



VI Congreso Iberoamericano de Acústica - FIA 2008
Buenos Aires, 5, 6 y 7 de noviembre de 2008

FIA2008-A194

Logatomos en español: una propuesta para medir la inteligibilidad

Jorge Sommerhoff^(a),
Claudia Rosas^(b),

(a) Instituto de Acústica, Facultad de Ciencias de la Ingeniería, Universidad Austral de Chile.
Campus Miraflores S/N, Valdivia, Chile. E-mail: jsommerh@uach.cl

(b) Instituto de Lingüística y Literatura, Facultad de Filosofía y Humanidades, Campus Teja,
S/N, Valdivia, Chile.

Abstract

This paper presents the procedure used to analyze the behaviour of 1200 logatoms with a CVC (Consonant + Vowel + Consonant) structure, that includes all possible Spanish phonetic CVC combinations, with data from the research Project DID UACH S-2005-50, subsidized by the Universidad Austral de Chile, in order to elaborate lists of logatoms that allow to measure the articulation percentage and it's correlation with STI.

Resumen

El presente trabajo presenta el procedimiento utilizado para analizar el comportamiento de 1200 logatomos de estructura CVC (Consonante + Vocal + Consonante), que incluye todas las combinaciones CVC fonéticamente posibles del español, con datos de un proyecto DID UACH S-2005-50, financiado por la Universidad Austral de Chile, con el objeto de elaborar listas de logatomos que permitan medir el porcentaje de articulación y su correlación con el STI.

1 Introducción

Se define la inteligibilidad del habla como una medición de la efectividad para comprender la voz hablada (Cf. Llisterri 1991: 118 y 119). Los procedimientos utilizados para ello se clasifican en dos tipos: subjetivos y objetivos. El método subjetivo consiste en contar los aciertos transcritos de un dictado de segmentos (sílabas, palabras o frases). El resultado de este procedimiento, conocido como test de articulación (Cf. Gavilán 1988:44), depende fundamentalmente de dos variables: el corpus —la lista de segmentos— y la administración de la prueba o test. (Cf. Egan, J.P. 1948: 963). El método objetivo, por su parte, utiliza máquinas, cuyos resultados se correlacionan con los obtenidos mediante el método subjetivo. Uno de los tests objetivos más utilizado corresponde al STI (Speech Transmission Index) desarrollado por Houtgast y Steeneken (1973). El STI es una técnica de medición que determina el grado de disminución que sufre la amplitud de la modulación de la señal sonora de prueba que se entrega a la sala, producto de la reverberación y ruido que existe en ella. Y esta disminución producida en el punto de recepción está relacionada con la pérdida de inteligibilidad, cuyos resultados se distribuyen en una escala que va desde 0 (ininteligibilidad absoluta) a 1 (inteligibilidad absoluta).

A continuación, la siguiente figura 1 muestra la relación obtenida por Houtgast y Steeneken (1985) en el idioma inglés entre el sistema de prueba y el porcentaje de inteligibilidad medido utilizando un listado logatomos con estructura CVC (Consonante + Vocal + Consonante) y STI.

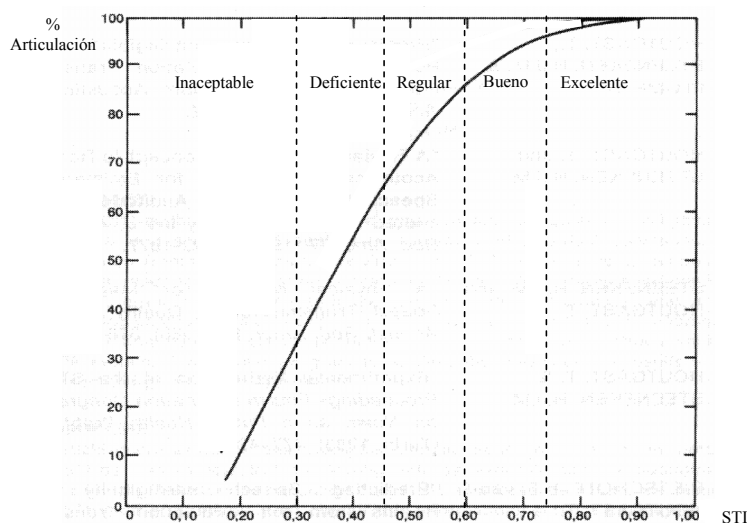


Figura 1. Relación entre prueba subjetiva de logatomos CVC en inglés y STI (Houtgast *et al.* 1985).

Dado que la prueba subjetiva de logatomos en el idioma español también es dependiente del corpus utilizado (Cf. Sommerhoff y Rosas 2007: 224), en este trabajo se elaboró un corpus ad hoc con el objeto de obtener resultados para el español, en función del STI, semejantes a los que se obtienen de las listas del idioma inglés, establecidas por Houtgast y Steenecken.

2 Metodología

2.1 El corpus

El corpus utilizado para medir la inteligibilidad está formado por 1200 logatomos de estructura CVC (Consonante + Vocal + Consonante) que incluye todas las combinaciones CVC fonéticamente posibles del español, representados en ortografía normal.

2.2 Materiales y métodos

La metodología general seguida para la realización del objetivo propuesto consideró tres aspectos: constitución del corpus inicial; administración de la prueba; y obtención del corpus final. Cada uno, a su vez, consideró una serie de tareas específicas como se describe a continuación.

2.2.1 Constitución del corpus inicial

2.2.1.1 Formación del corpus ad hoc

Se construyó un corpus de 1200 logatomos de estructura CVC (Consonante + Vocal + Consonante) que incluye todas las combinaciones fonéticas CVC posibles del español.

2.2.1.2 Distribución del corpus

El corpus inicial de 1200 logatomos se dividió en 12 listas de 100 unidades cada una.

2.2.1.3 Digitalización de las listas

Las listas ordenadas fueron grabadas en una sala anecoica con una separación de 1,5 s entre cada logatomo, a las que se agregaron frases introductorias (“pongan atención”, “ahora escuchen”, “escuchen a continuación”, etc.). Finalmente, se normalizó la potencia acústica, ajustándola a la de la voz de una persona hablando en forma normal (aproximadamente 60 dB a 1 m).

2.2.2 Administración de la prueba

2.2.2.1 Selección de los auditores

Se utilizó 30 estudiantes universitarios, de ambos sexos, sanos auditivamente y elegidos al azar, los cuales se distribuyeron en 6 grupos de 5 individuos cada uno. Se les explicó el procedimiento y propósito de la tarea, como se detalla más adelante.

2.2.2.2 Selección de los recintos

Se eligió recintos con un STI entre 0.38 y 0.7, rango de valores donde más claramente se manifiestan las diferencias de la relación entre STI y % de articulación, del corpus utilizado. Se les explicó el procedimiento y propósito de la tarea, como se detalla más adelante. Los recintos utilizados ad hoc utilizados se encuentran en Campus Miraflores y en el Campus Teja de la Universidad Austral de Chile.

2.2.2.3 Selección de los instrumentos de medición

Para medir el STI se utilizó el medidor NTI Acoustilyzer junto con la fuente sonora NTI Talk Box que imita la voz humana (tanto en su potencia acústica como en su directividad) y que, gracias a su tarjeta Compact Flash, permite el almacenamiento de las señales que emite, en nuestro caso: STI y corpus de logatomos.

2.2.2.4 Aplicación de la prueba

Los auditores se distribuyeron en distintos puntos de recepción de las salas desde donde escribieron las palabras escuchadas, en plantillas adaptadas al propósito. En esos mismos lugares se midió el STI de acuerdo al montaje que se muestra en la figura 2.

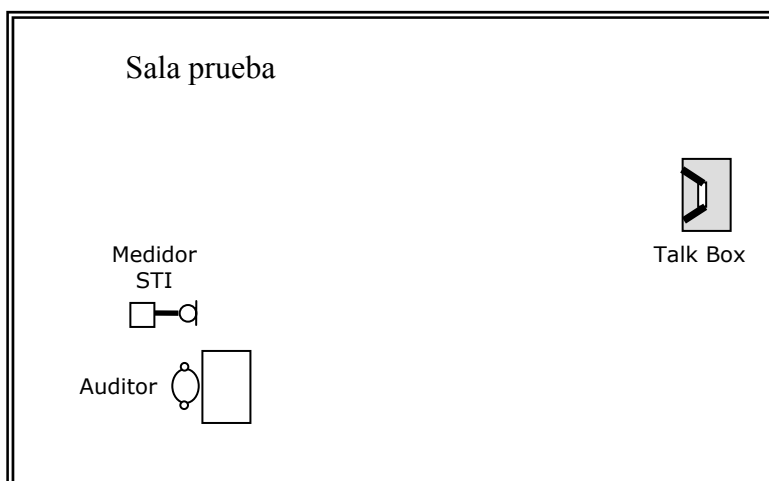


Figura 2. Montaje para las mediciones en sala de prueba.

2.2.3 Obtención del corpus

2.2.3.1 Análisis comparativo de las mediciones

Los porcentajes de inteligibilidad obtenidos en la prueba de articulación fueron pareados con los resultados del STI.

2.2.3.2 Afinación del corpus final

Se fueron eliminando sucesivamente los logatomo más difíciles de escuchar hasta lograr un nivel semejante al obtenido por Houtgast y Steenecken en la curva de relación entre STI y % de articulación, con un total de 850 logatomo.

2.2.3.3 Presentación del corpus

El corpus final obtenido se divide en 17 listas de 50 logatomo balanceadas en cuanto a criterios de composición fonética CVC y grado de dificultad observado

3 Resultados

En los casilleros de tablas 1, 2, 3, 4 y 5 se indica el porcentaje global de aciertos, considerando la totalidad de las pruebas realizadas. Este resultado es indicativo de la dificultad relativa que existe entre cada logatomo CVC de ser escuchado con claridad en distintas condiciones acústicas. Así por ejemplo, el logatomo JAB (con un 29% de aciertos) es más difícil de escucharlo en condiciones acústicamente adversas que el logatomo LUM (con un 89% de aciertos).

Los errores observados que se cometen al escuchar los logatomo pueden ser de 5 tipos, los cuales se identifican en la tabla 6.

Tabla 6. Clasificación de los errores

TIPO DE ERROR	DESCRIPCIÓN
Error 1	Escuchar incorrectamente la primera consonante de CVC
Error 2	Escuchar incorrectamente la vocal de CVC
Error 3	Escuchar incorrectamente la segunda consonante de CVC
Error 4	Escuchar incorrectamente las dos consonante de CVC
Error 5	Escuchar incorrectamente la dos consonantes y vocal de CVC

De acuerdo a esta clasificación, el porcentaje de error cometido en cada categoría de distribución del corpus seleccionado se muestra en la tabla 7. En esta tabla se evidencia que el gran porcentaje de los errores ocurre al escuchar las consonantes y, principalmente, al escuchar la segunda consonante del logatomo con estructura CVC.

Tabla 7. Distribución de los tipos de errores en porcentaje de ocurrencia

Error 1	Error 2	Error 3	Error 4	Error 5
21%	1,9%	61%	9%	7%

Este resultado es coherente con lo publicado por Peutz (1971: 23-27) donde se establece que la pérdida de articulación de las consonantes está relacionada en una forma simple con la acústica de la sala y la relación señal-ruido, y por lo tanto, puede ser utilizada como una medida confiable de la inteligibilidad del habla. También, está claro que el mayor número de errores ocurre al escuchar la segunda consonante de CVC debido a que está precedida por la primera consonante y vocal, que se mantienen en la sala debido al campo reverberante del recinto, enmascarándola más o menos según su tiempo de reverberación. De esto se concluye que las listas deben estar balanceadas fonéticamente con la segunda consonante.

La figura 3 muestra el resultado de la relación de entre el % de inteligibilidad y el STI medidos para el conjunto de los 1200 logatomos CVC, para el conjunto de 1000 logatomos de mayor acierto, para el conjunto 950 logatomos de mayor acierto, para el conjunto 900 logatomos de mayor acierto y para el conjunto 850 logatomos de mayor acierto, todos ellos en comparación con la curva de Houtgast y Steenecken.

Dado que la curva de resultado del conjunto de los 1200 logatomos está bajo la curva de Houtgast y Steenecken, se procedió a eliminar paulatinamente los logatomos con menor acierto, hasta lograr una curva con una altura semejante, como se muestra en la figura 3.

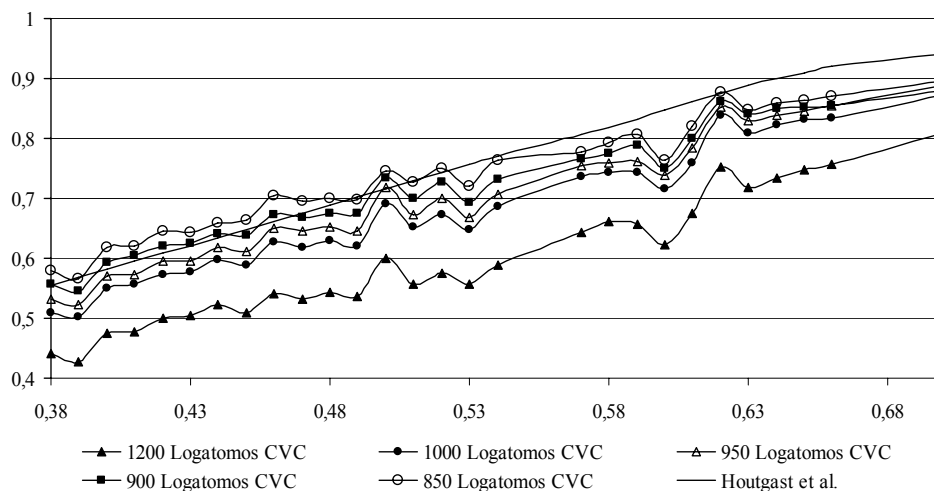


Figura 3. Relación de entre el % de inteligibilidad y el STI para distintos conjuntos de logatomos CVC de mayor acierto en comparación con la curva de Houtgast y Steenecken.

Así, se determinó que la curva que considera a los 850 logatomos con mayor acierto es la más cercana a la de Houtgast y Steenecken como se muestra en la figura 4.

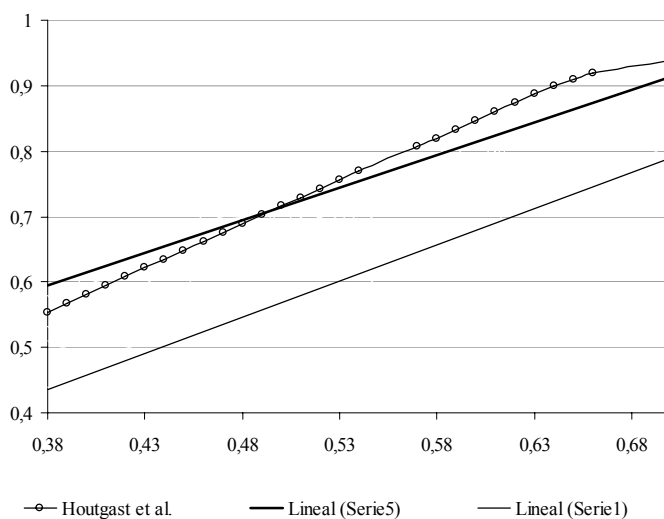


Figura 4. Curva lineal de mejor ajuste de las curvas de 1200 CVC y 850 CVC en comparación con la curva de Houtgast y Steenecken.

4 Balance

En este trabajo se ha aplicado la medición de la inteligibilidad con empleo de un corpus inicial adecuado al español panhispanico de 1200 logatomos que representan de manera exhaustiva las combinaciones posibles de estructura Consonante + Vocal + Consonante existentes en él. Los resultados de estas mediciones han sido pareados con los resultados de las mediciones de STI, junto con los porcentajes de medición subjetiva del inglés según las listas establecidas por Houtgast y Steenecken. Lo anterior ha permitido proponer un nuevo

corpus para evaluar la inteligibilidad del español de 850 logatomos que conserva la estructura CVC, clasificados en 17 listas fonéticamente balanceadas que tienen en cuenta, de acuerdo a lo constatado en este mismo trabajo, especialmente la importancia de la segunda consonante en el grado de efectividad de los resultados.

Finalmente, con el objeto de afinar esta propuesta y hacerla funcionalmente equivalente al STI para la medición acústica de los recintos, en español, se ha empezado a trabajar en la actualidad, en tareas de aplicación y evaluación del corpus diseñado.

Referencias

- Egan, FP. (1948). "Articulation testing methods". *Laryngoscope*, 58:955-991.
- Gavilán, J. (1988). "Fonoaudiología para educadores". Universidad Nacional de Educación a Distancia, Madrid, España.
- Houtgast, T. y J. M. Steeneken. (1973). "The modulation transfer function in room acoustics as a predictor of speech intelligibility". *Acustica*, 28, 1, 66-73.
- . (1985). "The Modulation Transfer Function MDT in room acoustics". *B & K Tech. Rev.*, 3, 3-29.
- Llisterri, J. (1991). "Introducción a la fonética: el método experimental". Anthropos, Barcelona, España.
- Peutz, V.M.A. (1971). "Articulation loss of consonants as a criterion for speech transmission in a room", *Journal of Audio Engineering Society*, 19, 11, 23-27.
- Sommerhoff, J.; Rosas, C. (2007). "Evaluación de la inteligibilidad del habla". *Estudios Filológicos* 42, 215-225.
- . (en preparación). "¿Existe la inteligibilidad absoluta? fonemas críticos en español".