

ESTUDIO ESTADÍSTICO DEL ESPAÑOL

Miguelina Guirao y María A. García Jurado
Laboratorio de Investigaciones Sensoriales
CONICET - Buenos Aires, Argentina
ISBN 950-3302-8, 276 pp.

Reseña: Patricia Reynal
FUNDALIS*

Esta obra presenta los resultados de un recuento estadístico de unidades de habla de Buenos Aires consideradas desde un punto de vista fonológico, fonético y ortográfico. La muestra abarca 43.306 palabras distribuidas en 74.460 sílabas y 163.861 fonemas. El inventario contiene la frecuencia de aparición, la distribución y las posibles combinaciones de cada unidad.

El corpus se extrajo de cinco piezas teatrales argentinas escritas en las mismas décadas y que, por su estilo realista, se pueden considerar documentales del habla del habitante medio de Buenos Aires.

Se verificó la confiabilidad estadística de esta muestra determinando funciones de error para las unidades computadas. Los programas de recuento

* Patricia Reynal, FUNDALIS M.T. de Alvear 2202 4o P.
Dirección Postal CC 53
1453 Buenos Aires

fueron diseñados especialmente y se procesaron mediante un sistema de computación.

El libro contiene una introducción al tema que abarca consideraciones generales sobre este tipo de estudio, antecedentes, descripción de las unidades y los procedimientos que se utilizaron para realizar el recuento. Los datos a su vez, se presentan en 11 tablas generales con subtítulos que elevan el total a 37 tablas.

El inventario presenta por separado las sílabas fonéticas, fonológicas y ortográficas. Los tipos silábicos consonante-vocal (CV), consonante-vocal-consonante (CVC) y otros, se agrupan en fonológicos y fonéticos fusionados y no fusionados. Se ordena el total de sílabas según el tipo silábico.

El estudio abarca además el recuento de fonemas y su distribución en la sílaba. Se analizan las combinaciones de sílabas y tipos silábicos, distinguiendo los que preceden y siguen a las más frecuentes. También se computaron los caracteres gráficos, las palabras ortográficas (expresión gráfica) y fonológicas (unidad léxica). Y las sílabas en posición inicial y final de ambas clases de palabras. Además se contaron las sílabas fonológicas que forman el mayor número de palabras diferentes en posición inicial, media y final.

Cada tabla permite acceder a cualquier dato relacionado con la frecuencia y probabilidad de aparición de cada una de las unidades de la muestra. Hay dos tipos de listado: 1) por orden de frecuencia, y 2) por orden alfabético. Conjuntamente con el listado y número de orden de las unidades, las columnas contienen el número de apariciones, el porcentaje de aparición, y el porcentaje acumulativo.

La edición de este material fue concebida como un aporte a varias disciplinas relacionadas con la lengua española. Provee material básico para confeccionar programas para el procesamiento de señales, inteligencia artificial, análisis y síntesis de habla y un sinnúmero de otras aplicaciones que se basan en la información probabilística.

En el campo de la educación proporciona datos para el desarrollo de programas de computación orientados a los procesos de aprendizaje del habla y de la ortografía, y para la enseñanza del español como lengua extranjera.

Se aplican además en las especialidades médicas relacionadas con las anomalías del habla, de la audición y de la comunicación verbal y patológica.

También interesa a los estudiosos de la Fonética, Fonología, Lingüística y para la comparación con otras lenguas.