

LA ESTANDARIZACION DE LA PRONUNCIACION Y LA VARIABILIDAD LINGÜÍSTICA

Santesteban/Donostia, 25-IX-1998

Dolors Poch Olivé,
Departamento de Filología Española,
Universidad Autónoma de Barcelona

1. LAS DESCRIPCIONES FONÉTICAS TRADICIONALES

La concepción imperante, hasta hace muy pocos años, del estudio del lenguaje arranca de Saussure (1916) y de su distinción entre *lengua* y *habla*. Para Saussure, el centro de interés de la lingüística lo constituye la lengua, el sistema en tanto producto social. En efecto, cuando este autor afirma que "*Le tout du langage est inconnaissable*" (1916, 1980: 38) lo que hace, en realidad, es dirigir el interés de la lingüística hacia la "*lengua*" negando, de forma explícita, la posibilidad de estudiar el "*habla*". El estudio del lenguaje se concibe así como un interés por su aspecto "colectivo", por aquello que es producto del "pacto" entre los hablantes o, dicho en otras palabras, por aquello que es "invariable" en las lenguas. Se suele aceptar, por tanto, que el objetivo de cada una de las diferentes ramas de la lingüística consiste en determinar la estructura de los diversos aspectos de las lenguas: fonética, fonológica, morfológica, sintáctica, semántica, etc., siempre desde el punto de vista del sistema. Las dos grandes escuelas lingüísticas del siglo XX, el estructuralismo y la gramática generativa, comparten este punto de vista aunque las soluciones que aporten y las perspectivas en las que se sitúan para llegar a alcanzar dicho objetivo sean muy diferentes. En ambos casos, los lingüistas no toman en cuenta el *habla* (en el sentido saussureano del término) sino que el interés de los diversos autores se centra en la lengua.

Este interés por la lengua y no por el habla ha condicionado fuertemente, durante casi todo el siglo XX, la metodología de trabajo de la lingüística, los corpus estudiados y el tipo de razonamiento efectuado a partir de los datos recogidos. Por lo que respecta a la fonética, el método de trabajo y el tipo de argumentación que podemos denominar "tradicional" consiste, en una primera etapa, en la preparación de una lista de sílabas, palabras o pequeñas frases que contienen aquel sonido o combinaciones de sonidos que constituyen el objetivo del estudio. A continuación, se pide a los informantes previamente seleccio-

nados que lean esa lista, en una habitación debidamente preparada por lo que a las condiciones acústicas se refiere, y “con naturalidad”, con vistas a obtener una lectura lo más parecida posible a las realizaciones “reales”. Esta forma de trabajar, es decir con corpus que podemos denominar “de laboratorio”, permite controlar cuántas veces van a leer los informantes los sonidos estudiados, en qué entorno fónico están situados dichos sonidos, en qué posición se encuentran con respecto al acento, etc. A partir de los datos así obtenidos y, obviamente después de proceder a su análisis —articulatorio y/o acústico—, y a aplicarles el tratamiento estadístico adecuado, el siguiente paso consiste en la elaboración de explicaciones o de modelos que, se supone, den cuenta del proceso de producción y percepción.

Dichos modelos tienen como objetivo común tratar de determinar los indicios acústicos de los fonemas, es decir, aquellas características de la onda sonora que serían supuestamente invariables, que estarían siempre presentes en ella y que los hablantes asociarían con las distintas unidades del sistema fonológico. La teoría de los rasgos distintivos de Jakobson-Fant-Halle (1952) constituye una manifestación clásica de los modelos de este tipo. La forma en que estos autores enfocan las relaciones entre la fonología y la fonética consiste en establecer de qué forma se manifiestan los rasgos distintivos en los espectrogramas y en los movimientos articulatorios que se realizan para la producción de los distintos sonidos. Así por ejemplo, un rasgo como *grave-agudo* se manifestaría acústicamente mediante una concentración de energía en las zonas bajas o altas del espectrograma (por lo que respecta a la caracterización acústica) y en una configuración articulatoria que favoreciera la producción de un tipo u otro de sonidos. El rasgo distintivo *consonántico-vocálico* se manifestaría articulatoriamente como existencia o no de una interrupción de la salida del aire por el canal fonatorio y acústicamente como presencia constante de ondas periódicas o aparición de interrupciones o de ruidos turbulentos (Jakobson-Fant-Halle, 1952). En cualquier caso, lo que aquí interesa destacar es que, como concepción de base, se considera que cada vez que se realiza un sonido determinado por parte de cualquier hablante y en cualquier situación de habla, los participantes en la conversación son supuestamente capaces de rastrear en la onda sonora los indicios así definidos.

Un planteamiento de este tipo asume que los datos obtenidos a partir de corpus “de laboratorio” son homologables a los que se obtendrían a partir de cualquier otra clase de muestras de lengua, de forma que una hipótesis sobre el funcionamiento del lenguaje elaborada a partir de dichos datos puede hacerse extensiva a la lengua en general. Este tipo de razonamiento es el corolario lógico de la concepción saussureana del estudio del lenguaje: si lo que interesa a la lingüística es la lengua y no el habla, si lo que se va a buscar es lo invariable, la mejor forma de proceder para alcanzar este objetivo consiste en elaborar un corpus ad hoc, recogerlo en las mejores condiciones posibles para

facilitar su posterior análisis y considerar que las conclusiones que se obtienen mediante este procedimiento son aplicables a la lengua en general.

Hay que señalar que, con el paso del tiempo, por lo menos en Fonética, la forma de concebir lo invariable va cambiando. Así por ejemplo, trabajos posteriores al modelo que hemos comentado, como son los de Stevens y Blumstein (1981), polemizan sobre cuáles son los indicios acústicos presentes en la onda sonora y la propuesta de estos autores es distinta a la de Jakobson-Fant-Halle, pero lo importante es que no se discute, en ningún momento, su existencia. Dicho en otras palabras, no cambia en absoluto la idea de que, en la onda sonora, están siempre presentes unos elementos invariables que los oyentes son capaces de detectar y los hablantes son capaces de emitir y que constituirían los indicios acústicos de los sonidos del habla.

2. LAS LIMITACIONES DE LAS DESCRIPCIONES FONÉTICAS TRADICIONALES

No obstante, el desarrollo de instrumentos de análisis de la señal acústica cada vez más precisos, ha provocado la confrontación del investigador con la realidad física de los sonidos del habla. Se ha puesto así de manifiesto una de las características fundamentales de dichos sonidos: su irreproductibilidad. Diversas realizaciones de un único corpus, incluso producidas por un mismo locutor, se manifiestan como fenómenos diferentes y, si dichas realizaciones proceden de locutores distintos, las diferencias son aún mayores. La existencia misma de estas variaciones es perturbadora puesto que parecen estar en contradicción con la existencia de invariantes, es decir, con la idea de *lengua* en el sentido saussureano del término.

En efecto, las fuentes de variación de la señal del habla, tradicionalmente aceptadas por la Lingüística, están relacionadas con el origen geográfico de los hablantes y/o con el origen social de los mismos. Se configuran así dos tipos distintos de variación que es posible, no obstante, relacionar con el modelo saussureano de *lengua / habla*. Según esta perspectiva, sería también posible establecer los invariantes de una determinada habla regional y también los invariantes de una determinada forma de hablar condicionada socialmente, pues se trataría de buscar las regularidades presentes en las mismas, de buscar lo "socialmente consensuado" en cada uno de estos casos. En cambio, las variaciones que se mencionan en el párrafo anterior, no parecen encontrar su lugar en un modelo de estas características.

En esta *confrontación con la realidad física de los sonidos del habla*, que mencionábamos más arriba, han jugado un papel fundamental las aplicaciones tecnológicas que implican la voz, desarrolladas especialmente en la segunda mitad del siglo XX. Dichas aplicaciones han estado ancladas durante muchos

años en los datos proporcionados por el análisis de corpus “de laboratorio”, es decir, de corpus como los presentados más arriba. Con respecto a la síntesis de habla se han obtenido resultados muy aceptables aunque, siempre, a estos sistemas, cuya inteligibilidad no se ve perturbada, “les falta naturalidad” o “suenan un poco raro”. En cambio, los sistemas que implican un proceso de reconocimiento automático del habla por parte de un programa informático plantean muchas más dificultades. La más importante de ellas consiste en el hecho de que la información que se suele suministrar a dichos sistemas, procedente del resultado del análisis de corpus “de laboratorio”, no se ajusta a los fenómenos con los que se enfrenta el sistema cuando debe reconocer una serie de frases realizadas de forma espontánea por un locutor cualquiera. El problema que aquí aparece de una manera tan clara es el de la variabilidad que, mediante un planteamiento simplista, puede formularse en los siguientes términos: un sistema de reconocimiento automático, basado en una información lingüística procedente de corpus “de laboratorio”, no puede relacionar el tipo de información acústica que le proporcionan los hablantes, cuando se expresan de forma espontánea, con la información acústica que dicho sistema tiene almacenada. Esta última, extraída de unas muestras de habla que se sitúan lejos de lo que es la forma habitual de hablar de los locutores, es difícil de relacionar con la información que se puede detectar en una onda sonora, que sí corresponde a las realizaciones “normales” de los hablantes. Ya en 1983, y a propósito de este problema, señalaba G. Fant que *“The bottleneck for the development of speech-based information systems is neither technological nor economic but is created mainly by the fact that our fundamental theoretical understanding of human speech processes is still limited”*.

3. EL ORIGEN DE LA VARIABILIDAD

Enfrentado, pues, con este problema, el investigador que necesita estudiar la señal de habla —sea con fines estrictamente lingüístico-teóricos, sea con vistas a las aplicaciones tecnológicas— debe comenzar a considerar que las características acústicas de la señal de habla son producto de dos fuentes de variación distintas:

1. La primera, de naturaleza fonológica, conduce al hablante a diferenciar los sonidos emitidos de tal manera que los contrastes que se produzcan entre ellos remitan a los contrastes que oponen entre sí las representaciones de dichos sonidos.

2. La segunda está relacionada con un importante número de factores difícilmente aislables tales como el efecto del contexto fónico, de la acentuación, de las características intrínsecas inherentes al individuo (sexo, edad, origen, etc.), etc. Parece natural considerar todo este conjunto de factores como un

bloque que comprende todo un conjunto aleatorio de variaciones de la señal de habla.

Desde una perspectiva epistemológica, las dos fuentes de variación poseen estatus comparables a los de los componentes saussureanos del lenguaje: la investigación en general procura alejarse de la segunda (del orden del *habla*) para tratar de comprender mejor la primera (del orden de la *lengua*).

Con el tiempo y de forma gradual, en el terreno de la Fonética, se han llevado a cabo experiencias sistemáticas focalizadas sobre el efecto de estas variables específicas, que ponen de manifiesto la dificultad de establecer invariantes. Algunos autores estudian, por ejemplo, la influencia de la prosodia (Delattre, 1969) sobre el timbre o la articulación de las vocales. Situado en esta perspectiva, hay que destacar el estudio que Lindblom (1963) consagra a la reducción vocálica ya que presenta un interés muy especial. La principal variable independiente cuyos efectos estudia es la duración. Lindblom demuestra que cuando la duración disminuye, el hablante, obligado a producir en menos tiempo los gestos articulatorios que necesita, no consigue realizarlos con la misma eficacia. En el plano acústico, los objetivos constituidos por los valores formánticos normales cuando la duración es relativamente larga, no pueden alcanzarse. Este fenómeno recibe el nombre de *undershoot*: el locutor no alcanza el objetivo que se ha propuesto.

En el plano epistemológico, el interés de esta experiencia reside sobre todo en la especificidad de la variable independiente. La propia variable sufre los efectos de la situación de comunicación. En el habla cotidiana, las condiciones de expresión influyen sobre la velocidad de elocución: el comentarista deportivo, por ejemplo, que está comentando, a través de un micrófono, un partido de cualquier deporte para hablantes que no están frente a él, habla a una velocidad muy distinta de la de una madre de familia que enseña a su hijo a recitar poesía.

La inferencia que puede extraerse de estas constataciones es la siguiente: está comprobado que la duración influye en el timbre, pero la duración misma depende de la situación de comunicación, por tanto la situación de comunicación influye en el timbre.

4. LA VARIABILIDAD Y LA SITUACION DE COMUNICACION

Esta idea de que las circunstancias influyen sobre las variables responsables de las modificaciones de las características de la señal se encuentra en la base de una cierta desconfianza hacia una situación ya mencionada más arriba: aquella en la que se encuentra el hablante sometido a una experiencia lingüística clásica, una experiencia en la que se utiliza el habla que hemos denominado "de laboratorio". Esta desconfianza, unida a las evidencias proporcionadas

das por las aplicaciones tecnológicas que se han mencionado anteriormente, ha dado lugar, entrada ya la década de 1980, a una serie de estudios que se interesan por el *habla continua* (*connected speech*) que intentan, de alguna forma, validar los resultados de las investigaciones anteriores (Barry, 1984). Así, en estos trabajos, los autores, en lugar de centrarse sobre producciones de sílabas sin sentido, de pseudo-palabras o de palabras aisladas, piden al locutor que realice secuencias contínuas similares a las que realizaría en la conversación normal.

La propia existencia de estas iniciativas consagra implícitamente la idea de que existe un *estilo de habla* propio a la producción de corpus destinados a los estudios lingüísticos, los corpus que antes hemos denominado *habla de laboratorio*. Los primeros trabajos sobre *habla continua* no buscan estudiar este *estilo de laboratorio*, sino que intentan neutralizar sus posibles efectos. Estudios posteriores, en cambio, utilizarán el estilo de laboratorio para estimar, por comparación con el habla espontánea, el grado de validez de las investigaciones realizadas anteriormente mediante métodos más clásicos (Harmegnies & Poch, 1992).

La bibliografía actual sobre estos temas presenta numerosas contribuciones basadas en el estudio de gran número de situaciones de comunicación que, en principio, exigen distintos estilos de habla: conversación de un adulto con un niño (Lindblom et al., 1992), solicitud de información telefónica (Eskénazi & Isard, 1991), descripción de un itinerario a partir de un mapa (Mc Allister et al., 1990), etc. Estos trabajos, en su conjunto, muestran una importante influencia de los estilos implicados sobre los diversos parámetros acústicos que dan cuenta de la señal del habla tales como la duración, la frecuencia fundamental o el timbre.

Desde un punto de vista estrictamente estadístico, tomar en cuenta los *estilos de habla* significa añadir a las dos fuentes de variación tradicionalmente consideradas (la geográfica y la social) una tercera, la *estilística*, hasta ahora incluida en el conjunto de factores de variación calificados como aleatorios. Este reconocimiento de la existencia de nuevos determinantes específicos de la señal vocal rompe de forma clara el modelo saussureano. Tal vez haya que entender en este sentido esta afirmación de Jakobson procedente de su trabajo "*Lingüística y poética*": *Mientras estemos interesados en invariantes fonológicas, las vocales inglesas /i/ e /i:/ se presentan como simples variantes de uno y del mismo fonema; pero si lo que nos preocupa son las unidades emotivas, la relación entre invariante y variante se revierte: la longitud y brevedad son invariantes ejecutadas por fonemas variables* (Jakobson, 1960, 1974: 132). Hay que señalar, además, que la consideración de la variación estilística deja, hoy por hoy, todavía muchas cuestiones abiertas, como por ejemplo, las diferencias entre estilos, los problemas que plantea la clasificación de los estilos, etc.

5. LA CUESTIÓN DE LA ESTANDARIZACIÓN

La operación de estandarizar una lengua determinada plantea el problema del establecimiento de modelos, el problema de la igualación. Y para efectuar dicha operación es necesario establecer criterios que permitan decidir cuáles son las formas que deben ser tomadas como modelo y ello en todos los aspectos de la lengua: morfológico, léxico, sintáctico y, obviamente, fonético. La cuestión de la pronunciación es extremadamente compleja porque, según las formas que se elijan como modelo, se corre el peligro de que los hablantes no se sientan reflejados en ellas y se produzca, por tanto, una reacción de rechazo frente a los modos de pronunciar que se quieren extender a toda la comunidad lingüística implicada. No es posible tampoco establecer, a priori, una lista de criterios que sirvan para determinar cuál debe ser la pronunciación estándar de una lengua determinada: cada comunidad lingüística, en función de sus características en el momento en que se lleva a cabo la operación de estandarización y también en función de factores de tipo histórico, necesita que dicha operación se realice pensando en la especificidad de cada caso. Las únicas directrices que es posible establecer, en lo referente al proceso de estandarización, están relacionadas con los diversos aspectos de la lengua que se deben tomar en cuenta en dicho proceso: no hay que olvidar en ningún momento que las lenguas son, por definición, variables y que, por tanto, cuando se deciden los modelos hay que sopesar previamente los factores de variación (geográficos, sociales y estilísticos) y decidir de qué manera deben ser considerados en dicha operación de estandarización.

REFERENCIAS BIBLIOGRÁFICAS

- BARRY, M. (1984): "Connected speech: processes, motivations, models", *Cambridge Papers in Phonetics*, 3, 2: 1-16.
- DELATTRE, P. (1969): An acoustic and articulatory study of vowel reduction in four languages, *International Review Applied Linguistics*, 7: 295-325.
- ESKENAZI, M.; ISARD, A. (1991): "Characterising the change from casual to careful style in spontaneous speech", *JASA*, 90.4: 2363-2364.
- FANT, G. (1983): "Phonetics and speech technology", en van den Boriecke, M.; & Cohen, A. (eds.), *Proceedings of the tenth international congress of phonetic sciences*, Dordrecht-Foris, pp. 13-24.
- HARMEGNIES, B.; POCH-OLIVÉ, D. (1992): "A study of style-induced vowel variability: laboratory versus spontaneous speech in spanish", *Speech Communication*, 11: 429-437.
- JAKOBSON, R. (1960): "Linguistics and Poetics", en Sebeok, T. (eds.), *Style in Language*, Cambridge, MIT Press, 350-377.

- JAKOBSON, R.; FANT, G.; HALLE, M. (1952): *Prelimaries to Speech Analysis*, Cambridge, MIT Press.
- LINDBLOM, B. (1963): "Spectrographic study of vowel reduction", *JASA*, 35: 1771-1781.
- LINDBLOM, B.; BROWNLEE, S.; DAVIS, B.; MOON, S.J. (1992): "Speech Transforms", *Speech Communication*, 11, 4-5: 357-368.
- McALLISTER, J.; SOTILO, C.; BARD, E.G.; ANDERSON, A. (1990): *Using the map task to investigate variability in speech*, Dep. of Linguistics, Univ. of Edinburgh.
- SAUSSURE, F. de (1916): *Cours de Linguistique Générale*, Edición de 1980 a cargo de Tullio de Mauro, Paris, Payot.
- STEVENS, K.N.; BLUMSTEIN, S. (1981): "The search for invariant acoustic correlates of phonetic features", en Eimas, P.; Miller, J.L. (eds.), *Perspectives on the study of speech*, New Jersey, Earlbaum, pp. 1-38.