

# SOSTÉN DE LA ATENCIÓN Y MODELO CROSS MODAL UN ESTUDIO DE CASO A LOS 12 Y 24 MESES

Silvia Malbrán - Julio Menéndez

## Introducción

En este trabajo se presentan avances relativos a un estudio de la performance de una niña a los doce y veinticuatro meses de edad; consiste en el registro fílmico de su actuación como espectadora de un video de la línea comercial Baby Einstein (“*My first signs*”). Las acciones de la niña se analizan en términos de 1) las modalidades de respuesta a diferentes recursos utilizados en la imagen (movimiento, palabra, música, teatro de títeres, juguetes mecánicos) 2) el sostén atencivo.

## 1.- La Imagen

El pensamiento de ficción y la representación abstracta propias del arte se consideran sellos de la cognición humana (Levitin & Tirovolas, 2009). El mundo ficcional compromete sistemas de emociones y genera las mismas respuestas emocionales que las provocadas en la vida real. Pareciera que los seres humanos estamos dotados de una maquinaria cognitiva que nos permite entrar y participar en mundos imaginados (Tooby & Cosmides, 2001).

Reybrouck (2005), postula que la música es escuchada y “*enactivada*” más que meramente imaginada o representada. Aún en los casos en que no es producida por el sujeto, le genera una suerte de *resonancia ideomotora* con el sonido como si hubiera estado comprometido en su producción. Esto sugiere que la intención de la acción es tan movilizadora como el comportamiento manifiesto. El autor lo explica como fenómeno endógeno localizado en el sistema nervioso central: un componente motor hace que las mismas áreas cerebrales que son activadas en la acción sean activadas cuando son observadas en otros o solamente imaginadas (neuronas espejo). Aclara que si bien no toda percepción puede reducirse a componentes motores, ellos están siempre comprometidos y son parte integral de la percepción; a la imagen virtual así conformada en la mente la denomina *simulación ideomotora*.

El mundo es inherentemente multimodal; recibimos simultáneamente información visual auditiva, olfativa y/o gustativa.

Forma parte del conocimiento folk la importancia de brindar a los niños en la primera infancia una diversidad de estímulos sensoriales. Resulta relativamente nuevo, en cambio, que los estímulos que activan simultáneamente diferentes canales sensitivos ponen en juego neuronas multisensoriales que se activan solamente por el *input* múltiple y que permanecen inactivas ante estímulos unimodales (Wright, 2002).

Estos datos resultan de particular importancia para la educación en la primera infancia. Deneve y Pouget. (2004) consideran que la integración multisensorial entre música e imágenes filmadas aporta a la profundidad en la percepción, a la discriminación en la acción, a la comprensión del lenguaje y a la identificación del objeto.

Estas visiones se inscriben en el modelo *cross-modal* para las ciencias cognitivas, que estudia la integración multisensorial suscitada por discursos que conjugan disciplinas artísticas diferentes, particularmente los de las artes temporales.

Una cuestión de vieja data en la Psicología Experimental es la relativa a la percepción de eventos simultáneos. En particular, ha sido motivo de interés lo que sucede cuando una información activa los receptores sensoriales provenientes de diferentes modalidades sensitivas. En tal sentido, los neurocientistas han obtenido importantes progresos; el uso de las técnicas de lectura de imágenes en movimiento ha aportado datos a los que nunca antes se había podido acceder.

La masa crítica acumulada muestra acuerdo en que los sentidos requieren muy variados montos de tiempo para procesar la información receptada y en que el

procesamiento de la información incluye complejos mecanismos tales como, anticipación, comparación, retroalimentación y recursión (Levitin et al, 2000).

Los teóricos del film coinciden en que la *saliencia perceptiva* concurrente de rasgos de la música y del film ayuda a la claridad narrativa y a la adhesión emocional; para estudiar la saliencia se analiza la frecuencia de la ocurrencia de eventos sobresalientes de la música congruentes con momentos significativos de la escena. Bolívar et al. (1994) sugieren que la integración de modalidades sensoriales incrementa la saliencia y decrementa la ambigüedad.

En el modelo *cross modal* la congruencia es entendida como *alineamiento* entre las sintaxis de diferentes artes; interesa en particular el *alineamiento temporal* entre música y escena, esto es, la correspondencia de la distribución temporal y/o de los agrupamientos formales de la imagen con los agrupamientos, niveles de pulso o patrones rítmicos de la música (Bolívar et al. 1994; Lipscomb & Kendall, 1994).

En la escena, los gestos, palabras, frases o sonidos vocales son usados para enfatizar la carga emocional del episodio; cuando estos recursos son concurrentes con el track de sonido obtienen una *congruencia empática* entre contenido dramático y música (Chion, 1993). Cohen (2005) ha mostrado que estos son factores determinantes para movilizar la emoción del espectador.

Tanto en la escena como en la música un recurso compositivo frecuente es la *redundancia*, entendida como la reiteración de ideas. En la música, es una ayuda para la segmentación discursiva, ya que el auditor entiende que el punto inmediato anterior a la repetición marca el límite para que comience otro grupo. En la escena es de uso tan frecuente como en la música y se observa en la reiteración de gestos, situaciones, palabras, entre otros. “Un cierto monto de redundancia es considerado esencial para la comunicación: contribuye a la claridad, predictibilidad, y eficiencia del mensaje (...) [Sin embargo], un exceso de redundancia puede ser interpretado como lugar común, excesivamente predecible o banal” (Wingstedt, 2005: 20-21).

Las artes del tiempo comparten además un criterio de segmentación de la trama que permite presentar ordenadamente las ideas de la secuencia discursiva, en torno a nudos argumentales o unidades de ideas; la denominación común es *agrupamiento*. En la música los auditores perciben el comienzo de un grupo o de una frase musical como proposición de un nuevo material (o repetición de uno ya presentado); las cadencias les permiten advertir los momentos de cierre. Del mismo modo los perceptores pueden señalar los grupos internos a una coreografía o a una escena dramática. El agrupamiento es un importante recurso para la congruencia entre las sintaxis de discursos artísticos diferentes.

Trehub (2003:5) considera que “los principios del agrupamiento perceptual en música son operativos en la infancia; los niños agrupan secuencias basándose en la similitud de sonoridad, timbre y altura”.

Mc Pherson (2006:15) expresa que “la percepción compromete la construcción de representaciones cognitivas de objetos del entorno al comparar input sensoriales en diferentes modalidades de los sentidos”

Según Vines et al. (2006) los procesos mentales subyacentes a las interacciones *cross-modal* son de central importancia para nuestra comprensión de las relaciones entre los sentidos, la organización perceptual y el modo en el cual las representaciones del mundo son configuradas.

Sintetizando, cuando música e imagen comparten rasgos tales como saliencia perceptiva, alineamiento temporal, congruencia empática, redundancia y agrupamiento contribuyen a la “lectura” de los productos fílmicos y videográficos.

Estos modelos son soporte suficiente para justificar el uso de productos videográficos de valor como parte de las experiencias de niñ@s en la primera infancia.

## 2.- La Atención

La atención ha sido un objeto de estudio prioritario de la Psicología Cognitiva; se la considera un “proceso de alto nivel sumamente elaborado, pese a a la simplicidad e inmediatez aparente de nuestra experiencia perceptiva” (de Vega, 1998:67).

McPherson (2006: 15) define psicológicamente la atención “como un estado intenso de alerta en respuesta a los estímulos acoplados a la percepción selectiva de los estímulos. (...) Atención no es lo mismo que conciencia pero está comprometida en la selección y el mantenimiento de contenidos concientes”.

En el ingreso perceptivo se codifican en paralelo y de manera automática las propiedades del input tales como color, forma y dirección del movimiento (Malbrán, s/f<sup>a</sup>). El proceso culmina en una integración sintética de los rasgos del input al alcanzar una configuración significativa, lo que requiere atención controlada (de Vega, 1997).

Según Kahneman (1997:95) “en la asignación de atención podemos fijarnos tanto en aspectos relacionados con el momento como en aspectos relacionados con la secuencia. (...) El acto de mirar es un ejemplo elemental de esta organización secuencial de la atención selectiva”.

“Las opiniones coinciden en que la atención selectiva desempeña un papel fundamental en la percepción, la acción y la memoria: es un factor decisivo para la unidad de la experiencia conciente (...). La capacidad del cerebro para procesar la información sensorial es más limitada que la capacidad de sus receptores para captar el entorno” (Kandel, 2007: 361).

Bolívar et al (1994) consideran que diversos mecanismos por los cuales resulta convocada la atención se basan en similitudes estructurales (principio Gestáltico de agrupamiento por similitud) y en mecanismos asociacionistas (principio de contigüidad temporal).

Son antecedentes de la autora en este tema diversos estudios previos (Malbrán, 2008; Malbrán, s/f<sup>a</sup> y s/f<sup>b</sup> Malbrán & Menéndez, 2009) y una experiencia de Extensión Universitaria en un Jardín Maternal de la Universidad Nacional de La Plata. En dicho proyecto tres educadoras musicales aplicaron la técnica de estímulos concordantes (Regla, 1999) con un grupo de niños entre las edades 6 y 26 meses. Las prácticas integraban la música con otras expresiones artísticas y recursos, tales como narraciones dramatizadas, títeres, juegos de luces y juguetes en íntimo acuerdo con el discurso musical. La iniciativa incluyó la producción de instrumentos musicales especialmente diseñados en relación con las disponibilidades motrices de los niños (Malbrán et al. 2000). En general la ciencia aporta evidencias útiles para la pedagogía; en este caso la experiencia docente basada en las peculiaridades de los niños en la primera infancia, nos sugirió la aplicación de estímulos concordantes. Hoy la ciencia fundamenta tal decisión.

## **Metodología**

El estudio se realizó con una niña argentina que vive en Francia y que por azar estuvo en contacto con un video de enseñanza del idioma inglés por el sistema de señas.

Se tomó una primera toma a los 12 meses. El segundo registro se realizó a los 24 meses. Antes de la primera toma la niña fue espectadora del video en seis oportunidades. Para la segunda toma la niña no había vuelto a contactarse con el video hasta el momento de la filmación.

Para analizar los registros se estimaron los tiempos de atención atendiendo a los cambios de foco de la mirada (rotar la cabeza, mirar en otra dirección). Se consideró que el mantenimiento de la mirada en la pantalla mostraba el sostén atento. Por ello, se calcularon los segundos de permanencia de la mirada en la cámara.

Se analizó también el tipo de recursos utilizados en el video que pudieran ser indicadores de estrategias para el sostén atento.

## **Resultados**

Se consignan a continuación los tiempos de desarrollo del video y la duración en segundos de los segmentos en que la niña permaneció con la mirada en la pantalla.

<b>1era toma (12 meses)</b>		<b>2da toma (24 meses)</b>	
Tiempo del video	Sostén de la atención (segundos)	Tiempo del video	Sostén de la atención (segundos)
		0.41	0.29
		1.13	1.33
1.18	0.28		
2.08	0.26		
2.39	0.20		
		2.49	0.54
3.31	0.31		
4.15	0.29		
4.38	0.28		
		4.45	0.42
		5.33	2.00
6.06	0.33		
6.47	0.28		
7.17	0.35		

***Diferencias atencionales entre la primera y segunda toma.***

<b>Descripción</b>	<b>Primera toma</b>	<b>Segunda toma</b>
Número de segmentos de atención sostenida	<b>9</b>	<b>5</b>
Segmento de mayor duración atenta	<b>35"</b>	<b>2 minutos</b>
Segmento de menor duración atenta	<b>20"</b>	<b>29"</b>
Iniciación de la atención al video	<b>1.18</b>	<b>0.41</b>

***Diferencias actitudinales entre la primera y segunda toma.***

<b>Primera toma</b>	<b>Segunda toma</b>
Emite palabras /sílabas	No emite palabras.
Realiza movimientos acordes a la trama	No realiza movimientos
Expresa gestos de beneplácito	Manifiesta sorpresa Sonríe y muestra "vergüenza"
Anticipa situaciones de la acción dramática	No anticipa movimientos, ni situaciones de la acción dramática

***Respuestas a características musicales del estímulo (primera toma)***

- \* anticipa el cierre de la escena y deja de atender, ante las cadencias armónicas
- \* vuelve a prestar atención cuando retorna la música

***Dispersión atenta (primera toma)***

- \* rotación de la cabeza ida y vuelta sin detenciones: 6 veces
- \* rotación de la cabeza con detención en la ida:: 12 veces
- \* miradas al entorno sin giro de cabeza (rejo) 9 veces.

***Recursos utilizados en el video***

- \* pantalla negra o blanca para corte de los segmentos formales
- \* uso de juguetes mecánicos que se desplazan
- \* uso de títeres
- \* reiteración de recursos (palabras, música, sonidos, gestos, movimientos).

## Conclusiones

La observación de los registros filmados permite adelantar algunas conclusiones preliminares.

**Primera toma:** la performance muestra que el material fílmico forma parte de los conocimientos previos de la niña. Puede i) anticipar la acción con los nombres de los personajes o los movimientos que los representan, ii) interactuar con el video durante la acción con movimientos similares a los de los actores o anticipando la palabra en inglés (daddy, .baby) y iii) desatender ante los previsibles cortes de la acción por salidas de los personajes de la escena o las cadencias musicales. Los períodos atencionales oscilan alrededor de los 28 segundos aunque los “cortes” de atención son muy breves y consisten en la rotación de la cabeza y las miradas al entorno.

Las escenas de mayor saliencia para la niña han sido i) la acción de acunar, a la que imita en acople con la imagen mientras verbaliza la palabra *baby* ii) la palabra *daddy* a la que acompaña con la seña respectiva (la mano en la cabeza). En ambos casos se trata de conocimiento enactivado: acciones corporales y expresión verbal alineadas temporalmente. Estas mismas imágenes son las más reiteradas en la imagen lo que ratifica el poder representacional y mnémico de la redundancia.

**Segunda toma:** la performance muestra que el material fílmico no le resulta familiar, aunque las actitudes de vergüenza y sorpresa podrían indicar que si bien no los recuerda como en la primera toma le significan un contacto con “algo conocido”. Las respuestas indican que i) inicia la atención al video más tempranamente que en la primera toma; ii) puede mantener la atención por períodos considerablemente más amplios iii) reduce el número de segmentos de atención ya que cada uno de ellos es más extenso.

**Ambas tomas:** en cuanto a la emoción, en la primera toma la niña muestra “sacudimientos” (*chills*) y en la segunda toma asume actitudes/expresiones de vergüenza; ambos indicadores han sido descritos como muestras de emoción (Juslin & Sloboda, 2001).

El poder de la asociación entre imagen-sonido como activación (*arousal*) y sostenimiento de la atención es claramente puesto en evidencia en las dos oportunidades. Los cortes de atención son de duración mínima y la niña mantiene el estado de alerta durante los 7.35 minutos del video.

En términos atencionales la niña muestra *respuestas de orientación*, ya sea hacia una nueva señal, hacia los juguetes que se mueven o con situaciones de humor de los protagonistas; así también *intenciones momentáneas* (Kahneman, 1997), al activar criterios selectivos tales como, dejar de atender ante la inminencia del cierre o ignorar los ruidos externos.

La actuación en su conjunto se encuadra en el concepto de *atención selectiva* (Kahneman, 1997) ya que efectúa una selección de los inputs focalizando la atención en estímulos que proceden de una fuente concreta y filtrando información irrelevante.

Los resultados obtenidos acuerdan con los marcos teóricos del modelo *crossmodal* y los estudios de los expertos acerca de la atención. La primera toma aporta mayores datos que la segunda, aunque el estudio comparado entre ambas sugiere el cambio sustantivo en el control atento entre los doce y los veinticuatro meses. El sostén de la atención, el acople corporal y vocal a la imagen, los momentos de emoción, la necesidad de compartir su disfrute con otros señalando la pantalla, resultan indicadores ilustrativos del atractivo de la imagen, de la viva interacción con el video y del universo de datos sobre su propio mundo que nos pueden proporcionar los niños en la primera infancia.

## Discusión

La práctica educativa en los primeros años de escolaridad señala habitualmente las dificultades para mantener la atención de los niños en estas edades. El registro fílmico

muestra la competencia atencional en los 12/24 meses y puede sugerir ciertas líneas de acción para el desarrollo de las potencialidades expresivas en dichas edades. Los agrupamientos temáticos de la imagen son de muy breve duración. Estas segmentaciones formales de la escena concluyen mediante pantallas neutras (negras o blancas). Las características comunes a los personajes son la géstica, el color, el humor y el movimiento asociado con motivos o discursos musicales. Las estrategias de uso de los recursos podrían sugerir pautas para la acción tales como: i) utilizar diversos motivos de trabajo para cada unidad clase; ii) valerse de recursos tales como juguetes mecánicos en movimiento, títeres y marionetas con acciones alineadas temporalmente con la música; iii) administrar los segmentos internos a la clase con mínimas variaciones dentro de la reiteración de motivos conocidos y con “cierres” claros de cada segmento como, por ejemplo, cambio de ubicación en el espacio físico, recursos diferentes en cada segmento, entre otros. Los niños en la primera infancia perciben el mundo de manera esencialmente multimodal. Las prácticas pedagógicas necesitan diseñarse de acuerdo con ello. Poner en marcha experiencias provocativas y gozosas para estas edades, sugiere la necesidad de un cambio de paradigma.

## Referencias

- Baby Einstein Co. (2007). *“My first signs” (DVD)*. Burbank, USA: Walt Disney Studios.
- Bolívar, V., Cohen, A. y Fentress, J. (1994). Semantic and formal congruency in music and motion pictures: effects on the interpretation of visual action. *Psychomusicology*, 13, 28-59
- Chion, M. (1990/ 1993). *La Audiovisión. Introducción a un análisis conjunto de la imagen y el sonido*. 1era edición en español. Paris: Ediciones Nathan
- Cohen, Annabel. ( 2001). Music as source of emotion in film. En P. Juslin y J. Sloboda: *Music and Emotion. Theory and Research*. Oxford: UP
- Cohen, Annabel J., (2005) ‘How Music Influences Film and Video: Approaches from Experimental Psychology’. En Kendall, Roger A., & Savage, Roger W. H.,(eds.). *Perspectives in Systematic Musicology*. California: UP
- Deneve, S. & Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *Elsevier. Journal of Physiology. Paris*
- de Vega, M. ( 1997). *Introducción a la Psicología Cognitiva*. Madrid : Alianza
- Kahneman, D. (1997). *Atención y esfuerzo*. Madrid: Biblioteca Nueva.
- Levitin, D., McLean, K., Mathews, M. & Chu, L., (2000). The Perception of Cross Modal Simultaneity. *Computing Anticipatory Systems: CAYSYS'99*. Vol. 517 Nro 1.
- Juslin, P. & Sloboda, J. (2001). *Music Emotion. Theory and Research*. Oxford: Oxford U.P.
- Kandel, E. (2007). *En busca de la memoria. El nacimiento de una nueva ciencia de la mente*. Buenos Aires: Katz
- Levitin, D. & Tirovolas, A. (2009). Current advances in the cognitive neuroscience of music. *The year in cognitive neuroscience 2009: Ann. N.Y. Acad. Sci.* 1156:211-231
- Lipscomb, S. & Kendall, R. (1994). Perceptual Judgment of the relationship between musical and visual components in film. *Psychomusicology*, 13, pp.60-98
- Mc Pherson, G. (2006). *The child as musician. A Handbook of Musical Development*. Oxford: Oxford U.P.

- Malbrán S. & Menéndez, J. (2009) Integración *cross modal* y música. *VII Encuentro Latinoamericano de Educación Musical (ISME)*. Buenos Aires.
- Malbrán S. (s/f <sup>a</sup>) Sonido y música en el arte de la imagen. Cap. 2 (en prensa). En M. Díaz y M.E. Riaño (eds.) *Voz, cuerpo y acción. Un espacio para la música*. Santander: PUbliCan-Ediciones.
- Malbrán S. (s/f <sup>b</sup>) El modelo *cross modal* aplicado a las Artes Temporales. Cap. 1 (en prensa). En M. Díaz y M.E. Riaño (eds.). *Voz, cuerpo y acción. Un espacio para la música*. Santander: PUbliCan-Ediciones
- Malbrán, S. (2008). El cruce de modalidades perceptivas. Un estudio sobre música y actuación. *Primer Congreso Internacional. Artes en cruce: Problemáticas teóricas actuales* Buenos Aires: Universidad de Buenos Aires
- Malbrán, S., Regla, I., García Malbrán, E, & Sauber, M. (2000). *Sonidos para los sentidos. Experiencias Integrales de Educación Musical con niños de 6 a 26 meses*. Secretaría de Extensión Universitaria. Facultad de Bellas Artes (UNLP)
- Regla, I. (1999). El Modelo de Interacción con Estímulos Concordantes; una propuesta educativa para la primera infancia. *Novedades Educativas*. Año 12 nro. 109
- Reybrouck, M. (2005). Body, mind and music: musical semantics between experiential cognition and cognitive economy. *Transcultural Music Review*. 9.
- Tooby, J & Cosmides, L. (2001). Does beauty adapted minds? Toward an Evolutionary Theory of Aesthetics, Fiction and the Arts. *SubStance*. Issue 94/95, 30(1)6-27
- Trehub, S. (2003). Music predisposition in infancy: an update. En I. Peretz, I y R. Zatorre (eds.). *The Cognitive Neuroscience of Music*. Oxford: Oxford U.P.
- Vines, B., Krumhansl, C., Wanderley, M. & Levitin, D. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101, 80-113.
- Wingstedt, J. (2005). Narrative Music: Towards an understanding of Musical Narrative Functions in Multimedia. Tesis de licenciatura. Suecia: *Musikhögskolan i Piteå*.
- Wright, S. (2002). Multimodality in a new key: the significance of the arts in research and education. *AARE. Conference. Papers*