



ARTÍCULO BREVE

Los sistemas de representación del habla: principios básicos para logopedas

Carmen Ávila Martín

Universidad de Granada

Introducción

El análisis del lenguaje (y de sus patologías) necesita de estudios que se basen en el análisis de casos reales. Para ese estudio es imprescindible la representación del lenguaje, que se inserta en una tradición lingüística a la que se suman en el momento actual intereses de disciplinas tan dispares como la psicología o las tecnologías del habla. En el ámbito interdisciplinario en el que nos encontramos, las técnicas empleadas en la representación del lenguaje tiene interés para el logopeda, que puede aprovechar las experiencias de transcripción de corpus para los propios fines, en este caso, la reproducción de alteraciones del lenguaje en un código escrito que pueda ser compartido por los profesionales.

La transcripción ortográfica

La escritura ha sido tradicionalmente el sistema de representación de lo oral. La ortografía constituye un ámbito de representación que pertenece a la tradición cultural de cada lengua y que ha sido extendida básicamente por los tipógrafos. Además, el estudio de las dificultades de aprendizaje de la escritura constituye un campo específico de observación y trabajo. En español, la norma ortográfica está establecida por la Real Academia Española, que modificó algunos signos en su última *Ortografía de la lengua española* (1999). Esta norma abarca todos los países hispanoamericanos y es, por tanto, panhispánica (rae.es).

La representación del lenguaje alterado precisa, sin embargo, de mecanismos mucho más precisos, aunque también utiliza la ortografía como forma más extendida de representación. El sistema utilizado precisa de justificaciones y de la

búsqueda del consenso entre los especialistas, por ejemplo, cómo representar la reducción de palabras, o si hay que usar mayúscula en los casos normativos, o la representación de vacilaciones u onomatopeyas.

La transcripción fonética

La realización fonética que se encuentra detrás de cada letra constituyó el punto de partida del que se considera el primer tratado de fonética del español. Fue escrito en 1620 por Juan Pablo Bonet con el título *Reducción de las letras y arte para enseñar a los mudos*, que tenía por objetivo la enseñanza del habla a los sordomudos (véase Martínez Celdrán y Romera, 2007).

La transcripción fonética ha constituido el sistema de análisis de los sonidos del habla desde el nacimiento de la lingüística moderna. Los sistemas de transcripción fonética tampoco están exentos de una tradición cultural, y así para el español podemos encontrar diferentes modos de escritura: el sistema de transcripción de la *Revista de Filología Española* se ha utilizado ampliamente en el ámbito hispánico, especialmente en estudios de carácter dialectal. Sin embargo, la necesidad de consenso internacional impone actualmente la adopción del sistema de transcripción de la Asociación Internacional de Fonética (AFI), que puede representar la transcripción de todos los sonidos de las lenguas del mundo, aunque no está especialmente pensado para la representación de los sonidos del español. La AFI ha desarrollado la de representación de sonidos alterados y que se denomina *extraIPA symbols for disordered speech* (<http://www.langsci.ucl.ac.uk/ipa/ExtIPAChart02.pdf>).

Los corpus

La recogida de materiales ha desarrollado en los últimos decenios una metodología específica en el acopio y la presentación de éstos. Las propuestas de transcripción y etiquetado de corpus orales se relacionan con distintos intereses, que van desde las tecnologías del habla, a propuestas de la lingüística conversacional para el estudio de lo oral. Un corpus para el reconocimiento automático de la voz no presenta las mismas necesidades de etiquetado y transcripción que uno para la realización de estudios pragmáticos del lenguaje. En el análisis lingüístico de lo oral, se ha puesto de manifiesto que es preciso recoger más datos de los que recoge la transcripción fonética, especialmente relacionado con aspectos pragmáticos y conversacionales. De este modo, se han desarrollado sistemas de transcripción en los que se intenta recoger el máximo de datos relevantes de la situación comunicativa, no necesariamente diferentes del aspecto fonético, como la energía articuladora, aunque también interesa la intencionalidad comunicativa o los aspectos gestuales.

La necesidad de contar con datos fiables en el análisis lingüístico ha hecho que se desarrolle una nueva disciplina que se conoce con el nombre de Lingüística de Corpus. No se trata sólo de recopilar materiales para el análisis posterior, sino también que se está desarrollando todo un protocolo de recogida de datos para la fiabilidad de las conclusiones, pues en muchas ocasiones la recogida del material condiciona los resultados obtenidos. En la elaboración de corpus es fundamental el objetivo que se pretende conseguir, para qué se elabora, independientemente de que se puedan utilizar datos con otros fines. Así, muchos de los corpus del español se han creado con fines lexicográficos (es decir, para la elaboración de diccionarios) como el CORDE y el CREA de la Real Academia Española (rae.es), mientras que otros se elaboran para el estudio del español coloquial, como el realizado por el grupo de investigación Valencia, Español Coloquial (Val.Es.Co.), dirigido por Antonio Briz.

La realización de corpus conversacionales con sistemas de transcripción consensuados para el estudio del español coloquial supone un avance en el estudio de las especificidades de la forma de comunicación oral y en el estudio del lenguaje alterado, por ejemplo, para el estudio del habla afásica (Gallardo, 2004).

En la representación de los corpus existen propuestas que van desde el nivel ortográfico, a niveles fonético-fonológicos, o pueden incluir también una transcripción fonológica alineada con una representación acústica del enunciado (Llisterri, 1999).

En la transcripción fonética, además del sistema AFI, también se utiliza SAMPA, y su ampliación X-SAMPA, cuando se requiere una transcripción fonológica y el intercambio electrónico de textos, y Woldlbet, cuando se necesita un sistema más detallado.

Otros sistemas de transcripción parten de la reproducción ortográfica con una serie de signos de diacríticos que recogen diversos aspectos de la producción lingüística. El CHILDES se desarrolló fundamentalmente para el estudio del lenguaje infantil. Los sistemas de transcripción utilizados por los psicólogos (el CHILDES y el SALT) presentan, en opinión de Beatriz Gallardo, problemas de claridad derivados fundamentalmente del uso excesivo de diacríticos, en el caso del CHILDES, y un excesivo condicionamiento del inglés, en el caso del SALT.

El corpus de habla afásica del proyecto PerLA (Percepción y Lenguaje afásico) que dirige Beatriz Gallardo descarta el uso de la transcripción fonética, porque se considera que la codificación fonética dificultaría la lectura y no resulta práctico en función de los intereses de los compiladores. Esos intereses se reflejan en el conjunto de signos diacríticos que utilizan en la transcripción del corpus, aunque el objetivo sea la transcripción de la forma más fiel posible de los rasgos del habla patológica. En este sentido, Gallardo señala: “La elección de los rasgos que van a registrarse encierra ya un buen número de asunciones teóricas y metodológicas previas” (p. 13).

Al igual que para el grupo Val.Es.Co, al corpus PerLa le interesa sobre todo el estudio de los aspectos pragmáticos. Le interesa así el estudio de la afasia contextualizado para la obtención de datos reales, para lo cual se graban conversaciones del paciente afásico con un “interlocutor clave” que forma parte de su entorno familiar.

El mayor problema de la transcripción de corpus orales es la estandarización para su aceptación por grupos amplios de investigadores. No es lo mismo un corpus para el reconocimiento del habla, que un corpus para la realización de estudios pragmáticos del lenguaje, o un corpus pensado para el estudio de aspectos de carácter fonológico. Aunque lo ideal es la creación de macrocorpus que puedan ser reutilizados por investigadores con diversos intereses.

Bibliografía general

- Gallardo Paúls, B (2004). *La transcripción del lenguaje afásico*. En Gallardo, B; Veyrat, M (Ed.). Estudios de Lingüística Clínica, II: Lingüística y patología. (pp. 83-114). Valencia: Universitat/AVaLCC.
- Llisterri, J (1999). *Transcripción, etiquetado y codificación de corpus orales*. En Gómez Ginovart, J; Lorenzo Suárez, A; Pérez Guerra, J; Álvarez Ladrón, A (Ed.). Panorama de la investigación en lingüística informática. *RESLA. Revista española de Lingüística Aplicada*. Volumen monográfico, 53-82.
- Martínez Celadrán, E; Romera Barrios, L (2007). *Historiografía de la fonética y fonología españolas*. En Dorta, J; Corrales, C; Cobella, D (Ed.). *Historiografía de la lingüística en el ámbito hispánico*. Madrid: Arco Libros.