



Influencia de la Compresión del Rango Dinámico en la Evaluación de Calidad de Grabaciones Musicales

Borup , Matias Bødker; Zabbal, Paul; Ordoñez, Rodrigo

Published in:

FIA2014 IX Congreso Iberoamericano de Acústica, Valdivia, Chile 1-3, Diciembre

Publication date:

2014

Document Version

Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Borup , M. B., Zabbal, P., & Ordoñez, R. (2014). Influencia de la Compresión del Rango Dinámico en la Evaluación de Calidad de Grabaciones Musicales. In J. P. Arenas (Ed.), *FIA2014 IX Congreso Iberoamericano de Acústica, Valdivia, Chile 1-3, Diciembre: Actas del Congreso* (pp. 830-838). Article 0407 Instituto de Acústica, Universidad Austral de Chile . <http://fia2014.cl>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Influencia de la Compresión del Rango Dinámico en la Evaluación de Calidad de Grabaciones Musicales

M. B. Borup, P. Zabbal, & Rodrigo Ordoñez

^aLaboratorio de Acústica, Instituto de Sistemas Electrónicos, Universidad de Aalborg, Aalborg, Dinamarca
rop@es.aau.dk

RESUMEN: Para aumentar la sonoridad (la intensidad percibida) de un registro de audio, una técnica bastante utilizada es la compresión del rango dinámico. Compresión en la música de hoy generalmente ocurre en las etapas de masterización y es bastante común encontrar grabaciones remasterizadas que tienen menor rango dinámico que las versiones anteriores. En el presente estudio, grabaciones musicales fueron procesadas con niveles distintos de compresión y luego evaluadas subjetivamente en relación a su profundidad, claridad, presencia de ruidos molestos o artefactos sonoros, calidad y preferencia. El nivel de reproducción de todos los estímulos fue ajustado para eliminar diferencias de sonoridad producidas por los distintos niveles de compresión. Los resultados obtenidos en este experimento no demostraron una relación significativa con los parámetros subjetivos estudiados para niveles de compresión pequeños. Es decir, que al eliminar diferencias de sonoridad, es muy difícil para sujetos no entrenados detectar diferencias en compresión dinámica leve. Al aumentar los niveles de compresión dinámica, sujetos son capaces de detectar las diferencias entre los estímulos y presentan una clara tendencia a preferir grabaciones con menor nivel de compresión.

KEYWORDS: Compresión dinámica, grabaciones musicales, evaluación subjetiva

1. INTRODUCCIÓN

Con la introducción de audio digital cambió en forma radical la manera en que la gente escucha música. Hoy en día una gran parte de la población escucha música a través de audífonos y plataformas móviles (reproductores de mp3, teléfonos, etc.), dando la posibilidad de escuchar música a cualquier hora y en cualquier lugar, incluso en ambientes con altos niveles de ruido. Para poder hacer que música se escuche “aceptablemente bien” en ambientes ruidosos, usuarios tienden a escoger niveles de reproducción más altos, que en ambientes sin ruido en un intento de aislarse del ruido existente para disfrutar de la música. En términos prácticos aumentar el nivel corresponde a aumentar la sonoridad (la sensación de intensidad sonora percibida) de la música.

Al reducir las diferencias entre los niveles máximos y mínimos a través de compresión dinámica, es posible aumentar el nivel equivalente de una grabación sin saturar la señal digital. Esto se puede apreciar en la Figura 1, donde se puede comparar los niveles máximos y equivalentes de la misma señal sin compresión, con compresión (umbral de -25 dBFS y una razón de compresión de 4:1) y la señal comprimida y normalizada para utilizar todo el rango dinámico del sistema.

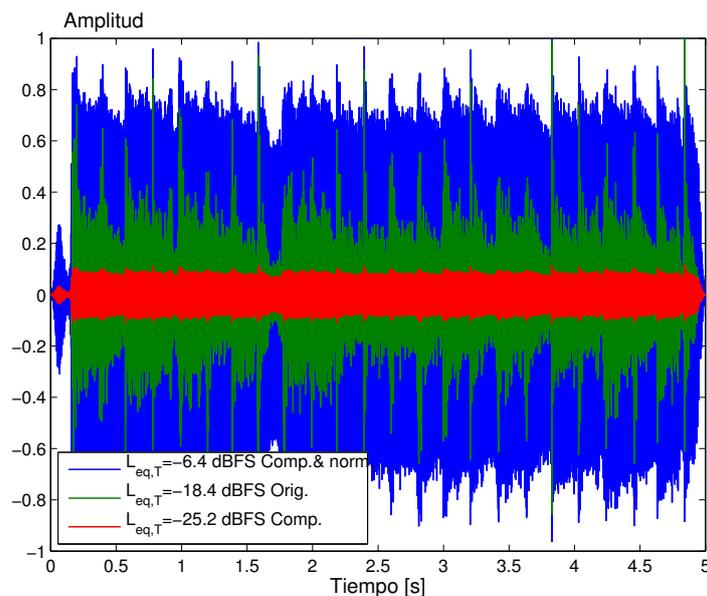


Figura 1: 5 segundos de una grabación musical de una banda de rock. Señal original sin compresión dinámica (color verde); señal comprimida con un umbral de -25 dBFS y una razón de compresión de 4:1 (color rojo); señal comprimida y normalizada para ocupar todo el rango dinámico del sistema digital (color azul). En la leyenda se pueden ver los niveles equivalentes ($L_{eq,T}$) en dB con la amplitud digital máxima de ± 1 como referencia, dBFS (dB Full Scale).

Durante las últimas décadas ha existido una tendencia a comprimir el rango dinámico de grabaciones musicales con el propósito de aumentar su sonoridad. Según Vickers [1], el rango dinámico promedio en grabaciones musicales se ha reducido de 12 dB hasta 8 dB en un período de 20 años (1990 - 2010). Basta comparar grabaciones originales editadas en los 80' con versiones

remasterizadas de los últimos años, para darse cuenta de las diferencias [2]. Esta tendencia en la industria musical de optar aumentar la sonoridad en la música a costas del rango dinámico se ha denominado la guerra de la sonoridad (Loudness wars), donde la premisa que no se ha investigado a fondo es: mientras más fuerte mejor!

La motivación para el presente artículo es investigar dicha premisa a través de una serie de experimentos psicoacústicos. La idea es investigar si la compresión del rango dinámico afecta la percepción de atributos asociados a la calidad de grabaciones musicales y si existe alguna correlación entre compresión dinámica y preferencia.

Los tres experimentos que se presentan a continuación han sido realizados por estudiantes del M.Sc. en Acústica y Tecnología de Audio y participantes del curso de verano Psicofísica: Del Estímulo a la Percepción de la Universidad de Aalborg.

2. MÉTODOS

Para la evaluación subjetiva de los efectos de la compresión dinámica se realizaron tres experimentos utilizando dos métodos psicoacústicos distintos. El primer experimento (Exp1) recogió evaluaciones directas en escalas de analogía visual (VAS visual analog scale); Los otros dos experimentos (Exp2 y Exp3) utilizaron comparaciones pareadas de escoja forzada (2-AFC two alternative forced choice).

Todos los experimentos se realizaron en cabinas para experimentos auditivos en el Laboratorio de Acústica de la Universidad de Aalborg. Todos los estímulos fueron presentados a través de audífonos (Beyerdynamic DT990). Todos los experimentos se realizaron con un programa hecho en MATLAB que controla la presentación de estímulos y registra las respuestas de los sujetos a través de una interfaz gráfica.

2.1. Objetivos

El objetivo principal de esta investigación es analizar los efectos de la compresión dinámica en la percepción de grabaciones musicales. A continuación se presentan los objetivos específicos de cada uno de los tres experimentos:

- **Exp1:** Investigar la relación entre distintas dimensiones perceptuales y compresión dinámica de estímulos musicales. Basados en estudios previos [3] y en un cuestionario que investigó características que sujetos consideran más importantes en la determinación de calidad de una grabación [4], las dimensiones psicoacústicas elegidas para el experimento fueron: profundidad (depth), claridad (clarity), sonidos molestos (disturbing sounds) y calidad (quality).
- **Exp2:** Corroborar los resultados de las evaluaciones directas del Exp1 con respecto a juicios de calidad. La idea es que al comparar los estímulos en pares, se puede determinar la proporción de respuestas que son al azar y no corresponden a juicios sistemáticos por parte de los sujetos, eliminando parte de la incertidumbre del proceso de evaluación y emisión de juicio asociada a los sujetos.
- **Exp3:** Expandir la cantidad de niveles de compresión dinámica, utilizada en Exp1 y Exp2, incluyendo niveles de compresión más extremos e investigar como varían juicios de preferencia al aumentar los efectos de la compresión dinámica.

2.2. Estímulos

Cuatro grabaciones musicales de distintos géneros fueron obtenidas de un estudio de grabación antes del proceso de masterización [5]. Todas las grabaciones eran mezclas estéreo en formato digital (frecuencia de muestreo $f_s = 48$ kHz, y resolución por muestra de 24 bits). Los géneros musicales incluían: 1) una banda de rock (Rock); 2) dos flautas travesas (Flauta); 3) una guitarra clásica (Guitarra); 4) un violín tocando música folclórica (Violín).

Para Exp1 y Exp2, los estímulos fueron procesados utilizando el sistema Finalizer 96k de TC-electronics. El Finalizer 96k tiene un compresor de tres bandas con la posibilidad de ajustar independientemente para cada banda el umbral, proporción de compresión, ataque y liberación (release) de la compresión. Para Exp3, sólo se utilizaron dos estímulos (Rock y Flauta) y fueron procesados con el sistema Gold-Channel de TC-electronics. El Gold-Channel es un amplificador de micrófonos y de señales de línea, que contiene un compresor de una sola banda, con la posibilidad de ajustar todos los parámetros mencionados anteriormente. A diferencia del Finalizer 96k, el Gold-Channel no introduce ganancia automática para compensar por la pérdida de energía debido a la compresión. El Finalizer 96k maximiza la salida del compresor para ocupar al máximo el rango dinámico del sistema, de esta forma las señales procesadas por el Finalizer 96k siempre tenían una sonoridad más alta que la señal original, y las señales procesadas con el Gold-Channel tenían una sonoridad más baja que la señal sin procesar.

Para investigar solamente los efectos de la compresión dinámica, todos los estímulos fueron normalizados con respecto a la sonoridad de las señales no comprimidas. La normalización se obtuvo mediante un proceso iterativo, calculando la fracción entre la sonoridad de la señal sin compresión y la señal comprimida. En cada iteración la sonoridad de la señal comprimida es aumentada (o disminuida) multiplicando la señal (amplitud en tiempo) por la fracción de las sonoridades. El proceso continúa hasta llegar a una fracción de sonoridad igual o menor a 0.01. La sonoridad de las señales fue calculada según el procedimiento descrito en la recomendación ITU-R BS.1770-3 [6].

Para Exp1 y Exp2 se utilizaron 3 niveles de compresión distintos obtenidos fijando la proporción de compresión a 2.5:1 y variando el umbral desde: 0 dBFS, señal original sin compresión; -12 dBFS; y -24 dBFS. Los valores de ataque y liberación del compresor se dejaron en los valores nominales del sistema. Para Exp3 se utilizaron 5 niveles distintos de compresión obtenidos fijando el umbral en -24 dBFS y variando la proporción de compresión desde: 1:1, señal original sin compresión; 1.6:1, 2:1, 8:1, ∞ :1 máxima compresión o limitador (limiter). El ataque y liberación del compresor se dejaron en los valores nominales del sistema. Para las evaluaciones directas del Exp1 se utilizaron segmentos de 15 s, y para las comparaciones pareadas segmentos de 5 s.

El lado izquierdo de la Figura 2 muestra los efectos de la compresión dinámica en las señales del Exp1 (cuatro géneros, tres niveles de compresión, 15 s de duración). El lado derecho de la Figura 2 muestra las señales utilizadas en Exp2 y la parte inferior de la Figura 2 muestra las señales utilizadas en el Exp3. La figura muestra histogramas de los niveles equivalentes de presión sonora calculados cada 0.1 s y normalizados al nivel RMS de cada señal. Cada recuadro de la Figura 2 muestra las señales para cada género musical utilizado y los niveles de compresión se muestran con los colores.

2.3. Sujetos

Todos los sujetos que participaron en los tres experimentos lo hicieron de forma voluntaria sin recibir ningún tipo de compensación. En Exp1 participaron 20 sujetos (9 mujeres y 11 hombres)

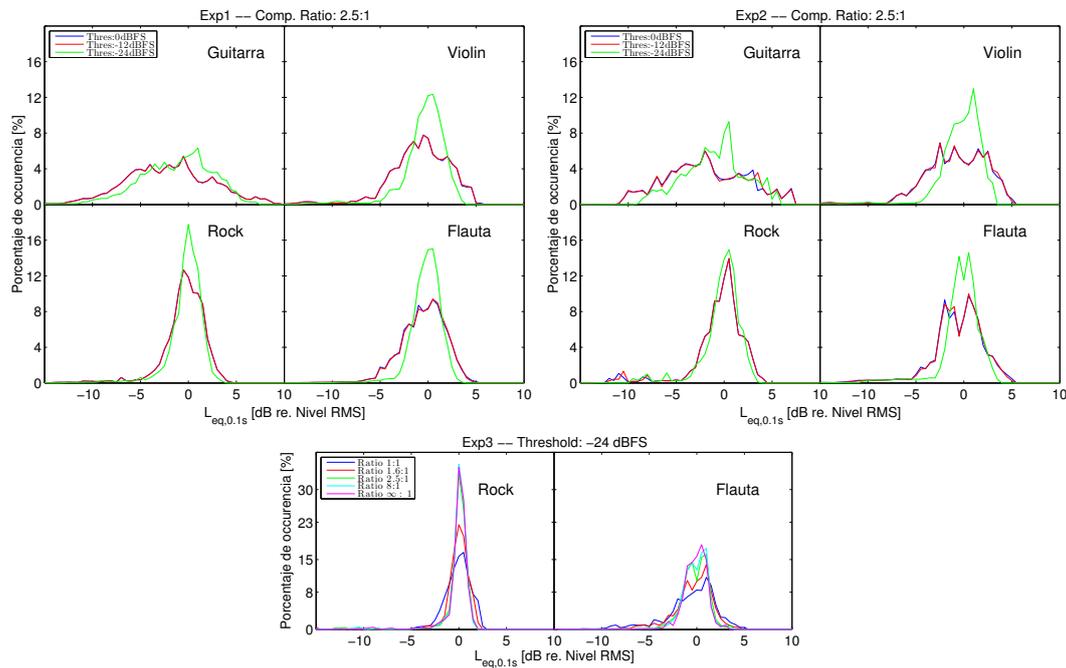


Figura 2: Histogramas de niveles equivalentes de presión sonora calculados con una constante de tiempo de 0.1s ($L_{eq,0.1s}$). Lado izquierdo, estímulos del Exp1. Lado derecho, estímulos del Exp2. Parte inferior, estímulos del Exp3. Cada recuadro en la figura muestra señales de un mismo género musical, y los niveles de compresión se muestran con los colores.

con audición normal y edades entre 18 y 26 años, excepto un sujeto hombre de 50 años de edad. Para el Exp2 participaron 8 sujetos hombres con audición normal y edades entre 28 y 42 años. Para los Exp1 y Exp2 los sujetos fueron reclutados entre estudiantes y empleados del Departamento de Sistemas Electrónicos de la Universidad de Aalborg. Para el Exp3 participaron 28 sujetos (8 mujeres y 20 hombres) con audición normal y edades entre 20 y 28 años. Sujetos del Exp3 fueron reclutados entre los participantes de los cursos de verano del Departamento de Sistemas Electrónicos de la Universidad de Aalborg. El estado de la audición de los sujetos fue evaluada mediante una audiometría de tonos puros en intervalos de octavas entre las frecuencias de 250 Hz y 8 kHz utilizando el método de límites con un audiómetro profesional (Madsen Orbiter 922 2 Clinical Audiometer). Todos los sujetos tenían niveles audiométricos mejores que 25 dBHL.

3. RESULTADOS

3.1. Evaluación Directa de Atributos Psicoacústicos (Exp1)

Los resultados del Exp1 se muestran en la Figura 3. Los resultados están presentados como el promedio, a través de sujetos, de las evaluaciones para cada atributo psicoacústico y género musical. Un ANOVA de medidas repetidas fue calculado con tres factores: género musical, nivel de compresión y atributos. Los resultados del análisis muestran efectos significativos entre evaluaciones debido a los atributos psicoacústicos ($F(3) = 19,37, p = 1,18 \times 10^{-8}$) y debido a los géneros musicales ($F(3) = 3,661, p = 0,0178$), pero con una interacción significativa entre am-

bos factores ($F(6) = 2,821, p = 0,00291$). Al comparar las evaluaciones obtenidas para “sonidos molestos” con los otros tres atributos, se puede ver que la diferencia significativa se debe al carácter negativo del atributo. La diferencia significativa correspondiente a los géneros musicales se debe a las evaluaciones correspondientes al estímulo de “Rock”. En los tres atributos positivos (profundidad, claridad y calidad), la señal de “Rock” fue evaluada más baja que todos los otros géneros musicales. El nivel de compresión no contribuye con diferencias subjetivas en ninguna de las evaluaciones.

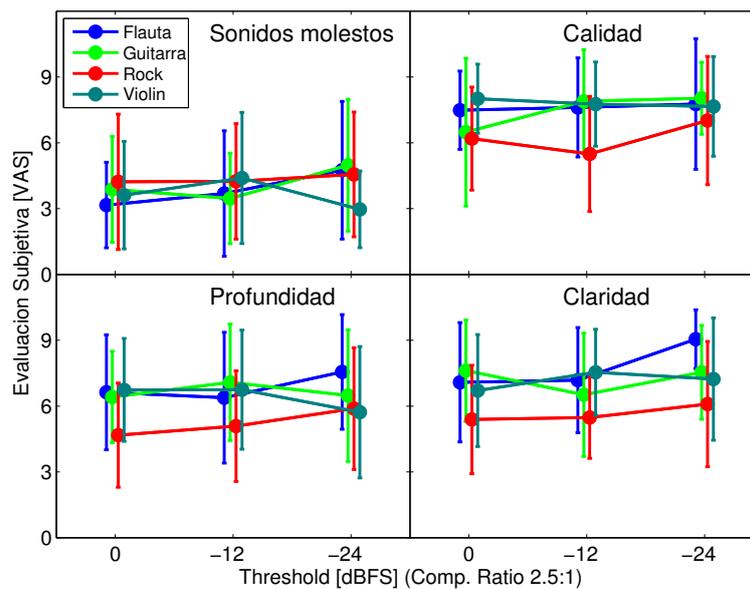


Figura 3: Evaluaciones subjetivas de las cuatro dimensiones psicoacústicas evaluadas en Exp1. Los símbolos representan el promedio de las evaluaciones de los 20 sujetos que participaron en el experimento. Las barras de error muestran $\pm DS$ (desviación estándar). Los cuatro géneros musicales se muestran con los distintos colores.

3.2. Comparaciones Pareadas de Calidad (Exp2)

Los resultados del Exp2 se muestran en la Figura 4. La figura muestra las escalas proporcionales normalizadas derivadas de las comparaciones pareadas utilizando la matriz de preferencias y un modelo probabilístico de opciones, de acuerdo al procedimiento descrito en [7]. La figura muestra 5 escalas normalizadas al valor de calidad obtenido para la señal sin compresión (Threshold 0 dBFS), una para cada género musical (símbolos de color) y una escala incluyendo todos los géneros musicales (símbolos negros) las barras de error muestran el error estándar del valor estimado para la escala a través del método de máxima probabilidad (maximum likelihood estimation). Los resultados del Exp2 confirman los resultados de las evaluaciones directas de calidad del Exp1, es decir, no hay diferencias entre las evaluaciones de calidad para los tres niveles de compresión utilizados en Exp1 y Exp2.

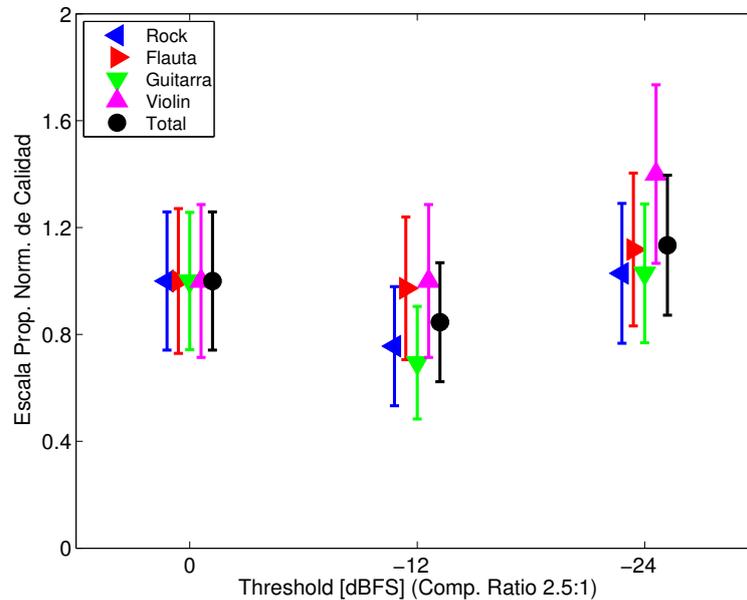


Figura 4: Escalas de proporción normalizadas de calidad, derivadas de las comparaciones pareadas del Exp2. Los resultados están basados en las respuestas de 8 sujetos y derivadas utilizando un modelo probabilístico de opciones. Las barras de error muestran el error estándar de las estimaciones. Todos los valores en la figura están normalizados al valor de las señales sin compresión (Threshold 0 dBFS)

3.3. Comparaciones Pareadas de Preferencia (Exp3)

Los resultados del Exp3 se muestran en la Figura 5, derivados de la misma forma que los resultados del Exp2 (Figura 5). Para este experimento se optó por reducir la cantidad de géneros musicales y aumentar los niveles de compresión dinámica. Los resultados muestran que a partir de una proporción de compresión (compression ratio) de 2.5:1 los sujetos comienzan sistemáticamente a preferir las señales sin compresión. También se puede apreciar como la preferencia por la señal de flautas decrece más rápido al aumentar el nivel de compresión que para la señal de rock.

4. CONCLUSIÓN

En el presente estudio, distintos niveles de compresión aplicados a cuatro géneros musicales fueron comparados subjetivamente por un grupo de personas. A niveles bajos de compresión los sujetos no fueron capaces de percibir diferencias entre los niveles de compresión. Esto significa que para la mayoría de los sujetos que participaron en estos experimentos las diferencias en rango dinámico de las señales no eran suficientes para producir cambios significativos en evaluaciones de calidad, profundidad, claridad o sonidos molestos. Inspección de la Figura 2, muestra que los niveles de compresión utilizados en los Exp1 y Exp2 (Figura 2), no produjeron cambios sustantivos en el rango dinámico de la señal, solamente el mayor nivel de compresión (Comp. Ratio 2.5:1, Threshold -24 dBFS) muestra una reducción del rango dinámico y normalización de los niveles de presión sonora al valor RMS de la señal (Figura 2, 0 dB). Considerando que los efectos de la compresión dinámica en la señal fueron pequeños los resultados de los experimentos Exp1

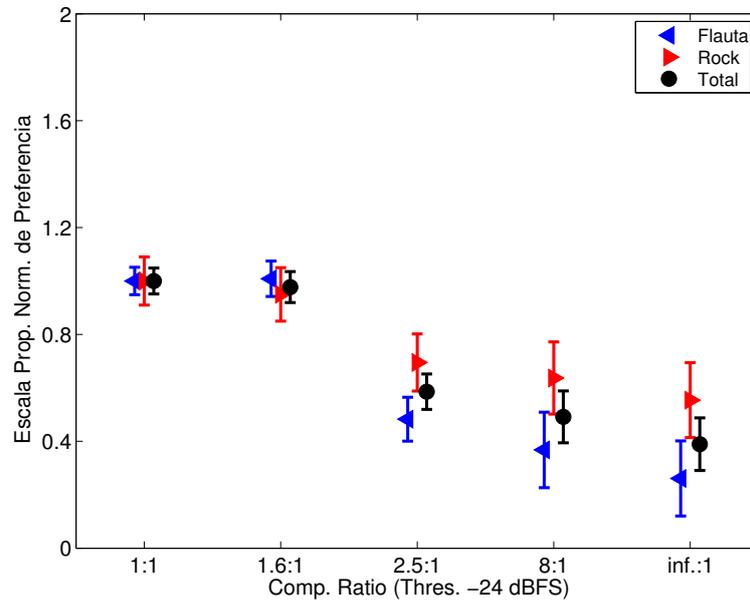


Figura 5: Escalas de proporción normalizadas de calidad, derivadas de las comparaciones pareadas del Exp3. Los resultados están basados en las respuestas de 28 sujetos y derivadas utilizando un modelo probabilístico de opciones. Las barras de error muestran el error estándar de las estimaciones. Todos los valores en la figura están normalizados al valor de las señales sin compresión (Comp. Ratio 1:1)

y Exp2 no son sorprendidos. Es necesario mencionar que se escogieron niveles de compresión que no produjeran artefactos audibles en la señal. Esta restricción redujo las posibilidades que ofrecía el compresor.

Para el Exp3 se utilizaron mayores niveles de compresión logrando evocar diferencias significativas en los juicios de preferencia de los sujetos (Figura 5). En este experimento también se puede observar una diferencia debida al género musical. Para la señal de Rock los sujetos tienden a tolerar mayores niveles de compresión que en el caso de la Flauta. Esto se debe en parte a que música rock tiene un menor rango dinámico que música clásica o instrumentos acústicos. Es más, distorsión es una parte integral de la música rock y artefactos producidos por compresión dinámica excesiva tienden a ser ocultados en la música.

Los resultados presentados aquí muestran que al eliminar diferencias en sonoridad producidas por compresión dinámica, la mayoría de los sujetos que participo en Exp1 no pudo escuchar diferencias en atributos psicoacústicos a través de evaluaciones directas en una escala de analogía visual (Exp1, Figura 3). Estos resultados fueron confirmados con un pequeño grupo de sujetos a través de comparaciones pareadas (Exp2, Figura 4). Diferencias en preferencia (Exp3, Figura 5) fueron más evidentes, pero solo para niveles de compresión dinámica, mayores que: Comp. Ratio. 2.5:1; Threshold -24 dBFS.

La sensibilidad del sistema auditivo depende no sólo de la frecuencia, sino también del nivel de sonido. A niveles más elevados las diferencias de sensibilidad entre frecuencias disminuye, dándole al sistema auditivo una respuesta más plana en frecuencia con respecto a la sensibilidad cercana al umbral auditivo. Al disminuir las diferencias de nivel entre los componentes de una señal de audio, como al aplicar compresión dinámica, se comienzan a igualar las intensidades

de las frecuencias de las señales produciendo una señal que se percibe más intensa y de mayor sonoridad. Considerando los resultados de estos experimentos, juicios de preferencia y calidad se mantienen y no decrecen con poca compresión, esto puede indicar que existe cierta tendencia a aceptar señales comprimidas. Considerando el uso masivo de reproductores musicales en lugares con altos niveles de ruido, cierto grado de compresión ayuda a mantener audibilidad en pasajes musicales de baja intensidad. Esto podría explicar, por lo menos en parte, la tendencia a aceptar un cierto grado de compresión dinámica en grabaciones musicales.

5. AGRADECIMENTOS

Los autores quieren agradecer a los alumnos del Colégio de Ingenieros ECE Paris: Marina Batros, Louise Brasseur, Zakaria Elamdizi, Antoine Seure and Charles Verron, por su ayuda en realizar el experimento 3. También nos gustaría agradecer a todos los voluntarios que nos facilitaron sus oídos y su tiempo participando en estos experimentos.

6. REFERENCIAS BIBLIOGRÁFICAS

- [1] E. Vickers, "The Loudness War: Background, Speculation and Recommendations", AES 129th Convention, San Francisco, CA, USA, November 4-7, (2010)
- [2] J. Hjørtkjær & M. Walther-Hansen, "Perceptual Effects of Dynamic Range Compression in Popular Music Recordings", J. Audio Eng. Soc., 62(1/2), 37-41, (2014).
- [3] N. B. H. Croghan & K. H. Arehart, "Quality and loudness judgments for music subjected to compression limiting", J. Acoust. Soc. Am., 132(2), 1177-1188 (2012)
- [4] M. B. Borup & P. Zabbal, "The role of compression in the subjective evaluation of music quality", M.Sc. Thesis in Acoustics and Audio Technology, Aalborg University, Denmark, September (2014).
- [5] R. Langvik, Recording Studio and Music Label: Ein Mangfaldig Kar
- [6] Recomendación UIT-R BS.1770-3, "Algoritmos para medir la sonoridad de los programas radiofónicos y el nivel de cresta de audio real", Unión Internacional de Telecomunicaciones, (2012)
- [7] F. Wickelmaier & C. Schmid, "A Matlab function to estimate choice model parameters from paired comparison data", Behav Res Meth Instrum Comput, 36(1), 29-40 (2004).