infants to respond on the basis of the number of actions seen, in this and all subsequent experiments videotape recordings of the experimental session were reviewed after each session by an observer who was blind to the infants' condition. Trials in which infants failed to observe at least half of each action (e.g., the infant glanced away just as the action commenced) were discarded.

Insert Table 1 about here

Procedure

The procedure was identical to Experiment 1 with the exception of the timing and the content of the sequences, as noted above. In addition, infants were now deprived of the key cue of motionless pauses between actions by having the puppet rotate approximately 20 degrees back and forth around its vertical axis between all actions and for 0.5 seconds following the last action in each trial. This made the puppet appear to be shaking its head and arms from side to side in a 'jiggling' motion[9].

Results and Discussion

A 2 (Familiarization condition-- 2 or 3 actions) by 2 (Control condition—test trials equated for duration or tempo) by 2 (Sex) by 2 (Trial Kind—2 or 3 actions) repeated-measures ANOVA conducted on infants' mean looking times revealed no significant effects or interactions. Infants who were familiarized to 2-action sequences looked at the 2-action test sequences for 5.2 seconds and the 3-action test sequences for 5.4 seconds;

infants familiarized to 3 actions looked 7.3 seconds at the 2-action test sequences and 6.8 seconds at the 3-action test sequences, $F(1,22) = .125$, $p > .7$. Analyzed in terms of looking times to new-number versus old-number test sequences, infants on average looked 6.0 second on new-number test trials and 6.3 seconds on old-number test trials (see Fig. 4). A non-parametric test confirmed the results: Overall, only 16 of the 30 infants looked longer at the novel-number trials, $p > .850$ by a Sign test. Infants thus failed to discriminate the old- from the new-number test sequences.

Insert Figure 4 about here

This result has a number of implications. First, it shows that large, presumably salient differences in extent are insufficient to enable infants to parse motion into discrete units: the jiggling was much smaller than either the jumps or falls, while the latter were of comparatively equal extent; thus if infants parse sequences based only on differences in salience, they should have succeeded in this experiment. Second, it reinforces the possibility that repeating patterns of motion are an important cue—recall that when presented, under very similar conditions, with a continuous, ***homogeneous*** sequence of actions (jumps interspersed with wagging in Wynn, 1996) infants successfully individuated and enumerated the jumps. However, before final conclusions are reached, two alternative explanations for infants' failure to discriminate the test trials in this experiment must be addressed.

First, it is possible that infants' difficulty in this experiment might have arisen not with *individuation* but with *enumeration*. In other words, their performance may not reflect problems in individuating distinct actions from a non-repeating pattern of on-going motion, but instead problems with enumerating heterogeneous sequences of actions. This could arise if 6-month-old infants fail to perceive the ontological similarity across different kinds of actions, and thus fail to conceive of a sequence of disparate actions as a set that can be counted. A somewhat analogous situation for an adult would be the unnaturalness of enumerating a set composed of 2 apples and a window—these comprise a grouping one would not usually make outside of unusual circumstances like poorly written word problems and laboratory experiments. This alternative is addressed in Experiment 3.

A second possible explanation arises from careful analysis of the stimuli used thus far in studies of infants' action individuation. In Wynn's 'jump-wagging' experiment the vertical jumps and horizontal wags offered a tangent discontinuity in the path of motion. However, this cue was not dependably present in the familiarization trials of Experiment 2 here, raising the possibility that infants may have failed not because of the lack of a structured pattern of motion but because of a lack of tangent discontinuities. Specifically, the jiggling motion in between the jumps and falls could have obscured the change in the overall path of motion, in effect occluding the discontinuity in path between the vertical jumps and sideways falls. In addition, if infants are like adults and perceive the jiggling as very similar to the wagging, then the jump/jiggle boundary would offer a tangent discontinuity but the fall/jiggle boundary would not (since both began with a sideways motion) (see Fig. 5). If this were the case, and infants were using tangent discontinuities as cues to action boundaries, it would be impossible for infants to respond appropriately in Experiment 2 because the number of units detected in a single familiarization condition would vary from trial to trial depending on the order of the sequence: A 'fall-jiggle-jump-jiggle' sequence, for example, would be parsed as 3 units (fall-jiggle/jump/jiggle), but a 'jump-jiggle-fall-jiggle' sequence would be parsed as 2 units (jump/jiggle-fall-jiggle). Although the larger numbers resulting from the use of this cue raise some questions regarding the likelihood of this hypothesis as an explanation for infants' success in the original 'jump-wagging' experiment (see the discussion above), it nonetheless needs to be addressed. Experiment 4 therefore tested whether tangent discontinuities in paths of motion are a necessary cue for infants.

Insert Figure 5 about here

EXPERIMENT 3

Infants were again presented with heterogeneous sequences of actions, except that now the key cue of temporal discontinuities between actions was restored (in all other aspects, the design and procedure of Experiment 3 was identical to Experiment 2). Given the importance of spatial and temporal discontinuities to individuation, this manipulation should be sufficient to allow infants to individuate the actions. In effect, the jumps and falls were now 'pre-individuated', and the task simply became a test of whether infants are able to enumerate a heterogeneous sequence.

Method

Subjects

Sixteen full-term 6-month-old infants participated (mean age 5 months 28 days, range 5 months 17 days to 6 months 17 days), evenly divided between boys and girls. An additional 3 infants were excluded after failing to complete at least 4 test trials because of fussiness.

Results and Discussion

A 2 (Familiarization condition-- 2 or 3 actions) by 2 (Control condition-- test trials equated for duration or tempo) by 2 (Sex) by 2 (Trial Kind—2 or 3 actions) repeated-measures ANOVA conducted on infants' mean looking times revealed no significant main effects but two significant interactions. First, there was a Trial Kind by Control interaction, $F(1,8)=9.147$, $p<.05$. Infants in the Tempo-equated condition looked longer at the 2-action test sequences (11.8 seconds versus 8.1 for the 3-action sequences), whereas infants in the Duration-equated condition looked longer at the 3-action test sequences (9.3 seconds versus 5.8 seconds for the 2-action sequences). Close analysis of Table 1 shows that this could reflect general preferences for sequences that are quicker and/or shorter.

More importantly, there was a significant Trial Kind by Familiarization interaction, $F(1,8)=12.94$, $p<.01$. Infants who were familiarized to 2-action sequences looked at the 2-action test sequences for 11.4 seconds and the 3-action test sequences for 7.0 seconds; infants familiarized to 3 actions looked 6.2 seconds at the 2-action test sequences and 10.4 seconds at the 3-action test sequences. Analyzed in terms of looking times to new-number versus old-number test sequences, infants on average gave 6.6 seconds of attention to new-number trials and 10.9 seconds to old-number trials (see Fig. 6). (For a discussion of this preference for the familiar number, see below.) A non-parametric test revealed the same pattern: overall, 13 of the 16 infants looked longer on average at the old-number test sequences, $p<.025$ by a Sign test. These results show that infants clearly discriminated the test sequences.

Insert Figure 6 about here

Interestingly, in contrast to the novelty preference often revealed by habituation studies, infants in this experiment who were familiarized to one number showed a strong preference for the *same* number in the test trials. There are two possible explanations for this.

First, infants in our study did not undergo a full habituation. Thus they may still have been actively processing the content of the familiarization trials, which could have led to a heightened interest in test trials with a similar content—in this case, numerical. Research specifically comparing the effects of a fixed-length familiarization versus an infant-controlled habituation confirms that, with the very same task and materials, a familiarization procedure can produce a familiarity preference in infants where a full habituation produces a novelty preference (Hunter, Ross, & Ames, 1982). However, an inspection of the data reveals that the three infants who met a habituation criterion in the six familiarization trials (i.e., whose total looking time on the last 3 trials was less than half of their looking time on the first 3 trials) *all* preferred the 'matching' number in the test trials, suggesting that an 'incomplete processing' account may not fully explain the infants' behavior.

A second explanation arises from research in discrepancy theory which shows that a moderate degree of novelty is sometimes preferred over either greater or lesser novelty, depending on which offers an optimal level of stimulation (Berlyne, 1968; Berlyne, 1970; Kagan, 1971; McCall & McGhee, 1977; see also Slater, Rose, & Morison, 1984). Given that the logical constraints of Experiments 2 and 3 required that the test trials be homogeneous (in contrast to the heterogeneous familiarization trials), an additional element of novelty was necessarily introduced. Infants in this experiment may thus have found the 'new number' test sequence (which also used a new sequence, tempo, and duration) *too* novel, and the 'old number' sequence more compelling, precisely because of the one element the latter shared with the familiarization sequences: the number of actions. This result is also consistent with other research showing that infants sometimes exhibit a preference for an abstract match between their initial exposure and test (Meltzoff & Borton, 1979; Spelke, 1981; Starkey, Spelke, & Gelman, 1990).

Experiment 3 thus shows that infants have no difficulty enumerating a heterogeneous sequence of actions when provided with the boundary cue of temporal discontinuity between actions. This reaffirms the fundamental importance of this cue. At the same time, it disproves one alternative explanation for infants' failure to discriminate the test sequences in Experiment 2: their difficulty was not due to an inability to enumerate a heterogeneous sequence of actions.

EXPERIMENT 4

This experiment addresses the second alternative explanation for infants' failure in Experiment 2, testing whether tangent discontinuities in motion are a necessary cue for infants. Infants were presented with continuous sequences of 'falls' followed by jiggling. This experiment was thus directly comparable to the 'jump-wagging' experiment of Wynn (1996) in that both the cues of repeating cycles and large differences in extent of motion were present; however, the cue of tangent discontinuities between the paths of motion of different actions was no longer available. If infants' failure in Experiment 2 is due to the lack of this cue, they should fail here as well. On the other hand, if infants rely on repeating cycles of motion, they should succeed here as they did in the 'jump-wagging' experiment. Thus, in combination with the previous studies, this experiment provides an exceptionally clear test between the alternative hypotheses.

Method

Subjects

Fifteen full-term 6-month-old infants participated (mean age 6 months 5 days, range 5 months 18 days to 6 months 21 days), almost evenly divided among boys and girls. An additional nineteen infants were excluded after failing to complete at least 4 test trials due to fussiness (15) , experimental error (2), exceeding the mean preference by more than 2.5 standard deviations (1 infant), and looking for the maximum looking time (30 seconds) on all 6 trials (1 infant).

Design

Infants were presented with falls separated by jiggling; the puppet also jiggled briefly (0.5 second) following the last fall in the sequence. In all other aspects the same overall design was used as in the previous studies. Half the infants were familiarized to 2 falls, and half to 3 (counterbalanced between sexes); all infants were then tested with 6 alternating trials of 2 or 3 falls embedded in jiggling (again, order and sex were counterbalanced). Tempo and duration were controlled for in the same manner as in Experiments 2 and 3.

Results and Discussion

Results were analyzed as before via a 2 (Familiarization condition- 2 or 3 falls) by 2 (Control condition—test trials equated for tempo or duration) by 2 (Sex) by 2 (Trial Kind- 2 or 3 falls) repeated-measures ANOVA on infants' mean looking times. There was a main effect of sex, $F(1, 7)=7.1$, $p<.05$, reflecting the fact that females looked longer overall than males (on average, 8.8 s and 5.3 s respectively). A four-way interaction of Trial Kind, Sex, Familiarization condition, and Control condition, which was difficult to interpret, just reached significance, $F(1,7)=5.8$, $p=.05$ (see Table 2.) No other main effects or interactions were significant. Infants who were familiarized to 2-action sequences looked at the 2-action test sequences for 7.4 seconds and the 3-action test sequences for 9.1 seconds; infants familiarized to 3 actions looked 5.4 seconds at the 2-action test sequences and 6.2 seconds at the 3-action test sequences, $F(1,7)= 1.2$, $p=.32$. Analyzed in terms of looking times to new-number versus old-number test sequences, infants on average gave 7.1 seconds of attention to new-number trials and 6.8 seconds to old-number trials (see Fig. 7). Non-parametrically, infants also completely failed to discriminate between the test sequences: 7 infants looked longer overall at the new sequence, and 8 at the old. Infants thus clearly and unequivocally failed to discriminate the test sequences, despite the presence of both large differences in extent of motion and a repeating cycle of motion.

Insert Table 2 and Figure 7 about here

This result, in combination with the results of the previous experiments, points to a number of conclusions. First, it reconfirms that infants are unlikely to be individuating actions based on large, presumably salient differences in extent: the falls in this experiment were just as relatively big, compared to the jiggling, as the jumps were to the wagging in the 'jump-wagging' experiment of Wynn, yet infants successfully parsed the latter but not the former from a backdrop of motion. Second, it shows that a repeating cycle of motion does not provide an adequate basis to infants for parsing a continuous sequence of motion. Infants failed to parse a sequence of falls and jiggles into units, despite the regular pattern. It remains possible, and indeed seems likely, that a repeating pattern is at least somewhat helpful to infants in providing the opportunity for extended observation of the same sequence of actions, but clearly it is not sufficient.

Third and most important, this pattern of results suggests that tangent discontinuities in paths of motion are an important cue to young infants' action individuation. Recall that infants successfully parsed a continuous sequence of actions in Wynn's 'jump-wagging' experiment, where the jumps and wags provided a tangent discontinuity in the paths of motion. The 'fall-jiggle' sequences of Experiment 4 were identical to the 'jump-wag' sequences in the tempo and duration of the sequences and in the presence of larger actions against a backdrop of smaller repeated movements; however, the falls unlike the jumps offered no tangent discontinuity with the intervening motion. This suggests that infants' success in the original 'jump-wagging' experiment can most parsimoniously be explained by the presence of tangent discontinuities rather than the presence of a repeating pattern. Conversely, their failure to parse a heterogeneous sequence of continuous motion in Experiment 2 here can be attributed to the absence of tangent discontinuities between the falls and jiggles, rather than the absence of a repeating cycle of motion.

It is important to note that infants' failure to individuate the falls from continuous motion is not due to having exceeded their perceptual capacities. From birth infants exhibit directionally-appropriate tracking, under the control of subcortical structures (Gorman, Cogan, & Gellis, 1957), and by 2 months show evidence of directionally-sensitive cortical neurons (Wattam-Bell, 1991). Further, although estimates of infants' velocity thresholds have varied depending on exact stimulus and testing characteristics, the velocity of both the jumps and falls lie well within infants' capacities on any estimate. Already by 3 months of age thresholds for slow and fast motion are estimated at around 1.4 to 5 degrees of visual angle per second for slow motion, to over 100 deg/s for rapid motion (Dannemiller & Freedland, 1989; Kaufmann, Stucki, & Kaufman-Hayoz, 1985). Thus, both the jump motion (at approximately 7.2 deg/s), and the fall (at approximately 36 deg/s) should be easily detectable by 6-month-olds. These rates are also not atypical of those used in other infant research. The issue therefore is not an inability to detect these differences but a disinclination to use them for the purposes of individuating actions. This is reminiscent of young infants' failure to use discriminable differences in color and shape for individuating objects. In the latter case, infants' conservatism arguably has some adaptive value. The usefulness of infants' conservatism in parsing action sequences (i.e., discounting directional changes that lack tangent discontinuities) is less apparent, but may still exist. One possibility is that it is adaptive during development, providing an initial entry into the task of action parsing. Infants may initially segment only the grossest units from an action stream, and then increase the 'resolution' of their action parsing with increasing experience. Or it may be that tangent discontinuities are simply one of the most reliable indicators of action boundaries. The true answer must await a much more fully developed understanding of the cues that mark action boundaries.

If infants require tangent discontinuities in paths of motion to individuate actions, they should perceive a sequence of falls embedded in jiggling as a bout of unstructured movement. That this was indeed the experience of infants in Experiment 4 is suggested by a second source of evidence, namely, the drop-out rate. Fully 48% of the infants originally tested in Experiment 4 failed to complete the required number of test trials due to fussiness; in comparison, the drop-out rates due to fussiness in the three previous experiments were 12%, 33%, and 16% respectively. Although indirect, the markedly higher failure to complete testing in Experiment 4 is consistent with the idea that infants became frustrated and/or bored when confronted with a display that lacked tangent discontinuities in the path of motion. Further, note that dropout rates were smallest in the two experiments (1 and 3) in which infants were provided the cue of temporal discontinuities between actions. This is again consistent with the argument that temporal discontinuities provide a fundamental and easily accessible basis for individuation. It is even conceivable that the roughly intermediate dropout rate in Experiment 2 (33%), where infants were familiarized to sequences of both jumps and falls embedded in jiggling, reflects the fact that infants were able to perceive some structure (and hence focus for interest) between the jumps and the jiggles, but became frustrated or bored when they were unable to resolve the bouts of falling and wagging into distinct actions. Although these last conjectures are of course speculative, the picture that emerges is consistent: Infants are able to individuate actions using temporal or tangent discontinuities, but lacking these cues, they may perceive only a portion of unstructured movement.

The result of Experiment 4 raises a question regarding the exact numerical distinction infants were making in the original 'jump-wagging' study of Wynn (1996). It could be that infants there applied only and strictly the cue of tangent discontinuities. In this case, the 2-jump sequences would have been parsed into 4 units (i.e., jump/wagging/jump/wagging) and the 3 -jump sequences similarly parsed into 6 units, and infants' differential looking in the test trials would indicate a successful discrimination between the numbers 4 and 6. However, given the equivocal evidence regarding infants ability to enumerate numbers above 3, such a conclusion would require further evidence. An alternative explanation is that infants used the cue of tangent discontinuities to segment the jumps from the bouts of wagging, but counted only the jumps—perhaps because the repeated nature of the wags may have made them less noteworthy, and thus made the jumps more salient. Although salience as indexed by large differences in extent of motion was insufficient on its own to support individuation of actions in the above experiments, it could still play a role in this fashion.

Chapter 4: Conclusions

The studies reported here augment the little that is known about infants' ability to individuate actions with a number of additional conclusions. Results showed that although 6-month-old infants can perceptually discriminate different kinds of actions (Experiment 1), and can individuate and enumerate a heterogeneous sequence of such actions when the action boundaries are specified by the key cue of temporal discontinuity (Experiment 3), they are unable to parse on-going behavior sequences into distinct units in a number of seemingly simple situations. Neither repeating patterns of motion (Experiment 4), nor large differences in the extent of motion (Experiments 2 and 4), are sufficient cues for infants at this age. However, tangent discontinuities in a path of motion may be a critical cue to infants' ability to parse actions from continuous motion (Experiment 4).

These results suggest a picture in which infants' ability to parse on-going activity around them is significantly limited. Although many actions are signaled by abrupt changes in direction of motion, many other actions are not. This raises the possibility that many scenes of everyday activity, such as a parent walking around a room tidying up, are perceived by infants at this age as inchoate portions of motion. It may be that with additional experience infants learn to use increasingly fine directional changes as a basis for segmenting sequences of action. This would be no different than the process adults undergo when encountering a novel form of activity (e.g., dressage, any number of competitive sports, etc.). Here also some accumulation of experience is often required before it is possible to notice and utilize increasingly detailed differences, and thus come to parse a continuous routine in the same manner as an expert.

Experience could also play a role in another way. Research shows that adults are much more likely to perceive that a given sequence of stimuli constitute a coherent entity *after experiencing the same sequence of stimuli across multiple contexts*. In one series of studies (Avrahami & Kareev, 1994) adults viewed short video clips of characters engaged in various activities; the films were compiled randomly such that, while individual acts (such as aiming a gun or throwing a dart) were present, the overall sequence of acts provided no plot or structure. Adults who viewed films in which the same (plot-less) sequence of 3 acts was repeated multiple times in varying contexts came to view that sequence as a coherent unit of its own: for example, they were more likely than control subjects to recall the target sequence in the proper order. Subjects who had an equal number of exposures to the same sequence but in an unvarying context did not show the same effects: Experience across multiple contexts was key.

The effect of repeated presentations across contexts was so strong it can even reverse baseline responses. In another study, control subjects watched another randomly compiled film and noted what point they perceived as the major point of division. The 3 acts surrounding this favored division point (agreed upon by 81% of the control subjects) were then embedded as a group at five points in another random film. Experimental subjects watched the resulting film first, *before* viewing the original film and noting where they thought it should be divided. The effect of viewing the second film before the original film was marked: only 22% of the experimental subjects segmented the test film at the same point which had been perceived by the control subjects as a point of major change. In short, how adults parsed the sequences was strongly influenced by the regularities they'd experienced in the input streams across different contexts. This effect is strongly reminiscent of that found by Saffran and colleagues (Saffran, Aslin, & Newport, 1996) in infant's language learning: just as infants' responses, in the absence of prosodic and other cues, were based on the transitional probabilities between sounds, adults' judgments in the absence of an overall plot or structure were based largely on the probabilities with which units occurred together.

Although the account of Avrahami and Hareev was developed to explain the individuation of events, as an essentially statistics-based mechanism—all that is required is a sequence of elements over which regularities can be calculated—it can be naturally extended to the individuation of specific acts (thus resurrecting the repeating cycles hypothesis in a somewhat different form). Applied to the infant studies reported here, their studies raise the possibility that the limiting factor was not a lack of sensitivity to motion regularities in the environment but a lack of sufficiently varied experience. That is, it is conceivable that infants in the above studies successfully parsed jumps but not falls from a continuous sequence of actions because they were more familiar with jump-like motions in a variety of settings (from seeing adults rise from a crouch beside their stroller, or an older sibling jumping on a bed) than with the somewhat unnatural, rigid keeling-over used in the falls; they may thus have only needed more experience with fall-like motions to parse the falls as well as the jumps from a continuous sequence.

A second route to more complete action individuation may be through learning language. Research reviewed above shows that infants and young children do not require language to attain fundamental ontological distinctions such as object, substance, or individual (Imai & Gentner, 1994; Soja, Carey, & Spelke, 1991). The studies reported here suggest that the same is true in the domain of actions: Infants were able to enumerate heterogeneous actions within a single count (provided the actions were pre-individuated for them by temporal discontinuities between actions), implying that they do possess a concept of ‘physical action’ which provides criteria for individuating and identifying such actions. This is not trivial; it is conceivable that 6-month-old infants would have failed to perceive the ontological similarity across different kinds of actions, and thus failed to conceive of a sequence of disparate actions as a set that can be counted. Infants’ ability to enumerate the heterogeneous sequences therefore suggests that they possess a concept of ‘physical action’ prior to acquiring language.

At the same time, however, infants’ inability to parse jumps and falls from an on-going sequence of actions suggests that they may lack concepts for these basic-level actions. For adults, individuation by categorization is common and effortless. Adults who see a cow wearing a hat immediately know that there are two entities, even though they are spatially contiguous, because they have a concept of “cow” and a concept of “hat”. Similarly, adults watching someone perform a leap followed seamlessly by a bow are aided in the perception of 2 actions by their concepts of "leap" and "bow". Infants appear to lack sortal concepts for these basic level actions[10].

It is here then that language could play a role, in helping infants acquire concepts for specific kinds of entities, such as cups or hugs. There is some tentative support for this in the domain of objects: a recent study found that the performance of 10-month-old infants on an object individuation task was predicted by the number of simple nouns they knew for objects used in the task (Xu & Carey, 1996).

In the domain of actions no directly analogous studies exist. There is research examining which components of actions (such as manner, instrument, or result) influence children's use and learning of verbs. This work could potentially illuminate children's initial, pre-linguistic biases for parsing actions as well as how language might influence the units children make. For example, it could clarify whether children assume that the scope of a single action is only the manner in which it occurs or also encompasses its outcome, and whether this assumption differs for linguistic versus non-linguistic use. Research thus far however has is inconsistent. In a study of children's generalizations of novel action verbs, Behrend (1990) found that changes in the result of the action were most important, changes in manner were of intermediate importance, and changes in instrument were relatively unimportant. Based on similar research, Forbes (1995) on the other hand argues that children by 3 years of age have developed a bias to attend to the manner of action when learning novel verbs. This later argument accords with previous research showing the importance of manner in children's representations of familiar action verbs (Bowerman, 1982; Gentner, 1978; Gropen, Pinker, Hollander, & Goldberg, 1991).

There are however some serious limitations in this research for the current purposes. First, these studies were not designed to address the question of how children individuate actions (test stimuli involved actions separated by clear breaks); thus any inferences on this question are necessarily tentative. Second, the evidence regarding whether the manner bias is in fact specific to verb-learning or a reflection of general linguistic or cognitive constraints is equivocal (Cartwright & Berhrend, 1995; Forbes & Farrar, 1993); thus it is impossible to evaluate whether it this is a case in which language influences concepts. Unfortunately then this literature can offer little guidance on the question of children's bases for individuation actions.

Similar limitations apply to two other lines of language research which could potentially inform the issue of infants individuation of actions: syntactic bootstrapping and the specificity hypothesis of Alison Gopnik and colleagues.

One study motivated by the syntactic bootstrapping hypothesis has already been described. In it (Naigles, 1990), 2-year-old children were found to attend differentially to two test videos (of arm-twirling or pushing/squatting), depending on the syntactic frame of a sentence they heard earlier while watching a video that combined both activities. This suggests that children had parsed the complex scene of motion in the original video into separate components. Unfortunately for our purposes, however, this study is uninformative regarding the relation between the linguistic input and children's individuation of actions. While it seems highly likely that the children were already capable of parsing the original video into two different kinds of action, and the syntactic frame simply directed their attention to one activity over the other, it remains theoretically possible that children failed to structure the depicted sequence of activity into distinct portions until hearing the sentence, which then directed their attention in a particular way. That is, if they heard the transitive frame "The rabbit is gorping the duck" their attention was drawn to look for an activity involving one entity acting on another, while the remainder of the scene remained unresolved. Thus the research on syntactic bootstrapping is indeterminate regarding linguistic influences on parsing action.

The third potential source of information is the work of Alison Gopnik and Melissa Bowerman (Choi & Bowerman, 1991; Gopnik & Choi, 1995; Gopnik & Meltzoff, 1986). Gopnik argues forcefully that differences in how languages emphasize and encode words can influence conceptual development. A study by Choi and Bowerman specifically examined how differences in the ways two languages lexicalize components of motion events influenced children's semantic organization (Choi & Bowerman, 1991). English and Korean differ significantly in how they treat different components of motion events: English characteristically conflates in the same verb motion information with cause and manner of motion, while expressing the path of motion separately through prepositions or particles such as 'up', 'down', or 'in'. Korean in contrast treats path differently across different situations. In intransitive clauses for spontaneous motion, path and manner are expressed through entirely separate verbs (for example, the main verb 'go' would be augmented with a verb specifying a path, such as 'ascending', 'along', etc.) In transitive clauses for caused motion, on the other hand, path is conflated in the same verb with both motion and ground (a reference point with respect to which the target entity moves). Thus one Korean verb (glossable as 'fit') is concerned only with creating a tight fit between items, regardless of whether the target item goes into, onto, over or together with the ground. In sum, while English isolates and thus emphasizes path information, regardless of type of motion or other details, Korean emphasizes the type of motion and treats path information less consistently.

Longitudinal records revealed that children showed sensitivity to language-specific patterns in the way they talked about motion from as early as 17-20 months. The English speakers quickly extended their earliest spatial words—path particles like 'up' and 'down'—to both spontaneous and caused changes of location, while Korean speakers kept spontaneous and caused motion strictly separated, and were relatively delayed in acquiring verbs that marked only path. This pattern suggests that English speakers more than Korean were prompted to classify events in terms of path information. However, as the authors themselves note, these findings are again inconclusive on the question of whether these differences in children's language use effected their *non*-linguistic processing of the events.

In sum, a substantial research tradition exists concerned with children's language development; many lines of inquiry in this area could potentially illuminate connections between children's developing command of language and their parsing of events in their surroundings. However to date these connections have not been explored. Clearly these are areas that call for more investigation.

Another possibility is that, rather than requiring more experience or the acquisition of language, infants simply require other cues not examined here, such as cues to the goals or intentions of the actor. An explosion of research in recent years has shown that infants from a remarkably early age are sensitive to stimuli from their environment that can be interpreted in intentional terms. For example, infants at 6 months are sensitive to the goal-directed motions of a hand: after habituation to a hand reaching for one of two toys, they looked longer when the hand reached for the second toy in the same location (a new goal) then when the hand reached along a new path (new spatio-temporal information) for the same toy (Woodward, 1998). At 9 months they interpret the actions of an abstract figure (a computer-animated circle) in terms of an intended goal, and by 12 months are able to infer an agent's goal without ever actually seeing it realized (Csibra & Gergely, 1996; Gergely, Nadasdy, Csibra, & Biro, 1995). These facts give prima facie plausibility to the possibility that infants may parse on-going scenes of action in terms of the goals and intentions of the actors involved[11], although to date this claim has not been directly investigated.

A related and intriguing proposal has recently been offered by Baldwin and Baird (Baldwin & Baird, 1997). They argue that the ability to extract intentionality from streams of actions is dissociable from the ability to analyze streams of actions into distinct acts, and in fact the latter is a prerequisite to the former. In other words, rather than extracting intentionality directly from motion information, infants under this account first analyze behavior streams into units which are 'commensurate' with intentional analysis. The basis for this analysis is a sensitivity to certain action features (such as changes in line-of-regard or contact with objects) which are assumed to covary reliably with transitions between intentions[12].

This proposal has the very appealing feature of providing a relatively rich and developed critique of the problem of action individuation. At this point however it remains largely a theoretical proposition. The authors admit that they lack evidence for the key claim that the units resulting from an analysis of action actually correspond to units defined by intentions; further, infants' sensitivity to covariation in movement patterns is merely inferred from sensitivity to the kinds of statistical regularities already discussed (e.g., in language learning).

One preliminary study from this perspective (Baird, Saylor, & Baldwin, 1999) has just been completed. In it, 10- to 11-month old infants were familiarized with two short (4 second) video clips of an adult engaging in everyday actions in which motion flowed continuously (i.e., a woman hanging a towel and a woman putting ice cream in the freezer). After each familiarization, infants were presented with two test sequences on alternating test trials. In one, the action was paused for 1.5 seconds just as the actor completed an intention (e.g., as she completed her grasp of the towel); in the other, the action was paused 'midstream', while the actor was still pursuing her goal or had just completed it (e.g., as she reached for the towel). Infants looked reliably longer at the intention-interrupting video then the intention-completing one. A control study established that infants had no baseline preference for either test video. Thus, the authors argue that 10-month-olds detected and responded to violations in the structure of intentional action.

This initial result opens up a promising line of inquiry. Clearly additional studies will be needed with a wider variety of stimuli to replicate and generalize the results, and confirmation that adults perceive the sequences as predicted (in terms of completed and uncompleted intentions) would be desirable. Tests should also be conducted to rule out the possibility that infants' responses to the test sequences cannot be explained by the concept of representational momentum (Freyd, 1987). Freyd argues that dynamic information is fundamental not just to perception but to mental representation, and is intrinsically incorporated into mental representations for dynamic events. Evidence comes from a large variety of studies in which subjects' mental representations are distorted in the direction of perceived movement. For example, adults who were shown static pictures of actors in the midst of completing an action which possessed a clear direction of motion (e.g., jumping off a wall) demonstrated a consistent distortion forward along the path of movement in their memories for the motion (Freyd, 1983). Representational momentum may be even greater in children than adults (Matzenbacher, Davis, & Hubbard, 1997), and there is some indication of it in infants as well (Hespos & Rochat, 1997). Thus, infants' longer looking times to the two intention-interrupting videos over the intention-completing one in the study by Baird and colleagues could conceivably have arisen because both of the former violated the momentum implicit in the actor's path of movement, whereas no violation in momentum occurred with the pause at the point of completing the action, since at that point the actor changed direction of movement. Nevertheless, this initial study opens up a highly promising avenue of inquiry.

The studies reported here have implications in one final area, namely the search for principles of individuation which are common across ontological domains. These studies make two contributions to this endeavor. First, they confirm the fundamental importance of breaks in spatio-temporal continuity: just as spatial gaps are important to object individuation, and temporal gaps play a role in individuation sounds, portions of motion separated by gaps in time tend to be perceived as distinct actions.

Second and more important, this research suggests a role for a factor not previously explored in action individuation, that of tangent discontinuities in the path of motion. Infants' failure to individuate actions from a continuous sequence when this cue was absent (Experiment 4) or unreliable (in Experiment 2), in contrast to their robust success when it was present (in the 'jump-wagging' experiment of Wynn, 1996), suggests that infants at 6 months may not only use this cue but may in fact require it, at least in situations where more definitive cues (such as temporal discontinuities) are lacking.

These results thus raise interesting questions for infants' object individuation. Tangent discontinuities are ubiquitous in a normal environment[13], and almost all of the displays used in this research have multiple tangent discontinuities (see Needham, 1997, for a helpful overview of stimuli). Since infants frequently fail to parse such displays into distinct units (as discussed in Chapter 2), it is apparent that tangent discontinuities are not a sufficient cue to support object individuation in infants. But the real question raised by the action results is not whether tangent discontinuities are sufficient, but whether they are *necessary* to infants' individuation of objects. The stimuli tested to date obviously were not designed to address this question, but research by Spelke and colleagues is at least suggestive. In these studies (Kestenbaum, Termine, and Spelke, 1987; von Hofsten and Spelke, 1985), young infants were presented with two blocks arranged in depth so that the smaller, closer one was centered in the larger, further one. At no point did the boundaries of the blocks cross, and so no tangent discontinuity was available to help define the boundaries between them. Results using both looking and reaching measures showed that unless the objects were separated in depth, infants perceived the two blocks as one unit, even when the blocks had different colors, textures, and substances. These results thus suggest that tangent discontinuities may play a key role in infants' parsing of multiple objects in a scene, as well as in the domain of action.

In all likelihood, infants bring multiple mechanisms to bear in converting continuous streams of behavior into distinct and meaningful parts. Perceptual, linguistic, and intentional cues could all contribute, perhaps with different impact at different levels. In addition, experience in specific social or cultural settings clearly plays a role in learning conventional ways of parsing some complex or ritualized events, such as a Japanese tea ceremony or a Greek Orthodox mass. Research has only just begun to explore how infants first gain entry into the process of individuating actions. Much more research is needed to discover the development and bases of this fundamental ability.

ENDNOTES

[1] Although Imai and Gentner did find evidence suggestive of a linguistic effect on children's responses to simple objects, their pattern of responses to substances versus complex objects was the same across languages. Soja et. al. found that in generalizing novel labels to objects, children respected both the shape and the number of the original referent, whereas in generalizing to substances they ignored both. This shows that children not only attended to the distinction between ontological types, but in the case of objects, individuated the entities as well.
[2] The term 'discontinuity' has a formal mathematical definition, discussed below. For present purposes, it is sufficient to understand discontinuities as abrupt changes in a perceived stimulus dimension.
[3] Infants like adults perceive phonemes categorically (Eimas, Miller, & Jusczyk, 1987). Interestingly, Aslin and colleagues have suggested that such categorical perception results from the auditory system's limited ability to resolve changes in the frequency transitions between consonants (Aslin, Jusczyk, & Pisoni, 1998). In other words, categorical perception may arise from limitations in the ability to detect discontinuities in frequency.
[4] As in the domain of objects, the mechanisms for auditory processing appears to have built-in adaptive features: Bregman points out that stream segregation takes at least 4 s to build up, and equally long to fade, a delay that could be useful in preventing uncontrolled switching between different ways of organizing an auditory scene.
[5] How 'complexity' is best defined has been a matter of some debate. For an overview, see Banks and Salpatek (1983).
[6] Alternatively, infants could arrive at the expected count of 2 or 3 by enumerating the actions in separate counts—for example, enumerating a sequence of 'jump-wag-jump-wag' as '2 jumps, 2 bouts of wagging' in comparison to the '3 jumps, 3 bouts of wagging' extracted from a sequence of 'jump-wag-jump-wag-jump-wag'. However, an unpublished study from our lab sheds doubt on this, as 6-month-old infants in that study were unable to maintain two separate counts for different kinds of familiar actions presented in a single sequence.
[7] In fact, previous research has shown that infants around 3 to 5 months can discriminate between different types of motion: they distinguish between rigid motions and deforming (non-rigid motions), translation and rotation, full rotation versus oscillation, as well as some combinations of these (Carroll & Gibson, 1986; Gibson, Owsley, & Johnston, 1978; Ruff, 1985). However, because these experiments generally gave infants extremely lengthy exposures (e.g., an object that translated 66 cm along a table edge, or completely rotated 11 times), it seemed prudent to confirm infants' abilities with the specific actions we wished to use.

[8]Pilot testing showed that familiarizing rather than fully habituating infants significantly increased their ability to last through the experiment, thereby markedly decreasing the drop-out rate.
[9] Note that this is a change in procedure from the wagging motion used in Wynn (1996). It was made necessary by the fact that, in order to offer a larger repertoire of actions (i.e., jumps and falls), the puppet was now manipulated by a stem from beneath, instead of a dowel through the beak as before. It should be noted that to adults the jiggling and wagging appear almost identical.
[10] It is also possible that infants possess these concepts but their representations are too weak to support individuation from a continuous background—just as it takes extensive practice for radiologists to learn to detect the boundaries of tumors in x-rays, infants could possess concepts of "jump" and "fall" but have difficulty detecting them in a noisy environment. This is however the less parsimonious account.
[11] This is not to claim that infants possess a complete representational theory of mind at the precocious age of 6 months; rather, their sensitivity to intentions is to be understood at the level of drives or goal-directedness.
[12] Studies with adults identified an additional possible action cue to intentions in the patterns of relative stability and instability in the motion: unstable motions (e.g., a leap) were perceived as the intention of a larger sequences of motions (e.g., a sequence of steps preceding the leap) (Lasher, 1981).
[13] Mathematically, tangent discontinuities are a generic phenomenon whenever the boundaries of objects meet; situations where boundaries meet with tangency are extremely rare. An example would be a ball resting on a table: Formally, the table is tangent to the ball at the point of contact, yet visually, tangent discontinuities are still perceived. The non-generic case is essentially unobservable.

APPENDICE A: TABLES

Table 1: Structure of Action Sequences in Experiments 2, 3, and 4

Familiarization Group	Test Sequences Matched for...	Trial Kind	Tempo*	Overall Duration of Action Sequence
2 Actions	Duration	Fam (2 actions)	2s	4.5s
		Test (2 actions)	3s	5.5s
		Test (3 actions)	1s	5.5s
2 Actions	Tempo	Fam (2 actions)	2s	4.5s
		Test (2 actions)	1s	3.5s
		Test (3 actions)	1s	5.5s
3 Actions	Duration	Fam (3 actions)	2s	7.5s
		Test (3 actions)	1s	5.5s
		Test (2 actions)	3s	5.5s
3 Actions	Tempo	Fam (3 actions)	.5s	4.5s
		Test (3 actions)	1s	5.5s
		Test (2 actions)	1s	3.5s

* Tempo= length of inter-action interval

To control for the possibility of infants responding in the test trials on the basis of the tempo or duration of the sequences, rather than number, half of each familiarization group received test trials that were matched for a new overall duration, with the tempo of the new-number and old-number test sequences differing equally (one faster, one slower) from that of the familiarized sequences. For the other half of each familiarization group, test sequences were similarly matched for tempo and controlled for duration

Table 2: Four-way interaction from Experiment 4

Habituation	Control	Sex	Test	Mean	Std. Dev.
3 Hab	Duration	M	2	4.5	4.3
3 Hab	Duration	M	3	2.9	1.2
3 Hab	Duration	F	2	7.4	1.2
3 Hab	Duration	F	3	11.7	4.0
3 Hab	Tempo	M	2	1.9	0.8
3 Hab	Tempo	M	3	2.8	2.6
3 Hab	Tempo	F	2	7.6	0.4
3 Hab	Tempo	F	3	7.5	0.4
2 Hab	Duration	M	2	6.2	5.1
2 Hab	Duration	M	3	7.3	1.2
2 Hab	Duration	F	2	7.0	0.6
2 Hab	Duration	F	3	8.5	2.2
2 Hab	Tempo	M	2	8.5	6.5
2 Hab	Tempo	M	3	8.3	3.6
2 Hab	Tempo	F	2	8.3	--
2 Hab	Tempo	F	3	15.7	--

The interaction was between trial kind (2 or 3 actions), sex, familiarization condition (2 or 3 actions) and control condition (test trials controlled for tempo or duration). The table shows mean looking time in seconds across test trials. N=2 in all cells except the last two, where N=1.

APPENDICE B: FIGURES

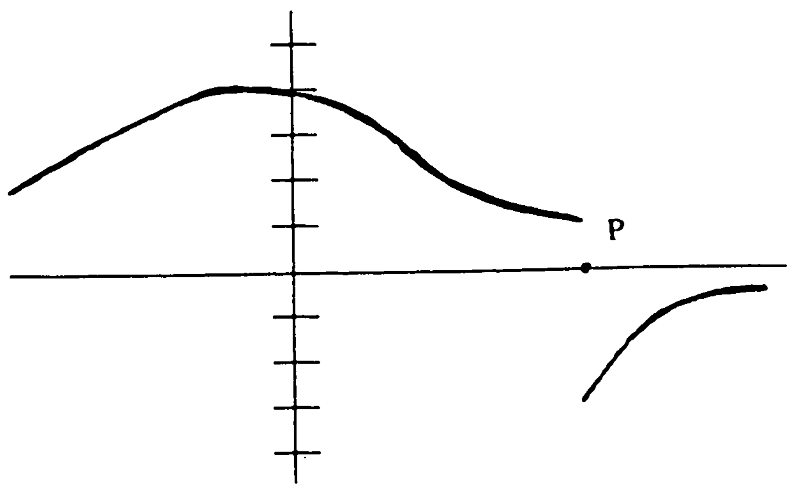

Figure 1A: Discontinuities in functions. A function F(x) has a discontinuity at x=P if F(x) approaches different values at P depending on the direction from which P is approached. For example, in Figure 1A above, F(x) approaches 1 as x approaches P from the left, but -3 as x approaches P from the right.

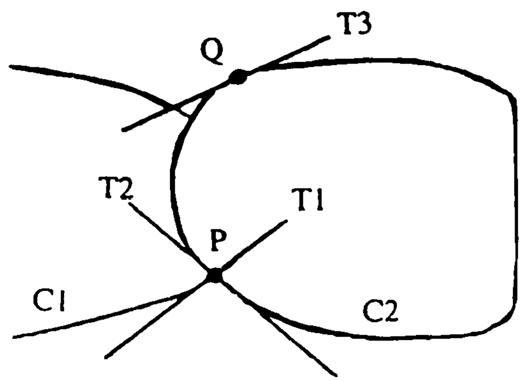

Figure 1B: Tangent discontinuities. A tangent discontinuity is a specific type of discontinuity, one in the tangent lines to a curve (or line). It arises when approaching a point from different directions yields two distinct tangent lines. For example, point P in Figure 1B yields tangent line T1 if approached along curve C1, and tangent line T2 if approached along curve C2. In contrast, point Q yields tangent T3 regardless of direction of approach, and it not a discontinuity.

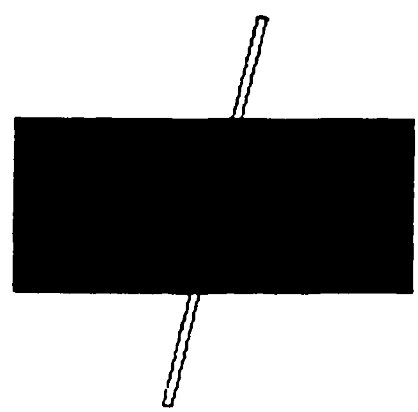

Figure 1C: Occlusion. A typical rod-and-box display has four tangent discontinuities at the points where the rod passes behind the occluder.

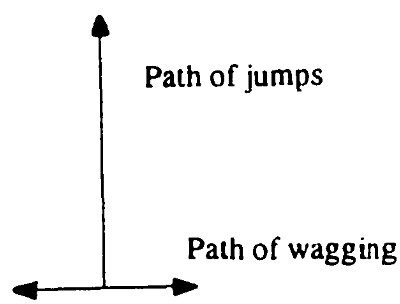

Figure 2: Discontinuities in path of motion. The switch between the vertical jumps and horizontal wagging in Experiment 2 of Wynn (1996) creates a tangent discontinuity in the overall path of motion.

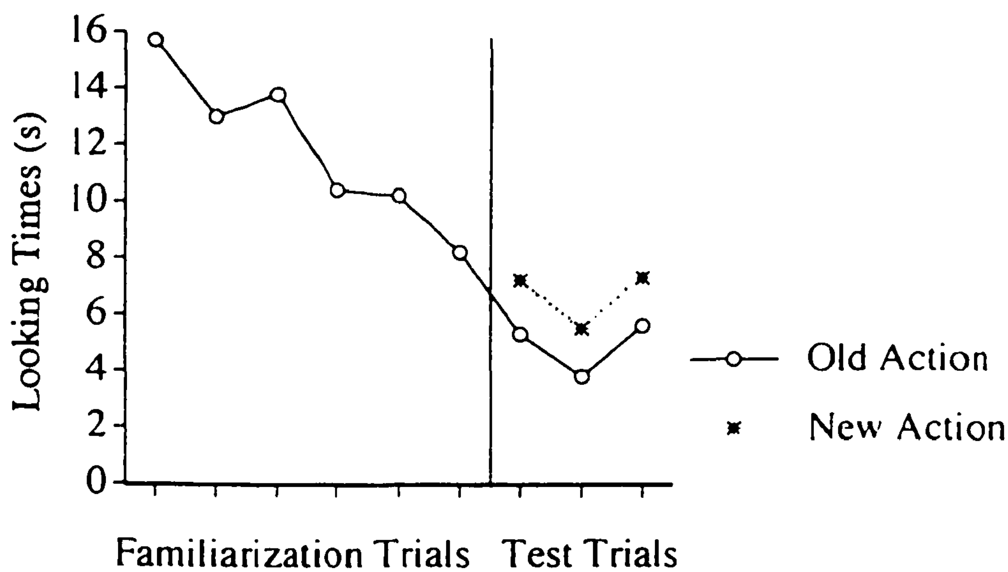

Figure 3: Infants' looking times to the familiar and new actions in Experiment 1.

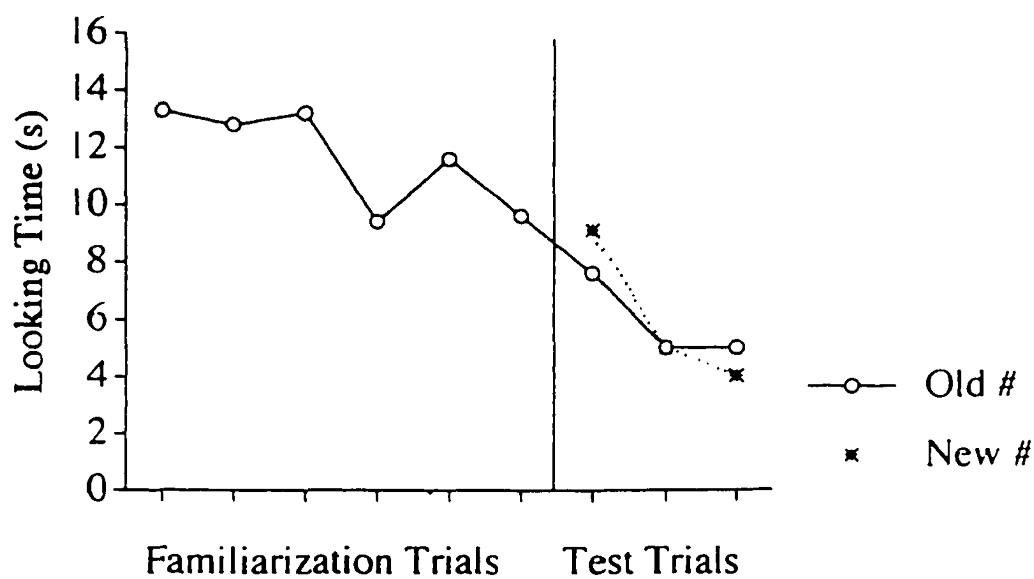

Figure 4: Infants' looking times to the familiar and new actions in Experiment 2.

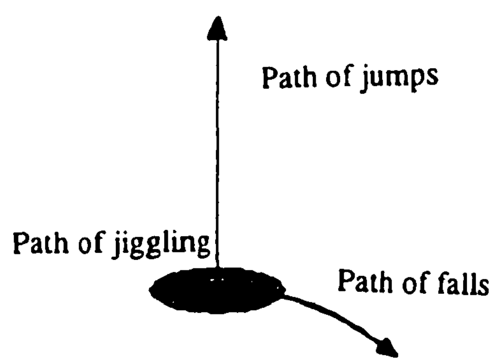

Figure 5: Path of motion in Experiment 2. The jiggling motion in between the jumps and falls in Experiment 2 may have obscured the tangent discontinuity in the overall path of motion. In addition, the falls and jiggles both began with a sideways motion and thus no tangent discontinuity was present to mark the boundary between them. This could cause a problem in parsing the sequences of Experiment 2 (see discussion in text).

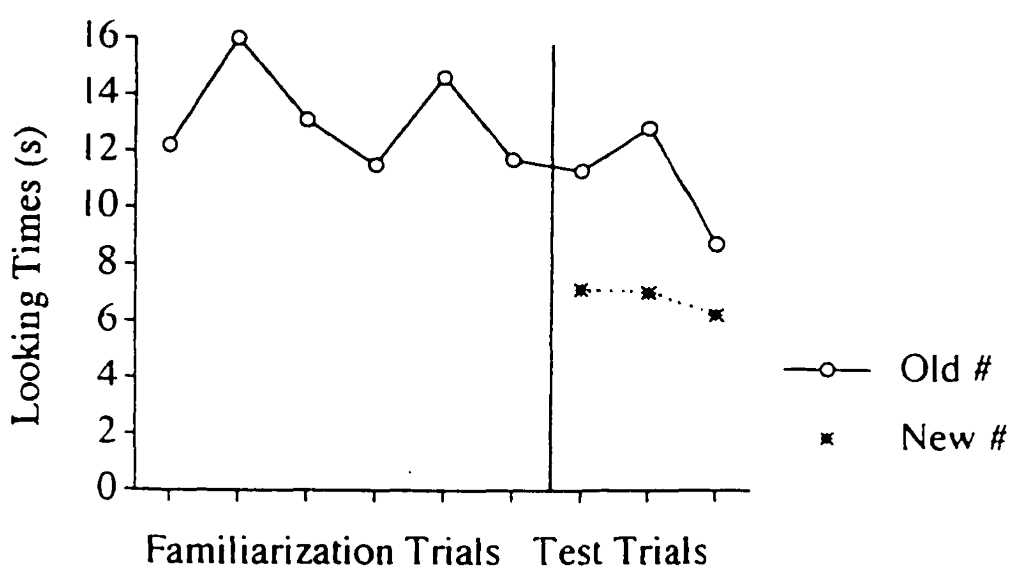

Figure 6: Infants' looking times to the familiar and new actions in Experiment 3.

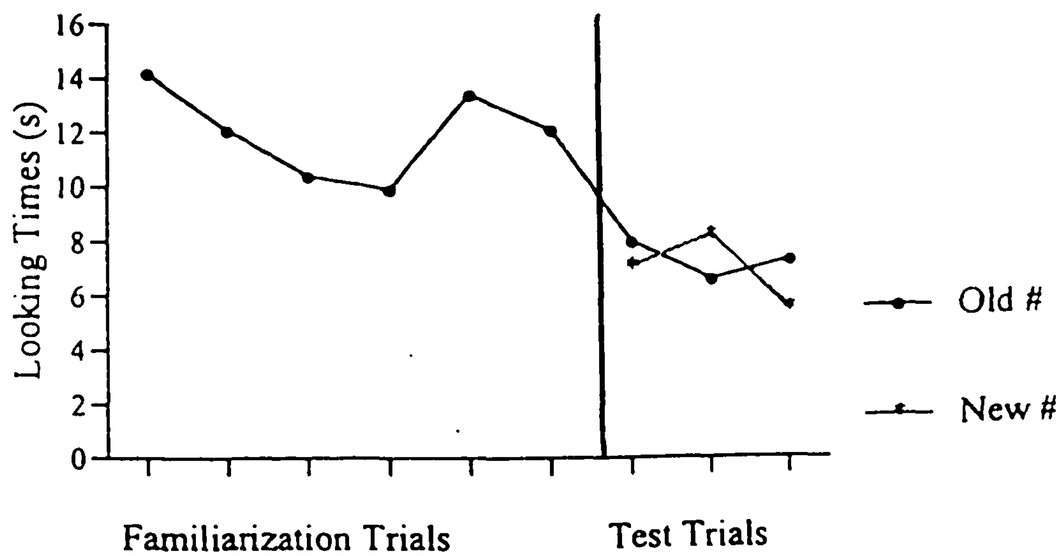

Figure 7: Infants' looking times to the familiar and new actions in Experiment 4.

REFERENCES

Antell, S., & Keating, D.P. (1983). Perception of numerical invariance in neonates. *Child Development, 54*, 695-701.

Arterberry, M.E., Craton, L.G., & Yonas, A. (1993). Infants' sensitivity to motion-carried information for depth and object properties. In C.E. Granrud (Ed.), *Visual perception and cognition in infancy, Vol. 23*. Hillsdale, NJ: Lawrence Earlbaum and Associates.

Aslin, R.N., Jusczyk, P.W., & Pisoni, D.B. (1998). Speech and auditory processing during infancy: constraints on and precursors to language. In D. Kuhn & R.S. Siegler (Eds.), *Cognition, Perception, and Language, Vol. 2* (pp. 147-198). New York: John Wiley and Sons.

Avrahami, J., & Kareev, Y. (1994). The emergence of events. *Cognition, 53*, 239-261.

Baird, J.A., Saylor, M.M., & Baldwin, D.A. (1999, April). Action parsing in infancy and the origins of intentional understanding. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Albuquerque, NM.

Baldwin, D., & Baird, J. (1997). Action analysis: a gateway to intentional inference. In P. Rochat (Ed.), *Early social cognition*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Banks, M.S., & Salapatek, P. (1983). Infant visual perception. In P.H. Mussen (Ed.), *Handbook of Child Psychology, Vol. 2: Infancy and Developmental Psychobiology* (pp. 436-550). New York: John Wiley and Sons.

Behrend, D. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development, 61*, 681-696.

Berlyne, D.E. (1958). The influence of the albedo and complexity of stimulus on visual fixation in the human infant. *Britsh Journal of Developmental Psychology, 56*, 315-318.

Berlyne, D.E. (1968). Determinants of subjective novelty. *Perception and Psychophysics, 3*(6), 415-423.

Berlyne, D.E. (1970). Novelty, complexity, and hedonic value. *Perception and psychophysics, 8*, 279-286.

Bertenthal, B.I., Campos, J.J., & Haith, M.M. (1980). Development of visual organization: The perception of subjective contours. *Child Development, 51*, 1072-1080.

Bertenthal, B.I., Profitt, D.R., & Kramer, S.J. (1987). Perception of biomechanical motions by infants: implementation of various processing constraints. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 577-585.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94.*

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner & L.R. Gleitmen (Eds.), *Language acquisition: the state of the art* (pp. 319-346). Cambridge: Cambridge University Press.

Bregman, A.S. (1990). *Auditory scene analysis: the perceptual analysis of sound.* Cambridge, MA: MIT Press/Bradford Books.

Canfield, R.L., & Haith, M.M. (1991). Active expectations in 2- and 3-month-old infants: complex event sequences. *Developmental Psychology, 32*, 269-279.

Canfield, R.L., & Smith, E.G. (1996). Number-based expectations and sequential enumeration by 5-month-old infants. *Developmental psychology, 32*, 269-279.

Carolyn, R.P. (1991). Discriminating between coherent and incoherent frequency modulation of complex tones. *Journal of the Acoustical Society of America, 89*, 329-340.

Caron, A.J., Caron, R.F., Caldwell, R.C., & Weiss, S.J. (1973). Infant perception of the structural properties of the face. *Developmental Psychology, 9*(3), 385-399.

Carroll, J.J., & Gibson, E.J. (1986). Infant perception of gestural contrasts: Prerequisites for the acquisition of a visually specified language. *Journal of Child Language, 13*, 31-49.

Cartwright, K.B., & Berhrend, D.A. (1995, April). Developmental changes in the nature of preschool children's verb-learning biases. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Indianapolis, IN.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*, 83-121.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical approach to the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America, 95*, 1570-1580.

Clifton, R.K., Rochat, P., Litovsky, R., & Perris, E. (1991). Object representation guides infants' reaching in the dark. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 323-329.

Columbo, J. (1985). Spectral complexity and infant attention. *Journal of Genetic Psychology, 146*, 519-526.

Columbo, J., O'Brien, M., Mitchell, D.W., & Horowitz, F.D. (1986). Stimulus salience and relational task performance. *Infant Behavior and Development, 9*, 377-380.

Csibra, G., & Gergely, G. (1996). Can one-year-olds infer the goal of an action without seeing it achieved?, *International Conference on Infant Studies*, Providence, RI.

Cutler, A., & Norris, D.G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Dannemiller, J.L., & Freedland, R.L. (1989). The detection of slow stimulus movement in 2- to 5-month-olds. *Journal of Experimental Child Psychology, 47*, 337-355.

Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development, 5*, 261-276.

Echols, C.H. (1996). A role for stress in early speech segmentation. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: bootstrapping from speech to grammar in early acquisition* (pp. 151-170). Mahwah, N.J.: Lawrence Earlbaum Associates.

Eimas, P.D., Miller, J.L., & Jusczyk, P. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical Perception* (pp.161-195). Cambridge: Cambridge University Press.

Fagan, R.L., & Fantz, J.F. (1975). Visual attention to size and number of pattern details by term and preterm infants during the first six months. *Child Development, 16*, 3-18.

Fantz, R.L., & Nevis, S. (1967). Pattern preferences and perceptual-cognitive development in infancy. *Merrill-Palmer Quarterly, 13*, 77-108.

Fassbender, C. (1993). *Auditory grouping and segregation processes in infancy.* Norderstedt: Kaste Verlag.

Fodor, J.A. (1981). *Representations*. Cambridge, MA: Bradford Books/MIT Press.

Forbes, J.N., & Farrar, M.J. (1993). Children's initial assumptions about the meaning of novel motion verbs: biased and conservative? *Cognitive Development, 8*, 273-290.

Forbes, J.N., & Farrar, M.J. (1995). Learning to represent word meaning: What initial training events reveal about children's developing action concepts. *Cognitive Development, 10*, 1-20.

Freyd, J.J. (1983). The mental representation of movement when viewing static stimuli. *Perception and Psychophysics, 33*, 575-581.

Freyd, J.J. (1987). Dynamic mental representations. *Psychological Review, 94*, 427-438.

Friederici, A.D., & Wessels, J.M.I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception and Psychophysics, 54*, 287-295.

Gentner, D. (1978). On relational meaning: the acquisition of verb meaning. *Child Development, 49*, 988-998.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165-193.

Gibson, E.J., Owsley, C.J., & Johnston, J. (1978). Perception of invariants by five-month-old infants: Differentiation of two types of motion. *Developmental Psychology, 14*, 407-415.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 174-223.

Gleitman, L.R., Gleitman, H., Landau, B., & Wanner, E. (1988). When learning begins: Initials representations for language learning. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey: Language: Psychological and Biological Processes, Vol. 3* (pp. 150-193). Cambridge, England: Cambridge University Press.

Goodsit, J.V., Morse, P.A., Ver Hoove, J.N., & Cowan, H. (1984). Infant speech perception in multisyllabic contexts. *Child Development, 55*, 903-910.

Gopnik, A., & Choi, S. (1995). Names, relational words, and cognitive development in English and Korean speakers: nouns are not always learned before verbs. In M.

Tomasello & W.E. Merriman (Eds.), *Beyond Names for Objects: Young Children's Acquisition of Verbs* (pp. 63-80). Hillsdale, NJ: Earlbaum.

Gopnik, A., & Meltzoff, A. (1986). Relations between semantic and cognitive development in the one-word stage: the specificity hypothesis. *Child Development, 57*, 1040-1053.

Gorman, J.J., Cogan, D.G., & Gellis, S.S. (1957). An apparatus for grading visual acuity on the basis of optokinetic nystagmus. *Pediatrics, 19*, 1088-1092.

Granrud, C.E., Haake, R.J., & Yonas, A. (1985). Infants' sensitivity to familiar size: The effect of memory on spatial perception. *Perception and Psychophysics, 37*, 459-466.

Granrud, C.E., & Yonas, A. (1984). Infants' perception of pictorially-specified interposition. *Journal of Experimental Child Psychology, 37*, 500-511.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acqusition of locative verbs. *Journal of Child Language, 18*, 115-151.

Held, R. (1983). Binocular vision-- Behavioral and neuronal development. In M.H. Johnson (Ed.), *Brain Development and Cognition: A Reader* (pp. 152-166). Cambridge, MA: Blackwell.

Hespos, S.J., & Rochat, P. (1997). Dynamic mental representation in infancy. *Cognition, 64*, 153-188.

Hummel, J.E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*, 480-517.

Hunter, M., Ross, H.S., & Ames, E.W. (1982). Preferences for familiar or novel toys: effects of familiarization time in 1-year-olds. *Developmental psychology, 18*, 519-529.

Imai, M., & Gentner, D. (1994). Linguistic relativity versus universal ontology: cross-linguistic studies of the object-substance distinction, *Proceedings of the Chicago Linguistic Society: What We Think, What We Mean, and How We Say It*, Vol. 2 (pp. 171-186).

Johnson, S.P., & Aslin, R.N. (1996). Perception of object unity in young infants: The roles of motion, depth, and orientation. *Cognitive Development, 11*, 161-180.

Johnson, S.P., & Nanez, J.E. (1995). Young infants' perception of object unity in two-dimensional displays. *Infant Behavior and Development, 18*, 133-143.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preferences for the predominant stress pattersn of English words. *Child Development, 64*, 675-687.

Jusczyk, P.W., Frederici, A.D., Wessels, J., Svenkerud, V.Y., & Jusczyk, A.M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*, 402-420.

Kagan, J. (1971). *Change and continuity in infancy*. New York: Wiley and Sons.

Karzon, R.G. (1982). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology, 39*, 326-34.

Kaufmann, F., Stucki, M., & Kaufman-Hayoz, R. (1985). Development of sensitivity for slow and rapid motions. *Infant Behavior and Development, 8*, 89-98.

Kaufmann, R., & Kaufmann, F. (1980). The face schema in 3- and 4-month-olds: The role of dynamic properties of the face. *Infant Behavior and Development, 3*, 331-339.

Kellman, P.J. (1996a). Ontogenesis of space and motion perception. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition* (pp. 327-364). New York: Academic Press.

Kellman, P.J. (1996b). The origins of object perception. In W. Epstein & S. Rogers (Eds.), *Handbook of Perception and Cognition* (pp. 3-48). New York: Academic Press.

Kellman, P.J., & Cohen, M.H. (1984). Kinematic subjective contours. *Perception and Psychophysics, 35*, 237-244.

Kellman, P.J., & Shipley, T.F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23*, 141-221.

Kellman, P.J., & Shipley, T.F. (1992). Perceiving objects across gaps in space and time. *Current Directions in Psychological Science, 1*, 193-199.

Kellman, P.J., & Spelke, E.S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology, 15*, 483-524.

Kellman, P.J., Spelke, E.S., & Short, K.R. (1986). Infant perception of object motion from translatory motion in depth and vertical translation. *Child Development, 57*, 72-86.

Kemler-Nelson, D.G., Hirsh-Pasek, K., Jusczyk, P., & Cassidy, K.W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*, 55-68.

Kestenbaum, R., Termine, N., & Spelke, E.S. (1987). Perception of objects and object boundaries by three-month-old infants. *British Journal of Developmental Psychology, 5*, 367-383.

Krumhansl, C.L., & Jusczyk, P.W. (1990). Infants' perception of phrase structure in music. *Psychological Science, 1*, 70-73.

Lasher, M.D. (1981). The cognitive representation of an event involving human motion. *Cognitive Psychology, 13*, 391-406.

Lawson, K.R., & Ruff, H.A. (1984). Infants' visual following: Effects of size and sound. *Developmental Psychology, 20*, 427-434.

Linn, S., Reznick, J.S., Kagan, J., & Hans, S. (1982). Salience of visual patterns in the human infant. *Developmental Psychology, 18*(5), 651-657.

Lowe, D.G. (1985). *Perceptual organization and visual recognition*. Boston: Kluwer Academic.

Matzenbacher, D.L., Davis, S.E., & Hubbard, T.L. (1997, April). Potential influences of surface form on representational momentum: Developmental considerations. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.

McCall, R.B., & McGhee, P.E. (1977). The discrepancy hypothesis of attention and affect in infants. In I.C. Uzgiris & F. Weisman (Eds.), *The structuring of experience*. NewYork: Plenum Press.

McKenzie, B.E., & Day, R.H. (1977). Infants' attention to stationary and moving objects at different distances. *Australian Journal of Psychology, 28*, 45-51.

Mehler, J., Dupoux, E., Nazzi, T., & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity: The infant's point of view. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 101-116). Hillsdale, NJ: Erlbaum.

Meltzoff, A.N., & Borton, R.W. (1979). Intermodal matching by human neonates. *Nature, 282*, 403-404.

Morgan, J.L. (1996). *From simple input to complex grammar.* Cambridge, MA: MIT Press.

Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development, 66*, 911-936.

Morrongiello, B.A., & Trehaub, S.E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology, 44*(3), 413-426.

Muir, D.W., Clifton, R.K., & Clarkson, M.G. (1989). The development of a human auditory localization response: A U-shaped function. *Canadian Journal of Psychology, 43*(2), 199-216.

Myers, K., Jusczyk, P.W., Kemler Nelson, D.G., Charles-Luce, J., Woodward, A., & Hirsh-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language, 23*, 1-30.

Naigles, L.G. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*, 357-374.

Needham, A., Baillargeon, R., & Kaufman, L. (1998). Object segregation in infancy. *Advances in Infancy Research, 11*, 2-44.

Nelson, C.A., & Horowitz, F.D. (1983). Visual motion perception in infancy: A review and synthesis. In P. Salapatek & L. Cohen (Eds.), *Handbook of Infant Perception: From Perception to Cognition, Vol. 2* (pp. 123-153). New York: Academic Press.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology, 28*(1), 28-38.

Newtson, D. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology, 35*(12), 847-862.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology, 12*, 436-450.

Newtson, D., Rindner, R., Miller, R., & LaCross, K. (1978). Effects of availability of feature changes on behavior segmentation. *Journal of Experimental Social Psychology, 14*, 379-388.

Piaget, J. (1954). *The construction of reality in the child.* NY: Basic Books.

Quine, W.V. (1981). *Theories and Things.* Cambridge, MA: Belknap Press, Harvard University Press.

Rochat, P., Morgan, R., & Carpenter, M. (1997). Young infants' sensitivity to movement information specifying social causality. *Cognitive Development, 12*(4), 441-465.

Ruff, H.A. (1985). Detection of information specifying the motion of objects by 3- and 5-month-old infants. *Developmental Psychology, 21*(2), 295-305.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-olds. *Science, 274,* 1926-1928.

Searle, J.R. (1983). *Intentionality: An essay in the philosophy of mind*: Cambridge University Press.

Shipley, T.F., & Kellman, P.J. (1990). The role of discontinuities in the perception of subjective figures. *Perception and Psychophysics, 48*(3), 259-270.

Shipley, T.F., & Kellman, P.J. (1992). Perception of partially occluded objects and illusory figures: Evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 106-120.

Slater, A., Rose, D., & Morison, V. (1984). New-born infants' perception of similarities and differences between two- and three-dimensional stimuli. *British Journal of Developmental Psychology, 2,* 287-294.

Soja, N.N., Carey, S., & Spelke, E.S. (1991). Ontological categories guide young children's inductions of work meanings: Object terms and substance terms. *Cognition, 38,* 179-211.

Spelke, E.S. (1981). Infants' acquisition of knowledge of bimodally specified objects. *Journal of Experimental Child Psychology, 31,* 279-299.

Spelke, E.S. (1988). Where perceiving ends and thinking begins: the apprehension of objects in infancy. In A. Yonas (Ed.), *Perceptual development in infancy, Vol. 20* (pp. 219-234). Hillsdale, NJ: Earlbaum.

Spelke, E.S. (1990). Principles of object perception. *Cognitive Science, 14,* 29-56.

Spelke, E.S. (1991). Physical knowledge in infancy: reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 133-169). Hillsdale, NJ: Earlbaum.

Spelke, E.S., Breilinger, K., Jacobsen, K., & Phillips, A. (1993). Gestalt relations and object perception: A developmental study. *Perception, 22*, 1483-1501.

Starkey, P., Spelke, E., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36*, 97-127.

Strauss, M.S., & Curtis, L.E. (1981). Infant perception of numerosity. *Child Development, 52*, 1146-1152.

Summerfield, Q., & Culling, J. (1992). Auditory segregation of competing voices: absence of effects of FM or AM coherence. *Philosophical Transactions of the Royal Society of London, Serial B 336*, 357-366.

Tan, L., & Bryant, P.E. (1996, April). Discrimination of simultaneously presented large-number arrays by 6-month-old infants. Poster presented at the 10th Biennial International Conference on Infant Studies, Providence, RI.

Termine, N., Hrynick, R., Kestenbaum, R., Gleitman, H., & Spelke, E.S. (1987). Perceptual completion of surfaces in infancy. *Journal of Experimental Child Psychology: Human Perception and Performance, 13*, 524-532.

Trehub, S.E., Schneider, B.A., & Henderson, J.L. (1995). Gap detection in infants, children, and adults. *Journal of the Acoustical Society of America, 98*(5), 2532-2541.

Trehub, S.E., & Trainor, L.J. (1990). Rules for listening in infancy. In J.T. Enns (Ed.), *The development of attention: Research and theory. Advances in Psychology, Vol. 69* (pp. 87-119). Amsterdam, Netherlands: North-Holland.

van Loosbroek, E., & Smitsman, A.W. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26*, 916-922.

von Hofsten, C. (1983). Foundations for perceptual development. *Advances in Infancy Research, 2*, 241-264.

von Hofsten, C., & Spelke, E.S. (1985). Object perception and object-directed reaching in infancy. *Journal of Experimental Psychology: General, 114*, 198-212.

Wattam-Bell, J. (1991). Development of motion-specific cortical responses in infancy. *Vision Research, 31*(2), 287-297.

Werker, J.F., & LaLonde, C.E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental Psychology, 24*, 672-683.

Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1-34.

Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science, 7*, 164-169.

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology, 30*, 111-153.

Xu, F., & Spelke, E. (1998, April). Large-number discrimination in 6-month infants: Evidence for an accumulator model. Poster presented at the 11th Biennial International Conference on Infant Studies, Atlanta, GA.

Yonas, A., Cleaves, W., & Petterson, L. (1978). Development of sensitivity to pictorial depth. *Science, 200*(4337), 77-79.

Yonas, A., Granrud, C.E., & Petterson, L. (1985). Infants' sensitivity to relative size information for distance. *Developmental Psychology, 21*(1), 161-167.