

El análisis de la entonación. Corpus oral de las variedades del español del norte

María Pilar BALLESTEROS PANIZO

Universidad de Barcelona

<http://www.ub.edu/lfa>

mapiballesteros@ub.edu

RESUMEN

El análisis melódico de las variedades del español es uno de los objetivos que se ha planteado el proyecto *Análisis Melódico de Habla*. Se trata de un proyecto iniciado en 2010 por un grupo de investigadores de la *Universidad de Barcelona* y la *Universidad Autónoma de Barcelona*, que busca el acceso a una mejor caracterización de la noción de *variedad*, a partir del conocimiento del papel que desempeña la entonación en cada dialecto.

En este artículo se presentan los fundamentos y las variables que se han considerado para la constitución del *corpus del español del norte de España* (en adelante CENE). También se aportan detalles relevantes acerca de su recolección y procedencia. Como el lector tendrá ocasión de comprobar, se trata de un corpus extenso y variado que permite hacer una descripción fidedigna de los patrones entonativos de las diferentes variedades del habla espontánea. Esto, a su vez, facilita la tarea de llegar a tener un conocimiento más profundo de las características dialectales del castellano. Así, la investigación tiene interés tanto para la entonología como para la dialectología. Sin embargo, CENE aspira a ser examinado por especialistas de diferentes ámbitos lingüísticos, pues el conocimiento completo y exacto del lenguaje y del fenómeno entonativo no es posible si se adopta una sola perspectiva.

Palabras clave: entonación, melodía, teoría, corpus.

ABSTRACT

A group of researchers from the Universidad de Barcelona and the Universidad Autónoma de Barcelona are seeking access to a better characterization of the notion of variety, from the knowledge of the role of intonation in each dialect.

This article presents the fundamentals and the variables that have been considered for the constitution of the corpus of the Spanish of the North of Spain (hereinafter CENE). It also provides relevant details about its collection. As the reader will have the opportunity to verify, CENE is an extensive and varied corpus that allows making a reliable description of the intonation of the different varieties of spontaneous speech patterns. This, in turn, makes it easy to have a deeper understanding of the dialect features. Thus, the research is of interest for both the field of intonation and for dialectology. However, CENE aspires to be examined by specialists in different language areas because a complete and accurate knowledge of the language and the intonational phenomenon is not possible from a single perspective.

Keywords: intonation, melody, theory, corpus.

Sumario

1. Sobre el concepto y el interés de la entonación. 2. Objetivos del trabajo. 3. Características de un corpus eficaz. 3.1. Procedimiento de elaboración del corpus. 3.2. El criterio de segmentación de las unidades de análisis. 3.3. La transcripción del corpus. 3.4 El etiquetado del corpus. 4. Resumen

1. Sobre el interés y el concepto de la entonación

La entonación es un fenómeno protagonista en la conversación (que es, a su vez, el acontecimiento lingüístico por antonomasia de la comunicación humana). Permite, por ejemplo, expresar e inferir significados llenos de matices (Bolinger, D.L. 1989); compartir y reconocer los estados mentales de los interlocutores y así, construir un universo compartido, una intersubjetividad que es, en última instancia, la que hace posible el entendimiento (Ladd, D.R., 1980; Escandell Vidal, M. V., 1999, 2005). Se trata, además, de uno de los elementos que más determinan el modo de ser de las lenguas y de sus variedades: los hábitos entonativos de la lengua materna tienden a transferirse a las lenguas segundas de manera automática como si existiera una resistencia especial por parte de los aprendices a despojarse de ellos (Cortés Moreno, M. 1999; Liu, Y-H 2005). Además, es un fenómeno que delata la procedencia geográfica de los hablantes. Principalmente, sabemos que una persona es asturiana, castellana, vasca o madrileña por la peculiar entonación con que habla, -aunque el melódico no es el único elemento que hace posible esta identificación- (Sosa, J.M., 1999).

Por eso, el modo en que se estructura fónicamente la emisión de voz es uno de los aspectos más interesantes en el conocimiento de una lengua (Hidalgo, A. 1997). La entonación no es un añadido, algo accesorio sino que forma parte de la identidad lingüística de las personas (Poyatos, F., 1994).

Se trata, por tanto, de un componente esencial de la explicación de la actividad lingüística y un ámbito de interés para todas las ciencias que estudian el lenguaje y la conducta lingüística. Aunque habitualmente se reconoce el importante papel que desempeña la entonación, ésta resulta difícil de describir porque, al actualizarse, se combina con el componente pragmático, gramatical y no verbal para construir el significado básico del mensaje y para completar su sentido. Por eso, el análisis de la entonación presenta especiales dificultades teóricas y metodológicas que han hecho que tradicionalmente haya tendido a ser soslayado.

Pero en las últimas décadas del siglo XX y con el desarrollo de medios instrumentales y software cada vez más eficaces, se han consolidado diferentes

escuelas¹ dedicadas al estudio de la entonación de diferentes lenguas. Sin embargo, cada grupo parte de enfoques teóricos y metodológicos distintos, y pocos especialistas son capaces de interpretar con soltura los resultados de las demás escuelas. De alguna manera, esto propicia que la entonología siga hoy en un estado paradójico de crecimiento acelerado y de infecundidad lamentable porque, aunque se trabaja y se publica mucho, los frutos de las investigaciones de cada grupo no circulan en la comunidad académica con el respaldo unánime de sus especialistas: las escuelas se han encerrado en su propia retórica.

Si buscamos la causa de esta crisis la encontramos en el concepto de entonación que subyace en cada uno de los modelos. Definir acertadamente qué es la entonación, establecer cuál es la unidad de análisis de los entonólogos, y determinar qué tipo de significado aporta, si es que lo hace, son cuestiones constitutivas que no pueden pasarse por alto si lo que se pretende es un estudio riguroso.

Para entender esta afirmación es necesario aclarar que en la mayoría de obras dedicadas al estudio de la entonación, los términos frecuencia fundamental, melodía y entonación, no tienen una clara definición y, en ocasiones, se emplean como sinónimos. Cantero (2002) aborda esta cuestión y aclara que la *frecuencia fundamental* es un parámetro acústico originado por las vibraciones de los pliegues que se mide en valores absolutos (Hz). En cambio, la *melodía* es la sucesión de tonos: las variaciones relativas de tono a lo largo de la emisión. La *entonación*, por su parte, es un fenómeno lingüístico, la interpretación fonológica de las variaciones relativas de tono (de la melodía).

Para ofrecer una definición legítima del concepto de entonación es necesario conocer bien su forma y sus funciones. Esto parece abocar a la necesidad de una colaboración interdisciplinar de amplio alcance. En cuanto a la forma hay que

¹ No analizaremos aquí todas las propuestas teóricas y procedimentales existentes, ni abordaremos la comparación de las consecuencias prácticas que surgen de todas las hipótesis explicativas que se han producido (para esta cuestión ver Hidalgo, 2003). Para establecer un boceto general de tendencias se pueden distinguir dos rutas paralelas en las que se han establecido diferentes escuelas:

Por un lado, el análisis por configuraciones, que considera la realidad entonativa como un todo distinto de las partes que lo componen y que aborda su descripción fonética. Es el enfoque es el que adoptan tanto los autores de la *escuela británica* como Navarro Tomás; y más recientemente, la *escuela holandesa* y el grupo AMPER

La segunda ruta recibe el nombre de análisis por niveles y aborda el análisis fonológico (no fonético) de la entonación y. Se trata de un enfoque que considera que los contornos entonativos están conformados por una serie de puntos distintivos y estáticos que funcionan como unidades contrastivas. Dentro de esta perspectiva pueden distinguirse un enfoque *por niveles melódico* que representa la *escuela norteamericana* estructuralista, y un enfoque *por niveles métrico*, que surge con la irrupción del generativismo y produce varios modelos; el *modelo métrico-autosegmental* y el *modelo Aix en Provence*.

aceptar que acento y entonación dependen del mismo parámetro físico: la frecuencia fundamental (Cantero, 2002). Acento y entonación, por tanto forman parte de una misma realidad de manera que la estructura acentual constituye la estructura entonativa. Son fenómenos que vienen informados por el mismo parámetro y que se funden para establecer una *jerarquía fónica*. Sin embargo, la relación entre acento y entonación es uno de los puntos que peor han abordado las distintas tradiciones.

Para llevar a cabo un análisis fonético (melódico) de la entonación, hay que extraer únicamente los valores frecuenciales de los segmentos vocálicos, que son los únicos que contienen información melódica relevante (más adelante nos detendremos en la explicación y la justificación de esta idea). Por tanto, una primera definición de entonación, compatible con la realidad física de la melodía puede ser la que ofrece Cantero (2002:18) y que tan ampliamente ha sido acogida en el ámbito de la fonética, a saber “las variaciones de F0 que cumplen una función lingüística (entiéndase fonológica) en una emisión de voz”. Ahora bien, ¿se agotan en esta definición todos los elementos conceptualmente relevantes para una teoría de la entonación? Desde luego, parece que para comprender el papel que la entonación desempeña en la comunicación humana no basta con analizar la forma que ésta presenta (dimensión formal), su materia prima. Es necesario desplazar la atención desde el análisis melódico, en sí mismo, hacia el usuario que hace uso de la entonación en sus intercambios comunicativos.

La entonación, como el lenguaje, tiene tres dimensiones que pueden ser estudiadas científicamente. La *dimensión formal*, o *estructural*, la *dimensión funcional*, y la *dimensión comportamental*. El estudio del componente o *dimensión formal* está directamente vinculado a la necesidad de conocer y dar cuenta de cómo es el sistema lingüístico en sí mismo; la descripción de la *dimensión funcional*, por su parte, ha de buscar explicar para qué le sirve al usuario y, finalmente, la *dimensión comportamental* ha de encaminarse a explicar cómo se interrelaciona la entonación con el resto de elementos que intervienen en la producción y la comprensión de mensajes.

Esta caracterización, que sólo sirve como punto de partida implica considerar la entonación a) como un fenómeno físico (tonal); b) como instrumento mediador con que, en última instancia, se pueden hacer cosas; y c) como un tipo de conducta que adopta dos modalidades, la producción y la recepción: interesa explicar cómo codifican y se descodifican los mensajes para determinar qué papel desempeña la entonación en su asignación.

La tendencia metodológica más generalizada consiste en estudiar la entonación a partir enunciados leídos o inducidos por el propio investigador: enunciados que, en definitiva, están extraídos de su contexto natural (el diálogo) y que, además, tienen las limitaciones derivadas de la consideración de la *oración* como unidad de análisis por excelencia. Como es sabido, la actividad lingüística es, en realidad, una actividad cooperativa que implica intercambio de creencias, conocimientos y

deseos, entre personas que se consideran en posesión de un estado mental determinado y, dichos intercambios, no exigen acudir a formas gramaticales concretas sino a unidades definidas por su fuerza ilocucionaria.

En este artículo se describe un corpus de habla espontánea elaborado para el estudio de la dimensión formal (análisis melódico) del español coloquial. Se trata de un corpus que consta de 1000 secuencias de voz emitidas en un contexto dialógico que permiten abordar la tarea más importante que ha de realizar la filología actual - el estudio de uso idiomático espontáneo-; pero también puede facilitar la indagación psicológica de la conducta lingüística. CENE se ha elaborado buscando el conocimiento de la dimensión estructural del fenómeno entonativo, porque consideramos que es necesario empezar analizando las realidades más materiales siempre que no se pierda de vista que ese estudio es sólo una fase preliminar de la tarea. En otras palabras, estudiar sólo los aspectos melódicos de las realizaciones orales es un reduccionismo, una abstracción porque la entonación no funciona de manera autónoma, independientemente de la intencionalidad y de las funciones comunicativas o como si le fueran ajenas las dimensiones léxica, sintáctica o pragmáticas.

Por ejemplo, y en un nivel de análisis puramente lingüístico, entonación y sintaxis funcionan conjuntamente y no se puede adjudicar valores de contenido que descansan en ambos niveles, a una o a otra de manera arbitraria. Si la sintaxis sólo considera las estructuras proposicionales transmitidas literalmente en los mensajes lingüísticos, no puede abordar el estudio de la atribución intencional que acompañan a la mayor parte de las emisiones de la conversación coloquial (forma prototípica de comunicación) que se manifiesta, muchas veces, en la entonación que emplean los usuarios de una lengua.

Por otro lado, y para una caracterización más amplia de la entonación, hay que destacar varias ideas entre las que destaca el carácter intencional de la actividad lingüística: se habla para algo. Siguiendo la tradición de los filósofos del lenguaje inspirados por la obra de Austin (1962), a las emisiones que realizan intenciones comunicativas se las denomina *actos de habla* (conductas verbales orientadas a un fin), cosa que implica la considerar la entonación como un instrumento lingüístico (fonológico) al servicio de un emisor inteligente que busca la consecución de un objetivo comunicativo, que no es lingüístico. Una segunda idea que haría falta estudiar para alcanzar una definición legítima de entonación es que la actividad lingüística es una actividad cooperativa. Los actos de habla sólo pueden realizarse eficazmente en ciertas condiciones de cooperación. De hecho, se ha demostrado empíricamente que los hablantes adaptan automáticamente ciertos parámetros acústicos en función de las necesidades de información que atribuyen a sus interlocutores (Clark y Wilkes-Gibbs, 1986). Por último, habrá que tener en cuenta una idea que ha cobrado mucha fuerza en las recientes teorías de la comunicación, a saber: que para que la actuación lingüística alcance el significado completo hay

que considerar el contexto porque éste ejerce influencia tanto en el contenido proposicional como en la forma lingüística (y entonativa) de las emisiones.

En la medida en que se puedan explicar y comprobar cómo interfieren estas ideas en el uso de la entonación se podrá ir elaborando una teoría de la entonación capaz de dar cuenta del papel que desempeña en la comunicación humana.

2. Objetivos del trabajo

Como se ha dicho, cada comunidad de hablantes de un idioma desarrolla algún tipo de identidad lingüística diferencial, un perfil melódico cuyo juego entre los acentos principales del enunciado determina cómo será su primer pico tonal, cómo declinará la melodía a lo largo del enunciado, y sobre todo, cómo se establecerá la última inflexión tonal del mismo.

Para hacer un estudio empírico de la entonación hay que contar con un corpus de habla espontánea que permita obtener los modelos entonativos reales y genuinos que, a su vez, haga posible la descripción de la realidad lingüística. Con la elaboración del corpus CENE se pretende una mejor caracterización de la noción de *variedad*, a partir del conocimiento del papel que desempeña la entonación en los dialectos. Así, el objetivo de esta investigación es ofrecer una descripción rigurosa de la entonación que presentan algunas variedades dialectales del español.

Ha sido necesario acotar el ámbito de estudio y hemos decidido centrarnos en el análisis de las variedades asturiana, castellano leonesa, navarra, madrileña y vasca: cinco comunidades autónomas de la mitad norte de la Península de las que se busca investigar las características entonativas.

3. Características de un corpus eficaz

Por corpus hay que entender una “colección de datos lingüísticos, ya sea de textos escritos o de transcripciones de habla grabada, los que pueden ser utilizados como punto de partida para descripciones lingüísticas o como un medio de verificación de hipótesis acerca de una lengua.” Crystal (1991, p. 32).

En nuestro ámbito, se trata de un conjunto de realizaciones sonoras que serán objeto de estudio para el entonólogo. Se trata de una herramienta muy valiosa para la indagación de fenómenos propios del registro idiomático coloquial, entre los cuales interesa especialmente el elemento entonativo.

Una de las mayores dificultades que presenta su elaboración es la de asegurar la calidad de las grabaciones, requisito imprescindible para proceder al análisis empírico de las realizaciones. Esto explica que la tendencia general de los estudiosos de la entonación sea el recurso a corpus creados *ad hoc* a partir de lo que suele llamarse *habla de laboratorio*. Se trata de obtener los datos a partir de enunciados aislados leídos o inducidos por el investigador, en un contexto estático y sin ruidos que asegure la nitidez de las grabaciones y que agiliza mucho el trabajo.

Este modo de proceder, sin duda, proporciona grabaciones técnicamente correctas y una cantidad de datos muy amplia en un espacio de tiempo muy

reducido; pero, a cambio, se trabaja con enunciados que minimizan la riqueza melódica de la conversación natural. Contornos alterados, por tanto, de cuyo análisis se obtendrán los resultados (que posiblemente también distorsionen la realidad). Nosotros entendemos que es la realidad la que norma el acceso metodológico y no al revés. Y, aunque somos conscientes de las dificultades que surgen al intentar describir el habla espontánea, no creemos que cejar en el empeño de asentar las bases para describir el habla vaya a proporcionar unos resultados deseables.

A continuación se explica el modo en que se ha trabajado para asegurar que CENE presente una serie de características que garanticen la fiabilidad de los resultados del estudio.

1- *Habla espontánea*. Si lo que se pretende es lograr la descripción de la entonación del habla espontánea de los diferentes dialectos del español, parece lógico que la base de la investigación sea un conjunto de emisiones de voz, reales y genuinas, de cada una de las zonas que se pretende estudiar. Hay que tener en cuenta que el marco natural de la entonación es el diálogo, por lo que, los enunciados entonativos deben extraerse de cada uno de los turnos que lo componen. Otro procedimiento, como el recurso a frases preparadas, resulta artificioso y proporciona resultados discutibles.

En la elaboración del corpus CENE se han desestimado las emisiones de los locutores profesionales porque se trata de personas que, además de dominar la técnica que requiere hablar detrás del micrófono y para las cámaras, no están hablando con finalidades interpersonales.

2- *Cantidad suficiente y equilibrada de contornos*. Se pretende analizar la entonación de muchas de las variedades del español, lo cual ya exige que el corpus sea extenso; sin embargo, cuando hablamos de “cantidad suficiente” nos referimos, no sólo a la suma de los datos totales del corpus, sino a la cantidad de contornos entonativos de cada variedad. Puede decirse que, en realidad, se han constituido 5 corpus independientes, uno por variedad estudiada: Asturias, Castilla León, Madrid, Navarra, y País Vasco. Pues bien, cada uno de los corpus se compone de 200 enunciados que suman un total de 1000 emisiones de voz en total.

3- *Buena calidad acústica*. Las dificultades que tradicionalmente se ha encontrado el dialectólogo a la hora de conseguir buenas grabaciones desde el punto de vista acústico, han quedado, en buena parte, superadas porque existe la posibilidad de acudir a la televisión. Este medio se ha convertido en una la fuente de datos idónea para el estudio de la entonación porque han proliferado en él los espacios dedicados a tertulias, debates y concursos. Se trata de una tipología de programas que interesa porque en ellos intervienen muchas personas que no son profesionales y que se expresan espontáneamente. Los participantes de estos programas no levantan la sospecha de tener sus intervenciones planificadas como sí ocurre, por ejemplo, en los *reality shows*. Las tertulias, los debates y los concursos

tienen otra ventaja: el perfil de los participantes es muy variado, lo cual facilita la labor de conseguir que el corpus sea compensado.

- *El corpus de Asturias* consta de 200 enunciados seleccionados de 13 horas de material audiovisual de tres programas que emite la Televisión del Principado de Asturias (TPA), a saber: *De tarde, Asturias en 25 y Conexión Asturias*. Se trata de programas tipo tertulia, reportaje y concurso respectivamente que han permitido, como interesaba, recoger los enunciados en un contexto de diálogo.
- *Los 200 enunciados que conforman el corpus de Navarra* proceden de tres programas que emite el Canal 6: *Cara a cara, Objetivo Navarra y Navarra Pregunta*. El primero es un programa de debate; el segundo, un espacio dedicado de actualidad donde se analizan las noticias más relevantes con tertulias y reportajes; el tercero es un programa sobre política en que los ciudadanos interrogan a los representantes de las agrupaciones ideológicas principales y merece un comentario más detallado.
- *El corpus del País Vasco* consta de 200 enunciados seleccionados de 27 horas de material audiovisual procedentes del programa *Objetivo Euskadi*.
- Para la confección del *corpus de Madrid* se han empleado 16 horas de grabación de los programas *Madrid en Directo* y *Pongamos que hablo de Madrid* que emite la televisión autonómica. Se trata de programas de análisis-reportaje y tertulia, respectivamente.
- El corpus de Castilla-León, por último, se ha elaborado a partir de la extracción de 200 enunciados del programa *Tal como somos* emitido en la televisión autonómica de Castilla y León. Se trata de un programa que reúne varios géneros audiovisuales: informativos, tertulias, concursos y debates.

4- *Número amplio y representativo de informantes*. Asimismo, interesaba que el número de informantes del corpus fuera amplio (nunca se han recogido más de diez emisiones por informante), y equilibrado en cuanto al sexo.

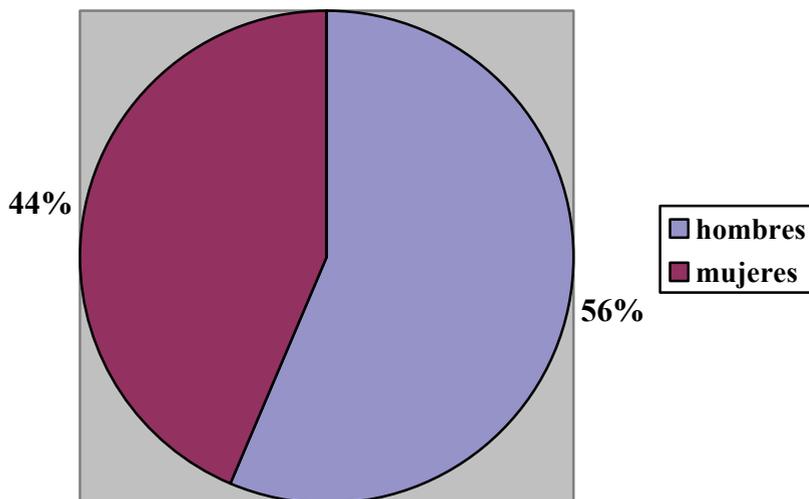
- Para el corpus de Asturias se han seleccionado 67 informantes en total donde el 56% son hombres y el 44% mujeres. Este desequilibrio entre los informantes de un sexo y otro se ha compensado seleccionando una cantidad mayor de enunciados emitidos por mujeres (sin superar nunca los diez enunciados por informante) de manera que del número total de enunciados un 59 % son producciones masculinas y el 41% restante, producciones femeninas.
- Para el corpus de Navarra se han seleccionado 63 informantes en total, 39 hombres y 24 mujeres. Se equilibró el corpus en cuanto al número de emisiones masculinas y femeninas y, finalmente, se ofrecen 99 enunciados producidos por hombres y 101, por mujeres

- En el corpus de Euskadi la desproporción entre la participación de hombres y mujeres volvía a ser considerable: 60% hombres frente a un 40% de mujeres. Por ello, se tuvo que exprimir al máximo las intervenciones de las mujeres con el objetivo de que el número de emisiones de cada uno de los sexos quedara más equilibrado. Así, 110 de los enunciados seleccionados son producciones de hombres y los 90 restantes, de mujeres.

-En el caso de Madrid ha sido posible conseguir un corpus bien compensado en lo que al sexo de los informantes se refiere. El número total de informantes es de 51 (26 hombres y 25 mujeres). El corpus consta, como el resto, de 200 enunciados, donde el 44% son emisiones producidas por mujeres y el 56 %, por hombres.

-Para la elaboración del corpus de enunciados de Castilla se ha contado con 36 hombres y 20 mujeres (56 en total) de los que se han seleccionado 124 y 86 enunciados, respectivamente.

Por tanto, se trata de un corpus equilibrado pues de los 1000 enunciados que lo conforman, 564 son producciones masculinas y 437, femeninas.



En resumen, se ha procedido recogiendo de forma masiva programas emitidos por las televisiones autonómicas de las diferentes zonas dialectales. El *Corpus del español del norte de España* está constituido por más de 50 horas de grabación y puede dividirse en cinco corpus independientes, cada uno de los cuales se compone de 200 enunciados emitidos por una media de 55 informantes, que suman un total 1000 emisiones de voz de 325 informantes en total. Posiblemente se trate del corpus el más amplio que se haya utilizado en investigaciones de este tipo. Pero la calidad más importante del corpus no tiene que ver tanto con sus dimensiones como con

que se trata de enunciados de habla espontánea con una calidad acústica muy notable, porque se han extraído de programas televisivos equipados con la tecnología más idónea para asegurarla.

Una muestra de este tipo es, seguramente, de interés para los estudiosos de la morfología, sintaxis, significación, realizaciones discursivas, pragmáticas y dialectológicas (porque la muestra no es la investigación; pero sí es necesaria para llevarla a cabo). Es muy de desear que el CENE sea explotado por especialistas de diferentes ámbitos lingüísticos y que éstos compartan los resultados de sus trabajos, pues el estudio del lenguaje no debe abordarse desde una sola perspectiva: esto daría lugar a teorías lingüísticas unilaterales. Un enfoque multidisciplinar, en cambio, permite conocer su complejo funcionamiento. Este es el espíritu con que la lingüística moderna ha de avanzar y, desde luego, la filosofía básica con que se ha elaborado el CENE: la entonación no es un sistema cerrado, una realidad material sin más, sino un lugar de confluencia: un ámbito; y el lenguaje, una red de ámbitos entrelazados que se condicionan, se relacionan, y que forman un todo superior a la simple yuxtaposición de las partes.

3.1. Procedimiento de elaboración del corpus

Con la finalidad de adaptar las grabaciones al formato requerido por el programa de análisis acústico PRAAT² convertimos los ficheros de vídeo con un programa gratuito y disponible en Internet, Format Factory 2.20 a un mismo formato destino, Windows Media Video (WMV). Tras la unificación de los formatos, se procedió al visionado de las grabaciones a fin de seleccionar sólo los archivos que interesaban.

3.2. El criterio de segmentación de las unidades de análisis

La unidad sobre la que se decide abordar la observación evidencia los objetivos de quien elabora el corpus. Nuestra atención se ha centrado en el estudio de los aspectos de índole melódica (dimensión formal de la entonación) cosa que queda reflejada, también, en el primer nivel de representación, como se explica en siguiente apartado.

Como ya se ha dicho, el objetivo de nuestro estudio consiste en describir el modo en que los hablantes de cada variedad desarrollan un perfil melódico característico cuyo juego entre los acentos principales del enunciado determina cómo será su primer pico tonal, cómo declinará la melodía a lo largo del enunciado, y cómo se establecerá la última inflexión tonal del mismo. Para ello (sobre todo si se trata de analizar producciones de habla espontánea) hay que segmentar el habla en unidades melódicas. Esto, a su vez, exige partir de un criterio formal: la presencia de una inflexión tonal que delimita el grupo fónico. Una vez

² <http://www.fon.hum.uva.nl/praat/>.

identificadas, las unidades melódicas (los grupos fónicos) han de recibir un tratamiento autónomo contengan o no unidades gramaticales enteras.

El Análisis Melódico del Habla se basa en un modelo teórico expuesto en Cantero (2002) que entiende que la cadena fónica no es sólo la sucesión de segmentos tímbricos, sino que estos se interrelacionan formando agrupaciones fónicas más complejas: unidades bien trabadas y jerarquizadas en tres niveles.

La primera jerarquización se da entre los elementos vocálicos y consonánticos y se establece por motivos articulatorios³. Las consonantes se agrupan en torno a las vocales, que son los *núcleos silábicos* y contienen información melódica relevante.

La segunda diferenciación jerárquica se da entre los propios segmentos vocálicos, que pueden ser tónicos o átonos. Los primeros tienen una función lexemática (núcleos acentuales) en torno a los cuales se agrupan los segmentos átonos de las palabras. Pero en la producción del habla los elementos conectores se agrupan entorno a una vocal tónica (*acento paradigmático*) que se diferencia del resto por su representatividad léxica, semántica o melódica. Estos elementos, agrupados entorno al acento paradigmático, contribuyen a la formación de una nueva agrupación jerárquicamente superior: las palabras fónicas. En la emisión oral, sólo tienen acento las palabras tónicas. Los artículos, verbos auxiliares, pronombres personales, preposiciones y conjunciones (las palabras átonas) se organizan alrededor de una vocal acentuada, que recibe el nombre de *acento paradigmático*, formando un bloque sonoro que se denomina *palabra fónica o grupo rítmico*. Por tanto, las palabras átonas, por su falta de independencia fónica no pueden considerarse palabras desde un punto de vista fónico aunque tengan independencia ortográfica⁴. Los conceptos de palabra fónica y palabra léxica no son, pues, equiparables: una palabra léxica es una unidad de análisis gramatical (*casa*), mientras que una palabra fónica es una unidad de análisis fonético y exclusivamente sonoro (*casa, lacasa, enlacasa, nuestracasa*).

La tercera jerarquización se establece porque los núcleos paradigmáticos se agrupan en torno a un núcleo o acento sintagmático constituyendo la unidad superior de la jerarquía fónica: *el grupo fónico*, que coincide con lo que se denomina *contorno entonativo* que se define como la sucesión de tonos organizados en torno a una inflexión tonal. El grupo fónico suele contener una estructura lexicogramatical coherente. No obstante, se trata unidad fónica por lo que esto no

³ Los segmentos consonánticos, frente a los vocálicos, presentan modos de obstrucción del aire según el punto de articulación por lo que, en sentido estricto, no pueden ser considerados sonidos, sino ruidos (v. Cantero 2002)

⁴ Navarro Tomás, en 1944, ya se dio cuenta de que “las diferencias de estructura melódica entre palabras agudas, llanas y esdrújulas, desaparecen en la línea musical de la oración” (1944:22)

ocurre siempre. Para definirlo hay que localizar la vocal tónica en la que tiene lugar la inflexión tonal que es, generalmente, la marca de final de grupo fónico.

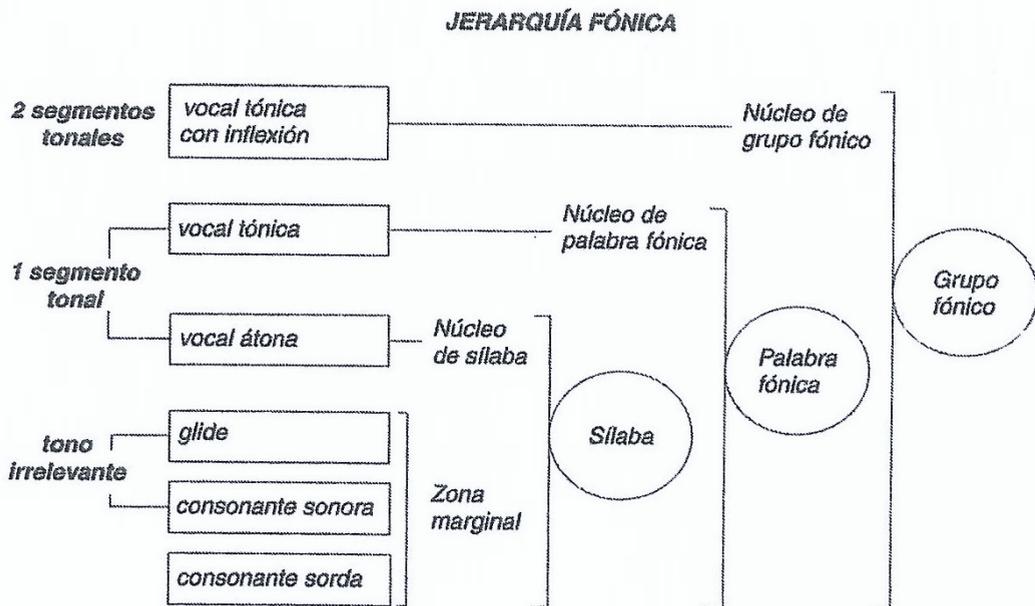


Figura 1 La jerarquía fónica (Cantero, 2002)

La figura 1 intenta representar gráficamente que el *grupo fónico* y el *grupo rítmico* (o *palabra fónica*) son unidades agrupadas en torno al acento sintagmático y paradigmático, respectivamente.

Desde esta perspectiva, el escrutinio analítico del habla espontánea se hace posible, pues ofrece un criterio de segmentación de las melodías del habla exclusivamente fónico, independiente de cualquier otro nivel de análisis.

Para analizar secuencias, lo primero que hay que hacer, partiendo de esta perspectiva, es determinar el valor absoluto (en Hz) de la frecuencia fundamental de cada segmento tonal (de cada vocal⁵), a continuación se procede a la estandarización, que consiste en la conversión de los valores obtenidos en un valores relativos expresados en porcentajes.

A través de un instrumento fiable de análisis acústico (como Praat), se identifican las vocales (mediante la observación del sonograma) y se anota su valor

⁵ Si se considera que, “el tono controlado conscientemente por el hablante ocurrirá en la vocal, que es el centro de la fonación” (Cantero, 2002: 147), el número de segmentos tonales del contorno equivale al número de sílabas (vocales) del grupo fónico

central (o bien la media de los valores de F0 de la vocal). Los valores que se obtengan constituirán, solamente, los datos en bruto de la melodía del contorno, luego, será necesario que se procesen, en un segundo paso de esta fase, porque la melodía del contorno es una sucesión de intervalos (no de valores absolutos).

Si se trata de una vocal tónica, hay que determinar el valor de los dos segmentos tonales (o tres, en el caso de la inflexión sea circunfleja) que constituyen la inflexión. Si la inflexión final acaba en consonante nasal o lateral, ésta puede constituir el último segmento tonal de la inflexión.

Para convertir el *continuum* de los valores de F0 en una sucesión de valores tonales relativos (la melodía) es necesario estandarizar el contorno. La estandarización es un procedimiento que permite comparar voces de altura tonal diferente. Habitualmente, la unidad de medida de estas descripciones es el semitono (escuela holandesa). Sin embargo, emplear una descripción por porcentajes⁶, como hace este modelo, tiene tres ventajas fundamentales: en primer lugar, resulta más intuitiva porque permite expresar linealmente un fenómeno algorítmico; además, es más sencilla de calcular y, por último, aleja de una perspectiva musical del habla.

Una vez relativizados los valores se expresan en positivo los ascensos y en negativo los descenso y se yuxtaponen; el algoritmo obtenido es la expresión de la melodía.

Enunciado	Es	muy	di	fí	cil
Hercios	177	230	176	279	160
Porcentaje	100%	29.9%	-23,5	58,5%	-42,6%

Figura 2 El algoritmo 100%+29%-23,5%+58%-42,6% es la expresión de la melodía “es muy difícil”.

Los valores estandarizados de la curva permiten hacer una representación gráfica de la curva que puede utilizarse para hacer clasificar los enunciados y para hacer comparaciones formales.

⁶ Hasta el año 2010, dicho el porcentaje de variación de cada valor absoluto respecto del anterior se calculaba mediante una regla de tres; pero los pasos posteriores a la identificación de las vocales han sido automatizados por Miguel Mateo (2010) de manera que, la captura de datos ha ganado en precisión y los gráficos de las curvas que se generan a continuación son más fiables.

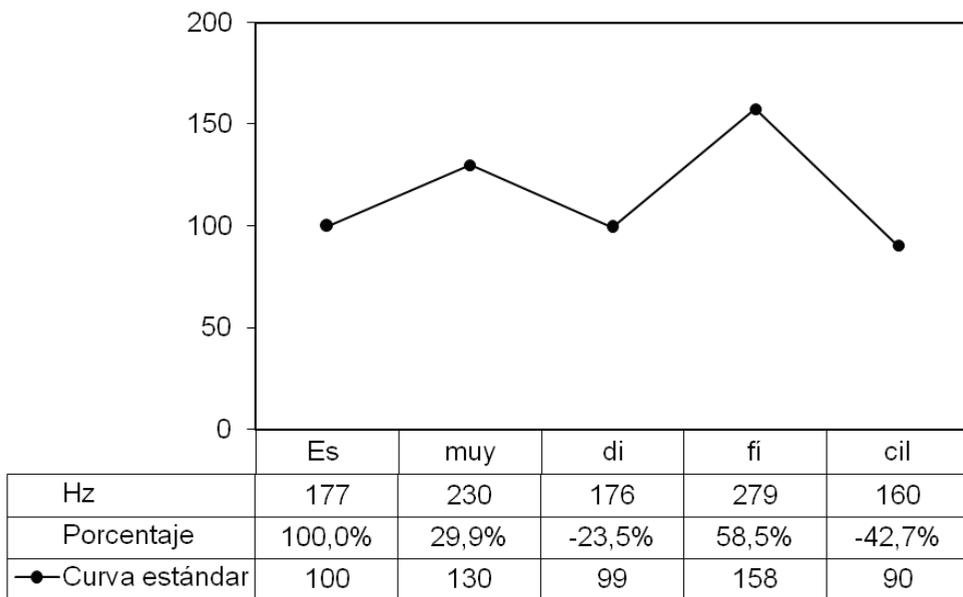


Figura 3 Representación gráfica de la melodía

En resumen, desde la perspectiva que empleamos, la unidad de análisis de la entonación es el grupo fónico (cuya melodía constituye un contorno) y la unidad de análisis del ritmo, la palabra fónica (cuya melodía también es relevante en el contorno). El análisis de ambos fenómenos (ritmo y entonación) se lleva a cabo mediante el análisis de la melodía: la sucesión de valores frecuenciales relevantes. La unidad de medición de la melodía, por su parte, es el *segmento tonal*: el valor tonal del núcleo silábico, a saber, la vocal.

3.3. La transcripción del corpus

Decíamos que la transliteración de un corpus oral revela los propósitos del investigador pues no hay forma de poner en el papel todos los matices, variaciones, énfasis e inflexiones del discurso oral (tratar de hacerlo, por otra parte, resultaría en un texto tan lleno de marcas, guías y señales que resultaría ilegible y, por eso, inútil).

El primer nivel de representación de un corpus (sea oral o no) es el de la representación ortográfica, su transliteración. En nuestro corpus, y siguiendo las recomendaciones básicas para la transcripción ortográfica de corpus orales que ofrece el informe del Grupo de Trabajo sobre Textos Orales de EAGLES- Expert

Advisory Group Engineering Standards⁷ (EAGLES 1996), se han utilizado, siempre que ha sido posible las formas ortográficas que aparecen en los diccionarios normativos. No obstante, se ha preferido representar algunas formas dialectales tal como son pronunciadas por el hablante: “Eso ye muy importante” (eso es muy importante).

También se ha tratado de representar de la forma más transparente posible las cuestiones relativas a la entonación aunque ello, en ocasiones, haya obligado a puntuar en contra de las prescripciones de la gramática normativa (como por ejemplo, separando mediante coma el sujeto y el predicado).

A continuación se ofrece una lista de diez enunciados extraídos de cada uno de los corpus:

As-5-4-2	Ten en cuenta que cuando menos lo esperabas, te meten un gol
As-5-4-3	En la segunda ya se vieron otras hechuras
As-5-5-1	Eso sería la parte inicial
As-5-5-2	Y la cuerda que puede ser de de diferente componente
As-5-7-2	Es un pixín pequeño, entonces con cinco, siete minutos ya estarían hechos
As-5-7-3	Los lomos de pixín más grandes los meteríamos pues unos quince minutos
As-5-8-2	Ese ye el problema
As-5-8-3	Exactamente
As-5-9-1	Aquí hacíamos lo que es el día del bollo
As-5-13-1	Conocer los rincones tan guapos que tiene Asturias

Enunciados corpus Asturias

⁷ El grupo EAGLES realiza recomendaciones o propuestas de estandarización con el fin de coordinar los trabajos que se realizan en las diferentes lenguas de Europa. Para ello, evalúa métodos y sistemas existentes y a partir de estos análisis realiza sus propuestas. Busca más bien entregar directrices que ayuden en el desarrollo de uno. Se ha propuesto, por ejemplo, tres criterios orientadores: a) flexibilidad, b) apertura teórica y c) búsqueda de consensos

PV-3-1-1	¿Qué te pongo más, cariño?
PV-3-2-1	Me puse yo con ella
PV-3-2-2	Yo calculo que un treinta por cierto...
PV-3-2-3	Y no piensa que igual el pescado le vale cinco,
PV-3-15-3	Estaba en la frontera de la manía y ya la, la obsesión, ¿no?
PV-3-15-4	Pero no, no lo era
PV-3-15-5	No, no, en absoluto
PV-3-15-6	Pero no deja de ser una manía
PV-3-16-3	Yo creo que es una manía que mucha gente tiene también
PV-3-17-7	Mira, oye, la gente tenemos una fecha de caducidad

Enunciados del corpus del País Vasco

CL-2-1-4	Porque el fin es ya no solamente conseguir el dinero...
CL-2-1-7	Que van a venir con los profesores
CL-2-1-9	Me imagino que saldrá mañana en la prensa
CL-2-2-1	Este dato está mal, es en Salamanca
CL-2-3-2	Eh, tendencias y moda hay en todos los sectores
CL-2-3-3	Pues de todo un poco ¿no?
CL-2-12-1	¡Somos seres humanos!
CL-2-12-2	¡Digo yo!
CL-2-14-1	¿Y por qué no se la dan?
CL-2-16-1	Tiene un problema el gobierno ahí

Enunciados del corpus de Castilla-León

N-4-1-1	¿Y qué cuentan?
N-4-1-2	No me lo creo

N-4-3-2	Podría ser una buena idea
N-4-3-4	Nos dimos cuenta de que había una persona que lo estaba haciendo en Australia
N-4-3-5:	El cliente decide si se queda la gallina o si la devuelve
N-4-4-2:	Pues nada, por probar a ver un mes a ver lo qué...
N-4-5-1	La verdad que mucha suerte de poder hacer cosas así
N-4-6-1	¿En realidad?
N-4-7-1	¿Tú sabes lo que es coger y meter la ropa y sacarla ya para tender?
N-4-8-1	Coño, porque es un disfrute

Enunciados corpus Navarra

M-2-2-3	Pues hará catorce años o por ahí que vive aquí
M-2-2-4	Es muy normal en el barrio
M-2-2-5	Que tú no le das mayor importancia
M-2-7-1	Yo estaba allí, tengo allí una finca
M-2-9-3	Y, sobre todo lo que sería el menudeo, el trapicheo de drogas
M-4-1-6	Lo de <i>sólo</i> , entre comillas
M-4-1-7	Y bueno, la verdad es que no paramos durante todos esos veinte días
M-4-1-9	Fueron utilizadas por los regímenes comunistas como lugares de deportación
M-4-2-2	La verdad que cada vez echamos más horas
M-4-3-1	Cada año se vuelve a repetir la jugada

Enunciados Corpus Madrid

3.4. El etiquetado del corpus

El etiquetado por el que se ha optado es sencillo e intuitivo. Ha consistido en asignar un código a cada informante y a cada enunciado y *en elaborar* unas tablas que contuvieran la información de cada uno de ellos. Dichos códigos aparecen encabezados por las iniciales de la comunidad a la que pertenece el informante. A

continuación se han asignado tres series de números secuenciales: el primer número identifica el número de fichero/grabación de esa variedad; el segundo, señala a un informante concreto de ese programa y, el tercero se refiere a los enunciados que ha producido. Así, los enunciados del corpus CENE tienen un formato como el siguiente:

AS-03-02-10

Que permite identificar que se trata del enunciado 10, emitido por el informante 02 del programa 03 de los seleccionados de la comunidad de Asturias.

Asimismo, se han elaborado unas tablas que resumen los datos relevantes del corpus de grabaciones y unas fichas que contienen los datos de cada informante y de cada enunciado. Las primeras recogen el nombre del programa, días y franjas de emisión, número de grabaciones recogidas, duración de cada programa y horas totales de grabación.

PV-1-15	Hombre	PV-1-15-1: Y luego, abajo, productos de mantenimiento de lentillas PV-1-15-2: Si que es verdad que podría venir uno y ponerse unas lentillas de una miopía de menos dos PV-1-15-3: ¡No hace falta ver que no se puede beber! PV-1-15-4: Sí, es evidente PV-1-15-5: Pero eso al final es como todo... PV-1-15-6: Antes la gente, había gente mayor que no la usaba nunca
---------	--------	--

Las segundas ofrecen información sobre el sexo de los participantes y los enunciados que se han seleccionado de cada informante, a saber: código identificativo, transcripción ortográfica del enunciado, contexto discursivo en que aparece, clasificación (interrogativo, enfático, suspendido), localización en el fichero original de la grabación (hh: mm: ss, ms)

Código	Transcripción	Contexto discursivo	Clasificación provisional	Localización
PV-1-1-2	¡A comer!	Explica las ventajas de aprender a cocinar ese plato	Declarativa	PV-1 Archivo 2 00:01:30

Código	Trascripción	Contexto discursivo	Clasificación provisional	Localización
PV-1-2-1	¿Ya estáis preparados?	Hace una pregunta	Interrogativa	PV-1 Archivo 2 00:02:20
PV-1-2-2	¿Y eso?	Hace una pregunta extrañada	Interrogativa	PV-1 Archivo 2 00:02:30
PV-1-2-3	¿Lo imprescindible qué es?	Hace una pregunta	Interrogativa	PV-1 Archivo 2 00:06:50
PV-1-3-1	Ya estamos preparados	Responde a al pregunta	Declarativa	PV-1 Archivo 2 00:02:20

En cuanto a la identidad de los informantes, su formación profesional, su procedencia sociocultural y su edad⁸ hay que decir que sabemos poco. Se tuvieron que hacer muchas estimaciones porque en su mayoría se trata de personas anónimas. Esas apreciaciones, que quizás sean bastante exactas porque al hacerlas se contaba con la imagen de la televisión, no nos parecieron del todo fiables para ser tomadas en consideración. Sin embargo, esto no ha de considerarse un inconveniente a pesar de la aparente falta de datos porque una de las fases del análisis consiste, precisamente, en eliminar el tipo de alteraciones micromelódicas que estas cuestiones pueden llevar consigo a través de la estandarización. Aquí puede ser útil recordar que se ha buscado intencionalmente el anonimato de los informantes porque parece lógico que, al tratarse de personas que son ignorantes de su condición de informantes en una investigación sobre la entonación, se minimice la posibilidad de que ésta se vea alterada por las circunstancias.

4. Resumen

En este artículo se ha descrito el corpus oral CENE: una fuente de datos empíricos que puede contribuir a que las teorías que construyen las ciencias del lenguaje sean más robustas, certeras dado que brindan sustento a la indagación de la lengua en uso

⁸ Al hacer un análisis estimativo del corpus de informantes se observa que contamos con 90% de personas de edad adulta (entre 25 y 65 años) y un 10 % de niños y ancianos.

(tal como es producida por hablantes reales en situaciones concretas y particulares). Se trata de un cuerpo de datos elaborado para la indagación melódica del habla con 1000 enunciados de habla espontánea transcritos y etiquetados de manera sencilla que facilita su acceso y uso a especialistas de diferentes ámbitos. Aunque se ha creado para estudiar la variación fonética de la melodía debería ser útil para los investigadores del discurso oral espontáneo. En el futuro se pretende ampliar la notación del corpus a fin de proporcionar datos como la variación fonológica de las emisiones y su clasificación gramatical y pragmática.

Obras citadas

- AUSTIN (1962). *How to do things with words*, Oxford, The Clarendon Press (Trad. cast. *Cómo hacer cosas con palabras*, Barcelona, Paidós, 1990)
- BALLESTEROS PANIZO, M. (2011). *La entonación del español del norte*. Tesis doctoral inédita. Dep. Filología hispánica. Universitat de Barcelona.
- BOLINGER, D.L. (1989). *D.L Intonation and its uses*, Stanford, Stanford University Press
- BRIZ, A. (1998). *El español coloquial en la conversación*, Barcelona, Ariel
- CANTERO SERENA, F.J. (1995). *Estructura de los modelos entonativos: interpretación fonológica del acento y la entonación en castellano*, tesis doctoral publicada en microforma (1997). Universitat de Barcelona.
- CANTERO SERENA, F. J. (2002). *Teoría y análisis de la entonación*, Barcelona, Edicions de la Universitat de Barcelona.
- CLARK Y WILKES-GIBBS (1986). “Referring as a collaborative process”, *Cognition* (22), 1-39
- CORTÉS MORENO, M. (1999). *Adquisición de la entonación española por parte de hablantes nativos de chino* (tesis doctoral). Universitat de Barcelona.
- CRYSTAL, D. (1991). *The Cambridge Encyclopedia of Language*, Cambridge, Cambridge. University Press.
- EAGLES (1996). *Synopsis and comparison of morphosyntactic phenomena encoded in lexicons and corpora. A common proposal and applications to european languages*, Pisa, ILC-CNR.
- ESCANDELL VIDAL, M. V. (1999). “Los enunciados interrogativos: aspectos semánticos y pragmáticos”, en I. Bosque y V. Demonte (Dir.). 3929-3992.
- ESCANDELL VIDAL, M. V. (2005). *La comunicación*, Madrid, Gredos.
- HIDALGO, A. (1997). *La entonación coloquial. Función demarcativa y unidades de habla*, Valencia, Universitat de València, anejo 21 de la revista *Quaderns de Filologia*.
- LADD, D.R. (1980). *The structure of intonational meaning*, Bloomington, Indiana University Press.
- LIU, Y-H (2005). *La entonación del español hablado por taiwaneses*, Biblioteca Phonica, 2. <http://www.ub.es/lfa>

- LLISTERRI, J. (1999) “Transcripción, etiquetado y codificación de corpus orales”.
En GÓMEZ GUINOVART, J LORENZO SUÁREZ, A., GÓMEZ GUERRA, J. ÁLVAREZ
LUGRIS, A. (Eds), Panorama de la investigación en lingüística informática,
RELSA, Revista Española de Lingüística Aplicada, volumen monográfico,
pp.59-82
- NAVARRO TOMÁS, Tomás (1944). *Manual de entonación española*, New York,
Hispanic Society (1944) Madrid, Guadarrama.
- POYATOS, F. (1994). *La Comunicación No verbal I: Cultura, Lenguaje y
Comunicación*, Madrid, Istmo
- SOSA, J.M., (1999). *La entonación del español. Su estructura fónica, variabilidad y
dialectología*, Madrid, Cátedra.