

CARACTERIZACION ACUSTICA DE LA VARIACION FONETICA EN DISTINTOS ESTILOS DE HABLA EN ESPAÑOL APLICADA A SISTEMAS DE SINTESIS Y RECONOCIMIENTO AUTOMATICOS

Dolors Poch; Joan A. Argente; Joaquim Llisterra;
Lourdes Aguilar; Beatriz Bleuca; Carmen de la Mota
(Universitat Autònoma de Barcelona)

Josep Martí; Francesc Tarrés
(Escuela de Telecomunicaciones de La Salle, Barcelona)

Salvador Gutierrez; Manuel Iglesias; Marina Maquieira
(Universidad de León)

1. Resumen

El proyecto que aquí presentamos se inscribe en dos dominios diferentes, aunque interrelacionados, del estudio del habla: la caracterización acústica de los fenómenos de variación fonética y la tecnología de la voz. Desde el punto de vista de la descripción acústica, el proyecto se propone dar cuenta de las modificaciones que presentan las realizaciones de los distintos sonidos del español así como los patrones melódicos de dicha lengua que estén asociadas a los cambios de "estilo de habla" elaborando una serie de reglas que especifiquen dicha variación en cada "estilo". Desde el punto de vista de la tecnología de la voz, los resultados de esta caracterización fonética permitirán mejorar la calidad de los sistemas de síntesis y reconocimiento automáticos mediante la incorporación a los mismos de las reglas elaboradas en la etapa de descripción. Con el fin de valorar los resultados obtenidos mediante la utilización de dichas reglas el proyecto se propone, además, la elaboración de una serie de herramientas lingüísticas que permitan evaluar especialmente la naturalidad de los sistemas de conversión de texto a voz.

Este proyecto está financiado por la DGICYT (PB90-0704) y participan en él tres centros de investigación: La Universitat Autònoma de Barcelona; la Escuela de Telecomunicaciones de La Salle de Barcelona y la Universidad de León. Se ha iniciado en enero de 1992 y acabará en diciembre de 1994.

2. Introducción

Una de las tareas fundamentales de la investigación en fonética experimental es determinar los procesos mediante los cuales los individuos somos capaces de producir y entender un número infinito de mensajes, en una multiplicidad de situaciones de comunicación diferentes, las cuales conllevan una importante variabilidad en los mensajes emitidos. A pesar de estas variaciones, que afectan de manera importante la naturaleza acústica del enunciado emitido, los hablantes somos capaces de hacer abstracción de las mismas y reconocer la "misma" palabra en todos los casos. Una de las posibles formas de contribuir a llevar a cabo esta tarea consiste en establecer un corpus que permita el desarrollo de modelos computacionales de la variación fonética que puedan ser implementados en diversos sistemas de síntesis y reconocimiento automáticos. Dichos sistemas suelen estar basados en una serie de valores previamente fijados de los parámetros acústicos de los sonidos del habla lo cual hace que, en muchos casos, los sistemas de reconocimiento, muy especialmente, no alcancen los resultados esperados. Gunnar Fant, en 1984, llamaba ya la atención sobre estos problemas al afirmar que el problema más grave que se presenta para el desarrollo de sistemas de información basados en el habla no es tecnológico ni económico sino que proviene de que nuestro conocimiento del sistema de procesamiento del lenguaje es todavía limitado (Fant, 1984).

La mayoría de los estudios realizados hasta hace muy pocos años en fonética experimental, tanto acústica como articulatoria, se han centrado en lo que podemos denominar "habla de laboratorio": logatomas, palabras aisladas, palabras en "frases marco", etc. y todo este material ha sido recogido en las restringidas condiciones de naturalidad que impone un laboratorio de fonética. Las conclusiones que se han obtenido de los trabajos realizados sobre este tipo de material se han hecho extensivas al procesamiento del habla en general y en ellas se suelen basar las aplicaciones tecnológicas que se sitúan, en cambio, en el dominio del "habla continua": un buen ejemplo de ello son los sistemas de conversión de texto a voz. Lo cierto es que si bien el procesamiento de sílabas, palabras aisladas, palabras en "frases-marco", etc. mantiene una estrecha relación con los fenómenos que se producen en el habla "fuera del laboratorio" no podemos identificar completamente estos fenómenos. Hay que admitir, por tanto, que sabemos todavía pocas cosas acerca del habla continua y/o espontánea y que no disponemos de un sistema de reglas que permitan derivar las características fonéticas de distintos estilos de habla de las características del habla de laboratorio. Esta situación, obviamente, limita las capacidades de las aplicaciones tecnológicas.

Desde hace muy pocos años se está desarrollando, especialmente en Europa, una línea de investigación en fonética experimental que se propone la caracterización de la variación fonética asociada a los estilos de habla. Dicha orientación tiene su origen en las universidades de Estocolmo (Suecia) y Austin (Texas - EEUU) y nace de la mano de B. Lindblom quien, con su teoría de la H&H (Lindblom, 1990) sienta las bases de esta línea de trabajo. Una panorámica general de las distintas perspectivas de este campo de investigación puede encontrarse en las Actas del Workshop de ESCA (European Speech Communication Association) "Phonetics and Phonology of Speaking Styles" celebrado en octubre de 1991 (Llisterri - Poch, eds., 1991). El proyecto que aquí presentamos se inscribe en esta perspectiva del estudio del habla y se propone desarrollar estos aspectos del estudio de la lengua española.

3. Objetivos

Los objetivos que se propone alcanzar este proyecto son los siguientes:

- * Crear un corpus de habla que permita estudiar las variaciones en la pronunciación de los sonidos y en la realización de los elementos suprasegmentales en distintos estilos de habla en español;
- * Cuantificar las variaciones de los valores de los parámetros acústicos de los sonidos del español y de los componentes prosódicos de dicha lengua en función de los estilos de habla;
- * Formular una serie de reglas computacionales que especifiquen de qué forma están relacionadas unas con otras las distintas variaciones;
- * Implementar dicho sistema de reglas en un sistema de conversión de texto a voz.

4. Plan de trabajo

La organización del trabajo entre los miembros de este equipo de investigación se realiza en función de la especificidad de cada uno de los subequipos de los distintos centros:

- * La Universitat Autònoma de Barcelona y la Universidad de León elaboran y recogen el corpus de análisis, realizan el análisis acústico del mismo y elaboran conjuntamente las reglas de especificación de las variaciones. El equipo de Barcelona ha realizado ya algunos trabajos, previos a este proyecto, en el dominio del estudio de los estilos de habla que ponen de manifiesto algunos de los fenómenos de variación que aparecen en español producidos por los cambios de estilo (Poch, 1989) y que se proponen contribuir a la elaboración de una metodología de trabajo apta para

la cuantificación de las variaciones que se pretende estudiar (Harmegnies, Poch, 1991). El equipo de León, también por su experiencia previa en los dominios de la sintaxis y la semántica (Gutiérrez, 1989), se encarga especialmente de la definición del corpus de trabajo así como del estudio de las relaciones entre estructura sintáctica y prosodia.

* La Escuela de Telecomunicaciones de La Salle - Barcelona y la Universitat Autònoma de Barcelona se ocupan conjuntamente de la implementación de las reglas que explicitan las variaciones en un conversor de texto a voz y de evaluar la calidad del mismo. El equipo de La Salle dispone ya de un prototipo de conversor de Texto-Voz (Martí, 1987) que utiliza un circuito integrado específico para la síntesis por formantes (PCF8200 de Philips). En el marco del proyecto se procederá a la utilización de circuitos DSP específicos para el procesado de señal acústica que poseen una mayor versatilidad a partir de síntesis por parámetros LPC que supone una nueva obtención de las unidades básicas de síntesis a partir de voz natural. El equipo de Barcelona ha realizado ya algunos trabajos de evaluación de sistemas de conversión de texto a voz que consisten en la elaboración de una serie de tests lingüísticos que permiten cuantificar la calidad de estos sistemas (Llisterri, 1989; Llisterri, Poch, 1991).

5. Observaciones finales

Como se ha mencionado más arriba, el proyecto se inició en enero de 1992. Actualmente se encuentra, por una parte, en la fase de definición y elaboración del corpus de trabajo y, por otra, en la preparación de tests para la evaluación del conversor de texto a voz.

Referencias bibliográficas

Harmegnies, B.; Poch, D. (1991), "Vowel reduction in spontaneous speech in Spanish", en Llisterri, J.; Poch, D. (eds.), *Proceedings of the ESCA Workshop Phonetics and Phonology of Speaking Styles*, Barcelona, nº 25.

Fant, G. (1984), "Phonetics and Speech Technology", en Van den Broecke, M.; Cohen A. (eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences*, Dordrecht, Foris, pp. 13-24.

Gutiérrez, S. (1989), *Introducción a la semántica funcional*, Madrid, Síntesis.

Lindblom, B. (1990), "Explaining phonetic variation: a sketch of the H & H theory", en Hardcastle, W.; Marchal, A. (eds.), *Speech Production and Speech Modelling*, Dordrecht, Kluwer, pp. 143-167.

Llisterri, J. (1989), "Algunas propuestas para la evaluación de la inteligibilidad y naturalidad de los sistemas de conversión de texto a voz", XIX Simposio de la Sociedad Española de Lingüística, Universidad de Salamanca, *Revista Española de Lingüística*, 20,1 (1990), pp. 199-200.

Llisterri, J.; Poch, D. (eds.) (1991), Proceedings of the ESCA Workshop *Phonetics and Phonology of Speaking Styles*, Barcelona.

Llisterri, J.; Poch, D. (1991), "Caracterización fonética del bilingüismo, análisis acústico del habla espontánea y evaluación lingüística de sistemas de síntesis del habla", Actas del *Simposio de la Lengua Española, Ciencia y Tecnología*, organizado por el Pabellón de España, Barcelona, 7 - 11 de octubre de 1991.

Martí, J. (1987), "Síntesis del habla", Actas del III Congreso de Lenguajes Naturales y Lenguajes Formales, Publicaciones de la Universidad de Barcelona, pp. 213-237.

Poch, D. (1989), "Hacia una tipología de los fenómenos fonéticos en habla espontánea", XIX Simposio de la Sociedad Española de Lingüística, Universidad de Salamanca, *Revista Española de Lingüística*, 20,1 (1990), pp. 222.