

, 2021, pp. 1-35.

Recognition of intersensory relationships based on duration in infants four, seven and 10 months old (Reconocimiento de relaciones intersensoriales basadas en la duración en bebés de cuatro , siete y 10 meses).

Martínez, Mauricio, Español, Silvia y Igoa, José-Manuel.

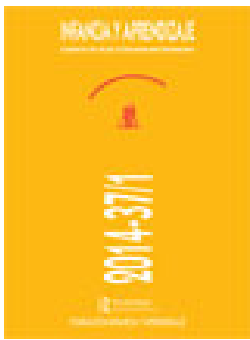
Cita:

Martínez, Mauricio, Español, Silvia y Igoa, José-Manuel (2021). *Recognition of intersensory relationships based on duration in infants four, seven and 10 months old (Reconocimiento de relaciones intersensoriales basadas en la duración en bebés de cuatro , siete y 10 meses)*. ,, 1-35.

Dirección estable: <https://www.aacademica.org/silvia.espanol/151>

ARK: <https://n2t.net/ark:/13683/pH0V/q8v>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.



Recognition of intersensory relationships based on duration in infants four, seven and 10 months old (*Reconocimiento de relaciones intersensoriales basadas en la duración en bebés de cuatro, siete y 10 meses*)

Mauricio Martínez, Silvia Español & José-Manuel Igoa

To cite this article: Mauricio Martínez, Silvia Español & José-Manuel Igoa (2021): Recognition of intersensory relationships based on duration in infants four, seven and 10 months old (*Reconocimiento de relaciones intersensoriales basadas en la duración en bebés de cuatro, siete y 10 meses*), Journal for the Study of Education and Development, DOI: [10.1080/02103702.2021.1925457](https://doi.org/10.1080/02103702.2021.1925457)

To link to this article: <https://doi.org/10.1080/02103702.2021.1925457>



Published online: 21 Jun 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



Recognition of intersensory relationships based on duration in infants four, seven and 10 months old (*Reconocimiento de relaciones intersensoriales basadas en la duración en bebés de cuatro, siete y 10 meses*)

Mauricio Martínez^{a,b}, Silvia Español^c and José-Manuel Igoa^b

^aUAI, Universidad Abierta Interamericana; ^bUAM, Universidad Autónoma de Madrid; ^cIICAL (Instituto de Investigaciones Sociales de America Latina), CONICET – FLACSO

ABSTRACT

Since birth, infants develop the ability to perceive a wide range of intersensory relations among various kinds of amodal temporal information. This study addresses the development of the ability to perceive duration-based intersensory relations. Three groups of infants, four, seven and 10 months old, participated in two trials of an intersensory 5preference task that used two audiovisual stimuli showing a woman executing two performances — with characteristics similar to the way adults interact with infants — consisting of sequences of sound-movement/silence-stillness of varying durations. The results show that infants at all three ages recognize the duration-based intersensory relation; however, recognition is expressed in different ways at each age according to the trial and the number of eye fixations recorded. These results are discussed in terms of the development of the ability to recognize duration-based intersensory relations and its relationship to between-subjects development.

RESUMEN

Desde que nacen, los bebés desarrollan la habilidad para percibir un amplio rango de relaciones intersensoriales entre diferentes tipos de información amodal temporal. En este estudio se indaga el desarrollo de la habilidad para percibir relaciones intersensoriales basadas en la duración. Tres grupos de bebés de cuatro, siete y 10 meses participaron en dos ensayos de preferencia intersensorial en los que se utilizaron dos estímulos audiovisuales donde se observa a una mujer ejecutando dos *performances* –con características similares a la que los adultos realizan frente a los bebés– compuestas de frases de sonido-movimiento/silencio-quietud de diferentes duraciones. Los resultados indican que los bebés de las tres edades reconocen la relación intersensorial basada en la duración. Sin embargo, en las tres edades estudiadas el reconocimiento se expresa de forma diferente en función del factor ensayo y de la cantidad de fijaciones oculares. Los resultados se discuten en términos del desarrollo de la capacidad para reconocer

ARTICLE HISTORY

Received 28 August 2019
Accepted 14 February 2020

KEYWORDS

intersensory development;
duration; intersensory
relations; social perception

PALABRAS CLAVE

desarrollo intersensorial;
duración; relaciones
intersensoriales; percepción
social

CONTACT Mauricio Martínez ✉ martinez_ms75@hotmail.com 📧 Facultad de Psicología y Relaciones Humanas, UAI, Chacabuco 90 P.6 – Ciudad Autónoma de Buenos Aires, Argentina.

English version: pp. 1-16 / *Versión en español*: pp.17-32

References / *Referencias*: pp. 32-35

Translation from Spanish / *Traducción del español*: Mary Black

© 2021 Fundación Infancia y Aprendizaje

relaciones intersensoriales basadas en la duración y su posible relación con el desarrollo intersubjetivo.

Starting in the first few months and throughout the entire first year of life, infants are involved with adults in dyadic interactions unmediated by objects, in which the topic of the interaction is pure social contact (Stern, 1983). Sometimes, too, infants participate in interactions in which adults handle objects in front of them in an ostensive manner (Moreno-Núñez et al., 2015). During this period, infants develop increasingly complex, subtle modes of between-subjects contact with their nurturers, and their modes of participation in the interaction change considerably: their express behaviours — smiles, vocalizations, movements — which are initially brief and isolated are gradually expressed in a more integrated fashion (Kaye, 1986). After two months of age, they use different modes (combining vocalizations, facial expressions and gestures) to express their intention to get the adult involved again when they break off the interaction dynamics (Melinder et al., 2010). At around six months old, they become active instigators of eliciting different responses in the adult (Reddy, 2008). And during the last quarter of their first year, they begin to coordinate joint praxic and interpersonal acts (Huble & Trevarthen, 1979), giving rise to the emergence of joint attention (Tomasello, 2013).

During dyadic interactions, the adult provides sensory information to several of the infant's perceptual systems (visual, auditory, tactile and others) in order to attract and sustain their attention (Stern et al., 1977). The information that the adult provides has particular features, such as the exaggeration of facial expressions; the slowness of changes in tone and the intensity of vocalizations (Stern, 1983); the presence of extended melodic features with varied intonations that are more pronounced than in usual speech among adults (Papoušek & Papoušek, 1981); the varied repetition of prototypical melodic contours (Papoušek, 1996); and the organization, coordination and multimodal concordance of sounds and movements (I. Martínez, 2014; Stern, 1983). The information provided by the adult is embedded in multimodal patterns from the very start: the melodic contours are closely related to the mother's movement patterns, and the regular synchronization of vocal and kinetic patterns — which include tactile information (when she caresses, touches or nudges the child), proprioceptive information (when she moves their hands or feet) and visual information (when greeting the infant with her head or shaking her head) — provides the infant with redundant multisensory information (Papoušek, 1996).

In recent years, adult interaction directed at infants has been researched using theoretical tools and analytical techniques from the temporal arts — music, dance and poetry. These studies have revealed the elaborate dynamic, temporal and multimodal moulding of some of these actions (Español & Shifres, 2015; I. Martínez, 2014; I. Martínez et al., 2018; Miall & Dissanayake, 2003; Schögler & Trevarthen, 2007; Shifres, 2014), which has led them to be described as *infant-directed artistic performances* (Español & Shifres, 2015). These performances are essentially comprised of sounds, shushing noises, vocalizations, touches and coordinated movements; they may contain speech, but they may also occur in the absence of speech (I. Martínez et al., 2018;

Ospina & Español, 2014). In the varied performances that adults improvise for infants, temporal information is presented to the infant's different perceptual systems. Temporal information (i.e., duration, rate [chronometric density], rhythm and intensity) is organized by means of the resources of the temporal arts, like the repetition-variation of sound-kinetic motifs and the uses of pauses and *ritardandi* (gradually getting slower) according to the hierarchical structure of the phrase, and the shaping of rhythmic patterns to achieve a state of balance or stability (Español & Shifres, 2015).

Even though the adult behaviour changes during the between-subjects encounter, the change that occurs is not structural in the sense that the capacity underlying the behaviour does not change: the adult only intuitively adjusts (Papoušek, 1996) their behaviour to the infant's participation, seeking to foster an optimal state of attention in her (Stern, 1983). However, as mentioned in the beginning, the infant's behaviour does change structurally during the first 12 years of life. The changes in the infant's behaviour may be partially sustained by changes in their ability to perceive the temporal information provided by the adult. The development of perceptual and between-subjects abilities might be best characterized in dialectical terms, rather than in a unidirectional fashion (Español, 2010; Rivière, 2003). In this sense, the infant's engagement in the adult performance will foster the development of intersensory perception, and the development of intersensory perception will have an impact on the infant's increasing engagement during the interaction. Given this hypothesis, the study of the development of intersensory perception within the temporal domain — synchrony, duration, rate, rhythm — is important not only for ascertaining the development of infant perception but also for understanding between-subjects development.

The duration of adult behaviour directed at the infant

The duration of adult behaviour in interaction with the infant has been studied in a range of contexts. Kaye (1977) showed that the variation in the duration of mothers' rocking while nursing or while feeding the infant with the bottle extends or reduces the pause that the infant takes between cycles of food sucking.

When comparing the behaviour of an adult in interaction with an infant and with another adult, we see that the phrases that the adult directs to the infant tend to last half as long as those they direct to another adult, while the pauses separating the phrases are almost twice as long (Stern et al., 1977). On the other hand, in relation to the speech directed at the infant, it has been found that adults stretch out the duration of vowels (Stern, 1983) and that the average length of the vocalization is longer than the pause (Jaffe et al., 2001; Stern, 1983).

Stern identified a particular interaction context which he called early social play. He considered it the social mode of interaction par excellence, as the focus of play is the social interaction itself. In this context, both the infant's behavioural and perceptual repertoire and the adult's action are used to share the intersubjective experience. Early social play is comprised of two episodes: one of participation and the other of pause. The mutual participation episodes are a sequence of social behaviours with different *durations*, while the pause episodes consist of behavioural silence (both sound and kinetic) whose *duration* is longer than that found in participation episodes (Stern, 1983). Likewise, the participation episodes are organized around phrases and series. A

phrase is a single vocal utterance or a burst of movement whose *duration* is generally less than one second, while a series is comprised of several phases whose content or *duration* are similar (Stern et al., 1977). The difference in durations between phrases, series (which comprise the interactive episode) and pause episodes illustrates the highly ordered temporal world that the mother provides her child (Schaffer, 1985). It would appear that mothers act at all times as if their infants could only assimilate much smaller bits of information than adults and needed more time to process each bit before receiving another one (Schaffer, 1989).

The temporal organization of adult actions in terms of duration can also be seen in the concurrence between the duration of vocalizations and the duration of the movements that adults make within the context of early social play (Stern, 1974; Stern et al., 1977). In turn, microanalyses of infant-directed performances show other kinds of organization based on duration. These performances have the structure of a phrase constructed by the varied repetition of brief sound-kinetic motifs. The motifs are diverse: raising and lowering the infant while holding her by the trunk, or the simple act of removing a blanket covering her, a sequence of shusher noises or a brief sequence of words. Variations come in the dynamic and temporal components of the motif (Español, 2014; I. Martínez et al., 2018). I. C. Martínez (2007) identified sound-kinetic phrases in which the variation lay in the duration of the repeated motifs. Español and Shifres (2015) identified phrases in which some features of the movement changed (such as its form) while the duration of a feature of the kinetic motif — its quality — remained constant in its repetition. The regularities and variations in the length of the motif reveal duration to be an organizer of adult action at different hierarchical levels. These variations and regularities in duration show the orderly yet changing, complex and highly stimulating world that adults provide the infant. For example, an adult directs a brief phrase at a seven-month-old infant starting with a kinetic motif: the move of removing the blanket covering the infant. The motif comes in the form of repetition-variation (it is repeated three times with variations) or coda (ending). The movement of getting closer to and further from the infant on the sagittal plane occurs with temporal regularity in all three repetitions. The adult accompanies the repetition of the motif with speech and vocalizations: ‘Let’s see ... should we take it off?’, ‘Tititi’¹, ‘Take it off-take it off-take it off!’ The words and vocalizations of the performance have musical qualities which finely blend with the qualities of the movement. For example, on the third repetition, the adult removes the blanket from the infant’s body with gliding (light, direct and sustained) and legato (bound) movements synchronized with the *staccato* (detached, choppy) sound of ‘Take it off-take it off-take it off!’ of a similar duration. The whispered words ‘Take it off-take it off-take it off!’ accompany the movement of removing the blanket, which is extended with a final *ritardando*, such that the sequence ends with a *ritardando* or prolongation and a *diminuendo* (reduction of the sound). Likewise, in the coda, the movement of getting closer and farther comes with the vocalization ‘u-uh-pate!’ which is produced right as the mother is moving towards the infant; this is the longest approach movement of all the repetitions of the motif and has the same extended duration as the vocalization ‘u-uh-pate!’ (simplified version of the analysis performed by Español & Shifres, 2015).

Finally, we should highlight the role played by duration in another characteristic adult behaviour: affect attunement. Mothers’ affect attunement responds to the infant’s

behaviour with a different behaviour that preserves and reflects the time pattern (duration, pulse and rhythm), the intensity (absolute or relative) and/or the spatial pattern of the infant's behaviour (Stern, 1991). For example, while an infant girl who is trying to reach a toy utters an exuberant 'ahhh!', her mother shrugs her shoulders and shakes her torso. 'The dance *lasts* as long as the girl's ahhh' (Stern, 1991, p. 175, emphasis added). 'The intensity and *duration* of the girl's voice pairs with the mother's body movements' (Stern, 1991, emphasis added). Once again, this reveals the constitutive role of duration: the affect attunement of the infant's behaviour via the affection of the behaviour displayed by the mother occurs by virtue of duration, along with other temporal elements. Affect attunement is an adult matching behaviour that occurs more frequently than imitation during early social play and is present after the second month of the infant's life (Español, Bordoni, Carretero Pérez et al., 2018). Affect attunement occurs in a time window of one second or less, and primarily the mother matches vocalizations to the infant's movements. The infant's visual orientation behaviour with regard to the affect attunement changes from six to 12 months. During the second half of the first year of life, the frequency of looks towards the adult rises when the adult imitates the infant and drops when the adult performs affective attunement towards her (Bordoni et al., 2016). In order for the affect attunement changes to impact the infant's behaviour, the infant must *recognize* — in the next section we shall explain why we specifically use this term — the intersensory relations between the duration of her behaviour and that of the adult which is expressed in a different sensory modality.

Multisensory perception

There is a delicate amalgamation between the adult's and the infant's behavioural repertoire, a delicate and not yet well explored relationship between the characteristics of the adult's behaviour and the infant's capacities to perceive not only the duration but also all the temporal information. Schaffer mentioned 'the impress precision with which, under normal circumstances, the stimulation produced by the adult matches the child's capacity to absorb and respond to that stimulation' (1989, p. 48). In a similar vein, Stern claimed that:

[...] it is difficult to understand how the infant could react as he does, as well as begin to comprehend his social universe, if he (or his nervous system if you prefer) were not capable of some fairly impressive time-estimating operations. (1983, pp. 112–113)

Given the multimodal nature of adult behaviour, we can assume that the infant's ability to perceive simultaneous information in a unified way is crucial for between-subjects contact. Multisensory perception is an organism's ability to perceive the stimuli that simultaneously provide information to different perceptual systems in a unified fashion (Bremner et al., 2012). It is possible to perceive different types of information from the same stimulus in a unified fashion: amodal information and modality-specific information (Bahrick & Lickliter, 2012). Amodal information is the kind that is perceived indistinctly via the different perceptual systems, such as duration, rate and rhythm in the temporal dimension and shape, texture and continuity in the spatial dimension. Modality-specific information is the kind that can only be perceived via a single perceptual system, such as colour, smell, heat or

timbre. Generally speaking, multimodal events are specified via a combination of amodal and modality-specific attributes (Lewkowicz & Kraebel, 2004).

In infant-directed performances, both types of information can be identified. Let us imagine, for example, a three-month-old infant lying before her mother. Imagine that the mother gently strokes the infant's chin while making a vibrating movement with her index finger and vocalizing the sound /trr/. The mother's quick, soft touches and the vocalizations take place repeatedly in a synchronous, simultaneous fashion. Both types of information are found in this situation: first, each of these elements — the timbre of the voice, the hand movements and the soft pressure of the mother's finger on the infant's chin — is perceived by a single perceptual system — sight, hearing and touch, respectively. They are clear examples of information in a single modality. Likewise, when the mother gently moves her index finger on the infant's chin in synchrony with the vocalization of the sound /trr/, the duration of the sound of the vocalization matches the duration of the movement of the mother's finger on the infant's chin. Duration is a clear example of amodal information. The mother is simultaneously providing the same information on duration via her vocalizations and her strokes of the infant's chin. Rhythm, duration, intensity and synchrony are cases of temporal amodal information. In the performances she directs at the infant, the adult is providing not only unimodal information but also a wide range of amodal temporal information.

In the field of multisensory perception, two capacities are predominantly studied: discrimination and recognition. *Discrimination* refers to the ability to distinguish between two or more elements, events or types of information, while *recognition* implies the ability to establish relationships (intersensory perception) between the information perceived simultaneously via two or more perceptual systems (Lewkowicz, 2000; Walker-Andrews, 1997). Discrimination is studied using the infant-controlled habituation procedure (Horowitz et al., 1972). The logic of the technique is that the cognitive system gradually pays decreased attention to a novel stimulus. When another new stimulus appears, attention to this new stimulus increases. The procedure consists of presenting a series of trials to get the participants accustomed to a certain visual or audiovisual stimulus. Then the testing trials are presented. If the infant fixes her gaze for a longer time, it is inferred that the infant is discriminating between the two stimuli. Recognition is studied using the 'intersensory preference technique' (Spelke, 1976), which usually consists of presenting two visual stimuli at the same time, next to each other, and a simultaneous auditory stimulus which is congruent with only one of the visual stimuli. During the infants' participation, the proportion of total looking time directed at the congruent visual stimulus is calculated. When this proportion is significantly higher (statistically) than 0.50, it can be inferred that the infant recognizes the relationship between the stimulus perceived by her visual system and the one perceived by her auditory system. This technique addresses the capacity to establish relationships (intersensory perception) between the information available to different perceptual systems — such as the equivalent duration of the information seen and heard. Even though the value of intersensory perception for social, emotional, cognitive and linguistic development is acknowledged (Bahrick & Lickliter, 2012), there are few studies that inquire into the different intersensory perception capacities associated with temporal features (synchrony, duration, rate and rhythm) of adult behaviour (M. Martínez,

2016). The overarching question guiding our research is how the infant perceives intersensory relations based on duration from all the amodal temporal information provided by the adult. In a recent study, M. Martínez et al. (2018) examined the recognition of intersensory relationships based on rhythm in infants aged four, seven and 10 months. They also inquired into the infant's ability to recognize the relationship between the same rhythmic pattern in hand movements (and the movement of the orofacial area made to articulate a vocal gesture) and the constant sound /ta/; both of them — sounds and movements — were executed by a woman mimicking prototypical adult behaviour when interacting with infants. The results showed that recognition of the intersensory relationship based on rhythm in a social stimulus resembling a performance emerges at around 10 months. The specific question we seek to address here is how the infant perceives the duration-based intersensory relations coming from the information available in the performance that the adult directs at them.

The development of the perception of duration

This section presents the available data on infants' capacities for perception associated with duration. There are several studies that inquire into the role of perception of duration with regard to speech perception (Ainsworth, 1972; Eilers et al., 1984; Jusczyk et al., 1983; Kuhl & Meltzoff, 1982; Miller & Liberman, 1979). All these studies were conducted with unimodal information. On the other hand, Lewkowicz (2000) claims that there is indirect evidence that after the first month of life, infants can discriminate differences in duration under 100 ms specified visually and auditorily. Morrongiello (1984) found that infants aged six and 12 months discriminated between series — in the sense that Stern (1983) used the term — comprised of three sound events and two silences (sound-silence-sound-silence-sound) when they differ in the durations of events. The minimum differences that the infants were able to discriminate were 40 and 60 ms. That is, they discriminated series of three sounds lasting 200 ms and two bouts of silence lasting 200 ms (200-200-200-200-200) from series of three sounds lasting 160 ms and two bouts of silence lasting 140 ms (160-140-160-140-160). On the other hand, Morrongiello and Trehub (1987) reported that six-month-old infants only discriminate changes in duration when they are longer than 20 ms. These studies also used unimodal information.

Some years ago, Lewkowicz (1986, 2000) stated that studies on infants' ability to establish duration-based intersensory relations were few and far between. Today, this has not changed much: if one examines the latest *Handbook on Multisensory Development* (Bremner et al., 2012), no specific studies on the intersensory perception of duration appear.

Lewkowicz (1986) exposed infants aged three, six and eight months to two types of intersensory preference trials (one where the difference in duration was 800 ms and another where it was 1,200 ms). In the first six trials, the infants simultaneously looked at two light tablets (located next to each other) which blinked at the same rate but with different durations. For the next 12 trials, the infants once again looked at the same tablets, but this time a sound was presented synchronously at the beginning and end of the light blink in one of them; that is, its duration was identical to the light blink of one of the tablets. In the respective trials, the tablets blinked at three possible durations

(400, 800 and 1,600 ms). We should note that the duration variable was implemented by means of synchrony between stimuli, such that both coincided at the beginning and end for equal durations and only at the beginning for different durations of both stimuli. The data obtained enabled him to conclude that the infants aged six and eight months old recognized duration-based intersensory relations when they were presented with the pairs 400–1,600 ms long and the pairs 800–1,600 ms long. Likewise, when the start of the sound was temporally displaced from the start of the light blink, the infants were not able to recognize duration-based intersensory relations. In this sense, one key factor in their recognition was the synchrony between the start and end of the duration of the information present in the different sensory systems. This might indicate that it is impossible for the infants to recognize duration-based relations in absolute terms. That is, they cannot recognize the equivalence of the duration of the information when it is not synchronous, that is, when its beginning and end do not match. Consequently, Lewkowicz (2000) claims that infants require signals corresponding to synchrony and duration to recognize intersensory equivalence based on duration and that duration by itself is not sufficient.

In summary, the studies reviewed show that infants aged six and eight months can recognize duration-based intersensory relations in simple (bimodal) stimuli. However, many questions clearly remain unanswered. We can wonder whether infants recognize duration-based intersensory relations in social stimuli, whether this recognition occurs before six months and whether changes in recognition can be observed throughout a relevant time period. The objective of this research is to experimentally study the ability of infants aged four, seven and 10 months to recognize duration-based intersensory relations in the information coming from the actions performed by a person mimicking an artistic performance directed at the infant. The interest of using a social stimulus arises from the implications that this study could have on understanding early between-subjects development.

Method

Participants

The participants were 48 infants divided into three groups, with 16 participants aged four months ($M_{\text{age}} = 125$ days, $SD =$ four days, 121–131; eight females), 16 participants aged seven months ($M_{\text{age}} = 220$ days, $SD =$ six days, 210–228; seven females) and 16 participants aged 10 months ($M_{\text{age}} = 307$ days, $SD =$ eight days, 283–316; nine females). The infants' participation was conditioned upon the following requirements: (i) full-term pregnancy; (ii) a score equal to or higher than 8 on the APGAR test; (iii) the general practitioner did not suspect or detect any serious health problem during the first month of life; and (iv) the absence of illnesses, impairments or developmental disorders before or during the period when the infant was participating in the experiment.

The data from three infants were excluded because they did not look at the two monitors used in the study to present the visual stimuli during the trials ($n = 3$ aged four months). Before participating, the parents signed an informed consent form.

Design

This study has a cross-sectional, ex post facto developmental design with two independent variables: (1) the intersensory relation between the duration of the hand movement phrases (visual component) and the sound issued (auditory component) as the within-subjects factor with two levels: equal duration (or congruent duration) and unequal duration (or incongruent duration); (2) the participants' age, with the levels four, seven and 10 months old as the between-subjects factor.

Apparatuses

Two 19-inch computer monitors (Samsung SyncMaster T1920) located next to each other and separated by a distance of 18 centimetres were used to present the visual stimuli. The monitors were surrounded by blue curtains so as to prevent distractions to the infants' attention. Between the monitors was an LED (red and blue) device that the researcher manipulated with the purpose of attracting the participants' visual attention to ensure that they all began the trials looking at the space between the two monitors. The visual stimuli were projected onto the monitors using remote control-operated notebooks (Asus Eee PC1015PX). The auditory stimulus was presented with two loudspeakers (Edifier R18USB) located between the two monitors just behind the LED device. For each of the trials in which the infants participated, a video of their face was recorded using a webcam (HP Deluxe Webcam KQ246AA) placed over the LED device in order to measure the amount of time their gaze was fixed on the visual stimuli.

Stimuli

Two audiovisual stimuli (films) were designed following the logic of artistic performances directed at infants (Español & Shifres, 2015). Both stimuli were comprised of a woman facing the camera against a blue background uttering the sound /trrr/ (created with the alveolar stop /t/ and the alveolar approximant /r/) while moving the fingers of both hands (as in a tickling gesture). Her hands were located at face height, with her palms facing forward. The woman moved the fingers of both hands synchronously with the production of the /trrr/ sound. The hand movement started and stopped with her fists closed. The woman made an iterative series of sound/movement phrases of differing durations followed by a pause (episodes of silence/stillness). The phrases comprised two series (Stern et al., 1977), one series per audiovisual stimulus: series A and B, which lasted 20 seconds, resembling prototypical adult action in natural interaction (Español, 2014). The independent variable duration was operationalized according to the temporal distance between the synchronous start and end of the sound and movement phrases made by the woman.

The alternation of multimodal sound-movement phrases and pauses (episodes of silence-stillness) was distributed such that it would be impossible to extract a rhythmic regularity from the series. Likewise, the proportion of sound-movement phrases and pauses was distributed in a balanced fashion. The proportion of sound-to-movement for the first series was 0.65, and for the second it was 0.63. The total duration of both audiovisual stimuli was 90 seconds. In each of them, the series was repeated four and a half times. When the filming

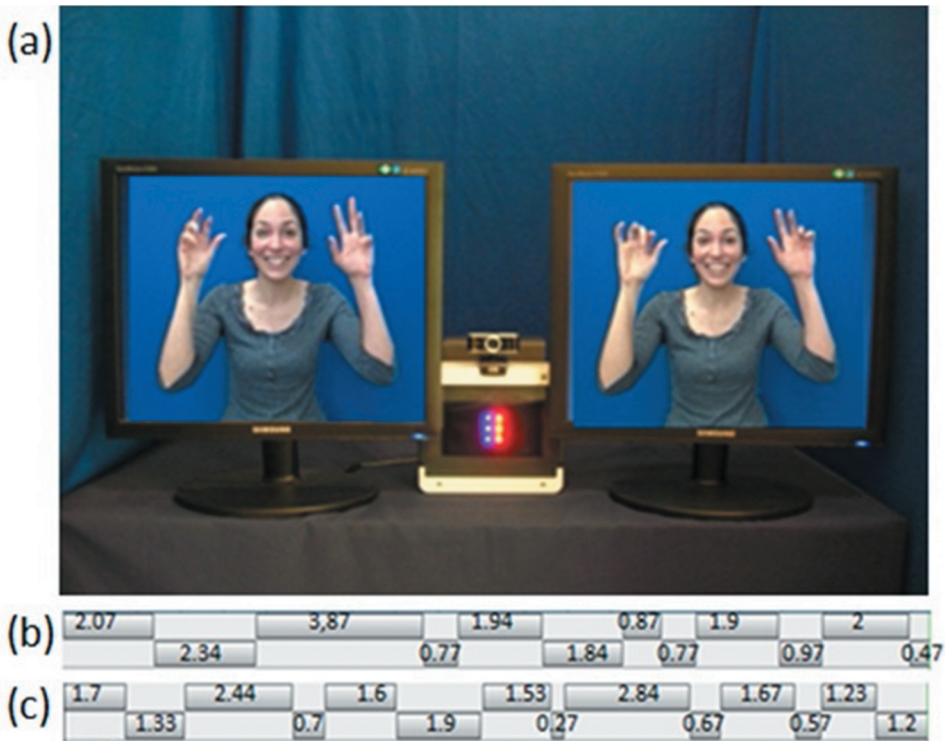


Figure 1. (a) Arrangement of the materials in the experimental setting. (b) and (c) Duration (in ms) of the sound/movement phrases (upper row) and silence/stillness (lower row) comprising each of the episodes used as stimuli.

started, the onset of each of the patterns was synchronous. As both stimuli are being projected, the series starts synchronously on five occasions (Figure 1).

Even though series with intersensory redundancy (between sound and movement, for example) are used to study intersensory relations, the stimulus designed for this study also has intrasensory redundancy (between sound: vocalization /trrr/, and movement: orofacial region, and movement: tickling gesture with the fingers). That is, there is synchrony between the start and the end of the vocalizations and between the start and the end of both movements. There is synchrony in the visual information of both stimuli between the start and the end of the hand movements and the orofacial region. Recently, Bahrck et al. (2015) proved that intrasensory redundancy attracts infants' attention to amodal information, just as intersensory redundancy does. Thus, intrasensory redundancy adds ecological validity (resembling the stimulus of an artistic performance directed at the infant) without generating undesirable effects caused by recognition of the intersensory relationships based, in this case, on duration.

Procedure

The intersensory preference technique was used (Spelke, 1976), which consists of simultaneously presenting two visual stimuli and one synchronous auditory stimulus (the duration

pattern starts and ends at the same time in the visual and auditory information) congruent with one of the two visual stimuli. The infants were seated on their mother's or father's lap around 60 centimetres away from both monitors. The parents were previously instructed on how to hold their child and not to interfere during the session. They were asked to sit comfortably with their back resting against the back of the chair set up for this purpose and to hold their child comfortably on their lap. They were asked to refrain from interacting with their infants (to not direct gestures, verbalizations or facial expressions to them) while the stimuli were being presented. They were also told that if at any point during the procedure they sensed that their child was uncomfortable, they could end it by getting up from the chair.

Each infant participated in two trials (lasting 90 seconds each) in which they simultaneously looked at both visual stimuli as they listened to the sound (auditory stimulus) congruent with one of the visual stimuli. The lateral position of the visual stimuli remained constant on each trial, and from one trial to the next the congruent auditory stimulus varied. The lateral position and the presentation of the auditory pattern were counterbalanced among all participants. During the procedure, the video camera located between both monitors recorded the infant's face in order to subsequently calculate the amount of time their gaze was fixated on the visual stimuli projected on the monitors. Gaze fixation time was measured using ANVIL video-annotation software (Kipp, 2008).

Results

With the aim of finding out at what age infants recognize duration-based intersensory relations, the Proportion of Total looking Time (PTLT) was calculated. The infant's PTLT is the most widely used measurement to verify recognition of intersensory relationships (Flom et al., 2009). It is calculated by dividing the total time of the gaze at the monitor displaying the visual stimulus congruent with the sound by the total time of the gaze directed at both monitors. In order to determine when the infants recognize the duration-based intersensory relation, the mean PTLT of each age group was compared by adding up the data from both trials (T1 + T2) against the value 0.50. The results indicate that infants at all three ages in our study — four months ($M = 0.57$, $SD = 0.10$) $t(15) = 2.759$, $p < .05$, seven months ($M = 0.57$, $SD = 0.06$) $t(15) = 4.582$, $p < .05$ and 10 months ($M = 0.54$, $SD = 0.07$) $t(15) = 2.239$, $p < .05$ — recognized the duration-based intersensory relations. A second analysis was performed to check whether there was a difference in means among the groups according to age. To do so, a one-factor ANOVA was performed grouping the data from both trials (T1 + T2). The result of the analysis did not reveal a significant effect of the age factor ($p > .05$). These first two analyses enable us to assert that infants aged four, seven and 10 months recognize duration-based intersensory relations (Figure 2).

A second series of analyses revealed that the infants did not show any preference for the series (A or B) or lateral visual dominance during each of the trials. To evaluate whether the infants had a preference in terms of *prägnanz* qualities of the stimulus for any of the series, a repeated measures ANOVA was performed (with series A and B as the within-subjects factor and age as the between-subjects factor) over total gaze time that the infants directed at each of them, regardless of the sound they were hearing. No main effects were found for these factors. The result of the analysis revealed that there

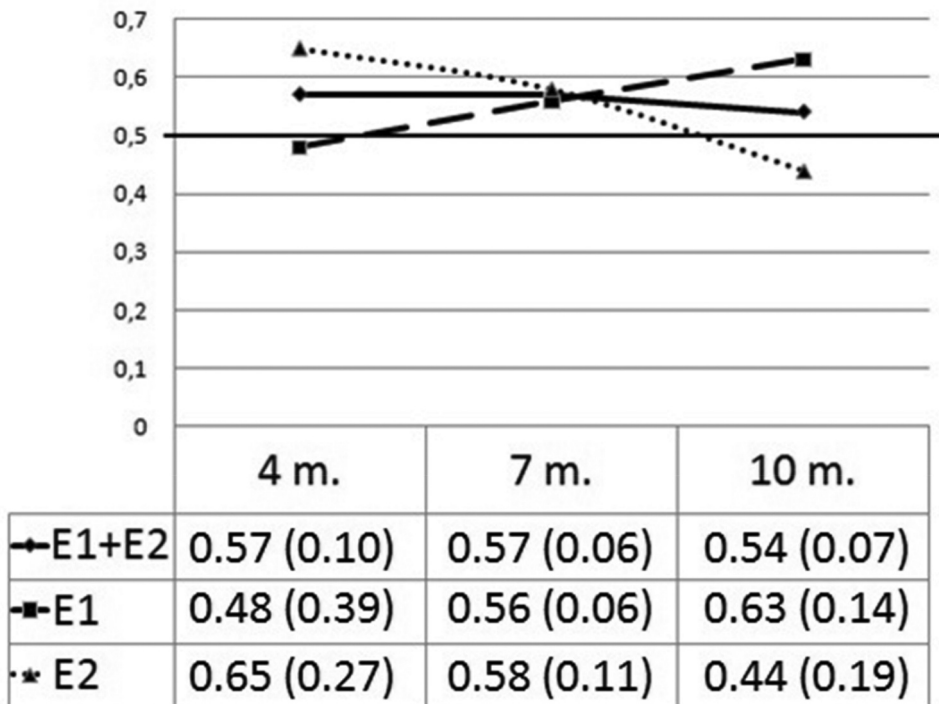


Figure 2. Means of the Proportion of Total Looking Time (and their corresponding standard deviations) directed at the visual stimulus congruent with the sound stimulus at four, seven and 10 months.

are no significant interactions between series (A or B) and age (all ps were $> .05$). The same was done with lateral visual dominance. A repeated-measures ANOVA was conducted (with the monitor, left or right, as the within-subjects factor and age as the between-subjects factor) on the total gaze time the infants directed to the left and right monitors, regardless of the sound they were hearing. No main effects of these factors were found. The results of this analysis did not reveal significant interactions between monitor (right vs left) and age (all the ps were $> .05$). So far, our results indicate that infants aged four, seven and 10 months recognize duration-based intersensory relations.

Given that the infants at all three ages recognize the intersensory relation studied, a third series of analyses was performed to inquire into whether there was any other effect of age on the capacity to recognize intersensory relations.

Specifically, we asked whether recognition of duration-based intersensory relation was expressed the same way at the different ages. E. Gibson (1969) posited that the development of perception entails not only the acquisition of a particular ability but also an increase in the organism's effectiveness at extracting information. Consequently, we first independently analysed the data from each of the trials in which the infants participated. The PTLT measure for each age group in each of the trials (T1 and T2) was checked separately against the value 0.50. The results showed that at four months the infants only recognized the duration-based intersensory relation at T2 ($M = 0.65$,

$SD = 0.27$) $t(15) = 2.184$, $p < .05$. At seven months, the infants recognized the duration-based intersensory relation in both trials, T1 ($M = 0.56$, $SD = 0.6$) $t(15) = 3.967$, $p < .05$ and T2 ($M = 0.58$, $SD = 0.11$) $t(15) = 2.678$, $p < .05$. And at 10 months, the infants recognized the duration-based intersensory relation only at T1 ($M = 0.63$, $SD = 0.14$) $t(15) = 3.468$, $p < .05$. We then performed a two-factor ANOVA (age and trial number: T1 and T2). No main effects were found of either age or trial number on PTLT ($p > .05$). Conversely, an interaction between these two factors was found ($F(2, 33) = 8.098$; $p < .05$, partial eta squared = .15).

The analysis seems to indicate that four-month-old infants need more time to identify the intersensory relation (the recognition occurs in T2) than seven-month-old infants (where recognition occurs between T1 and T2). At 10 months, we also seem to find a fatigue effect (recognition only occurs in T1).

According to E. Gibson (1969), an increase in the ability to extract information depends on development of the voluntary control of exploratory activity. Ruff and Rothbart (1996) believe that the increase in the gaze alternation between two stimuli is an indicator of increased exploratory activity. In order to study the increase in the infants' effectiveness at recognizing duration-based intersensory relations, we performed two analyses. In the first one, we tested the effect of age on the number of gaze alternations between the monitors showing the stimuli, examining both trials together. In the second one, we inquired into the difference between the amount of gaze alternations between monitors in trials T1 and T2. For the first analysis, a one-factor ANOVA was performed, which revealed a significant effect of age ($F(2, 45) = 5.776$; $p < .05$, partial eta squared = .51) on the number of gaze alternations.

The post-hoc Tukey *b* test showed that the mean of all gaze alternations was different at four months ($M = 13.31$, $SD = 7.48$), seven months ($M = 20.91$, $SD = 8.65$) and 10 months ($M = 20.80$, $SD = 5.16$). These data indicate that at seven and 10 months, infants alternate their gazes between the monitors more than at four months.

With regard to the differences between trials, the analysis showed that at four months there is a statistically significant difference with the two other ages in the number of gaze alternations in the second trial ($M = 7.81$, $SD = 4.19$) when compared to the first trial ($M = 5.5$, $SD = 4.12$) $t(15) = 2.569$, $p < .05$, $d = 0.56$. We should recall that at four months, the infants only recognized the intersensory relation during the second trial. At seven months, there is no statistically significant difference between the number of gaze alternations during the first trial ($M = 10.59$, $SD = 4.17$) and the second trial ($M = 10.31$, $SD = 5.11$) $t(15) = 0.324$, $p > .05$. We should recall that at seven months, the infants made the intersensory correspondence in both trials. Finally, at 10 months, there is a statistically significant difference between the number of gaze alternations during the first trial ($M = 11.13$, $SD = 3.48$) and the second trial ($M = 9.67$, $SD = 2.44$) $t(15) = 1.901$ $p < .05$, $d = 0.49$. We should recall that at 10 months the infants only recognized the intersensory relation during the first trial.

Discussion

The results of our study indicate that infants at four, seven and 10 months of age recognize duration-based intersensory relations coming from a stimulus that mimics the multisensory performance that adults tend to make for infants during between-

subjects exchanges. Results from previous studies indicated that infants aged six and eight months, but not those aged three months, recognized intersensory relationships based on different durations (800 and 1,200 ms) in stimuli comprised of light and sound. Our results expand our knowledge of infants' ability to recognize duration-based intersensory relations. First, they enable us to state that infants recognize duration-based intersensory relations with information coming from people. Secondly, we can assert that four-month-old infants recognize duration-based relations. Thirdly, our results reveal that infants' ability to recognize is not expressed identically between four and 10 months.

This last piece of information is relevant in the face of one of the most frequent criticisms of the studies that examine the different perceptual skills. This criticism claims that most of the research only focuses on identifying at what age (the earlier, the better) infants show some kind of perceptual capacity without providing further information on how this perceptual capacity develops (Mariscal et al., 2012). Our data show how the effectiveness of infants' ability to recognize duration-based intersensory relations improves. The notion of the increase in a skill stems from the tradition developed by J. Gibson (1966) and E. Gibson (1969; E. Gibson & Pick, 2000). This increase is revealed by a reduction in the amount of time needed to extract information (E. Gibson, 1969). Our data show that infants aged four months only detect the relationship in the second trial, while at seven months they detect it in both trials and at 10 months they detect it only in the first trial.

Several studies prior to ours show that after five months old, infants tend to exhibit the skill of recognizing various intersensory relationships in a different fashion along the trials in which they participate (Schmuckler & Jewell, 2007; Soken & Pick, 1992; Walker, 1982). Generally speaking, when infants show recognition in the first trial(s) but not in the last one(s), this difference is attributed to several factors such as boredom, fatigue or attention to other properties of the events (Vaillant-Molina et al., 2013). In our case, this occurred with the 10-month-old infants, so it is possible that at this age the stimulus is no longer interesting for them. We should mention that this pattern has been attributed to other factors, such as perceptual narrowing (Kubicek et al., 2014) or violation of the infant's expectations (Imafuku et al., 2019). However, future research should determine whether perceptual narrowing and violation of expectations play a role in the development of the perception of temporal amodal information.

On the other hand, the infants aged seven and 10 months alternated their gazes more than the four-month-olds in both trials. Likewise, at all three ages, the infants alternated their gazes more than in the trials in which they recognized the intersensory relation: at four months old the infants alternated their gazes more in the second trial, at seven months they made the same number of gaze alternations in both trials, and the infants aged 10 months alternated their gazes more in the first trial. Both the increase and the decrease in the number of gaze alternations tend to be interpreted as the expression of the development and voluntary control of visual attention (Ruff & Rothbart, 1996). As a whole, the data enable us to identify a pattern of change in recognition of duration-based intersensory relations between four and 10 months.

In the introduction, we hypothesized about the possible dialectical link between perceptual and between-subjects development during the first year of life, we suggested

that some of the changes observed in the infant's behaviour may be due to the development of their perceptual capacities associated with duration, and we listed a series of studies which highlight the importance of duration in between-subjects encounters between adults and infants.

The relationship between perceptual and between-subjects development can be seen in a variety of contexts. Stern (1991) exemplifies it in the interactive context in which the adult makes a performance: 'If the duration of the contoured stroke and the pauses between strokes had the same absolute and relative durations as the patterns of vocalisation and pause, the infant will experience similar activation profiles' (p. 58). The similarity in duration allows the infant to connect the phenomena, recognize the other as a unified whole. This is because, according to Stern (1991), all the stimuli that issue from a person have a temporal coherence (duration, rhythm). And their uniqueness (the particular features of the sounds and movements) will enable the infant to recognize and distinguish among people because each person has a particular style of stroking and speaking, or, in more technical terms, each person uniquely expresses the forms of vitality (Stern, 2010). Likewise, the relationship between perceptual and intersubjective development can be observed in the emergence of the infant's ability to visually alternate between their own and others' actions (basic skill for joint reference). The infant's recognition of the very temporal pattern of rhythm in her own and the adult's behaviour sustains the emergence of gaze alternation. For example, when the adult and the infant play at beating out the same rhythmic pattern on a table (activity achieved by the adult's adjustment to the infant's rhythmic pattern), the infant can then alternate her gaze between her own and the adult's action (Fogel & DeKoeper-Laros, 2007).

In this sense, we believe that the data obtained are encouraging to continue working on the dialectic link speculatively suggested by Rivière (2003) and Español (2010). To do so, the occasional emergence of different between-subjects capacities should be connected with the detection, discrimination and recognition of multimodal patterns of temporal amodal information. For example, M. Martínez (2019) links the development of multisensory perception (detection, discrimination and recognition) of rhythmic patterns to changes in the infant's social behaviour during the first year of life. On the other hand, it is essential to raise theoretical proposals that explain the ontogenesis of the multisensory perception of temporal amodal information according to the infant's engagement in intersubjective exchanges with the adult. M. Martínez (2021) outlines an initial approach to the ontogenesis of multisensory perception of temporal amodal information based on a Vygotskian perspective, identifying possible social mediation systems that enable the genesis of the infant's autonomous exploratory activity needed to extract perceptual invariants by means of the reorganization of natural and combined systems.

On the other hand, the data provided in this study, along with those reported by M. Martínez et al. (2018) on the recognition of intersensory relationships based on rhythm, are an important contribution to our knowledge of the development of intersensory perception of temporal amodal information. Lewkowicz (2000) proposes a progress sequence for the development of the perception of temporal information in which the ability to perceive synchrony, duration, rate and finally rhythm emerge successively. Our data generally support this sequence. However, Lewkowicz's proposal does not consider a series of questions that we believe are relevant, such as whether infants

recognize duration-based intersensory relations in information afforded by people, whether the recognition of these relations predates recognition of intersensory relationships based on rhythm (in information from people) and whether there are developmental changes associated with the effectiveness in recognizing duration-based intersensory relations during the first year of life. First, the data reported in this study enable us to assert that at four, seven and 10 months, infants recognize duration-based intersensory relations in information from people. Secondly, the data reported here, coupled with the data reported by M. Martínez et al. (2018), allow us to assert that the recognition of duration-based intersensory relations ontogenetically precedes the recognition of rhythm-based intersensory relationships in information from people. Finally, this study provides empirical evidence that supports the claim that during the span from four to 10 months, infants gain greater effectiveness at recognizing duration-based intersensory relationships.

Reconocimiento de relaciones intersensoriales basadas en la duración en bebés de cuatro, siete y 10 meses

Desde los primeros meses y a lo largo del primer año de vida, los bebés se involucran con los adultos en interacciones diádicas, no mediadas por objetos, en las que el tema de la interacción es el puro contacto social (Stern, 1983). En ocasiones también, los bebés participan en interacciones en las cuales los adultos, de manera ostensiva, manipulan objetos frente a ellos (Moreno-Núñez et al., 2015). Durante este periodo, los bebés desarrollan modos cada vez más complejos y sutiles de contacto intersubjetivo con sus figuras de crianza. Y sus modos de participación en la interacción cambian ostensiblemente: sus conductas expresivas — sonrisas, vocalizaciones, movimientos — inicialmente breves y aisladas, van tendiendo progresivamente a manifestarse de manera integrada (Kaye, 1986); a partir de los dos meses, expresan de diversos modos (conjugando vocalizaciones, expresiones faciales y gestos) su intención de involucrar nuevamente al adulto cuando éste quiebra la dinámica de la interacción (Melinder et al., 2010); hacia los seis meses, se tornan activos provocadores elicitando diferentes respuestas en el adulto (Reddy, 2008); y durante el último trimestre del primer año, comienzan a coordinar actos práxicos conjuntos y actos interpersonales (Hubley & Trevarthen, 1979) dando lugar a la emergencia de la atención conjunta (Tomasello, 2013).

Durante las interacciones diádicas, con el objeto de atraer y sostener la atención del bebé, el adulto brinda información sensorial a varios de sus sistemas perceptivos (visual, auditivo, táctil, entre otros) (Stern et al., 1977). La información que provee el adulto presenta rasgos particulares, como la exageración de expresiones faciales y la lentitud en los cambios en el tono e intensidad de las vocalizaciones (Stern, 1983), la presencia de rasgos melódicos extendidos con variaciones entonativas más acusadas que las del habla habitual entre adultos (Papoušek & Papoušek, 1981), la repetición-variada de contornos melódicos prototípicos (Papoušek, 1996), así como la organización, coordinación y concordancia multimodal de sonidos y movimientos (I. Martínez, 2014; Stern, 1983). La información que provee el adulto está, desde el inicio, incrustada en patrones multimodales: los contornos melódicos están estrechamente relacionados con los patrones de movimiento de la madre, la sincronización regular de patrones vocales y kinéticos — que incluyen información táctil (cuando lo acaricia, lo toca o lo empuja), propioceptiva (cuando le mueve las manos o los pies) y visual (al saludarlo con la cabeza o cuando agita la cabeza) — provee al bebé de información multisensorial redundante (Papoušek, 1996).

En los últimos años, se ha investigado la actuación adulta dirigida al bebé con herramientas teóricas y técnicas de análisis provenientes de las artes temporales — música, danza y poesía. Estos estudios pusieron en evidencia el elaborado moldeado dinámico, temporal y multimodal de algunas de estas actuaciones (Español & Shifres, 2015; I. Martínez, 2014; I. Martínez et al., 2018; Miall & Dissanayake, 2003; Schögler &

Trevarthen, 2007; Shifres, 2014), lo que llevó a describirlas como *performances artísticas dirigidas al bebé* (Español & Shifres, 2015). Las *performances* están constituidas esencialmente por sonidos, chistidos y vocalizaciones, toques y movimientos coordinados; pueden contener habla, pero pueden ocurrir también en ausencia de habla (I. Martínez et al., 2018; Ospina & Español, 2014). En las variadas *performances* que los adultos improvisan para el bebé se presenta información temporal para los diferentes sistemas perceptivos del bebé. La información temporal (como la duración, la tasa (densidad cromométrica), el ritmo y la intensidad) se organiza en ellas con recursos de las artes temporales, como la forma repetición-variación de motivos sonoro-kinéticos, el uso de pausas y *ritardandi* (progresivamente más lento) de acuerdo a la estructura jerárquica de la frase, y la conformación de patrones rítmicos para alcanzar un estado de balance o estabilidad (Español & Shifres, 2015).

Si bien la conducta adulta cambia durante el encuentro intersubjetivo, el cambio que acontece no es estructural, en el sentido de que la capacidad que subyace a la conducta no cambia: el adulto sólo ajusta intuitivamente (Papoušek, 1996) su conducta frente a la participación del bebé, buscando propiciar un estado óptimo de atención en el bebé (Stern, 1983). Sin embargo, como se mencionó al inicio, la conducta del bebé sí cambia estructuralmente durante los primeros 12 meses de vida. Es posible que los cambios en la conducta del bebé estén parcialmente sustentados por los cambios en su capacidad de percibir la información temporal que le brinda el adulto. El desarrollo de las capacidades perceptivas e intersubjetivas puede pensarse en términos dialécticos más que de determinación unidireccional (Español, 2010; Rivière, 2003). En tal sentido, el involucramiento del bebé con la *performance* adulta propiciaría el desarrollo de la percepción intersensorial, y el desarrollo de la percepción intersensorial impactaría en el involucramiento creciente del bebé durante la interacción. Bajo esta hipótesis, el estudio del desarrollo de la percepción intersensorial vinculada al dominio temporal — sincronía, duración, tasa, ritmo — es relevante no sólo para el conocimiento del desarrollo de la percepción infantil, sino también para la comprensión del desarrollo intersubjetivo.

La duración de la conducta adulta dirigida al bebé

El estudio de la duración de las conductas de los adultos en interacción con el bebé se ha realizado en diversos contextos. Kaye (1977) mostró que la variación de la duración del mecimiento que realizan las madres durante el amamantamiento, o durante la alimentación con biberón, prolonga o disminuye la pausa que realiza el bebé entre los ciclos de succión alimenticia.

Al comparar la conducta de un adulto en interacción con un bebé y en interacción con otro adulto, puede observarse que las frases que los adultos dirigen al bebé suelen durar la mitad de lo que duran las que dirigen a otro adulto, mientras que las pausas que separan a las frases duran casi el doble (Stern et al., 1977). Por otro lado, en relación con el habla dirigida al bebé, se observó que los adultos estiran la duración de las vocales (Stern, 1983) y que la duración promedio de la vocalización es mayor que la de la pausa (Jaffe et al., 2001; Stern, 1983).

Stern identificó un contexto de interacción particular al que denominó juego social temprano. Lo consideró el modo de interacción social por antonomasia, ya que el foco del juego es la interacción social misma. En este contexto, tanto el repertorio

conductual y perceptivo del bebé como la actuación adulta se encuentran al servicio de compartir la experiencia intersubjetiva. El juego social temprano se compone de dos episodios: uno de participación y otro de pausa. Los episodios de participación mutua son una secuencia de comportamientos sociales de diversa *duración*, mientras que los episodios de pausa consisten en un silencio comportamental (sonoro y kinético) cuya *duración* es más prolongada que la de los episodios de participación (Stern, 1983). A su vez, los episodios de participación se organizan a partir de frases y series. La frase es una emisión única de voz o ráfaga de movimientos que, por lo general, *dura* menos de un segundo; la serie se forma por varias frases de contenido o *duración* semejantes (Stern et al., 1977). La diferencia de las duraciones entre frases, series (que conforman el episodio interactivo) y los episodios de pausa ilustra el mundo temporal altamente ordenado que la madre proporciona a su hijo (Schaffer, 1985). Pareciera que las madres actúan en todo momento como si los bebés sólo pudieran asimilar trozos de información mucho más pequeños que los adultos y como si necesitasen más tiempo para procesar cada uno antes de recibir otro (Schaffer, 1989).

La organización temporal de la actuación adulta a través de la duración se observa también en la concordancia existente entre la duración de las vocalizaciones y la duración de los movimientos que realizan los adultos en el contexto del juego social temprano (Stern, 1974; Stern et al., 1977). A su vez, los microanálisis de las *performances* dirigidas a bebés muestran otros modos de organización basados en la duración. Las *performances* tienen la estructura de una frase construida mediante la repetición variada de motivos sonoros-kinéticos breves. Los motivos son diversos: hacer subir y bajar al bebé tomándolo desde su tronco, o la acción simple de sacar una manta que lo cubre, una secuencia de chistidos o una secuencia breve de palabras. Las variaciones suelen recaer en los componentes dinámicos y temporales del motivo (Español, 2014; I. Martínez et al., 2018). I. C. Martínez (2007) identificó frases sonoro-kinéticas en las que la variación radica en la duración de los motivos que se repiten. Español y Shifres (2015) identificaron frases en las que, mientras varios rasgos del movimiento cambian (su forma, por ejemplo), la duración de un rasgo del motivo kinético — la cualidad — se mantiene constante en su repetición. Las regularidades y variaciones de la duración del motivo muestran a la duración como un organizador de la actuación adulta en distintos niveles jerárquicos. Estas variaciones y regularidades de la duración ilustran el mundo ordenado, pero cambiante, complejo y altamente estimulante que los adultos proporcionan al bebé. Por ejemplo, una adulta elabora una breve frase frente a un bebé de siete meses, a partir de un motivo kinético: el movimiento ocasional de retirar la manta que abrigaba al bebé. El motivo entra en la forma repetición-variación (se repite tres veces con variaciones) y *coda* (o cierre). El movimiento de acercarse y alejarse del bebé en el plano sagital sucede con regularidad temporal en las tres repeticiones. El adulto acompaña las repeticiones del motivo con habla y vocalizaciones: ‘A ver ¿la sacamos?’, ‘Tititi’, ‘La saco la saco la saco la saco’. Las palabras y vocalizaciones de la *performance* presentan cualidades musicales que se combinan finamente con las cualidades del movimiento. Por ejemplo, en la tercera repetición, el deslizante y ligado movimiento de retirar la manta del cuerpo del bebé se combina con el sonido *staccato* (picado), de duración semejante, de la emisión ‘la saco la saco la saco’. Las palabras ‘la saco la saco la saco’ susurradas acompañan el último movimiento de sacar la manta que se extiende en un ritardando final, de modo que la secuencia termina con un *ritardando*

o alargamiento y un *diminuendo* (disminución de la sonoridad). Así también, en la coda, el movimiento de acercarse alejarse se acompaña con la vocalización ‘ju-uh-pate!’ que se produce justo en el momento de avanzar hacia el bebé; éste es el movimiento de avance más largo de todas las repeticiones del motivo y tiene la misma duración extendida que la vocalización ju-uh-pate!’ (Versión simplificada del análisis realizado por Español & Shifres, 2015).

Finalmente, cabe destacar el rol que desempeña la duración en otra conducta característica de los adultos: el entonamiento afectivo. En el entonamiento afectivo las madres responden a la conducta del bebé con un comportamiento diferente, pero en el cual conservan y reflejan la pauta temporal (duración, pulso y ritmo), la intensidad (absoluta o relativa) y/o la pauta espacial de la conducta del bebé (Stern, 1991). Por ejemplo, mientras una niña que intenta alcanzar un juguete emite un exuberante ‘jahhhh!’, su madre junta sus hombros y realiza un movimiento sacudido con su torso. ‘El baile *dura* lo mismo que el jahhhh! de la niña’ (Stern, 1991, p. 175, el destacado nos pertenece). ‘La intensidad y la *duración* de la voz de la niña forman pareja con los movimientos corporales de la madre’ (Stern, 1991, p. 177, el destacado nos pertenece). De nuevo, se pone en evidencia el rol constitutivo de la duración: el entonamiento del afecto de la conducta del bebé a través del afecto de la conducta que exhibe la madre se produce por intermediación de, entre otros elementos temporales, la duración. El entonamiento afectivo es una conducta de emparejamiento adulta que ocurre con mayor frecuencia que la imitación durante el juego social temprano y se encuentra presente desde el segundo mes de vida del bebé (Español, Bordoni, Carretero Pérez et al., 2018). El entonamiento afectivo ocurre en una ventana temporal de 1 segundo o menos y, preponderantemente, la madre empareja los movimientos del bebé con vocalizaciones. La conducta de orientación visual del bebé frente al entonamiento afectivo cambia de los seis a los 12 meses. Durante la segunda mitad del primer año de vida aumenta la frecuencia de miradas dirigidas al adulto cuando este lo imita y disminuye la frecuencia de miradas hacia el adulto cuando lo entona afectivamente (Bordoni et al., 2016). Para que el entonamiento afectivo impacte en la conducta del bebé, es preciso que éste *reconozca* — en el próximo apartado explicamos por qué utilizamos específicamente este término — la relación intersensorial existente entre la duración de su conducta y la duración de la conducta del adulto que se expresa en otra modalidad de información sensorial.

La percepción multisensorial

Existe un delicado amalgamamiento entre el repertorio conductual del adulto y el del bebé, una delicada y aún no bien explorada relación entre las características de la conducta del adulto y las capacidades perceptivas del bebé relativa no sólo a la duración, sino, a toda la información temporal. Schaffer mencionó ‘la impresionante precisión con la que, en circunstancias normales, la estimulación producida por el adulto casa con las capacidades del niño para absorber y responder a esa estimulación’ (1989, p. 48). En la misma línea, Stern afirmó que:

Nos resulta difícil comprender cómo puede reaccionar el lactante tal como lo hace y cómo comienza a comprender su universo social, de no admitir que él, o su sistema nervioso (si

así lo preferimos), no fuese capaz de realizar determinadas operaciones, bastante notables, de evaluación del tiempo. (1983, p. 145)

Cabe suponer, dado el carácter multimodal de la conducta del adulto, que la habilidad del bebé para percibir de manera unificada información simultánea es crucial para el contacto intersubjetivo. La percepción multisensorial es la capacidad de un organismo para percibir de manera unificada estímulos que proporcionan información simultáneamente a diferentes sistemas perceptivos (Bremner et al., 2012). Es posible percibir de manera unificada distintos tipos de información proveniente de un mismo estímulo: información amodal e información de modalidad única (Bahrack & Lickliter, 2012). La información amodal es aquella que se percibe de manera indistinta a través de los diferentes sistemas perceptivos, por ejemplo: duración, tasa, ritmo, en la dimensión temporal; y forma, textura y continuidad, en la dimensión espacial. La información de modalidad única es aquella que sólo puede ser percibida a través de un único sistema perceptivo, por ejemplo color, aroma, calor, timbre. Por lo general, los eventos multimodales se especifican mediante una combinación de atributos amodales y de modalidad única (Lewkowicz & Kraebel, 2004).

En las *performances* dirigidas a bebés pueden identificarse los dos tipos de información. Imaginemos, por ejemplo, un bebé de tres meses, recostado delante de su mamá. Supongamos que la mamá toca suavemente su mentón, mientras realiza un movimiento vibratorio, con su dedo índice al tiempo que vocaliza el sonido /trr/. Los breves y suaves toques y las vocalizaciones de la madre se producen reiteradamente de manera sincrónica y simultánea. En esta situación se encuentran presentes ambos tipos de información: por un lado, cada uno de estos elementos — el timbre de la voz, los movimientos de la mano y la suave presión del dedo de la mamá sobre el mentón del bebé — se percibe por un único sistema perceptivo — visión, audición, y tacto, respectivamente. Son claros ejemplos de información de modalidad única. Asimismo, cuando la mamá mueve suavemente su dedo índice sobre el mentón del bebé en forma sincrónica con la vocalización del sonido /trr/, la duración del sonido de la vocalización coincide con la duración del movimiento del dedo de la mamá sobre el mentón. La duración es un claro ejemplo de información amodal. La mamá está brindando simultáneamente la misma información sobre la duración a través de sus vocalizaciones y del toque sobre el mentón. El ritmo, la duración, la intensidad y la sincronía son casos de información amodal temporal. En las *performances* que dirige al bebé, el adulto brinda, además de información unimodal, una amplia gama de información amodal temporal.

En el ámbito de la percepción multisensorial se estudian predominantemente dos capacidades: la discriminación y el reconocimiento. La *discriminación* refiere a la habilidad para diferenciar entre dos o más elementos, eventos o tipos de información, mientras que el *reconocimiento* implica la capacidad para establecer relaciones (percepción intersensorial) entre la información percibida simultáneamente a través de dos o más sistemas perceptivos (Lewkowicz, 2000; Walker-Andrews, 1997). La discriminación se estudia utilizando la técnica de habituación controlada por el bebé (Horowitz et al., 1972). La lógica de la técnica es que, ante un estímulo novedoso, el sistema cognitivo va reduciendo progresivamente la atención que se presta al mismo. Ante la aparición de otra novedad estimular, aumenta la atención hacia el estímulo nuevo. El procedimiento consiste en la presentación de una serie de ensayos para

habituar a los participantes a un estímulo visual o audiovisual. A continuación, se presentan los ensayos de testeo. Si frente al nuevo estímulo el bebé exhibe un tiempo mayor de fijación de la mirada, se infiere que el bebé discrimina entre ambos estímulos. El reconocimiento se estudia a través de la ‘técnica de preferencia intersensorial’ (Spelke, 1976), que consiste en la presentación de, generalmente, dos estímulos visuales expuestos al mismo tiempo uno al lado del otro y simultáneamente con un estímulo auditivo, que es congruente sólo con uno de los dos estímulos visuales. Durante la participación de los bebés se calcula el tiempo de proporción total de la mirada dirigido hacia el estímulo visual congruente. Cuando esta proporción es significativamente superior (en términos estadísticos) a 0.50, se infiere que el bebé reconoce la relación que existe entre el estímulo que percibe por el sistema visual y el que percibe por el sistema auditivo. La capacidad para establecer relaciones (percepción intersensorial) entre la información disponible para diferentes sistemas perceptivos — por ejemplo, la equivalencia en la duración de la información vista y escuchada — se estudia mediante esta técnica. Si bien se reconoce el valor de la percepción intersensorial para el desarrollo social, emocional, cognitivo y lingüístico (Bahrick & Lickliter, 2012), existen pocos estudios que indaguen las diferentes capacidades de percepción intersensorial vinculadas a los aspectos temporales (sincronía, duración, tasa y ritmo) de la actuación adulta (M. Martínez, 2016). La pregunta general que orienta nuestros trabajos es qué percibe el bebé del conjunto de información amodal temporal que el adulto le brinda. En un trabajo reciente, M. Martínez et al. (2018) estudiaron el reconocimiento de relaciones intersensoriales basadas en el ritmo, en bebés de cuatro, siete y 10 meses. Puntualmente, indagaron la capacidad del bebé para reconocer la relación entre el mismo patrón rítmico presente en los movimientos de las manos (y en el movimiento de la zona orofacial movimiento realizado para producir el sonido vocálico) y en el sonido consonántico/ta/; todos ellos — sonidos y movimientos — ejecutados por una mujer que remeda la conducta prototípica de los adultos cuando interactúan con los bebés. Los resultados mostraron que el reconocimiento de la relación intersensorial basada en el ritmo, en un estímulo social que semeja una *performance*, emerge alrededor de los 10 meses. La pregunta específica que se pretende abordar aquí es cómo percibe el bebé las relaciones intersensoriales basadas en la duración proveniente de la información disponible en la *performance* que el adulto le dirige.

El desarrollo de la percepción de la duración

En este apartado se presentan los datos disponibles sobre las capacidades de percepción del bebé vinculadas a la duración. Existen varios estudios que indaguen el rol de la percepción de la duración respecto de la percepción del habla (Ainsworth, 1972; Eilers et al., 1984; Jusczyk et al., 1983; Kuhl & Meltzoff, 1982; Miller & Liberman, 1979). Todos estos estudios se realizaron con información unimodal. Por otro lado, Lewkowicz (2000) sostiene que existe evidencia indirecta que sugiere que a partir del primer mes de vida los bebés pueden discriminar diferencias de duración especificadas visual y auditivamente por debajo de los 100 ms. Morrongiello (1984) descubrió que los bebés de seis y 12 meses discriminaban entre series — en el sentido en el que Stern (1983) usa el término — conformadas por tres eventos de sonido y dos de silencio (sonido-silenciosonido-silenciosonido) cuando estas diferían en las duraciones de los eventos. Las

diferencias mínimas que discriminaron los bebés fueron de 40 y 60 ms. Es decir, discriminaban series de tres sonidos de 200 ms y dos silencios de 200 ms (200-200-200) de series de tres sonidos de 160 ms y dos silencios de 140 ms (160-140-160-140-160). Por otra parte, Morrongiello y Trehub (1987) informaron que los bebés de seis meses sólo discriminan cambios en la duración cuando estos son mayores a 20 ms. Estos estudios también son sobre información unimodal.

Hace tiempo, Lewkowicz (1986, 2000) indicó que los estudios sobre la capacidad de los bebés para establecer relaciones intersensoriales basadas en la duración eran notablemente escasos. En los últimos tiempos el panorama no parece haber cambiado: si se revisa el último *Handbook sobre Multisensory Development* (Bremner et al., 2012) no se encuentran estudios específicos sobre la precepción intersensorial de la duración.

Lewkowicz (1986) expuso a bebés tres, seis y ocho meses de edad a dos tipos de ensayos (uno donde la diferencia de la duración era de 800 ms y otro donde la diferencia era de 1,200 ms) de preferencia intersensorial. Durante los primeros seis ensayos los bebés observaban simultáneamente dos tableros lumínicos (ubicados una al lado del otro) que destellaban con idéntica tasa, pero con duraciones diferentes. Durante los 12 ensayos siguientes los bebés observaban nuevamente los mismos tableros pero esta vez se presentaba además, de manera sincrónica a uno de ellos, un sonido cuyo inicio y fin, es decir su duración, era idéntica al destello de uno de los dos tableros luminosos. En los respectivos ensayos los tableros destellaban a tres posibles duraciones (400, 800 y 1,600 ms). Cabe destacar que la variable duración se operativiza por medio de la sincronía entre estímulos, haciendo que coincidan ambos en su inicio y final para duraciones iguales y solo en el inicio para duraciones distintas de ambos estímulos. Los datos obtenidos permitieron afirmar que los bebés de seis y ocho meses reconocen relaciones intersensoriales basadas en la duración cuando se presentan los pares de 400–1,600 ms y los pares de 800–1,600 ms. Asimismo, cuando se desplaza temporalmente el inicio del sonido respecto del inicio del destello lumínico, los bebés no pueden reconocer relaciones intersensoriales basadas en la duración. En tal sentido, un factor clave para que se produzca el reconocimiento es la sincronía del inicio y fin de la duración de la información presente para los diferentes sistemas sensoriales. Esto indicaría que para los bebés no es posible el reconocimiento, en términos absolutos, de relaciones basadas en la duración. Es decir, no pueden reconocer la equivalencia de la duración de la información cuando esta no es sincrónica, cuando no coincide su inicio y su fin. Consecuentemente, Lewkowicz (2000) sostiene que los bebés requieren señales correspondientes a la sincronía y a la duración para reconocer la equivalencia intersensorial basada en la duración, la duración por sí sola no es suficiente.

En resumen, los estudios reseñados muestran que los bebés de seis y ocho meses pueden reconocer relaciones intersensoriales basadas en la duración en estímulos simples (bimodales). No obstante, quedan ciertamente muchas cuestiones por indagar. Cabe preguntarse si los bebés reconocen relaciones intersensoriales basadas en la duración en estímulos sociales, si ese reconocimiento ocurre antes de los seis meses, si pueden observarse cambios en el reconocimiento a lo largo de un período de tiempo relevante. El objetivo de este trabajo es estudiar experimentalmente la habilidad de bebés de cuatro, siete y 10 meses para reconocer relaciones intersensoriales basadas en la duración en información proveniente de las acciones que realiza una persona remedando una *performance* artística dirigida al bebé. El interés por usar un estímulo

social proviene de las implicaciones que el estudio pueda tener para la comprensión del desarrollo intersubjetivo temprano.

Método

Participantes

Cuarenta y ocho bebés, divididos en tres grupos, con 16 participantes de cuatro meses ($M_{\text{edad}} = 125$ días, $SD =$ cuatro días, 121–131; ocho mujeres), 16 participantes de siete meses ($M_{\text{edad}} = 220$ días, $SD =$ seis días, 210–228; siete mujeres), y 16 participantes de 10 meses ($M_{\text{edad}} = 307$ días, $SD =$ ocho días, 283–316; nueve mujeres). La participación de los bebés quedó condicionada a los siguientes requisitos: (i) embarazo a término, (ii) una puntuación igual o mayor a ocho en la prueba de APGAR, (iii) que el médico de cabecera no haya sospechado o detectado algún problema de salud grave durante el primer mes de vida, y (iv) la ausencia de enfermedades, alteraciones o trastornos del desarrollo antes o durante el período en el cual el bebé se encuentra participando en el experimento.

Los datos de tres bebés fueron excluidos debido a que no miraron a los dos monitores empleados en el estudio para presentar los estímulos visuales durante los ensayos ($n = 3$ a los cuatro meses). Antes de participar los padres firmaron un consentimiento informado.

Diseño

El estudio presenta un diseño evolutivo transversal ex post facto con dos variables independientes: (1) la relación intersensorial entre la duración de las frases de movimiento de las manos (componente visual) y del sonido emitido (componente auditivo), como factor intrasujeto, con dos niveles: igual duración (o duración congruente) y desigual duración (o duración incongruente) (2) la edad de los participantes, con los niveles cuatro, siete y 10 meses como factor intersujeto.

Aparatos

Para la presentación de los estímulos visuales se utilizaron dos monitores de computadora de 19 pulgadas (Samsung SyncMaster E1920) ubicados uno al lado del otro separados a una distancia de 18 centímetros. Los monitores se encontraban rodeados por cortinas azules con el objetivo de no distraer la atención de los bebés. Entre los monitores se ubicó un dispositivo LED (rojo y azul) que el investigador manipuló con el objeto de atraer la atención visual de los participantes para que todos comenzaran los ensayos mirando hacia el centro de ambos monitores. La proyección de los estímulos visuales en los monitores se realizó a través de notebooks (Asus Eee PC1015PX), operadas con el mismo control remoto. El estímulo auditivo se presentó con dos altoparlantes (Edifier R18USB) ubicados entre ambos monitores justo detrás del dispositivo LED. Para cada uno de los ensayos en los que participaron los bebés se obtuvo un registro videograbado de su cara con una webcam (HP Deluxe Webcam KQ246AA)

ubicada sobre el dispositivo LED con el objeto de medir los tiempos de fijación de la mirada sobre los estímulos visuales.

tímulos

Se diseñaron dos estímulos audiovisuales (películas) de acuerdo a la lógica de la *performance* artística dirigida al bebé (Español & Shifres, 2015). Ambos estímulos están compuestos por una mujer vista de frente delante de un fondo azul que emite el sonido /trrr/ (creado con los sonidos consonántico pico dental/t/y consonántico linguoalveolar/r/) mientras mueve los dedos de ambas manos (gesto de hacer cosquillas). Las manos se encuentran ubicadas a la altura de la cara con las palmas hacia el frente. En forma sincrónica a la producción del sonido /trrr/ la mujer mueve los dedos de ambas manos. El movimiento de las manos inicia y finaliza con los puños cerrados. La mujer realiza una serie iterativa de frases (Stern et al., 1977) de sonido/movimiento de diversa duración seguida de una pausa (episodios de silencio/quietud). Las frases conforman dos series (Stern et al., 1977), una serie por cada estímulo audiovisual: serie A y B, cuya duración es de 20 segundos semejando la actuación adulta prototípica en interacción natural (Español, 2014). La variable independiente duración se operativizó en función de la distancia temporal que separa el inicio y fin sincrónicos de las frases de sonido y movimiento, realizadas por la mujer.

La alternancia de frases multimodales de sonido-movimiento y las pausas (episodios de silencio-quietud) fue distribuida de modo tal que fuese imposible extraer de la serie una regularidad rítmica. Asimismo, se distribuyó de manera balanceada la proporción de frases de sonido-movimiento y de pausas. La proporción de sonido-movimiento para la primera serie es de 0.65 y de 0.63 para la segunda. La duración total de ambos estímulos audiovisuales es de 90 segundos. En cada uno de ellos, la serie se repite cuatro veces y media. Cuando empieza la proyección de las películas, el inicio de cada uno de los patrones es sincrónico. Durante la proyección de ambos estímulos, la serie inicia de manera sincrónica en cinco oportunidades (Figura 1).

Si bien para el estudio de las relaciones intersensoriales se utilizan estímulos que presentan redundancia intersensorial (entre sonido y movimiento, por ejemplo), el estímulo diseñado para el presente estudio presenta además redundancia intrasensorial (entre sonido: vocalización/tr/, movimiento: zona orofacial, y movimiento: gesto de hacer cosquillas de los dedos de las manos). Es decir, hay sincronía entre el inicio y el fin de las vocalizaciones y entre el inicio y fin de ambos movimientos. En la información visual de ambos estímulos existe sincronía entre el inicio y finalización del movimiento de las manos y el de la zona orofacial. Recientemente Bahrck et al. (2015) demostraron que la redundancia intrasensorial atrae la atención del bebé hacia la información amodal, de la misma manera que la redundancia intersensorial. En tal sentido, la redundancia intrasensorial añade validez ecológica (asemejando el estímulo a una *performance* artística dirigida al bebé) sin generar efectos indeseables con relación al reconocimiento de las relaciones intersensoriales basadas, en este caso, en la duración.

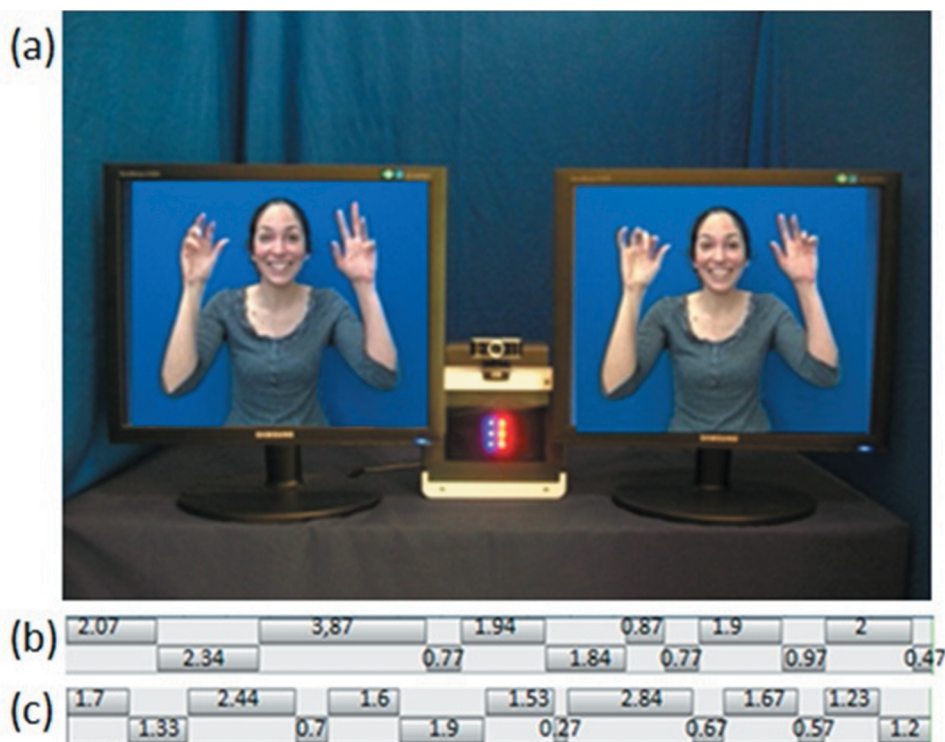


Figura 1. (a) Disposición de los materiales en el dispositivo experimental. (b) y (c) Duración de las frases (en ms) de sonido/movimiento (fila superior) y silencio/quietud (fila inferior) que componen cada uno de los dos episodios utilizados como estímulos.

Procedimiento

Se utilizó la técnica de Preferencia Intersensorial (Spelke, 1976), que consiste en la presentación en forma simultánea de dos estímulos visuales y uno auditivo sincrónico (el patrón de duración inicia y finaliza el mismo tiempo en la información visual y audible) y congruente con uno de los dos estímulos visuales. Los bebés se sentaron sobre el regazo de la mamá o del papá delante de ambos monitores a una distancia de 60 centímetros. Los padres fueron instruidos previamente sobre cómo sostener a su hijo y no interferir durante la sesión. Se los invitó a que se dispusieran de manera cómoda, con la espalda recostada en el respaldo, en la silla dispuesta a tal fin y que sostuvieran a su hijo sobre su regazo confortablemente. Se les solicitó evitar interactuar con sus bebés (no dirigir gestos, verbalizaciones o expresiones faciales) durante la presentación de los estímulos. Asimismo, se les informó que, si en algún momento sentían que el bebé pudiera estar incómodo durante el procedimiento, podían finalizarlo levantándose de la silla.

Cada bebé participó en dos ensayos (de 90 segundos de duración cada uno) en los cuales observaron en forma simultánea ambos estímulos visuales mientras escuchaban el sonido (estímulo auditivo) congruente con uno de ellos. En cada ensayo se mantuvo constante la posición lateral del estímulo visual, y de uno a otro ensayo se varió el

estímulo auditivo congruente. La posición lateral y la presentación del patrón auditivo se contrabalanceó entre todos los participantes. Durante el procedimiento la cámara de video ubicada entre ambos monitores registró la cara del bebé a fin de calcular, posteriormente, el tiempo de fijación de la mirada sobre los estímulos visuales proyectados en los monitores. La medición del tiempo de fijación de la mirada se realizó utilizando el software de video-notación ANVIL (Kipp, 2008).

Resultados

Con el objetivo de determinar a qué edad los bebés reconocen relaciones intersensoriales basadas en la duración, se calculó la Proporción del Tiempo Total de Mirada (PTTM). La PTTM del bebé es la medida más utilizada para corroborar el reconocimiento de relaciones intersensoriales (Flom et al., 2009). La misma se calcula dividiendo el tiempo total de la mirada dirigida hacia el monitor que exhibe el estímulo visual congruente con el sonoro por el tiempo total de mirada dirigida hacia ambos monitores. A fin de establecer cuándo los bebés reconocen la relación intersensorial basada en la duración, se contrastó la media de la PTTM de cada grupo de edad agrupando los datos correspondientes a ambos ensayos (E1 + E2) contra el valor 0.50. Los resultados indican que los bebés de las tres edades estudiadas — cuatro meses ($M = 0.57$, $DS = 0.10$) $t(15) = 2.759$, $p < .05$, siete meses ($M = 0.57$, $DS = 0.06$) $t(15) = 4.582$, $p < .05$, y 10 meses ($M = 0.54$, $DS = 0.07$) $t(15) = 2.239$, $p < .05$ — reconocieron la relación intersensorial basada en la duración. Se realizó un segundo análisis con el fin de averiguar si existía diferencia de medias entre los grupos en función de la edad. Para ello se realizó un ANOVA de un factor agrupando los datos correspondientes a ambos ensayos (E1 + E2). El resultado del análisis no reveló un efecto significativo del factor edad ($p > .05$). Estos dos primeros análisis permiten afirmar que los bebés de cuatro, siete y 10 meses reconocen relaciones intersensoriales basadas en la duración (Figura 2).

Una segunda serie de análisis mostró que los bebés no presentaron preferencia respecto de las series (A o B) ni dominancia lateral visual durante cada uno de los ensayos. Para evaluar si existía por parte de los bebés preferencia (en términos de *pregnancia estimular*) por alguna de las series se realizó un ANOVA de medidas repetidas (con la serie A y B como factor intra-sujeto y la edad como factor inter-sujeto) sobre el tiempo total de mirada que los bebés dirigieron a cada uno de ellos, con independencia del sonido que estuvieran escuchando. No se hallaron efectos principales para estos factores. El resultado del análisis reveló que no existen interacciones significativas entre los factores serie (A y B) y edad (todos los valores p fueron $> .05$). De manera similar se procedió con la dominancia lateral visual. Se realizó un ANOVA de medidas repetidas (con el monitor, izquierdo o derecho, como factor intra-sujeto y la edad como factor inter-sujeto) sobre el tiempo total de mirada que los bebés dirigieron hacia el monitor izquierdo y hacia el derecho, independientemente del sonido que estuvieran escuchando. No se hallaron efectos principales para estos factores. El resultado de este análisis no reveló interacciones significativas entre los factores (monitor izquierdo, monitor derecho) y edad (todos los valores p fueron $> .05$). Hasta aquí, nuestros resultados indican que los bebés de cuatro, siete y 10 meses reconocen relaciones intersensoriales basadas en la duración.

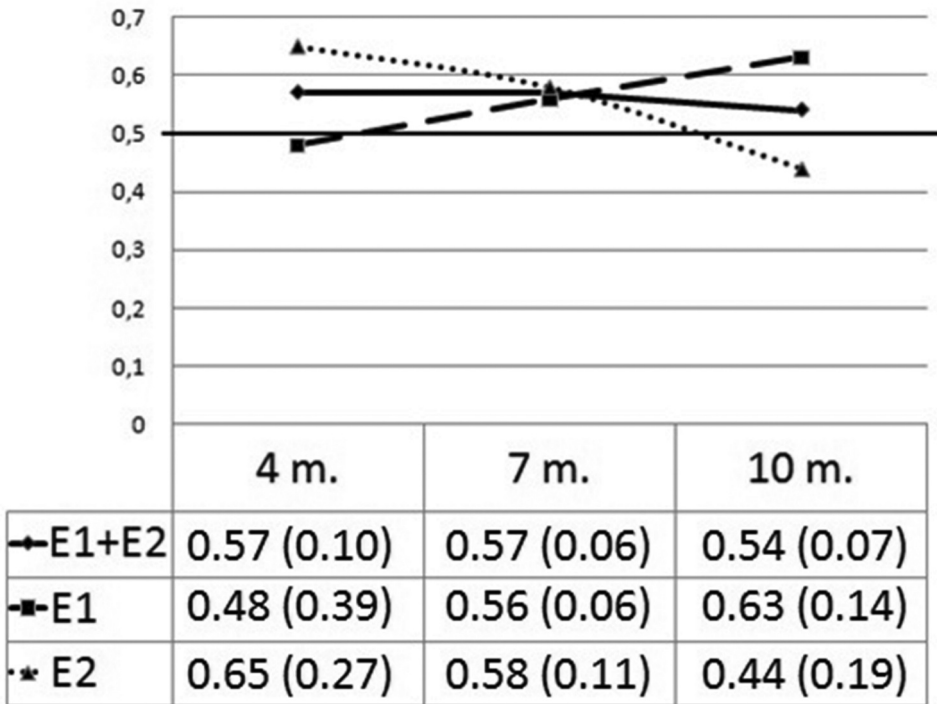


Figura 2. Medias de la Proporción del Tiempo Total de Mirada (y su correspondiente desviación estándar) orientada hacia el estímulo visual congruente con el estímulo sonoro a los cuatro, siete y 10 meses.

En virtud de que los bebés de las tres edades reconocen la relación intersensorial estudiada, se realizó una tercera serie de análisis para indagar si existe algún otro efecto de la edad con relación a la capacidad de reconocimiento de relaciones intersensoriales.

Específicamente, se indagó si el reconocimiento de la relación intersensorial basada en la duración se manifiesta de la misma manera en las diferentes edades. E. Gibson (1969) postuló que el desarrollo perceptivo implica, además de la adquisición de una habilidad particular, un incremento en la efectividad del organismo para extraer información. Consecuentemente, en primer lugar, se procedió a analizar de manera independiente los datos correspondientes a cada uno de los ensayos en que participaron los bebés. Se contrastó la media de la PTTM de cada grupo de edad correspondiente a cada uno de los ensayos (E1 y E2) por separado contra el valor 0.50. Los resultados indicaron que a los cuatro meses sólo en el E2 los bebés reconocen la relación intersensorial basada en la duración ($M = 0.65$, $DS = 0.27$) $t(15) = 2.184$, $p < .05$. A los siete meses en ambos ensayos, E1 ($M = 0.56$, $DS = 0.6$) $t(15) = 3.967$, $p < .05$ y E2 ($M = 0.58$, $DS = 0.11$) $t(15) = 2.678$, $p < .05$, los bebés reconocen la relación intersensorial basada en la duración. Y, a los 10 meses sólo en el E1 ($M = 0.63$, $DS = 0.14$) $t(15) = 3.468$, $p < .05$, los bebés reconocen la relación intersensorial basada en la duración. A continuación se realizó un ANOVA de dos factores (edad y número de ensayo: E1 y E2). No se hallaron efectos principales de los factores edad y número de ensayo sobre la PTTM ($p > .05$). Por el contrario, se halló un efecto de

interacción entre los factores edad y número de ensayo ($F(2, 33) = 8.098$; $p < .05$, *partial eta squared* = .15).

El análisis parece indicar que los bebés de cuatro meses necesitan más tiempo para identificar la relación intersensorial (el reconocimiento ocurre en el E2) que los bebés de siete (en los que el reconocimiento ocurre en el E1 y E2). A los 10 meses, parece observarse, además, un efecto fatiga (el reconocimiento ocurre solo en el E1).

Según E. Gibson (1969), el incremento en la habilidad para extraer información depende del desarrollo del control voluntario de la actividad exploratoria. Ruff y Rothbart (1996) consideran que el incremento de la alternancia de mirada entre dos estímulos es un indicador del incremento de la actividad exploratoria. Con el objetivo de indagar el incremento en la efectividad de los bebés para reconocer la relación intersensorial basada en la duración, realizamos dos análisis. En el primero, se indagó el efecto de la variable edad sobre la cantidad de alternancias de miradas entre los monitores que presentaban los estímulos contemplando ambos ensayos de manera conjunta. En el segundo, se indagó la diferencia entre la cantidad de alternancia entre monitores en los ensayos E1 y E2. Para el primer análisis, se realizó un ANOVA de un factor el cual mostró un efecto significativo del factor edad ($F(2, 45) = 5.776$; $p < .05$, *partial eta squared* = .51) sobre la cantidad de alternancias de miradas.

El contraste posthoc Tukey *b* mostró que la media del total de alternancias de miradas fue diferente a los cuatro meses ($M = 13.31$, $SD = 7.48$), a los siete meses ($M = 20.91$, $SD = 8.65$) y a los 10 meses ($M = 20.80$, $SD = 5.16$). Estos datos indican que a los siete y 10 meses los bebés realizan más cantidad de alternancia de miradas entre los monitores que a los cuatro meses.

Respecto a la diferencia entre ensayos, el análisis mostró que a los cuatro meses existe una diferencia estadísticamente significativa, con respecto a las otras dos edades, en la cantidad de alternancia de miradas durante el segundo ensayo ($M = 7.81$, $SD = 4.19$) en comparación con el primer ensayo ($M = 5.5$, $SD = 4.12$) $t(15) = 2.569$, $p < .05$, $d = 0.56$. Ha de recordarse que a los cuatro meses los bebés sólo reconocen la relación intersensorial durante el segundo ensayo. A los siete meses, no existe diferencia estadísticamente significativa en la cantidad de alternancia de mirada durante el primer ensayo ($M = 10.59$, $SD = 4.17$) y el segundo ensayo ($M = 10.31$, $SD = 5.11$) $t(15) = 0.324$, $p > .05$. Ha de recordarse que a los siete meses los bebés realizan la correspondencia intersensorial en ambos ensayos. Finalmente, a los 10 meses, existe una diferencia estadísticamente significativa en la cantidad de alternancia de miradas durante el primer ensayo ($M = 11.13$, $SD = 3.48$) en comparación con el segundo ensayo ($M = 9.67$, $SD = 2.44$) $t(15) = 1.901$ $p < .05$, $d = 0.49$. Ha de recordarse que a los 10 meses los bebés sólo reconocen la relación intersensorial durante el primer ensayo.

Discusión

Los resultados de nuestro estudio indican que los bebés de cuatro, siete y 10 meses reconocen relaciones intersensoriales basadas en la duración proveniente de un estímulo que remeda la *performance* multisensorial que los adultos suelen brindar a los bebés durante los intercambios intersubjetivos. Los datos previos existentes indicaban que los bebés de seis y ocho meses, no así los de tres meses, reconocen relaciones intersensoriales basadas en diferentes duraciones (800 y 1,200 ms) en estímulos

compuestos por luz y sonido. Nuestros resultados permiten ampliar el conocimiento sobre la capacidad de los bebés para reconocer relaciones intersensoriales basadas en la duración. En primer lugar, permiten afirmar que los bebés reconocen la relación intersensorial basada en la duración en información proveniente de las personas. En segundo lugar, indican que los bebés de cuatro meses reconocen relaciones basadas en la duración. En tercer lugar, nuestros resultados dan cuenta de que la habilidad de los bebés para el reconocimiento no se expresa de manera idéntica entre los cuatro y 10 meses.

Este último dato resulta relevante en función de una de las críticas más frecuente a los trabajos que estudian el desarrollo de diferentes habilidades perceptivas. La crítica sostiene que la mayor parte de la investigación sólo se focaliza en identificar a qué edad, cuanto más temprano mejor, los bebés muestran alguna capacidad perceptiva sin brindar mayor información sobre el desarrollo de la misma (Mariscal et al., 2012). Nuestros datos muestran cómo se incrementa la efectividad de la habilidad de los bebés para reconocer las relaciones intersensoriales basadas en la duración. La noción de incremento de una habilidad responde a la tradición desarrollada por J. Gibson (1966) y E. Gibson (1969; E. Gibson & Pick, 2000). El incremento se evidencia a partir de la reducción del tiempo necesario para extraer la información (E. Gibson, 1969). Nuestros datos muestran que los bebés de cuatro meses sólo detectan la relación durante el segundo ensayo, mientras que a los siete meses la relación se detecta en ambos ensayos y, a los 10 meses la detectan sólo durante el primer ensayo.

Varios estudios previos muestran que a partir de los cinco meses los bebés suelen exhibir la habilidad de reconocimiento de diferentes relaciones intersensoriales de manera diferencial durante los distintos ensayos en los que participan (Schmuckler & Jewell, 2007; Soken & Pick, 1992; Walker, 1982). Por lo general, cuando los bebés dan cuenta del reconocimiento durante el o los primeros ensayos, y no durante el o los últimos, la diferencia se atribuye a diversos factores: aburrimiento, fatiga, atención a otras propiedades de eventos, etc. (Vaillant-Molina et al., 2013). En nuestro caso esto ocurrió con los bebés de 10 meses, por lo que es posible que a esta edad el estímulo ya no les resulte tan interesante. Cabe mencionar que este patrón también fue atribuido a otros factores: el estrechamiento perceptivo (Kubicek et al., 2014) o la transgresión a las expectativas del bebé (Imafuku et al., 2019). Sin embargo, queda a futuro averiguar si el estrechamiento perceptivo y la transgresión a las expectativas juegan algún rol en el desarrollo de la percepción de la información amodal temporal.

Por otra parte, los bebés de siete y 10 meses realizaron una mayor cantidad de alternancias de miradas que los bebés de cuatro meses durante ambos ensayos. Asimismo, en las tres edades estudiadas los bebés realizaron la mayor cantidad de alternancia de miradas en aquellos ensayos en los cuales reconocían la relación intersensorial: a los cuatro meses los bebés realizaron la mayor cantidad de alternancia de mirada durante el segundo ensayo, los bebés de siete meses realizaron la misma cantidad de alternancia en ambos ensayos y los bebés de 10 meses realizaron mayor cantidad de alternancia de miradas durante el primer ensayo. Tanto el incremento como la merma de la cantidad de alternancias de mirada suele interpretarse como la expresión del desarrollo y control voluntario de la atención visual (Ruff & Rothbart, 1996). En su conjunto, los datos expuestos nos permiten identificar un patrón de

cambio, entre los cuatro y los 10 meses, en el reconocimiento de relaciones intersensoriales basadas en la duración.

En la introducción hipotetizamos sobre el posible vínculo dialéctico entre el desarrollo perceptivo e intersubjetivo durante el primer año de vida, propusimos que algunos de los cambios que se observan en la conducta del bebé pueden deberse al desarrollo de sus capacidades perceptivas vinculadas a la duración, y enumeramos una serie de estudios en los cuales se resalta la importancia de la duración para el encuentro intersubjetivo del adulto y del bebé.

La relación entre el desarrollo perceptivo e intersubjetivo puede observarse en distintos contextos. Stern (1991) lo ejemplifica en el contexto interactivo en el que el adulto realiza una *performance*: ‘Si la duración de la caricia perfilada y las pausas entre caricias tienen la misma duración absoluta y relativa que las pautas de vocalización y pausa, el infante experimentará perfiles de activación similares’ (p. 81). La semejanza en la duración le permitirá al bebé aunar el fenómeno, reconocer al otro como unidad. Porque todos los estímulos que emanan de una persona tienen, de acuerdo con Stern (1991), una coherencia temporal (de duración, de ritmo). Y su singularidad (los rasgos particulares de los sonidos y movimientos) permitirá el reconocimiento y la distinción entre las personas. Porque cada persona tiene un estilo particular de acariciar y hablar o, en términos más técnicos, cada persona expresa singularmente las formas de la vitalidad (Stern, 2010). Asimismo, la relación entre el desarrollo perceptivo e intersubjetivo puede observarse en la emergencia de la habilidad del bebé para alternar visualmente entre sus propias acciones y las acciones de los demás (habilidad básica para la referencia conjunta). El reconocimiento, por parte del bebé, del mismo patrón temporal de ritmo en su conducta y en la del adulto sustenta la emergencia de la alternancia de mirada. Por ejemplo, cuando adulto y bebé juegan a golpear con el mismo patrón rítmico una mesa (actividad lograda por el ajuste del adulto al patrón rítmico realizado por el bebé), el bebé puede entonces alternar la mirada entre su acción y la del adulto (Fogel & DeKoeper-Laros, 2007).

En tal sentido, nos parece que los datos obtenidos alientan a continuar trabajando sobre el vínculo dialéctico planteado de modo especulativo por Rivière (2003) y Español (2010). Para ello debería enlazarse la emergencia de diferentes capacidades intersubjetivas puntualmente con la detección, discriminación y reconocimiento de patrones multimodales de información amodal temporal. Por ejemplo, M. Martínez (2019) enlaza el desarrollo de la percepción multisensorial (detección, discriminación y reconocimiento) de patrones rítmicos con cambios en la conducta social del bebé durante el primer año de vida. Por otro lado, resulta indispensable construir propuestas teóricas que expliquen la ontogénesis de la percepción multisensorial de información amodal temporal en función de la participación del bebé en intercambios intersubjetivos con el adulto. M. Martínez (2021) esboza una primera aproximación a la ontogénesis de la percepción multisensorial de información amodal temporal desde una perspectiva Vygotskiana identificando posibles sistemas de mediación social que permiten, a partir de la reorganización de sistemas naturales y combinados, la génesis de la actividad exploratoria autónoma del bebé necesaria para extraer invariantes perceptivas.

Por otra parte, los datos aportados en este trabajo, conjuntamente con los informados en M. Martínez et al. (2018) sobre el reconocimiento de relaciones intersensoriales basadas en el ritmo, constituyen un aporte relevante al conocimiento del desarrollo de

la percepción intersensorial de la información amodal temporal. Lewkowicz (2000) propuso una secuencia progresiva para el desarrollo de la percepción de información temporal, según la cual las capacidades de percibir la sincronía, la duración, la tasa y finalmente el ritmo emergerían sucesivamente. Nuestros datos apoyan en términos generales esta secuencia. Sin embargo, la propuesta de Lewkowicz no contempla una serie de cuestiones que consideramos relevantes. Por ejemplo, si los bebés reconocen relaciones intersensoriales basadas en la duración en información proveniente de las personas, si el reconocimiento de dichas relaciones antecede al reconocimientos de relaciones intersensoriales basadas en el ritmo (en información proveniente de personas), y si existen cambios evolutivos vinculados con la efectividad para reconocer relaciones intersensoriales basadas en la duración durante el primer año de vida. En primer lugar, los datos informados en el presente trabajo permiten sostener que a los cuatro, siete y 10 meses los bebés reconocen relaciones intersensoriales basadas en la duración en información proveniente de personas. En segundo lugar, los datos aquí informados contemplados de manera conjunta con los datos informados por M. Martínez et al. (2018) permiten afirmar que el reconocimiento de relaciones intersensoriales basadas en la duración precede ontogenéticamente al reconocimiento de las relaciones intersensoriales basadas en el ritmo en la información proveniente de personas. Finalmente, el presente trabajo aporta evidencia empírica que permite sostener que durante el período comprendido entre los cuatro y los 10 meses los bebés incrementan su efectividad para reconocer relaciones intersensoriales basadas en la duración.

Note

1. In this paper, the original transcription of the non-semantic sounds in Spanish made by the adult participation in the research were left intact. Nota

Disclosure statement *Conflicto de intereses*

No potential conflict of interest was reported by the authors. /Los autores no han referido ningún potencial conflicto de interés en relación con este artículo.

References / Referencias

- Ainsworth, W. A. (1972). Duration as a cue in the recognition of synthetic vowels. *The Journal of the Acoustical Society of America*, 51(2B), 648–651. <http://dx.doi.org/10.1121/1.1912889>
- Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 183–205). Oxford University Press.
- Bahrick, L. E., Lickliter, R., Castellanos, I., & Todd, J. T. (2015). Intrasensory redundancy facilitates infant detection of tempo: Extending predictions of the intersensory redundancy hypothesis. *Infancy*, 20(4), 377–404. <https://doi.org/10.1111/infa.12081>
- Bordoni, M., Español, S., & De Grande, P. (2016). La incidencia del entonamiento afectivo y la imitación en el involucramiento visual-social temprano. *Avances en Psicología Latinoamericana*, 34(3), 487–503. <https://doi.org/10.12804/apl34.3.2016.04>
- Bremner, A. J., Lewkowicz, D. J., & Spence, C. (2012). *Multisensory development*. Oxford University Press.

- Eilers, R. E., Bull, D. H., Oller, D. K., & Lewis, D. C. (1984). The discrimination of vowel duration by infants. *The Journal of the Acoustical Society of America*, 75(4), 1213–1218. <http://10.1121/1.390773>
- Español, S. (2010). Interazione precoce. Una prospettiva vygotskijana a partire dagli schemi di Piaget, 25 anni dopo. *Metis*, 17(1), 67–91.
- Español, S. (2014). La forma repetición-variación: Una estrategia para la reciprocidad. In S. Español (Comp.), *Psicología de la música y del desarrollo. Una exploración interdisciplinario sobre la musicalidad humana* (pp. 157–192). Paidós.
- Español, S., Bordoni, M., Carretero Pérez, S., Martínez, M., & Camarasa, R. (2018). Imitación y entonamiento afectivo en el juego social temprano. *Interdisciplinaria: Revista de Psicología y Afines*, 35(2), 291–305. <http://www.ciipme-conicet.gov.ar/ojs/index.php?journal=interdisciplinaria&page=article&op=view&path%5B%5D=interd.2020.37.1.5&path%5B%5D=html>
- Español, S., & Shifres, F. (2015). The artistic infant directed performance: A microanalysis of the adult's movements and sounds. *Integrative Psychological and Behavioral Science*, 49(3), 371–397. <https://dx.doi.org/10.1007/s12124-015-9308-4>
- Fogel, R., Whipple, A., & DeKoeyer-Laros, I. (2007). The developmental transition to secondary intersubjectivity in the second half year: A microgenetic case study. *Journal of Developmental Processes*, 2(2), 63–90.
- Flom, R., Whipple, H., & Hyde, D. (2009). Infants' intermodal perception of canine (*Canis familiaris*) facial expressions and vocalizations. *Developmental Psychology*, 45(4), 1143–1150. <https://dx.doi.org/10.1037/a0015367>
- Gibson, E. (1969). *Principles of perceptual learning and development*. Appleton.
- Gibson, E., & Pick, A. (2000). *An ecological approach to perceptual learning and development*. Oxford University Press.
- Gibson, J. (1966). *The senses considered as perceptual system*. Houghton-Mifflin.
- Horowitz, F., Paden, L., Bahana, K., & Self, P. (1972). An infant-control procedure for studying infant visual fixations. *Developmental Psychology*, 7(1), 90. <http://dx.doi.org/10.1037/h0032855>
- Huble, P., & Trevarthen, C. (1979). Sharing a task in infancy. In I. C. Uzgiris (Ed.), *Social interaction and communication during infancy* (pp. 57–80). Jossey Bass.
- Imafuku, M., Kawai, M., Niwa, F., Shinya, Y., & Myowa, M. (2019). Audiovisual speech perception and language acquisition in preterm infants: A longitudinal study. *Early Human Development*, 128, 93–100. <https://doi.org/10.1016/j.earlhumdev.2018.11.001>
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, 66(2), i–132. <https://www.jstor.org/stable/3181589?seq=1>
- Jusczyk, P. W., Pisoni, D. B., Reed, M. A., Fernald, A., & Myers, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. *Science*, 222(4620), 175–177. <http://10.1126/science.6623067>
- Kaye, K. (1977). Toward the origin of dialogue. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction: Proceedings of the Loch Lomond Symposium* (pp. 89–117). Academic Press.
- Kaye, K. (1986). *La vida mental y social del bebé. De cómo los padres crean personas*. Paidós.
- Kipp, M. (2008). Spatiotemporal Coding in ANVIL. In *Proceedings of the 6th international conference on language resources and evaluation (LREC-08)*.
- Kubicek, C., De Boisferon, A. H., Dupierrix, E., Pascalis, O., Loevenbruck, H., Gervain, J., & Schwarzer, G. (2014). Cross-modal matching of audio-visual German and French fluent speech in infancy. *PLoS One*, 9(2), e89275. <https://doi.org/10.1371/journal.pone.0089275>
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218(4577), 1138–1141. <http://dx.doi.org/10.1126/science.7146899>
- Lewkowicz, D. J. (1986). Developmental changes in infants' bisensory response to synchronous durations. *Infant Behavior and Development*, 9(3), 335–353. [http://dx.doi.org/10.1016/0163-6383\(86\)90008-1](http://dx.doi.org/10.1016/0163-6383(86)90008-1)

- Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, 126(2), 281–308. <https://dx.doi.org/10.1037/0033-2909.126.2.281>
- Lewkowicz, D. J., & Kraebel, K. (2004). The value of multimodal redundancy in the development of intersensory perception. In G. Calvert, C. Spence, & B. Stein (Eds.), *Handbook of multi-sensory processing* (pp. 655–678). MIT Press.
- Mariscal, S., Casla, M., Rujas, I., & Aguado-Orea, J. (2012). Los métodos basados en la duración de la mirada: ¿una ventana a la cognición temprana? *Estudios de Psicología*, 33(3), 277–292. <https://doi.org/10.1174/021093912803758219>
- Martínez, I. (2014). La base corporeizada del significado musical. In S. Español (Comp.), *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 71–110). Paidós.
- Martínez, I., Español, & Pérez, D. (2018). The interactive origin and the aesthetic modelling of image-schemas and primary metaphors. *Integrative Psychological and Behavioral Science*, 52(4), 647–671. <https://doi.org/10.1007/s12124-018-9432-z>
- Martínez, I. C. (2007). La composicionalidad de la performance adulta en la parentalidad intuitiva. In M. De La, P. Jacquier, & A. Pereira Ghiena (Eds.), *Música y Bienestar Humano. Actas de la VI Reunión de SACCoM* (pp. 25–34). SACCoM.
- Martínez, M. (2016). Desarrollo de la percepción intersensorial. Aspectos metodológicos, empíricos y conceptuales. In E. Huare Inacio, Á. Elgier, & G. Clerici (Comps.), *Pensar la niñez: Psicología del desarrollo desde una perspectiva americana* (pp. 39–59). Editora y Librería Jurídica Grijley EIRL.
- Martínez, M. (2019). Desarrollo intersubjetivo y perceptivo: El ritmo como ejemplo de su enlace durante el primer año de vida. In C. Scotto, F. Rodríguez, & I. Audisio (Comps.), *Los signos del cuerpo. Enfoques multimodales de la mente y el cuerpo* (pp. 273–325). TESEO-UAI.
- Martínez, M. (2021). *El desarrollo de la percepción de información amodal temporal y del contacto intersubjetivo durante el primer año de vida* [Tesis doctoral en prep.]. Universidad Autónoma de Madrid.
- Martínez, M., Español, S., & E Igoa, J. M. (2018). Reconocimiento de relaciones intersensoriales basadas en el ritmo a los 4, 7 y 10 meses. *Epistemus, Revista de Estudios en Música, Cognición y Cultura*, 6(2), 10–32. <https://doi.org/10.21932/epistemus.6.6233.2>
- Melinder, A., Forbes, D., Tronick, E., Fikke, L., & Gredebäck, G. (2010). The development of the still-face effect: Mothers do matter. *Infant Behavior and Development*, 33(4), 472–481. <http://dx.doi.org/10.1016/j.infbeh.2010.05.003>
- Miall, D., & Dissanayake, E. (2003). The poetics of Babytalk. *Human Nature*, 14(4), 337–364. <https://dx.doi.org/10.1007/s12110-003-1010-4>
- Miller, J. L., & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception and Psychophysics*, 25(6), 457–465. <https://doi.org/10.3758/BF03213823>
- Moreno-Núñez, A., Rodríguez, C., & Del Olmo, M. J. (2015). The rhythmic, sonorous and melodic components of adult-child-object interactions between 2 and 6 months old. *Integrative Psychological and Behavioral Science*, 49(4), 737–756. <http://doi:10.1007/s12124-015-9298-2>
- Morrongiello, B. A. (1984). Auditory temporal pattern perception in 6-and 12-month-old infants. *Developmental Psychology*, 20(3), 441–448. <http://dx.doi.org/10.1037/0012-1649.20.3.441>
- Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology*, 44(3), 413–426. [http://dx.doi.org/10.1016/0022-0965\(87\)90043-9](http://dx.doi.org/10.1016/0022-0965(87)90043-9)
- Ospina, V., & Español, S. (2014). El movimiento y el sí mismo. In S. Español (Ed.), *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 111–155). Paidós.

- Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliège, & J. Sloboda (Eds.), *Musical beginnings. Origins and development of musical competence* (pp. 88–112). Oxford University Press.
- Papoušek, M., & Papoušek, H. (1981). Musical elements in the infant's vocalizations: Their significance for communication, cognition and creativity. In L. P. Lipsitt (Ed.), *Advances in infancy research* (Vol. 1, pp. 163–224). Ablex Norwood.
- Reddy, V. (2008). *How infants know minds*. Harvard University Press.
- Rivière, A. (2003). Interacción precoz. Una perspectiva vygotskiana a partir de los esquemas de Piaget. En M. Belinchón, A. Rosa, M. Sotillo, & I. Marichalar (Comps.), *Ángel Rivière. Obras Escogidas* (Vol. II, pp. 109–142). Panamericana.
- Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development. Themes and variations*. Oxford University Press.
- Schaffer, H. R. (1985). *Ser Madre*. Morata.
- Schaffer, H. R. (1989). *Interacción y socialización*. Visor.
- Schmuckler, M. A., & Jewell, D. T. (2007). Infants' intermodal perception with imperfect contingency information. *Developmental Psychobiology*, 49(4), 387–398. <http://dx.doi.org/10.1002/dev.20214>
- Schögler, B., & Trevarthen, C. (2007). To sing and dance together. In S. Bråten (Ed.), *On being moved: From mirror neurons to empathy* (pp. 281–302). John Benjamins.
- Shifres, F. (2014). Algo más sobre el enlace entre la infancia temprana y la música: El poder expresivo del rubato. In S. Español (Comp.), *Psicología de la música y del desarrollo. Una exploración interdisciplinario sobre la musicalidad humana* (pp. 21–70). Paidós.
- Soken, N. H., & Pick, A. D. (1992). Intermodal perception of happy and angry expressive behaviors by seven month old infants. *Child Development*, 63(4), 787–795. <http://dx.doi.org/10.2307/1131233>
- Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, 8(4), 553–560. [https://doi.org/10.1016/0010-0285\(76\)90018-9](https://doi.org/10.1016/0010-0285(76)90018-9)
- Stern, D. N. (1974). The goal and structure of mother infant play. *Journal of the American Academy of Child Psychiatry*, 13(3), 402–421. [https://doi:10.1016/S0002-7138\(09\)61348-0](https://doi:10.1016/S0002-7138(09)61348-0)
- Stern, D. N. (1983). *La primera relación madre-hijo*. Morata. [The first relationship. Infant and mother. Harvard University Press, 2002].
- Stern, D. N. (1991). *El mundo interpersonal del infante. Una perspectiva desde el psicoanálisis y la psicología evolutiva*. Paidós. [The interpersonal World of The Infant: A View from Psychoanalysis and Developmental Psychology. Basic Books, 1985].
- Stern, D. N., Beebe, B., Jaffe, J., & Bennett, S. L. (1977). The infant's stimulus world during social interaction. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction: Proceedings of the Loch Lomond Symposium* (pp. 177–202). Academic Press.
- Stern, D. N. (2010). *Forms of vitality: Exploring dynamic experience in psychology, the Arts, psychotherapy, and development*. Oxford University Press.
- Tomasello, M. (2013). *Los orígenes de la comunicación humana*. Katz.
- Vaillant-Molina, M., Bahrick, L. E., & Flom, R. (2013). Young infants match facial and vocal emotional expressions of other infants. *Infancy*, 18(Suppl. 1), E97–E111. <http://dx.doi.org/10.1111/infa.12017>
- Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, 33(3), 514–535. [http://dx.doi.org/10.1016/0022-0965\(82\)90063-7](http://dx.doi.org/10.1016/0022-0965(82)90063-7)
- Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, 121(3), 437. <http://dx.doi.org/10.1037/0033-2909.121.3.437>