

Medición Mediante Autoevaluación de Percepción de Música Electroacústica Acusmática

Juan Ignacio Mendoza Garay

Departamento de Música,
Universidad de Jyväskylä, Finlandia

Noviembre de 2014

Resumen

La Música Electroacústica Acusmática es una forma de arte en la que se considera que cualquier sonido tiene el potencial para ser utilizado como una unidad básica en la construcción de un discurso musical. El presente estudio aborda la medición de tres aspectos de la percepción de esta música: Segmentación, Mímesis y Percepción de Emociones. Los resultados sugieren que las estrategias para segmentar secuencias de sonidos pueden corresponder tanto a sus características acústicas como a su significado inferido del reconocimiento de fuente y acción productora. Este reconocimiento es variable y respuestas narrativas sugieren que los sonidos evocan sentido en lugar de tener sentido por sí mismos. Las mediciones a posteriori y continuas de emociones percibidas dieron resultados muy similares, lo que sugiere que las emociones percibidas en sonidos individuales son constructos establecidos que podrían no cambiar considerablemente cuando los mismos sonidos están en contexto. También se muestra que la medición continua de segmentación y emociones percibidas es fiable sólo para sonidos con una duración de más de 4 segundos.

Abstract

Acousmatic Electroacoustic Music is a form of art in which any sound has the potential to be used as a basic unit for the construction of a musical discourse. This study addresses the measurement of three aspects of perception of this music: Segmentation, Mimesis and Perception of Emotions. Results suggest that strategies to segment sequences of sounds can correspond both to their acoustic characteristics and to their meaning inferred from the recognition of the source and action that produced the sound. These variable and narrative responses suggest that sounds evoke meaning rather than having meaning by themselves. Post-hoc and continuous measurements of perceived emotions yielded very similar results, which suggests that perceived emotions in discrete sounds are established constructs that might not change considerably when they are in context. It is also shown that continuous measurement of segmentation and perceived emotions is reliable only for sounds with a duration over 4 seconds.

1 INTRODUCCIÓN

Este texto da cuenta de una investigación realizada entre los años 2012 y 2014 como parte de mis estudios de maestría en la Universidad de Jyväskylä, Finlandia. La investigación (Mendoza Garay, 2014) se centró en la medición de tres aspectos de la percepción de Música Electroacústica Acusmática (véase Landy, 1999; Truax, 1999): Segmentación, Mímesis y Percepción de Emociones. Una motivación importante para estudiar este tipo de música es que hoy parece no haber consenso entre académicos y artistas, sobre las estrategias perceptivas generales que dan sentido a la música electroacústica acusmática y por lo tanto de explicar desde un punto de vista objetivo el funcionamiento de la experiencia estética en la escucha de esta forma de arte. Quizás la manera más establecida de abordar el análisis de música electroacústica es primero segmentar la pieza en partes, hasta llegar a la unidad significativa más pequeña que se ha llamado "objeto sonoro" (Schaeffer, 1966; Schaeffer, 1967; Chion, 1983). A partir de esta unidad básica estructuras más complejas pueden ser percibidas y surge significado, principalmente por la identificación de la naturaleza del sonido y luego por la percepción de sus posibles cualidades vinculadas a la influencia de las emociones. La música electroacústica se compone de sonidos a menudo manipulados de manera que su resultado evoca experiencias que trascienden la experiencia terrenal provocando reacciones análogas en los oyentes. Entonces, el estudio de la misma desde una perspectiva científica no sólo ayuda a comprender la experiencia estética, sino que también podría revelar aspectos de la percepción humana de una manera que quizás no sería posible o fácil con otro tipo de estímulo. Autoevaluación ha sido la estrategia escogida para medir los tres aspectos mencionados, ya que necesita introspección por parte del oyente así proporcionando acceso a los mecanismos psicológicos de alto nivel que tienen lugar en la percepción.

2 MÉTODOS

2.1 Registro de Datos y Análisis

Se realizaron cuatro experimentos, en los que se registraron y analizaron las respuestas de una veintena de participantes, estudiantes y profesores de la Universidad de Jyväskylä con edades entre 20 y 40 años. Para el registro de datos se diseñó un sistema que capturó las respuestas de los participantes mediante un controlador "Nintendo Wii Remote" (también conocido como Wiimote), transmitiendo datos de botones y movimiento a una computadora personal. También se usó un micrófono para grabar respuestas verbales y el teclado de la computadora para responder cuestionarios. El núcleo de este sistema es software desarrollado como parte de este proyecto en el entorno Pure Data (véase Puckette, 1996) teniendo en cuenta que debía ser efectivo y fácil de usar. Se ha publicado bajo la licencia de software libre GNU versión 3 (Free Software Foundation, 2007) y se puede descargar de manera gratuita desde el siguiente sitio de internet: <http://juanignaciomendoza.weebly.com>. Asimismo, se diseñó una herramienta de análisis numérico para medir el nivel de similitud entre series de datos obtenidos en los experimentos que se describen a continuación.

2.2 Experimento 1: Segmentación de un Fragmento de Música Acusmática

El objetivo de este experimento fue estudiar el comportamiento de participantes humanos al segmentar Música Electroacústica Acusmática. A los participantes se les presentó cinco veces un estímulo consistente en un fragmento de "Ciguri" (Otondo, 2008) del compositor chileno Felipe Otondo. Se les pidió que presionaran un botón del Wiimote cada vez que escucharan *cambios en la música* mientras escuchaban el extracto. El análisis de datos incluyó una comparación entre los datos de segmentación obtenidos de los participantes y datos de segmentación efectuados por computación del archivo de audio mediante un algoritmo de autosimilitud (véase Foote & Cooper, 2003; Klien, Grill & Flexer, 2012; Lartillot & Toiviainen, 2007). También por medio de un análisis semántico se observaron las estrategias de los participantes para llevar a cabo la segmentación, de modo de inferir las características acústicas o musicales relevantes relacionadas con los puntos de segmentación

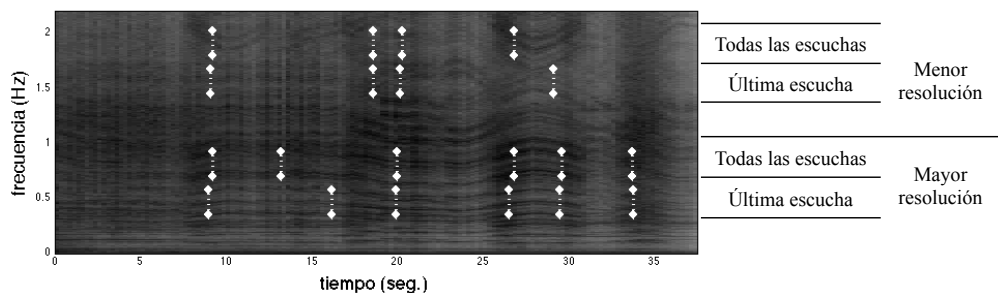


ILUSTRACIÓN 1: Espectrograma con puntos de segmentación de los participantes superpuestos en diferentes resoluciones. Se pueden apreciar que los puntos de segmentación son cercanos a zonas sombreadas que representan cambios en la señal de audio.

percibidos. Los resultados del análisis sugieren que algunos participantes evaluaron la segmentación a un nivel más detallado que otros y que las principales características para la segmentación se definen por un cambio general en altura y nivel energía. Los participantes no dieron respuestas más coherentes con más escuchas, lo que no apoya la hipótesis de que las respuestas son más precisas a medida que el estímulo se escucha más veces, como sugieren las investigaciones de Bailes y Dean (Bailes & Dean, 2007a; Bailes & Dean, 2007b). Una posibilidad es que a medida que transcurren las escuchas se ponga atención a diferentes características del sonido y que algunas de ellas se consoliden con el tiempo, al mismo tiempo que otras características surgen como más relevantes. Sin embargo el experimento no logró mostrar claramente qué sonidos en el estímulo estaban estableciendo los puntos de segmentación, porque el estímulo está compuesto por más de una capa de sonidos. Ver la Ilustración 1.

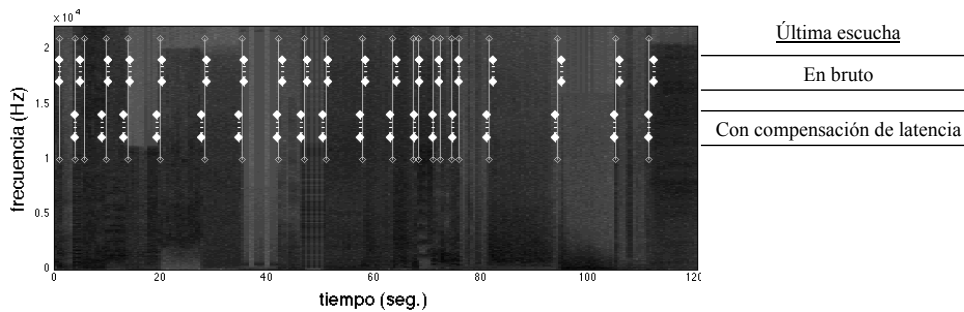


ILUSTRACIÓN 2: Espectrograma con puntos de segmentación superpuestos, correspondientes a la última escucha de los participantes, tanto en bruto como con compensación de latencia. Las líneas delgadas de color gris muestran los límites de cada sonido en el estímulo. Se observa que para sonidos con duraciones mayores a 2 segundos los puntos de segmentación de los participantes son cercanos a las ubicaciones de los sonidos en el estímulo. Sin embargo en el caso de sonidos de menor duración no es posible establecer con claridad a qué sonido corresponden los puntos de segmentación, con o sin compensación de latencia.

2.3 Experimento 2: Segmentación de un Estímulo Acusmático de Una Capa

Este experimento se diseñó después del Experimento 1 para enfocar la observación de la segmentación de los participantes. En lugar de usar una pieza musical acusmática se usó un estímulo compuesto de 24 sonidos concatenados. Esto proporcionó una sola capa de sonidos simplificando el problema de saber cuál los participantes usaron para segmentar, como ocurrió en el Experimento 1. A los participantes se les presentó el estímulo 3 veces y se les pidió que presionaran un botón del Wiimote cuando escucharan un cambio en la música. El resultado del análisis de datos muestra que la última escucha alcanzó mayores valores de similitud con las ubicaciones de los sonidos en el estímulo. Sin embargo no todos los sonidos fueron segmentados con precisión. 19 sonidos fueron segmentados de manera muy precisa por los participantes, teniendo éstos sobre 2 segundos de duración. Los datos de 5 sonidos, con duraciones menores a 2 segundos mostraron imprecisión. Quizás esto se debe a la latencia, es decir, el tiempo necesario para escuchar el sonido, evaluarlo conscientemente y accionar el botón del Wiimote para indicar la respuesta. Además, los participantes no dieron respuestas más coherentes en las últimas escuchas. Debido a esto todavía no es posible probar la hipótesis de que la precisión de la segmentación mejora con la cantidad de escuchas del estímulo. Ver la Ilustración 2.

2.4 Experimento 3: Reconocimiento de Agente y Acción en Sonidos

En este experimento los participantes escucharon 24 sonidos por separado. Después de cada sonido respondieron un cuestionario en el que indicaron el origen (o "agente") del sonido y la acción que lo produjo. Mediante un análisis semántico se agruparon las respuestas (ver la Ilustración 3), lo que reveló que la discriminación entre la fuente de origen y la acción en muchos casos no fue completamente clara y la identificación de ellas tuvo una distancia semántica variada (véase Rips, Shoben & Smith, 1973). Por ejemplo, para un sonido la mayoría de los participantes respondió "arma - disparando", "arma - guerra" o "ametralladora - disparando", mientras que uno de los participantes respondió "soldado - disparando una ametralladora". Las respuestas de este sonido se pueden agrupar con el par origen-acción general de "arma - disparando". Sin embargo algunos sonidos presentaron menos coherencia. Por ejemplo, para un sonido la mayoría de los participantes describió la acción "susurrando" pero un participante respondió "el demonio dice ¿Dónde estás?", mientras que otro participante mencionó personajes de ficción como "Gollum", el personaje de *El Señor de los Anillos* (Tolkien, 2002), que hace un "horrible ruido al tragar en su garganta". Otro participante mencionó a "Voldemort", el malvado mago que aparece en *Harry Potter* de J.K. Rowling. Ambos personajes de ficción han aparecido con sus voces características en las versiones cinematográficas de las novelas. Para el mismo sonido otros participantes respondieron "soplete" como su origen. Conviene mencionar que este sonido fue producido mediante modulación de ruido blanco con un filtro que aplica formantes de vocales. Otro tipo de respuestas que cabe destacar es aquel en los que los participantes no se limitaron a mencionar el origen y la acción, elaborando narrativas. Por ejemplo: "grupo de pájaros pequeños" y "los pájaros pelean por un trozo de pan en una orilla"; "una persona que acosa a alguien llamándolo" y "el acosador respirando pesadamente en el teléfono"; "una persona nadando" y "alguien nadando lentamente en agua quieta". Este comportamiento refuerza la noción de que las ideas e imágenes extraídas de la experiencia auditiva son un enlace entre las experiencias personales y colectivas.

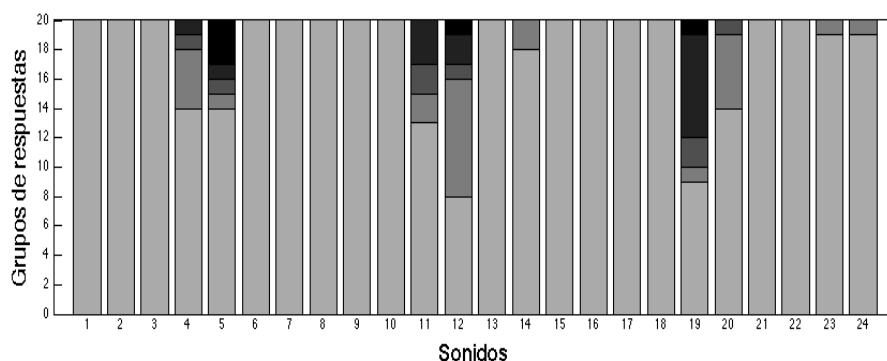


ILUSTRACIÓN 3: Agrupaciones semánticas para cada sonido. El eje vertical corresponde a los participantes. Cada grupo se representa con un sombreado diferente.

2.5 Experimento 4: Percepción Continua y A Posteriori de Emociones en un Estímulo Acusmático de Una Capa

Este experimento fue enfocado a la medición de emociones percibidas en dos condiciones. La primera condición fue *continua*, en la que los participantes evaluaron emociones percibidas mientras escuchaban un estímulo compuesto de varios sonidos concatenados. Para evaluar valencia emocional a los participantes se les pidió que movieran la punta del Wiimote hacia su posición vertical si consideraban el sonido como positivo y para evaluar el sonido como negativo debían mover el aparato a la posición vertical opuesta. Mantener el Wiimote en posición horizontal corresponde a valencia neutra. A los participantes se les pidió usar el rango completo entre las posiciones verticales para evaluar la magnitud de la emoción percibida. Del mismo modo los participantes evaluaron actividad. La posición vertical hacia abajo corresponde a actividad nula mientras que la posición opuesta corresponde a máxima actividad. La segunda condición fue *a posteriori*, en la que los participantes evaluaron los sonidos usados para componer el estímulo de la condición continua, después de escuchar cada uno de ellos por separado, mediante un cuestionario con escalas numéricas (véase Likert, 1932). El análisis de datos mostró que las evaluaciones a posteriori y continuas de emociones percibidas son muy parecidas. Asimismo la mayor correlación entre mediciones a posteriori y continuas ocurre hacia las últimas escuchas. Esto sugiere que los constructos emocionales se consolidan con escuchas sucesivas. La alta similitud entre las respuestas continuas y a posteriori sugiere que el contexto no influye en las

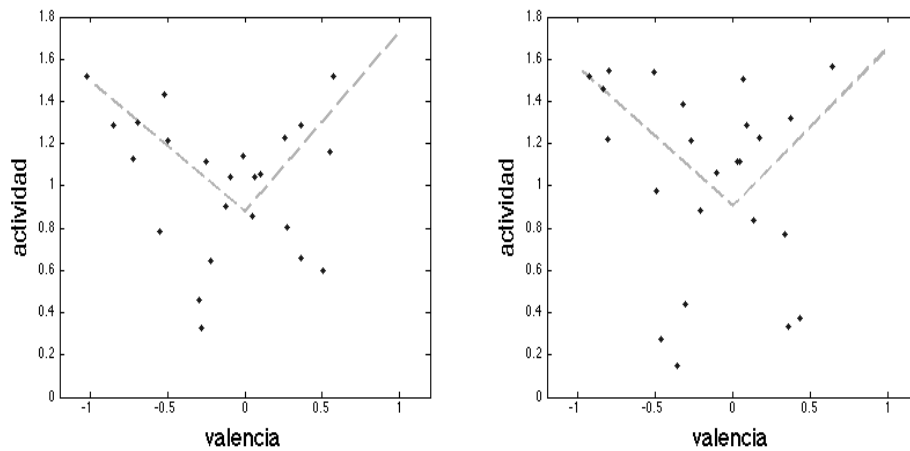


ILUSTRACIÓN 4: Evaluaciones cuantitativas de los sonidos en el "espacio afectivo" (Posner, Russell & Peterson, 2005). El panel a la izquierda muestra un promedio de las mediciones a posteriori y el panel a la derecha muestra un promedio de mediciones continuas. Cada punto indica un sonido. La línea segmentada representa el mejor ajuste polinómico.

emociones percibidas. Sin embargo, este punto de vista todavía es simplista en tanto este experimento por sí solo no revela si los sonidos concatenados proporcionaron un contexto efectivo pues no fueron organizados intencionalmente, de manera musical. Los participantes podrían haber entendido los sonidos tan sólo como una secuencia en lugar de una narrativa que genera un contexto. Finalmente, mediante ajuste por regresión polinómica de mínimos cuadrados (véase Mathews & Fink, 1999) aplicado a las respuestas, se encontró la misma relación de actividad como función de valencia encontrada en estudios previos (Bradley & Lang, 2000; Bradley & Lang, 2007). Véase la Ilustración 4.

3 COMENTARIO

Los resultados obtenidos en todos los experimentos se compararon para tener una visión general de las relaciones entre los tres conceptos medidos, aunque no se observaron relaciones relevantes. De todos modos es necesario repetir estos experimentos con un número mayor de participantes para obtener resultados más generalizables. Sin embargo, a partir de los resultados obtenidos en los Experimentos 1, 2 y 4, es posible deducir que sólo son fiables los análisis de mediciones de eventos con una duración del doble del evento más corto medible, cuando se realiza medición continua por medio de autoevaluación. Por ello sólo son fiables las mediciones de sonidos de más de 4 segundos. Esto está en línea con conclusiones emanadas de trabajos anteriores (Krumhansl, 1996; Schubert, 2004). Asimismo es necesario realizar nuevas investigaciones sobre las técnicas para acceder a la zona de tiempo por debajo de 4 segundos pues es en esta zona donde habita la mayor parte de los objetos sonoros descritos fenomenológicamente. Aunque la medición mediante autoevaluación está dirigida a obtener una idea de representaciones de alto nivel de los mecanismos de percepción que se pretende observar, la medición continua de emociones percibidas podría ser usada en conjunción con técnicas más precisas de evaluación de emociones inducidas, tales como mediciones fisiológicas (véase Hodges, 2010) o mediciones de actividad cerebral (véase Koelsch, 2010). Los resultados del Experimento 3 sugieren que el énfasis de escucha en el par origen-acción del sonido o en sus cualidades espectrales es una opción tomada por el oyente. En general los resultados de los experimentos sugieren que los estímulos usados actúan como un desencadenante de significado individual desde un significado más general y que los sonidos no contienen un significado universal en sí mismos.

CRÉDITO

Este texto ha sido financiado por el Consejo Nacional de la Cultura y las Artes del Gobierno de Chile, mediante el Fondo para el Fomento de la Música Nacional.

REFERENCIAS

- Bailes, F. & Dean, R.T. (2007). Facilitation and coherence between the dynamic and retrospective perception of segmentation in computer-generated music. *Empirical Musicology Review*, 2(3).
- Bailes, F., & Dean, R. T. (2007). Listener detection of segmentation in computer-generated sound: An exploratory experimental study. *Journal of New Music Research*, 36(2), 83-93.
- Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37(2), 204-215.
- Bradley, M. M., & Lang, P. J. (2007). The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual. *University of Florida, Gainesville, FL, Tech. Rep. B-3*.
- Chion, M. (1983). *Guide des objets sonores*. Buchet/Chastel.
- Foote, J. T., & Cooper, M. L. (2003). Media segmentation using self-similarity decomposition. En *Electronic Imaging 2003* (pp. 167-175). International Society for Optics and Photonics.
- Free Software Foundation (2007) GNU general public license, version 3. Recuperado el 4 de Mayo de 2014 desde <https://www.gnu.org/copyleft/gpl.html>
- Hodges, D. (2010). Psychophysiological measures. *Handbook of music and emotion: Theory, research, applications*, 279-311.
- Klien, V., Grill, T., & Flexer, A. (2012). On automated annotation of acousmatic music. *Journal of New Music Research*, 41(2), 153-173.
- Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in cognitive sciences*, 14(3), 131-137.
- Krumhansl, C. L. (1996). A perceptive analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, 401-432.
- Landy, L. (1999). Reviewing the musicology of electroacoustic music: a plea for greater triangulation. *Organised Sound*, 4(1), 61-70.
- Lartillot, O., & Toivainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction From Audio. En *International Conference on Digital Audio Effects* (pp. 237-244), Bordeaux.

- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Mathews, J. H., & Fink, K. D. (1999). *Numerical methods using MATLAB* (Vol. 31). Upper Saddle River, NJ: Prentice hall, pp. 270-273.
- Mendoza Garay, J. (2014). Self-Report Measurement of Segmentation, Mimesis and Perceived Emotions in Acousmatic Electroacoustic Music. Recuperado el 7 de Noviembre de 2014 desde <https://jyx.jyu.fi/dspace/handle/123456789/43811>
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(03), 715-734
- Puckette, M. (1996). Pure Data: another integrated computer music environment. *Proceedings of the Second Intercollege Computer Music Concerts*, 37-41.
- Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of verbal learning and verbal behavior*, 12(1), 1-20.
- Schaeffer, P. (1966). *Traité des objets musicaux, essai interdisciplines*. Paris: Éditions du Seuil.
- Schaeffer, P. (1967). *Solfège de l'objet Sonore*. Paris: INA/GRM.
- Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music perception*, 21(4), 561-585.
- Tolkien, J. R. R. (2002). Anderson, Douglas A., (Ed.) *The Annotated Hobbit*. Boston: Houghton Mifflin Harcourt.
- Truax, B. (1978). *Handbook for Acoustic Ecology*, Originalmente publicado por The World Soundscape Project, Simon Fraser University.