

## RUIDO MODULADO VERSUS MLS PARA LA MEDIDA DE PARÁMETROS DE INTELIGIBILIDAD: PRIMERA APROXIMACIÓN

PACS: 43.55.Hy

Teófilo Zamarreño, Miguel Galindo, Sara Girón  
Departamento de Física Aplicada II  
E. T. S. de Arquitectura - Universidad de Sevilla  
Avda. Reina Mercedes 2  
41012 Sevilla  
Tel: 954 556 612  
Fax: 954 557 892  
E-mail: teofilo@cica.es

### ABSTRACT

This paper is a first approximation to compare two different methods of measuring Modulation Transfer Functions (MTF) and the corresponding derived STI-RASTI indices to assessing intelligibility. The first method uses as test signal a band filtered noise (500 and 2000 Hz) amplitude modulated. The second one uses MLS signals to obtain the impulse response of the system; further post processing allows calculation of MTF from this response.

### RESUMEN

Esta comunicación pretende ser una primera aproximación para comparar dos métodos diferentes para medir la Función de Transferencia de Modulación (FTM) y los índices STI y RASTI, de ella derivados, para calificar la inteligibilidad de un canal de transmisión. El primer método utiliza como señal de prueba un ruido filtrado (bandas de octava de 500 y 2000 Hz) modulado en amplitud. El segundo utiliza señales MLS para obtener la respuesta al impulso y calcula las FTM mediante el adecuado posprocesado de esta respuesta.

### ANTECEDENTES TEÓRICOS

Los índices STI (Speech Transmisión Index) y RASTI (Rapid Speech Transmisión Index) fueron propuestos por Houtgast y Steeneken [1]-[5] para valorar la inteligibilidad del discurso hablado; siendo el segundo una versión reducida del primero. Ambos se calculan a partir de la Función de Transferencia de Modulación (FTM), definida por el cociente entre la amplitud de modulación de la intensidad de la señal percibida y la correspondiente a la señal emitida:

$$m(F) = \frac{m_{per}}{m_{emi}} ;$$

los valores de  $m(F)$  se obtienen para cada frecuencia de modulación  $F$ , en incrementos de 1/3 de octava, comprendidas entre 0.63 Hz y 12.5 Hz (14 en total), que cubren el rango de las

modulaciones de la voz humana. Por otro lado, la señal portadora es filtrada en bandas de octava en el intervalo comprendido entre 125 Hz y 8 kHz. De este modo se dispone de  $7 \times 14 = 98$  valores de  $m(F)$ . Para el cálculo del índice STI, estos factores de reducción de la modulación se convierten en relaciones señal/ruido aparentes y éstas se ponderan y promedian adecuadamente [5] para obtener los valores STI. La Fig. 1 muestra un esquema posible para calcular las FTM. El método está implementado, por ejemplo, en el programa STI de Lexington School for the Deaf [6], que funciona sobre plataforma PC con Windows y tarjeta de sonido (se puede descargar libremente de la red). La señal excitadora utiliza como portadora un ruido blanco de banda ancha modulado por señales sinusoidales de baja frecuencia. Cada medida requiere unos dos minutos, para aplicar al sistema, sucesivamente, todas las frecuencias de modulación.

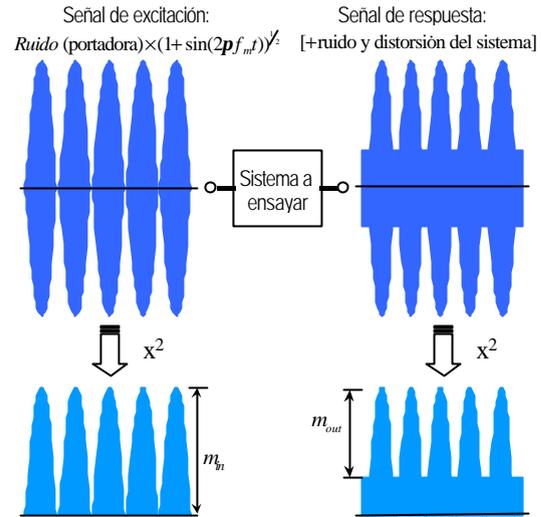


Fig. 1.- Esquema conceptual de la FTM.

Los propios Houtgast y Steeneken pusieron de manifiesto [4] que no era necesario obtener un conjunto de FTM tan detallado para calificar la inteligibilidad de espacios cerrados; de hecho definieron un parámetro reducido, el RASTI. El procedimiento es muy similar, pero se reduce el número de bandas para la portadora (500 Hz y 2 kHz) y el número de frecuencias de modulación (4 para la banda de 500 Hz y 5 para la de 2 kHz). Está implementado, por ejemplo, en el equipo Speech Transmisión Meter de Brüel&Kjær 3361 [7] y, más recientemente, la firma Brüel Acoustics ha desarrollado, para uso propio según ellos, el equipo RASTI 44BA basado en un concepto similar con algunas modificaciones significativas [8].

Una alternativa al procedimiento de la Fig. 1 para determinar los valores de  $m(F)$  fue propuesta por Schroeder [10] al indicar que se pueden obtener directamente a partir de la respuesta al impulso  $h(t)$ , siempre que el sistema sea lineal, pasivo e invariante con el tiempo. Para ello,  $h(t)$  ha de filtrarse previamente en bandas de octava para seleccionar la banda portadora,  $f$ . A partir de la señal filtrada  $h_f(t)$  podemos calcular  $m(F)$  mediante la expresión:

$$m(F) = \frac{\int_0^{\infty} h_f^2(t) \cdot \exp(-j2\pi ft) dt}{\int_0^{\infty} h_f^2(t) dt};$$

Para obtener la respuesta al impulso podemos utilizar una señal impulsiva (detonación, disparo...), pero ello tiene el inconveniente de que no podemos contabilizar el efecto del ruido de fondo, dado que el espectro de la señal excitadora no se corresponde con el espectro normalizado de la voz definido por Houtgast y Steeneken [1]-[5] ni con los propuestos en [9], que se muestran en la tabla 1, donde los valores de la voz están normalizados respecto del valor del espectro de referencia en la banda de 250 Hz y la función de transferencia del filtro

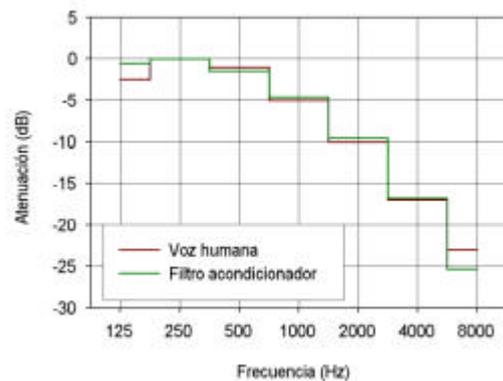


Fig 2.- Espectro de referencia de la voz y del filtro de acondicionamiento.

**Tabla1.-** Espectros IEC de la voz y función de transferencia del filtro utilizado en las medidas MLS.

Octava	Espec. referencia	Voz masculina	Voz femenina	Filtro de medida
125	-2,5	2,9	-200,0	-0,55
250	0,0	2,9	5,3	0,00
500	-1,0	-0,8	-1,9	-1,43
1000	-5,0	-6,8	-9,1	-4,60
2000	-10,0	-12,8	-15,8	-9,52
4000	-17,0	-18,8	-16,7	-16,80
8000	-23,0	-24,8	-18,0	-25,30

está normalizada respecto de su valor en la misma octava.

Este inconveniente se puede soslayar si se utilizan secuencias de máxima longitud como señal de excitación para determinar la respuesta al impulso, tal como puso de manifiesto Rife [11]. En tal caso es posible intercalar un filtro antes de alimentar la señal MLS al amplificador para acondicionar su espectro, la función de transferencia del que nosotros hemos utilizado se muestra en la última columna de la tabla 1 y se compara con el espectro de referencia de la voz humana en la Fig. 2.

Además de los requisitos generales para utilizar MLS han de tenerse en cuenta algunas recomendaciones para obtener medidas precisas de las FTM:

- Se ha de medir una respuesta al impulso suficientemente larga (al menos 1 s) con un ancho de banda mínimo de 12 kHz (3 kHz para medidas RASTI).
- No promediar para obtener la respuesta al impulso. Ello reduciría artificialmente la relación señal ruido (la relación S/R aumenta +3 dB cada vez que duplicamos le número de periodos LMS promediados). Si bien es posible sacar partido de esta situación para evaluar la incidencia del ruido de fondo frente a otros factores de degradación de la modulación.
- Utilizar una fuente acústica que reproduzca la señal MLS con la directividad de la cabeza humana si no se utiliza sistema electroacústico. En caso de utilizarlo, la señal se liberará a la audiencia a través del mismo.
- Si el tiempo de reverberación es muy grande (unos 7 s en nuestro sistema) debe reducirse el ancho de banda para evitar errores *time aliasing*. Reducir a la mitad el ancho de banda permite que se duplique el tiempo de reverberación para obtener medidas fiables (si bien se prescindirá de las correspondientes bandas al calcular el STI).

## ADQUISICIÓN DE DATOS Y RESULTADOS OBTENIDOS

Los datos que presentamos a continuación se tomaron en una iglesia barroca abovedada con planta de cruz latina y capillas laterales (ver Fig. 3) y que tiene los tiempos de reverberación mostrados en la Fig. 4. Los valores de EDT y  $T_{20}$  se han obtenido utilizando el analizador basado en PC MLSSA de DRA Laboratories [12], promediando 8 periodos MLS en cada medida para mejorar la relación señal ruido. Para capturar la señal se utilizó un micrófono multipatrón AT4050/CM5 de Audio-technica, configurado en su patrón omnidireccional, con el preamplificador Lab 1 de Earthworks. La elección del micrófono está justificada por el hecho de que, a partir de  $h(t)$  se derivan otros parámetros de calidad acústica, entre ellos la eficiencia lateral, para la cual se hace necesario usar un micrófono con patrón figura de ocho. De este modo el mismo micrófono sirve para hacer todos los registros cambiando el

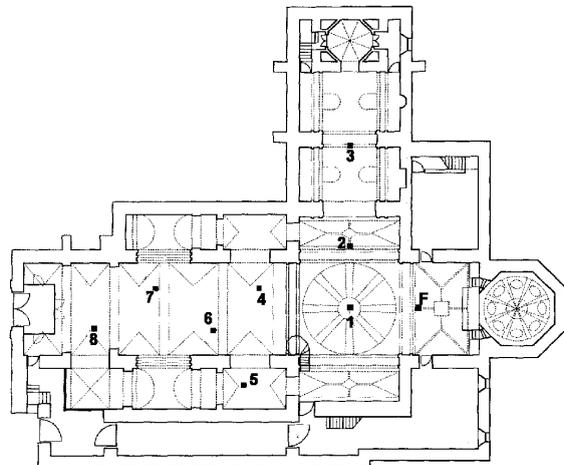


Fig 3.- Planta de la iglesia en la que se ha medido.

patrón de direccionalidad simplemente accionando un interruptor.

Para valorar la inteligibilidad se han adquirido datos utilizando dos equipos diferentes: en primer lugar el equipo B&K-3361 que utiliza la técnica del ruido modulado. El emisor se situó siempre en el lugar indicado con **F** en la Fig. 3. Se ajustó la emisión al nivel Ref+10 dB; esto implica un nivel de presión, a 1 m del emisor, de 69 dB en la banda de 500 Hz y 60 dB en la de 2kHz ( $\approx 67$  dB(A)). El periodo de cada medida se fijó en 16 s. Nuestro grupo está familiarizado con esta técnica con la que viene trabajando desde hace ya varios años.

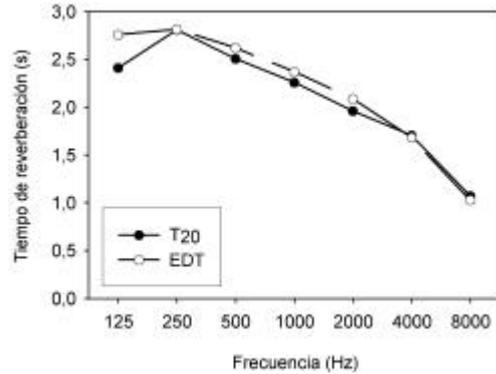


Fig 4.- Tiempo de reverberación de la iglesia.

En segundo lugar se utilizó el analizador MLSSA que utiliza la técnica MLS. La envolvente del espectro de esta señal está descrita por  $sen^2(x)/x^2$ . MLSSA ajusta la frecuencia de muestreo al triple del ancho de banda solicitado, de modo que, en tales condiciones, puede considerarse prácticamente plano su espectro. Por ello, la señal generada por el analizador es acondicionada por un filtro CWF-1 de One on One Technical Products antes de ser alimentada al amplificador. La función de transferencia del citado filtro se ha medido utilizando el propio analizador MLSSA y el resultado, en bandas de octava, se presenta en la Fig. 2. La salida del amplificador alimenta la fuente omnidireccional B&K-4296 colocada también en el punto **F** de la Fig. 3, y a la misma altura, que en el ensayo anterior. La señal se ajustó para producir un nivel de 67 dB(A) a 1 m de la fuente (según recomendación del fabricante) medido con un sonómetro B&K-2231.

Se utilizó el micrófono AT4050/CM5, por las razones ya indicadas, configurado en su patrón omnidireccional. La señal de salida de su preamplificador se alimenta a la entrada de la tarjeta MLSSA para obtener la respuesta al impulso. Al hacer la medida no se promedia para no aumentar la relación señal ruido artificialmente. Una vez obtenida  $h(t)$  el software del analizador permite calcular los valores  $m(F)$  para cada frecuencia de modulación y banda de octava, a partir de los cuales se obtiene el índice STI para el espectro de referencia (al que corresponden los datos aquí utilizados), al espectro de la voz masculina o al de la voz femenina, tras aplicar las correcciones adecuadas. Así mismo permite calcular el índice reducido RASTI.

La Fig. 5 muestra los resultados, para cada punto, de los valores STI-RASTI medidos con MLSSA y los obtenidos utilizando ruido modulado (B&K-3361). A medida que no alejamos de la fuente aumenta la discrepancia entre los valores de MLSSA y los de B&K. Los valores de MLSSA tienden a infravalorar la inteligibilidad respecto de los proporcionados por B&K.

La Fig. 6 muestra el grado de correlación lineal entre los valores medidos con MLSSA (RASTI-STI) y los medidos con B&K. Es cierto que son pocos datos para sacar conclusiones definitivas,

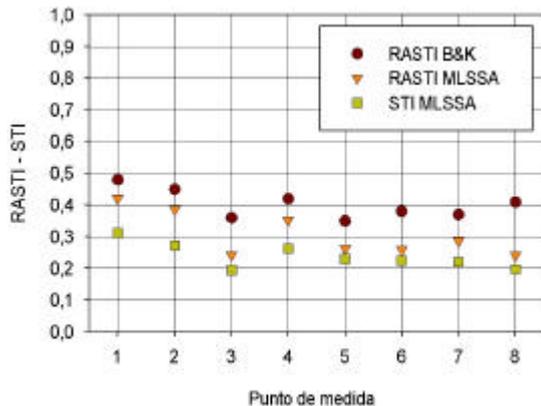


Fig 5.- Valores STI - RASTI para cada punto.

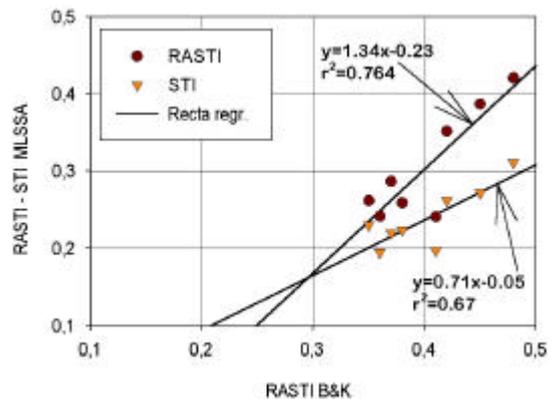


Fig 6.- Valores RASTI-STI dados por MLSSA versus valores medidos con B&K-3361.

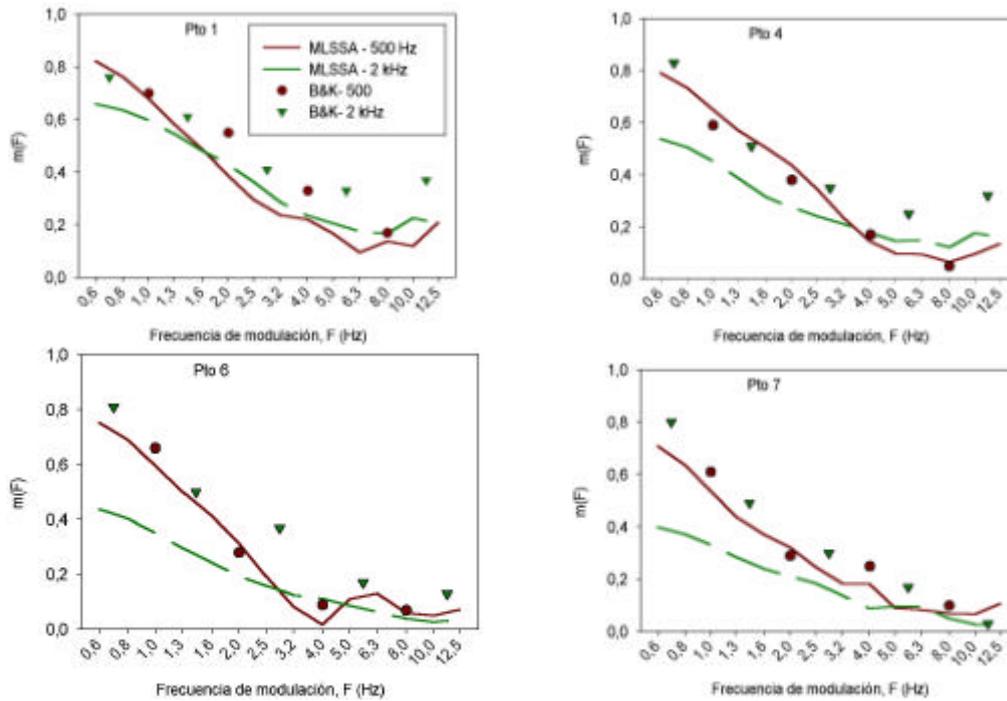


Fig 7.- Comparación de las FTM completas (MLSSA) y las reducidas (B&K-3361).

pero la recta de regresión para los valores RASTI-MLSSA presenta una desviación significativa respecto de la esperada (pendiente unidad y ordenada en el origen nula). La discrepancia aumenta para los valores STI, aunque aquí puede admitirse una justificación: probablemente los factores que degradan la modulación son múltiples y las bandas de 500 y 2000 Hz no son suficientes para dar cuenta de todos ellos.

En la Fig. 7 se pretende comparar visualmente hasta qué punto la descripción simplificada de las FTM dada por el B&K-3361 (4 frecuencias de modulación en la banda de 500 Hz y 5 en la de 2kHz) es aceptable frente a la descripción completa de MLSSA. Puede observarse que, sobre todo en la banda de 2 kHz, los resultados B&K dan valores bastante más elevados, sobre todo para los puntos más alejados de la fuente. Se han representado los datos correspondientes a los puntos de la nave central.

La Fig. 8 muestra las FTM completas para todas las bandas de octava para dos puntos: uno próximo a la fuente (el 1) y otro alejado y situado bajo el coro (el 8). Mientras que para el punto cercano su forma corresponde con la degradación debida a la reverberación, salvo para las bandas de 4 y 8 KHz, en el punto 8 se hace patente la degradación debida al ruido de fondo y la presencia de reflexiones aisladas retardadas que también degradan la modulación.

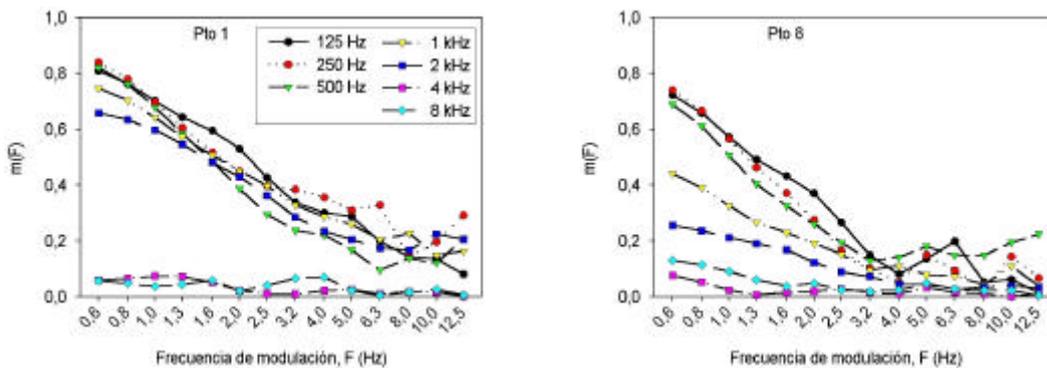


Fig 8.- Funciones de transferencia de modulación completas para los puntos 1 y 8.

Para valorar la influencia del ruido de fondo sobre las FTM, procesamos las respuestas al impulso que se habían obtenido promediando 8 periodos MLS (medida de otros parámetros acústicos), para obtener los valores de RASTI. Además de promediar, en este caso el nivel no se ajustó a los 67 dB(A) (de hecho era mucho mayor aunque no se midió) ni se insertó el filtro CWF-1 para acondicionar la señal. Por todo ello, la influencia del ruido de fondo debe ser despreciable. El resultado se muestra, correlacionado con los valores RASTI – B&K, en la Fig. 9. Nos sorprendió el alto grado de correlación. Disponíamos de respuestas al impulso promediadas correspondientes a otros espacios eclesiales, de modo que ampliamos con ellos el conjunto de datos con los valores RASTI y tratamos de correlacionarlos con los valores medidos con ruido modulado. El resultado, para todos los datos disponibles hasta el momento, aparece en la Fig. 9. La correlación es similar a la obtenida inicialmente y no tenemos una explicación plausible para esto, por el momento.

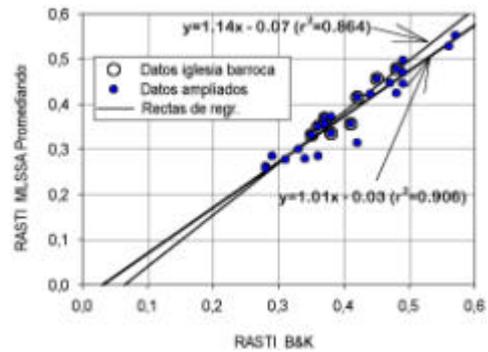


Fig 9.- Valores RASTI obtenidos a partir de respuestas promediadas.

## CONCLUSIONES Y TRABAJO FUTURO

Esta primera aproximación sugiere que se han de realizar más ensayos para tratar de cubrir un rango mayor de valores de los parámetros. Así mismo, trataremos de evidenciar el efecto de la direccionalidad de la fuente y las características del micrófono utilizado y el nivel global frente a la fuente.

Es también evidente que la descripción reducida de las FTM ha de tomarse con precaución, y que su descripción completa puede poner en evidencia diferencias relativas entre diferentes posiciones de un auditorio que no se manifiestan de forma tan clara con la descripción reducida.

## BIBLIOGRAFÍA

- [1] Steeneken, H.J.M. and Hougast, T. "The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility", *Acustica* 28, 66-73, 1973.
- [2] Hougast, T. and Steeneken, H.J.M. "A Physical Method for Measuring Speech Transmission Quality", *J. Acoust. Soc. Amer.* 67, 318-326, 1980.
- [3] Hougast, T, Steeneken, H.J.M. and Plomp R. "Predicting Speech Intelligibility in Rooms from the Modulations Transfer Function. I. General Room Acoustics", *Acustica* 46, 66-73, 1973.
- [4] Hougast, T. and Steeneken, H.J.M. "A Multi-Language Evaluation of the RASTI-Method for Estimating Speech Intelligibility in Auditoria", *Acustica* 54, 185-199, 1984.
- [5] Hougast, T. and Steeneken, H.J.M. "A Review of the MTF Concept in Room Acoustics and it's Use for Estimating Speech Intelligibility in Auditoria", *J. Acoust. Soc. Amer.* 67, 1060-1077, 1985.
- [6] <http://hearingresearch.org/STI.htm>
- [7] Manual del equipo *Speech Transmission Meter Type 3361* de Brüel&Kjær, DK-2850 Nærum, Denmark.
- [8] <http://www.bruel-ac.com> (E-mail: bruel@mclink.it)
- [9] Redraft of IEC 268-16, "Sound System Equipment", part 16: *The Objective Rating of Speech Intelligibility by the Speech Transmission Index*.

- [10] Schroeder, M. R. "*Modulation Transfer Function: Definition and Measurement*", *Acustica* 49, 179-182, 1981.
- [11] Rife D. D. "*Modulation Transfer Measurement with Maximum Length Sequences*", *Journal of the Audio Engineering Society* 40 (10), 779-790, 1992.
- [12] MLSSA, Maximun-Length Sequence System Analyzer, Reference Manual, version 10W. DRA laboratories.