

Enseñanza de la pronunciación, corrección fonética y nuevas tecnologías

Enseñanza de la pronunciación y nuevas tecnologías

JOAQUIM LLISTERRI. Universidad Autónoma de Barcelona (UAB)

La enseñanza de la pronunciación y la corrección fonética pueden beneficiarse, en el contexto de la enseñanza asistida por ordenador, del uso de procedimientos que permiten la visualización de la señal sonora y de tecnologías como el reconocimiento del habla, que facilitan la detección de errores. Es necesaria, sin embargo, una reflexión sobre el proceso de integración de los sonidos de una lengua extranjera para valorar adecuadamente las ventajas y las limitaciones ofrecidas por los avances técnicos.

Introducción

Pese al auge del aprendizaje asistido por ordenador al que hemos asistido en las últimas décadas, la enseñanza de la pronunciación y la corrección fonética no han recibido, tal como es también el caso en la enseñanza presencial, la misma atención que otras áreas como la gramática o el léxico. Por ello, este

trabajo se propone, en primer lugar, presentar de forma introductoria dos procedimientos empleados en algunos materiales didácticos multimedia para la enseñanza de la pronunciación: la visualización de la señal sonora y el reconocimiento del habla. Sin embargo, el innegable atractivo de determinadas tecnologías no debería dejar de lado una reflexión de carácter más fundamental sobre el proceso de integración del sistema fonético de una lengua extranjera. Asumir que la observación de una representación acústica o la obtención automática de una determinada puntuación correspondiente a la realización de un enunciado supone una enseñanza efectiva de la pronunciación o un procedimiento eficaz de corrección fonética implica, en cierto modo, ser negligentes respecto a aspectos esenciales del proceso de producción y percepción de los sonidos de una lengua extranjera. Por este motivo, la conclusión del trabajo pretende mostrar que, a pesar de sus indiscutibles ventajas, determinados usos de las técnicas de análisis acústico y de reconocimiento del habla no garantizan de forma automática la adecuada asimilación del sistema sonoro de una lengua extranjera.

Los procedimientos de visualización del habla

La utilización de procedimientos de visualización del habla como ayuda a la enseñanza de la pronunciación y a la corrección fonética no es patrimonio exclusivo de los sistemas de aprendizaje asistido por ordenador. Desde que la fonética experimental puso a disposición de los profesores de lenguas extranjeras métodos de representación del habla relativamente accesibles, se

han utilizado diversas técnicas que permiten, en lo esencial, que el estudiante observe una representación articuladora o acústica de su propia producción y pueda compararla con un modelo proporcionado por el profesor o por un hablante nativo. Estos mismos procedimientos se aplican también, con una metodología muy similar, a la reeducación de las patologías del habla. Sistemas ampliamente utilizados, como por ejemplo el *SpeechViewer* de IBM (Puyuelo, 1992; Martínez Ledesma *et al.*, 2000), se basan en los mismos principios de visualización e imitación de un modelo.

En los estudios de fonética acústica se emplean habitualmente diversos tipos de representación de la señal sonora: la evolución temporal de la amplitud del sonido (oscilograma), la representación de los cambios de frecuencia y amplitud a lo largo del tiempo (espectrograma), el análisis de la relación entre frecuencia y amplitud en un determinado instante en el tiempo (espectro), la variación temporal de la frecuencia fundamental (curva melódica) y la evolución en el tiempo de la intensidad (curva de intensidad) (Kent y Read, 1992; Ladefoged, 1996; Stevens, 1998). Cada una de estas representaciones aporta una información específica sobre las características fonéticas de un enunciado, tanto en la vertiente segmental (sonoridad, nasalidad, modo y lugar de articulación y timbre) como suprasegmental (entonación, acento y ritmo). Sin embargo, la interpretación de un documento de análisis acústico requiere unos buenos conocimientos de la materia, de modo que, sin una adecuada preparación, es sumamente difícil relacionar las propiedades

fonéticas de un enunciado con su correspondiente manifestación acústica representada en un documento.

A la vista de la figura 1, en la que se muestran el oscilograma y el espectrograma del enunciado *El año tiene cuatro estaciones* no parece obvio que un estudiante de lengua extranjera sin una preparación especial en fonética acústica pueda llegar a deducir, a través de este tipo de representación, ni los rasgos que debe imitar en un modelo ni los errores de su propia producción.

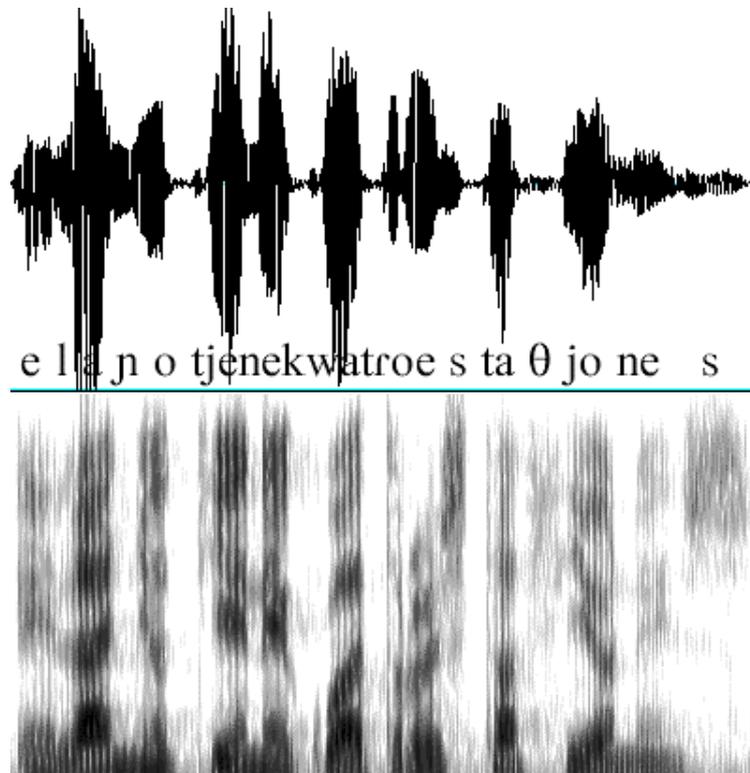


Figura 1: Oscilograma y espectrograma del enunciado *El año tiene cuatro estaciones* obtenidos mediante el programa *Praat*

Algo más intuitiva puede ser la información que aparece en una curva melódica o en una curva de intensidad, tal como puede apreciarse en las figuras 2 y 3, aunque no por ello pensamos que su interpretación sea inmediata.

e l a ñ o t j e n e k w a t r o e s t a θ j o n e s



Figura 2: Curva melódica del enunciado *El año tiene cuatro estaciones* obtenida mediante el programa *Praat*.

e l a ñ o t j e n e k w a t r o e s t a θ j o n e s

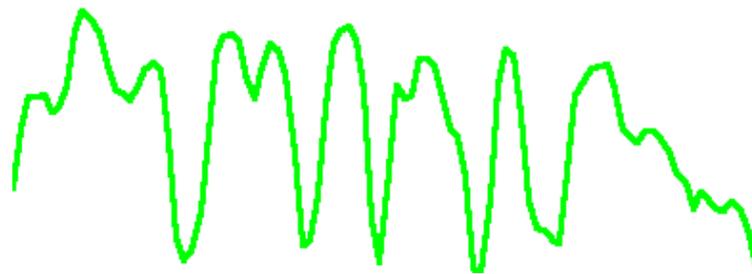


Figura 3: Curva de intensidad del enunciado *El año tiene cuatro estaciones* obtenida mediante el programa *Praat*.

Puede ser ilustrativo comparar estos documentos, realizados mediante un programa de análisis acústico del habla, y habituales en la investigación en

fonética experimental, con los que se proponen en el campo de la enseñanza de lenguas extranjeras. Por ejemplo, Germain y Martin (2000) presentan una herramienta para la enseñanza de la pronunciación en la que el estudiante trabaja con representaciones como las que se reproducen en la figura 4.

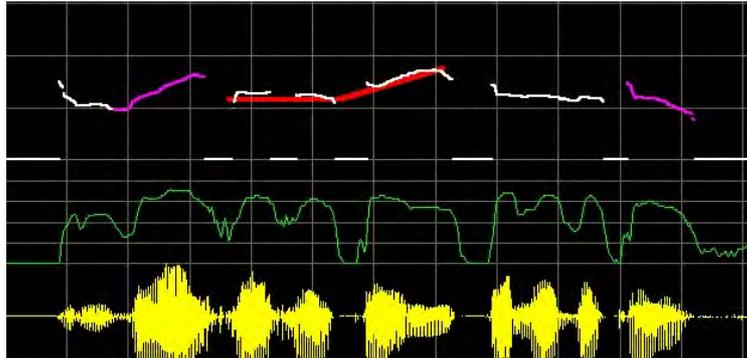


Figura 4: Oscilograma (parte inferior), curva de intensidad (parte central) y curva melódica (parte superior) obtenidos con el programa *WinPitch LTL* (Germain y Martin, 2000: figura 2) .

Es interesante destacar aquí que sobre la curva melódica real, trazada en blanco, se superponen otras líneas, como la que aparece en rojo, que muestran los movimientos entonativos esenciales, prescindiendo del detalle fonético; estas representaciones estilizadas son también habituales en el análisis prosódico del habla (Estruch *et al.*, 1996). Las curvas originales, producidas por el profesor y por el estudiante, se utilizan, tal como se muestra en la figura 5, para que el alumno pueda comparar su pronunciación con el modelo nativo.

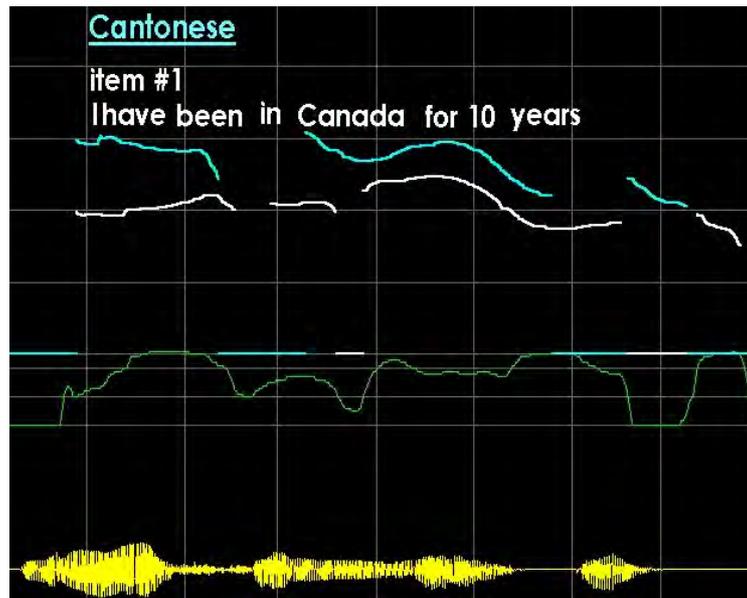


Figura 5: Curva melódica del enunciado modelo (en blanco) y de la producción del estudiante (en azul) obtenidas con programa *WinPitch LTL* (Germain y Martin, 2000: figura 7).

También Cazade (1999) recoge ejemplos del uso de representaciones estilizadas de la curva melódica y de la curva de intensidad en el programa *Speaker* (v.3).

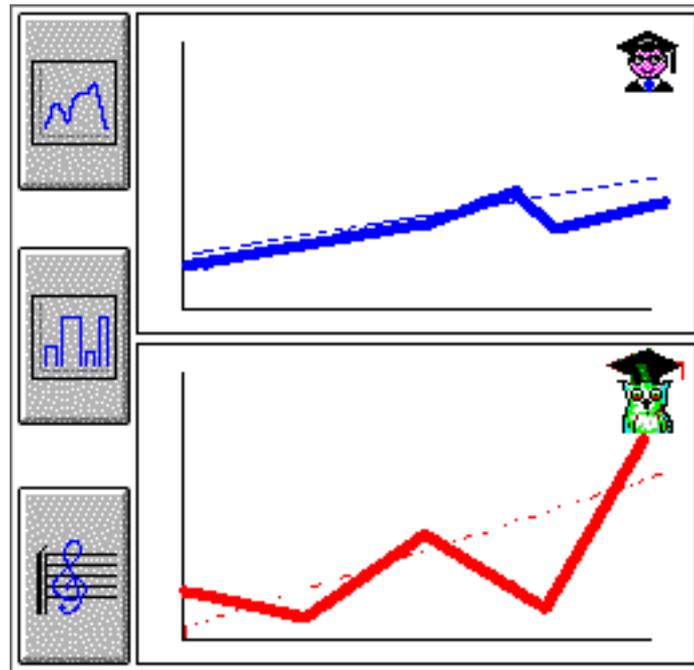


Figura 6: Representación estilizada de la curva melódica en la que se muestra el modelo propuesto por el profesor (panel inferior) y la producción del alumno (panel superior) obtenidas con el programa *Speaker* (v.3) (Cazade, 1999: figura 3).

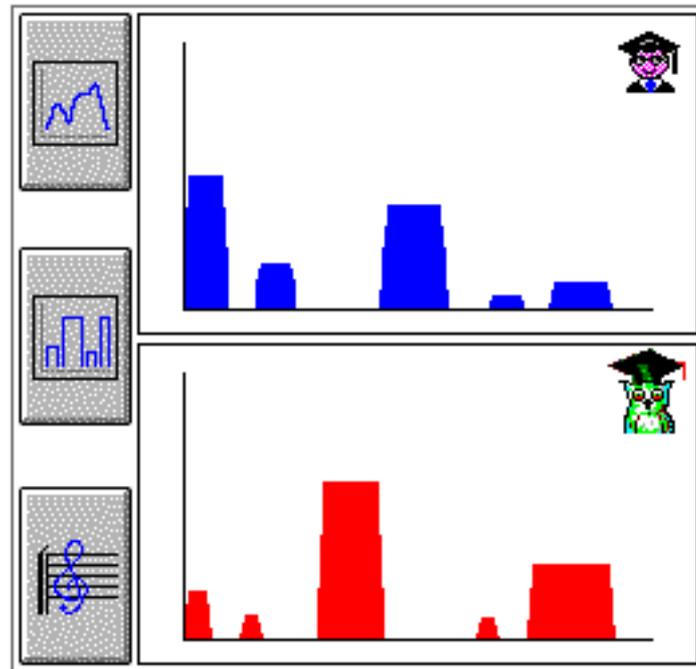


Figura 7: Representación estilizada de la curva de intensidad en la que se muestra el modelo propuesto por el profesor (panel inferior) y la producción del alumno (panel superior) obtenidas con el programa *Speaker* (v.3) (Cazade, 1999: figura 2).

Si, como indicábamos al principio, parece que estas representaciones pueden interpretarse de un modo relativamente inmediato, más difícil es, en nuestra opinión, justificar el uso de oscilogramas en cursos como *Tell Me More*, reproducidos igualmente en Cazade (1999).

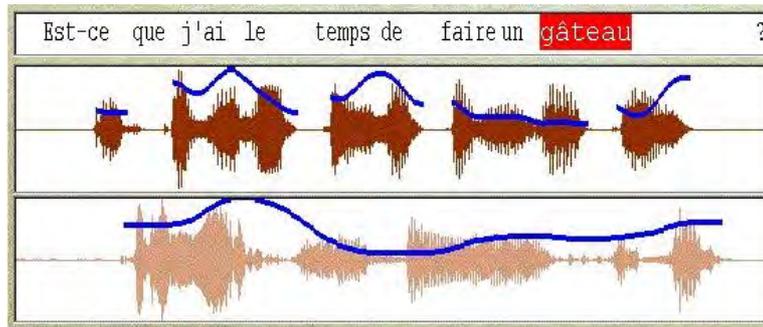


Figura 8: Oscilograma (en rojo) y curva melódica superpuesta (en azul) del modelo de pronunciación (panel superior) y de la producción del alumno (panel inferior) en el curso *Tell Me More* (Cazade, 1999: figura 5).

Esta selección de ejemplos no pretende sino mostrar las diversas posibilidades de visualización del habla y algunas de las modificaciones que es posible realizar sobre las representaciones habitualmente empleadas en fonética para dotarlas de un carácter más pedagógico. Sin embargo, Germain y Martin (2000: 64-65) señalan acertadamente que uno de los principales problemas del uso de representaciones visuales es la falta de indicaciones que recibe el estudiante respecto a los aspectos de su producción que deben corregirse: aunque el estudiante realice un análisis de las diferencias de forma entre su realización y el modelo propuesto, parece realista pensar que no siempre encontrará aquellas diferencias que son lingüísticamente relevantes y que le pueden aproximar a una producción nativa. En este sentido, las representaciones estilizadas, como las que se proponen en Germain y Martin (2000), representan un importante avance, y superan las limitaciones de los primeros sistemas de visualización señaladas, entre otros, por Chun (1998).

Por otra parte, el empleo de representaciones gráficas de la onda sonora asume que el estudiante es capaz de establecer la relación que existe, por un lado, entre la forma sonora y la forma gráfica, y por otro, entre la percepción del modelo que recibe y la imitación que él mismo realiza. Esto presenta el inconveniente de que, como se indica más adelante, la interacción entre producción y percepción es aún una cuestión debatida en lo que respecta a la adquisición de una lengua extranjera. También la capacidad de asociar representaciones acústicas con representaciones visuales es un aspecto que tal vez no cuenta todavía con los suficientes estudios empíricos que precisen en qué mecanismos se apoya y posiblemente sea necesaria una investigación en profundidad con el fin de evaluar adecuadamente su uso.

Finalmente, otra cuestión que plantea el uso de este tipo de representaciones se relaciona con un aspecto más fundamental: es bien sabido en el campo de la fonética que dos producciones sonoras de diferentes hablantes, o incluso de un mismo hablante en diferentes situaciones, no tienen idénticas características, por lo que su representación acústica presentará variaciones tanto en el dominio del tiempo como en los de la frecuencia y la intensidad. La figura 9 muestra la palabra catalana *hola* producida por dos locutores: pueden observarse fácilmente las diferencias en la duración de los sonidos y en la amplitud de la onda sonora, sin que por ello ninguna de las dos realizaciones plantee problemas de comprensión a un catalanohablante.

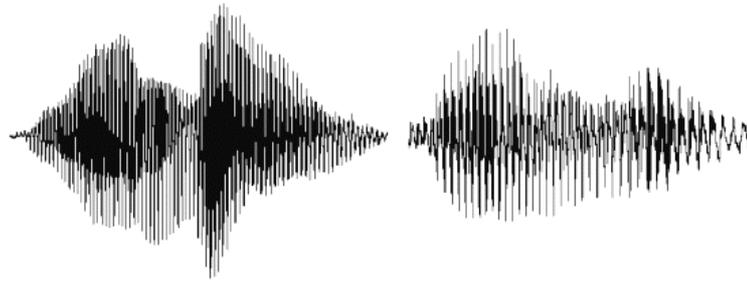


Figura 9: Oscilograma de la palabra catalana *hola* producida por un hablante femenino (izquierda) y un hablante masculino (derecha).

Cabe pues pensar si la exacta imitación de la forma sonora de un enunciado por parte de un estudiante no es un objetivo que, en cierta manera, obvia el hecho de que en la comunicación real tiene más peso la percepción por parte del receptor que los detalles individuales de la manifestación fonética de un enunciado.

Aunque en el presente trabajo no abordemos las cuestiones articulatorias, vale la pena mencionar que, en lo que respecta a la producción del habla, puede aplicarse el mismo razonamiento. El fenómeno conocido en fonética como “compensación articulatoria” _la posibilidad de llegar al mismo resultado acústico mediante diferentes posiciones del aparato fonador_ llevó, en cierto modo, a cuestionar los métodos de enseñanza de la pronunciación basados en la imitación de una determinada configuración del tracto vocal (Renard, 1979: 29-38). Sin embargo, como se aprecia en la figura 10, métodos recientes de enseñanza asistida por ordenador como *Talk to Me* recurren también a la representación articulatoria “modélica” para complementar la visualización de la onda sonora.

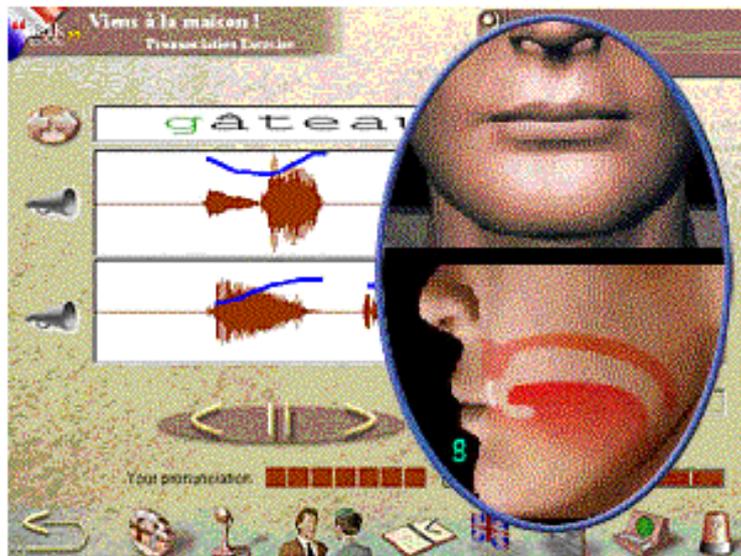


Figura 10: Ejemplo de representación articuladora del método *Talk to Me*. (<http://www.auralog.com/en/talktome.html>, 18/112001)

A pesar de lo dicho hasta aquí, las representaciones acústicas del habla tienen una indudable utilidad en la enseñanza de la pronunciación y en la corrección fonética, pues aportan una información adicional _la denominada “retroalimentación visual” _ que puede resultar interesante para muchos estudiantes si se entiende como un complemento y no como el único medio de enseñanza o de corrección. Trabajos como los de James (1976, 1977, 1979), mencionados por Germain y Martin (2000), mostraron ya en su momento la eficacia del uso de una representación visual de la curva melódica en la enseñanza de una lengua extranjera, y en la misma línea pueden citarse, entre otras, las aportaciones de Abberton (1972), Léon y Martin (1972), Weltens y Bot (1984), Chun (1989, 1998), Anderson-Hsieh (1992, 1994), Bot (1993), Wieringen y Abberton (1994), Bagshaw (1994), Stibbard (1996), Öster (1998), Rocca (1998), Germain y Martin (1998), y Taniguchi y Abberton (1999)

en lo que respecta a los elementos suprasegmentales, y las de Flege (1988), Lambacher (1996, 1999) y Murawaka y Lambacher (1996) para los elementos segmentales.

El uso del reconocimiento del habla

Una segunda posibilidad que representan las nuevas tecnologías es el uso del reconocimiento del habla en la enseñanza de la pronunciación y en la corrección fonética (Hiller *et al.*, 1993, 1994; Coniam, 1998, 1999; Eskénazi, 1999 a, b; Menzel *et al.*, 2000 a).

Las técnicas de reconocimiento de habla se basan, en los métodos más sencillos, en una comparación entre los patrones acústicos previamente almacenados en el sistema y los enunciados producidos por el hablante (Bernstein y Franco, 1996; Kurzweil, 1998; Levinson y Liberman, 1981; Ehsani y Knodt, 1998). Puesto que debe establecerse un criterio de similitud o umbral de aceptación, como aparece en la figura 15_, de modo que el reconocedor acepte o rechace enunciados más o menos parecidos al original, es posible aprovechar este criterio para proporcionar al alumno una puntuación, que no es otra cosa que la “distancia” entre el modelo que ha éste intentado imitar y su propia producción.

Tal estrategia se utiliza, por ejemplo, en *Talk to Me*, tal como muestra la figura 11, en la que encontramos la representación oscilográfica y la curva melódica junto con una puntuación en el panel señalado con el número 3.

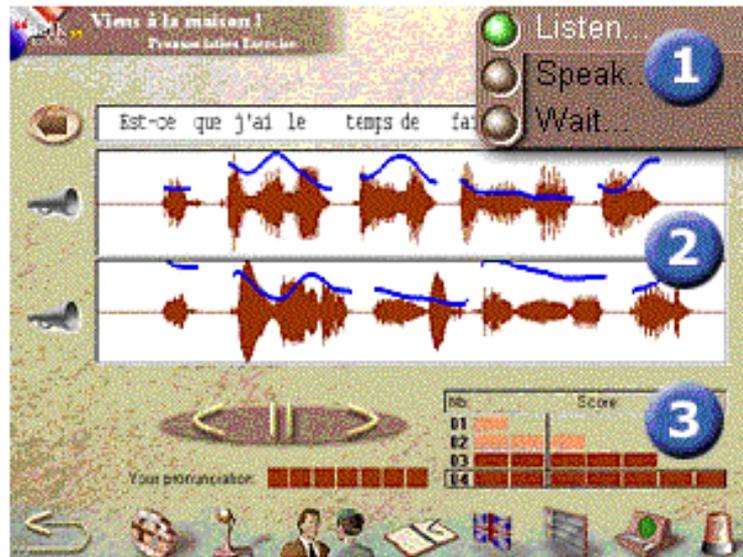


Figura 11: Ejemplo de uso del reconocimiento del habla para asignar una puntuación a la producción del estudiante en el método *Talk to Me*. (<http://www.auralog.com/en/talktome.html>, 18/112001)

La comparación entre el enunciado almacenado y el original requiere también una deducción, por parte del sistema de reconocimiento, de los elementos que forman el enunciado, de modo que no únicamente se realiza una comparación global, sino que se lleva a cabo lo que se conoce como una “alineación” _es decir, una sincronización_ entre el texto correspondiente al enunciado y la señal sonora. Esto permite señalar al estudiante aquellos fragmentos de su producción en los que se produce una discrepancia más fuerte con respecto al modelo, de modo que pueda concentrarse en ellos a la hora de repetir para mejorar la semejanza con el enunciado original. La figura 12 ilustra este procedimiento.



Figura 12: Ejemplo de uso del reconocimiento del habla para señalar los errores de pronunciación en el método *Talk to Me*. (<http://www.auralog.com/en/talktome.html>, 18/112001)

Las actuales tecnologías de reconocimiento permiten también que la detección de los errores se realice en el nivel subléxico, indicando al estudiante la puntuación obtenida en cada uno de los alófonos que componen el enunciado. Un ejemplo de esta posibilidad puede verse en el sistema WebGrader diseñado para evaluar la pronunciación a través de la *web* (Neumeyer *et al.*, 1998). Como se observa en la figura 13, se obtiene en primer lugar una representación del enunciado, en la que se señalan, en una escala de 0 a 100, las diferencias entre el modelo nativo y la producción del estudiante, junto con la puntuación global; la figura 14 muestra, en cambio, la puntuación obtenida en cada uno de los elementos segmentales.

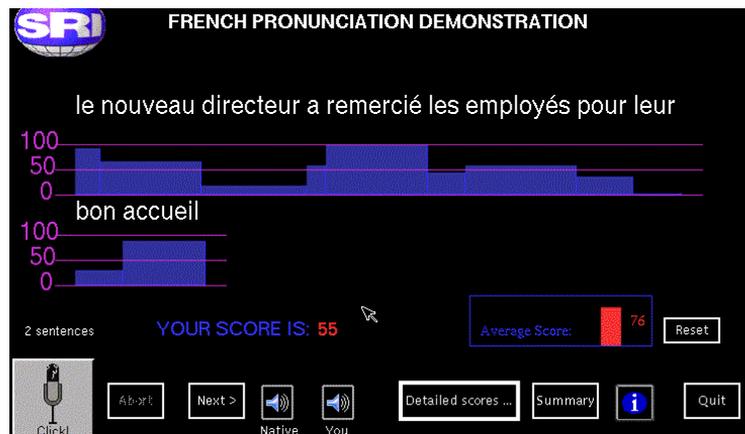


Figura 13: Ejemplo de uso del reconocimiento del habla para mostrar las diferencias globales entre el enunciado del estudiante y el modelo en el sistema *WebGrader*.

(<ftp://ftp.speech.sri.com/pub/people/julia/videos/webgrad.avi.gz>, 18/11/2001)

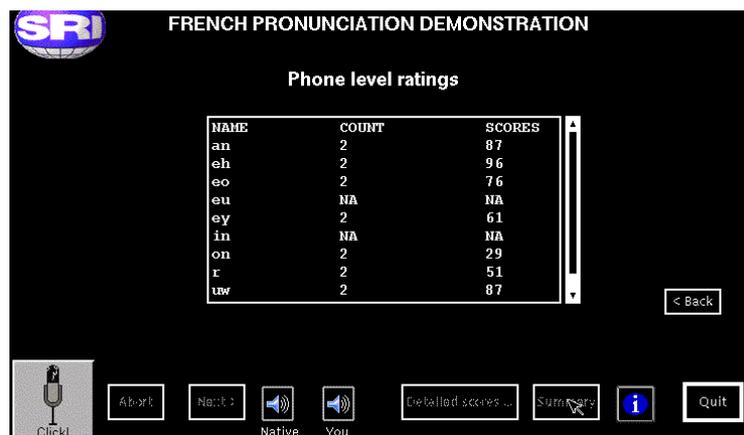


Figura 14: Ejemplo de uso del reconocimiento del habla para proporcionar la puntuación correspondiente a cada alófono en el sistema *WebGrader*.

(<ftp://ftp.speech.sri.com/pub/people/julia/videos/webgrad.avi.gz>, 18/11/2001)

Una de las mayores dificultades en la utilización del reconocimiento del habla para la enseñanza de la pronunciación radica en la propia tecnología del reconocimiento, basada, como hemos indicado, en la comparación entre patrones. La figura 15 presenta un esquema simplificado de los pasos que www.esespasa.com 17

seguiría un reconocedor para asignar una puntuación a la producción de un estudiante. Como puede observarse, las características fonéticas que se extraen del enunciado no nativo se comparan con los modelos fonéticos almacenados en el sistema, contando para ello con la ayuda de un diccionario fonético que, estableciendo la relación entre la onda sonora y los elementos segmentales que la constituyen, contribuye al proceso de alineación entre el texto y la señal sonora anteriormente comentado.

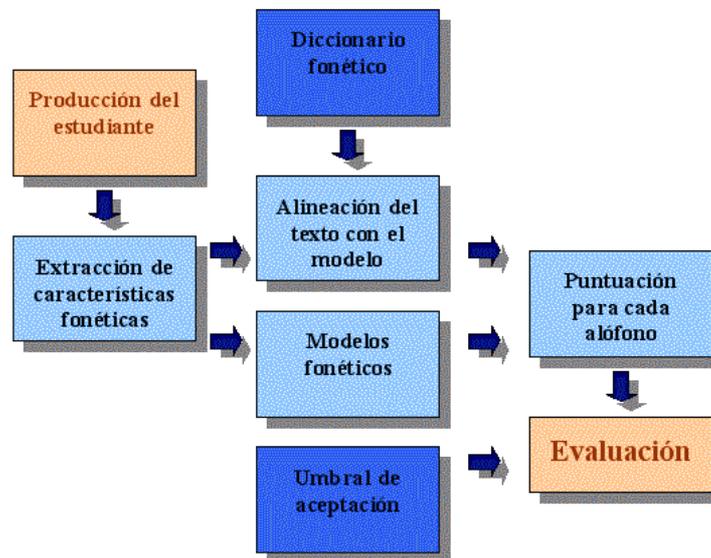


Figura 15: Esquema simplificado de un modelo de reconocimiento de habla aplicado a la evaluación de la pronunciación en una lengua extranjera. (Adaptado de Witt y Young 2000: 97)

La creación de modelos fonéticos con los que comparar los enunciados que llegan al reconocedor requiere, normalmente, el entrenamiento con un número muy elevado de muestras de habla procedente de diversos locutores, para lo cual se han constituido diversas bases de datos como las que distribuyen ELRA (*European Language Resources Association*) y LDC (*Linguistic Data*

Consortium). Sin embargo, por el momento son escasas las bases de datos que recogen amplias muestras de habla no nativa, por lo que cada sistema debe entrenarse con producciones recogidas *ad hoc*, teniendo en cuenta la lengua materna del estudiante y la lengua extranjera que se enseña.

El segundo problema al que se enfrenta la aplicación del reconocimiento del habla a la corrección fonética es la identificación de los errores _es decir, de las discrepancias entre los modelos fonéticos y el enunciado producido por el alumno que no se deben simplemente a la variabilidad propia del habla sino al acento no nativo_ (Herron *et al.*, 1999; Menzel *et al.*, 2000b; Bonaventura *et al.*, 2000) y la evaluación automática de la pronunciación de los estudiantes con el fin de asignar, en función del criterio de aceptación establecido en el sistema, una puntuación global o detallada (Witt y Young, 1997, 2000; Witt, 1999; Cucchiarini *et al.*, 2000a; Franco *et al.*, 2000; Neumeyer *et al.*, 2000). En este sentido, para el desarrollo de algunos de los sistemas actuales se han comparado las puntuaciones obtenidas automáticamente con las proporcionadas por expertos (Cucchiarini *et al.*, 2000 b), intentando emular de la forma más exacta posible la actividad que realizaría el profesor.

En conjunto, el reconocimiento del habla facilita al estudiante una evaluación global de la pronunciación y una indicación de los fragmentos de su producción en la que se encuentran mayores discrepancias respecto al modelo propuesto. Con ello se asume, en cierta manera, que el alumno va a ser capaz de modificar su propia producción, acercándose por aproximaciones sucesivas a una reproducción óptima del original. Sin embargo, no se le

proporciona ninguna indicación sobre cómo llegar a corregir sus errores, más allá de señalar aquellos puntos en los que el estudiante se separa de la pronunciación “modélica”. Parece, por tanto, que la estrategia pedagógica que se contempla en este caso consiste en señalar los errores y ofrecer un modelo que se repite y se puntúa tantas veces como sea necesario.

La integración del sistema fonético de una lengua extranjera

Desde las observaciones pioneras de Trubetzkoy (1939) y Polivanov (1931) se conoce, en lo que respecta a la adquisición de los sonidos de una lengua extranjera, el papel de la interferencia del sistema fonológico de la lengua materna. Partiendo de estas bases, corrientes como el método verbo-tonal (Renard, 1979) insisten en la importancia de la percepción y en la necesidad de lograr una auténtica reestructuración del sistema fonético con el fin de minimizar los efectos de la interferencia. Aunque se trate de una cuestión en la que no existe una total unanimidad, diversos autores apuntan a la prioridad de la percepción sobre la producción _un estado de la cuestión se presenta en Llisterra (1995) _, de modo que difícilmente puede un estudiante producir un contraste fonético que no percibe adecuadamente debido a la influencia de su lengua materna. Precisamente en este aspecto se basa la crítica a los métodos de enseñanza de la pronunciación basados en la repetición y en la imitación que se encuentra en los trabajos que aceptan los postulados del sistema verbo-tonal.

En este sentido, los procedimientos basados en la repetición de determinados enunciados no resolverían _al igual que no lo harían los clásicos laboratorios de idiomas (Renard, 1979: 38-40) _ el problema fundamental de la interferencia perceptiva de los propios patrones fonéticos y fonológicos, aunque se apoyen en representaciones visuales o proporcionen una puntuación que indique la similitud entre el modelo y su imitación. Los sistemas que, de algún modo, señalan la localización del error o enfatizan aquellas partes lingüísticamente significativas del enunciado visualizado presentan ventajas sobre los que ofrecen únicamente la posibilidad de repetir un modelo y constituyen, probablemente, una de las vías por las que podría avanzar la aplicación del reconocimiento del habla a la enseñanza de la pronunciación.

Incluso así, la tecnología actual se encuentra aún relativamente lejos de ofrecer estrategias de corrección similares a las que un profesor con una buena formación en fonética integra en la enseñanza presencial, diagnosticando en cada caso la causa del error, ofreciendo un modelo que ponga de relieve aquellos parámetros fonéticos que el estudiante no percibe y facilitando la producción recurriendo, por ejemplo, a los elementos suprasegmentales, a la fonética sintáctica, o a la pronunciación matizada (Renard, 1979: 69-96).

En todo caso, las herramientas descritas en este trabajo representan una importantísima ayuda en el contexto del autoaprendizaje y, a pesar de sus limitaciones prácticas y conceptuales, ofrecen al estudiante un apoyo

indiscutible. Corresponde a los profesionales de la enseñanza, integrados en un equipo multidisciplinar, evaluar y mejorar los nuevos prototipos, productos y métodos para intentar conjugar los avances de la tecnología con los conocimientos actuales sobre la integración del sistema fonético de una lengua extranjera.

Conclusión

En un opúsculo titulado *La phonétique expérimentale appliquée à l'enseignement des langues vivantes*, publicado por la Alliance Française de París en 1898, Adolphe Zünd-Burguet, investigador en el Laboratorio de Fonética Experimental del Collège de France, reflexionaba con estas palabras sobre el uso del llamado “cadran-indicateur”, un aparato diseñado para mostrar la posición de los articuladores en la producción de un sonido:

“Pour que l'on ne se méprenne pas sur la portée de ce que je viens de dire au sujet du cadran-indicateur je tiens à ajouter que cet appareil, destiné a rendre de grands services *sous la direction d'un maître* [cursiva del autor], n'offre pas les même avantages à l'élève non dirigé. Le cadran n'enseigne pas d'une façon directe la bonne prononciation. Il a été construit pour aider l'élève dans l'éducation de son ouïe et des organes de la parole en le renseignant sur la nature de certains efforts et mouvements organiques. Encore ces renseignements sont-ils relatifs, vu qu'il n'y a rien d'absolu dans la production des sons du langage. Presque tous peuvent, en effet, être prononcés de deux

ou plusieurs manières sans que notre oreille ne soit ni choquée ni même avertie” (Zünd-Burguet, 1898: 32).

No deja de ser curioso que hace más de un siglo Zünd-Burguet planteara varias cuestiones que aún siguen siendo relevantes en la actualidad: la necesidad de educar el oído _es decir, el papel primordial de la percepción_, el papel del profesor como guía del proceso de aprendizaje, las limitaciones del autoaprendizaje en la enseñanza de la pronunciación, la utilidad de las ayudas técnicas y la negligencia de la variabilidad inherente a la producción del habla en los métodos basados en la imitación de determinado tipo de representaciones, articulatorias en el caso de Zünd-Burguet o acústicas en los ejemplos que hemos mostrado en este trabajo. Parece pues que las nuevas tecnologías hacen resurgir los viejos problemas y que, una vez más, obligan a plantearse la actuación del profesor en lo que se refiere a la adquisición del sistema fonético de una lengua extranjera.

La consideración final de Zünd-Burguet en el librito mencionado resulta también de sorprendente actualidad y con ella, substituyendo la “fonética experimental” por las “nuevas tecnologías” utilizadas en la enseñanza de la pronunciación y la corrección fonética, podría ser adecuado concluir este trabajo:

“La méthode scientifique de la phonétique expérimentale est certainement la plus efficace et la plus rationnelle; mais mal comprise et mal employée, elle ne

pourrait que nous induire en erreur ou donner une fausse sécurité” (Zünd-Burguet, 1898: 32).

BIBLIOGRAFÍA

ABBERTON, E. (1972): "Visual Feedback and Intonation Learning". En Rigault, A. - Charbonneau, R. (eds.): *Proceedings of the Seventh International Congress of Phonetic Sciences*. The Hague, Mouton.

ANDERSON-HSIEH, J. (1992): "Using electronic visual feedback to teach suprasegmentals". *System*, 20, I, pp. 51-62.

ANDERSON-HSIEH, J. (1994): "Interpreting visual feedback on suprasegmentals in computer assisted pronunciation instruction". *CALICO Journal*, 11, IV, pp. 5-22.

BAGSHAW, P. (1994): *Automatic prosodic analysis for computer aided pronunciation teaching*. PhD Thesis. Edinburgh, Center for Speech Technology Research, University of Edinburgh. <http://www.cstr.ed.ac.uk/~pcb/thesis.html> (18/11/2001).

BERNSTEIN, J. y FRANCO, H. (1996): "Speech recognition by computer". En Lass, N. J. (ed.): *Principles of Experimental Phonetics*, pp. 408-434. St Louis, Mosby.

BONAVENTURA, P.; HERRON, D. y MENZEL, W. (2000): "Phonetic rules for diagnosis of pronunciation errors". En *Tagungsband 5. Konferenz Verarbeitung natürlicher Sprache, Konvens 2000*, Ilmenau, pp. 225-230.
<http://nats-www.informatik.uni-hamburg.de/~wolfgang/papers/konvens2000.ps.gz> (18/11/2001).

BOT, K. de (1983): "Visual feedback of intonation I: Effectiveness and induced practice behavior". *Language and Speech*, 26, IV, pp. 331-350.

CAZADE, A. (1999): "De l'usage des courbes sonores et autres supports graphiques pour aider l'apprenant en langues". *ALSIC, Apprentissage des Langues et Systèmes d'Information et de Communication*, 2, II, pp. 3-32.
<http://alsic.univ-fcomte.fr/Num4/cazade/default.htm> (18/11/2001).

CONIAM, D. (1998): "The use of speech recognition software as an English language oral assessment instrument: An exploratory study". *CALICO Journal*, 15, IV, pp. 7-23.

CONIAM, D. (1999): "Voice Recognition Software Accuracy with Second Language Speakers of English". *System*, 27, I, pp. 49-64.

CUCCHIARINI, C.; STRIK, H. y BOVES, L. (2000a): "Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology". *Journal of the Acoustical Society of America*, 107, II, pp. 989-999.

CUCCHIARINI, C.; STRIK, H. y BOVES, L. (2000b): "Different aspects of expert pronunciation quality ratings and their relation to scores produced by speech recognition algorithms". *Speech Communication*, 30, II-III, pp. 109-120.

CHUN, D. M. (1989): "Teaching tone and intonation with microcomputers". *CALICO Journal*, 7, I, pp. 21-46.

CHUN, D. M. (1998): "Signal Analysis Software For Teaching Discourse Intonation". *LLT, Language Learning & Technology*, 2, I, pp. 61-77.
<http://polyglot.cal.msu.edu/llt/vol2num1/article4/index.html> (18/11/2001).

EHSANI, F. y KNODT, E. (1998): "Speech Technology in Computer-Aided Language Learning: Strengths and Limitations of a New CALL Paradigm", *LLT, Language Learning & Technology*, 2, I, pp. 45-60.
<http://llt.msu.edu/vol2num1/article3/> (18/11/2001).

ELRA, *European Language Resources Association*.
<http://www.elda.fr/catalog.html> (18/11/2001).

ESKÉNAZI, M. (1999a): "Using Automatic Speech Processing for Foreign Language Pronunciation Tutoring: Some Issues and a Prototype". *LLT, Language Learning & Technology*, 2, II, pp. 62-76.
<http://polyglot.cal.msu.edu/llt/vol2num2/article3/index.html> (18/11/2001).

ESKÉNAZI, M. (1999b): "Using Automatic Speech Processing for Foreign Language Pronunciation Tutoring: Some Issues and a Prototype", *LLT*,

Language Learning & Technology, 2, II, pp. 62-76.
<http://polyglot.cal.msu.edu/lt/vol2num2/article3/index.html> (18/11/2001).

ESTRUCH, M.; GARRIDO, J. M.; LLISTERRI, J. y RIERA, M. (1996-97): "Una aproximación fonética al estudio de la entonación". *Philologia Hispalensis*, 11, pp. 281-293.

FLEGE, J. E. (1988): "Using visual information to train foreign language vowel production". *Language Learning*, 38, III, pp. 365-407.

FRANCO, H.; NEUMEYER, L.; DIGALAKIS, V. y RONEN, O. (2000): "Combination of machine scores for automatic grading of pronunciation quality". *Speech Communication*, 30, II-III, pp. 121-130.

GERMAIN, A. y MARTIN, P.h. (1998): "WinPitch Language Teaching, Easy Learning (LTEL) : Un outil multimédia pour faciliter l'apprentissage de l'oral en langue seconde". En *Conférence virtuelle sur les applications des TIC dans l'enseignement du français langue Étrangère*. Agence intergouvernementale de la Francophonie. Du 7 au 18 décembre 1998.
<http://ciffad.francophonie.org/CONFFLE/RESSOURCES/contrib-martin.html>
(18/11/2001).

GERMAIN, A. y MARTIN, Ph. (2000): "Présentation d'un logiciel de visualisation pour l'apprentissage de l'oral en langue seconde". *ALSIC, Apprentissage des Langues et Systèmes d'Information et de Communication*,

3, 1, pp. 61-76. <http://alsic.univ-fcomte.fr/Num5/germain/default.htm>
(18/11/2001).

HERRON, D.; MENZEL, W.; ATWELL, E.; BISIANI, R.; DANELUZZI, F.;
MORTON, R. y SCHMIDT, J. A. (1999): "Automatic localization and diagnosis
of pronunciation errors for second-language learners of English". En
*Eurospeech'99. 6th European Conference on Speech Communication and
Technology*. September 5-9, 1999, Budapest, Hungary. pp. 855-858.
[http://nats-www.informatik.uni-
hamburg.de/~wolfgang/papers/eurospeech99.ps.gz](http://nats-www.informatik.uni-hamburg.de/~wolfgang/papers/eurospeech99.ps.gz) (18/11/2000).

HILLER, S.; ROONEY, E.; LAVER, J. y JACK, M. (1993): "SPELL: An
automated system for computer-aided pronunciation teaching". *Speech
Communication*, 13, pp. 463-473.

HILLER, S.; ROONEY, E.; VAUGHAN, R.; ECKERT, M.; LAVER, J. y JACK,
M. (1994): "An automated system for computer-aided pronunciation learning".
Computer Assisted Language Learning, 7,1, pp. 51-63.

JAMES, E. (1976): "The acquisition of prosodic features of speech using a
speech visualizer". *IRAL, International Review of Applied Linguistics*, 14, III,
pp. 227-243.

JAMES, E. (1977): "The acquisition of second language intonation using a
visualizer". *Canadian Modern Language Review*, 33, IV, pp. 503-506.

JAMES, E. (1979): "Intonation through visualization". En Hollien, H. y Hollien, P. (eds.): *Current Issues in the Phonetic Sciences*, pp. 295-301. Ámsterdam, John Benjamins.

KENT, R. D. y READ, Ch. (1992): *The Acoustic Analysis of Speech*. London - San Diego, Whurr Publishers - Singular Publishing Group.

KURZWEIL, R. (1998): "When Will HAL Understand What We Are Saying? Computer Speech Recognition and Understanding". En Stork, D. G. (ed.): *Hal's Legacy: 2001's Computer as Dream and Reality*. Cambridge, Mass., The MIT Press. <http://mitpress.mit.edu/e-books/Hal/chap7/seven1.html> (18/11/2001).

LADEFOGED, P. (1996): *Elements of Acoustic Phonetics*. Chicago – London, University of Chicago Press. Second Edition.

LAMBACHER, S. (1999): "A CALL tool for improving second language acquisition of English consonants by Japanese learners". *Computer Assisted Language Learning*, 12, II, pp. 137-156.

LAMBACHER, S. (1996): *Teaching English Pronunciation Using a Computer Visual Display*. <http://www.u-aizu.ac.jp/~steeve/york95.html> (18/11/2001).

LDC, *Linguistic Data Consortium*.

http://www ldc.upenn.edu/Catalog/by_type.html (18/11/2001).

LÉON, P. y MARTIN, P. (1972) "Applied linguistics and the teaching of intonation". *Modern Language Journal*, 56, pp. 139-44.

LEVINSON, S. E. y LIBERMAN, M. Y. (1981): "Speech Recognition by Computer", *Scientific American*, 244, pp. 64-76. Trad. cast. de R. Cerdà: "Reconocimiento del habla por medio de ordenadores", *Investigación y Ciencia*, 1981, pp. 38-51. En Agulló, J. (ed.) (1989): *Acústica musical*, pp. 106-121. Barcelona, Prensa Científica (Libros de Investigación y Ciencia).

LLISTERRI, J. (1995): "Relationships between Speech Production and Speech Perception in a Second Language". En Elenius, K. y Branderud, P. (eds.): *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Stockholm, Sweden, 13-19 August, 1995 [Vol. IV], pp. 92-99. Stockholm, KTH / S t o c k h o l m U n i v e r s i t y .
http://liceu.uab.es/~joaquim/publicacions/Prod_Percep.html (18/11/2001).

MARTÍNEZ LEDESMA, J.; GONZÁLEZ RUS, G. y LÓPEZ TORRECILLAS, M. (2000): "Tarjetas reconocedoras de voz. IBM Speechviewer III (Visualizador fonético)". *Boletín de AELFA*, 3, pp. 12-15.

MENZEL, W.; HERRON, D.; BONAVENTURA, P. y MORTON, R. (2000b): "Automatic detection and correction of non-native English pronunciation". En *InStil 2000, Proceedings of the Workshop Intergrating Speech Technology in the (Language) Learning and Assistive Interface*, pp. 49-56. Dundee, UK.
http://nats-www.informatik.uni-hamburg.de/~wolfgang/papers/instil2000_diagnosis.ps.gz (18/11/2001).

MENZEL, W.; HERRON, D.; MORTON, R.; PEZZOTTA, D.; BONAVENTURA, P. y HOWARTH, P. (2000a): "Interactive Pronunciation Training". *ReCALL*, 13, 1, pp. 67-78.

<http://nats-www.informatik.uni-hamburg.de/~wolfgang/papers/recall2000.ps.gz>

(18/11/2001).

MURAWAKA, H. y LAMBACHER, S. G. (1996): *Improving Japanese Pronunciation of American English [r] Using Electronic Visual Feedback*.

<http://www.u-aizu.ac.jp/~steve/rsound2.html> (18/11/2001).

NEUMEYER, L.; FRANCO, H.; ABRASH, V.; JULIA, L.; RONEN, O.; BRATT, H.; BING, J.; DIGALAKIS, V. y RYPA, M. (1998): "WebGrader™: A multilingual pronunciation practice tool". En *ESCA Workshop on Speech Technology in Language Learning (STiLL 98)*. Proceedings. Marholmen Conference Centre, Sweden, May 24-17, 1998. pp. 61-64. Stockholm: ESCA - Department of Speech, Music and Hearing, KTH.

<ftp://ftp.speech.sri.com/pub/people/julia/papers/still98.ps.gz> (18/11/2001).

NEUMEYER, L.; FRANCO, H.; DIGALAKIS, V. y WEINTRAUB, M. (2000): "Automatic scoring of pronunciation quality". *Speech Communication*, 30, 2-3, pp. 83-94.

ÖSTER, A.-M. (1998): "Spoken L2 Teaching with Contrastive Visual and Auditory Feedback". En *ICSLP 98, Proceedings of the 5th International Conference on Spoken Language Processing*. Sydney Convention Centre,

Sydney, Australia. 30th November - 4th December 1998. CD Rom edition.
Rundle Mall: Casual Production. Paper n. 256.

POLIVANOV, E. (1931): "La perception des sons d'une langue étrangère".
Travaux du Cercle Linguistique de Prague 4. En *Le Cercle de Prague*
(Change, 3) Paris, 1969. pp. 111-14.

Praat: doing phonetics by computer. Institute of Phonetic Sciences, University
of Amsterdam. <http://www.fon.hum.uva.nl/praat/> (18/11/2001).

PUYUELO, M. (1992): "Visualizador fonético Speech Viewer. Características".
Revista de Logopedia, Foniatría y Audiología, 12, 1, pp. 45-47.

RENARD, R. (1979): *Introduction à la méthode verbo-tonale de correction
phonétique*. Troisième édition entièrement refondue. Bruxelles / Centre
International de Phonétique Appliquée, Didier / Mons.

ROCCA, P.D.A. (1998): "The efficacy of computer-driven visual feedback in
the teaching of intonation to Brazilian learners of English". En *ESCA Workshop
on Speech Technology in Language Learning (STiLL 98)*. Proceedings.
Marholmen Conference Centre, Sweden, May 24-17, 1998. pp. 139-142.
Stockholm, ESCA - Department of Speech, Music and Hearing, KTH.

S p e a k e r *A u t e u r*. *N e u r o c o n c e p t*.
http://www.neuroconcept.com/pages/outil_speaker/outil_speak.htm
(18/11/2001).

SpeechViewer III. IBM Accessibility Center. <http://www-3.ibm.com/able/snsspv3.html> (18/11/2001)

STEVENS, K.N. (1998): *Acoustic Phonetics*. Cambridge, MA., The MIT Press (Current Studies in Linguistics, 30).

STIBBARD, R. (1996): "Teaching English intonation with a visual display of fundamental frequency". *The Internet TESOL Journal*, 2, VIII. <http://www.aitech.ac.jp/~iteslj/Articles/Stibbard-Intonation/> (18/11/2001)

Talk to Me. Auralog. <http://www.auralog.com/en/talktome.html> (18/11/2001).

TANIGUCHI, M. y ABBERTON, E. (1999): "Effect of interactive visual feedback on the improvement of English intonation of Japanese EFL learners". *Speech, Hearing and Language: work in progress* (University College London, Department of Phonetics and Linguistics), 11, pp. 76-89. http://www.phon.ucl.ac.uk/home/shl11/pdf_files/taniguchi.pdf (18/11/2001)

Tell me More. Auralog. <http://www.auralog.com/en/tellmemore.html> (18/11/2001).

TROUBETZKOY, N. S. (1939): "Grundzüge der Phonologie". *Travaux du Cercle Linguistique de Prague*, 7. Traducción francesa de J Cantineau: *Principes de Phonologie*. Paris, Klincksieck, 1949. Traducción castellana de D García Giordano, con la colaboración de L. J. Prieto: *Principios de fonología*. Madrid, Cincel, 1973.

WebGrader. SRI International.

<http://www-speech.sri.com/people/julia/webgrader.html>;

<http://www-speech.sri.com/people/julia/webgrader2.html> (18/11/2001).

WELTENS, B. y BOT, K. de (1984): "Visual feedback of intonation II: Feedback delay and quality of feedback". *Language and Speech*, 27, 1, pp. 79-88.

WIERINGEN, M. van y ABBERTON, E. (1994): "The use of computerized visual representation in L2 acquisition of intonation: a pilot study". *Speech, Hearing and Language, Work in Progress, 1994* (University College London, Department of Phonetics and Linguistics), 8, pp. 245-258.

WITT, S. M. (1999): *Use of Speech Recognition in Computer-assisted Language Learning*. PhD Thesis. Department of Engineering, University of Cambridge.

ftp://svr-ftp.eng.cam.ac.uk/pub/reports/witt_thesis.ps.gz (18/11/2001).

WITT, S. M. y YOUNG, S. (1997): "Language learning based on non-native speech recognition". En *Eurospeech'97. Proceedings of the 5th European Conference on Speech Communication and Technology*. pp. 633-636. Rhodes, Greece, 22-25 September 1997.

ftp://svr-ftp.eng.cam.ac.uk/pub/reports/witt_euro97.ps.Z

<http://svr-www.eng.cam.ac.uk/~smw24/eurospeech.ps> (118/11/2001).

WITT, S. M. y YOUNG, S. J. (2000): "Phone-level pronunciation scoring and assessment for interactive language learning". *Speech Communication*, 30, II-III, pp. 95-108.

ZÜND-BURGUET, A. (1898): *La Phonétique Expérimentale appliquée à l'enseignement des langues vivantes*. Paris, Alliance Française.

Una bibliografía más detallada se recoge en:

LLISTERRI, J. *Computer-assisted Pronunciation Teaching*. Departament de Filologia Espanyola, Universitat Autònoma de Barcelona. http://liceu.uab.es/~joaquim/applied_linguistics/L2_Phonetics/CALL_Pron_Bib.html (18/11/2001).

El lector interesado puede encontrar más información sobre la aplicación de las tecnologías del habla a la enseñanza de la pronunciación en las páginas de InSTIL (*Integrating Speech Technology in (Language) Learning, CALICO Special Interest Group*): <http://dbs.tay.ac.uk/instil/> (18/11/2001).