
Lengua y Habla

Revista del Centro de Investigación y Atención Lingüística C.I.A.L.

ISSN 1316-1180
Nº 8 2003-2004



Índice

| | |
|---|-----|
| Presentación editorial | 3 |
| Contribuciones | 5 |
| Breve descripción de la duración silábica como señal de frontera sintáctica en tres velocidades de elocución <i>María Alejandra Blondet, Jorge Méndez, Elsa Mora</i> | 9 |
| Estudio fisiológico de procesos neurolingüísticos mediante Potenciales Evocados: Una revisión con actualización de tecnología y análisis de datos <i>Carmen Elena Contreras, Michelle Delappe, Avilio Moreno, Emily Chávez, Pedro Coutin, Lourdes Pietrosevoli</i> | 17 |
| Historias de vida: entre la temporalidad y el aspecto <i>Lourdes A. Díaz Blanca</i> | 27 |
| Música y lenguaje <i>Drina Hocevar</i> | 50 |
| Estudio fonético-sintáctico de algunos marcadores discursivos: propuesta de formalización para desambiguarlos automática o informáticamente <i>Hernán Martínez, Lino Urdaneta, Carmen Luisa Domínguez</i> | 59 |
| Características del discurso oral pedagógico <i>Teresa Molina</i> | 80 |
| La necesidad de planificar una norma lingüística en lenguas de señas para usos académicos <i>Lionel Antonio Tovar</i> | 97 |
| Normas | 135 |

Lengua y Habla

Revista del Centro de Investigación y Atención Lingüística C.I.A.L.

ISSN 1316-1180

Nº 8 2003-2004

Índice

| | |
|--|------------|
| Presentación editorial | 3 |
| Contribuciones | 5 |
| Breve descripción de la duración silábica como señal de frontera sintáctica en tres velocidades de elocución <i>María Alejandra Blondet, Jorge Méndez, Elsa Mora</i> | 9 |
| Estudio fisiológico de procesos neurolingüísticos mediante Potenciales Evocados: Una revisión con actualización de tecnología y análisis de datos <i>Carmen Elena Contreras, Michelle Delappe, Avilio Moreno, Emily Chávez, Pedro Coutin, Lourdes Pietrosevoli</i> | 17 |
| Historias de vida: entre la temporalidad y el aspecto <i>Lourdes A. Díaz Blanca</i> | 27 |
| Música y lenguaje <i>Drina Hocevar</i> | 50 |
| Estudio fonético-sintáctico de algunos marcadores discursivos: propuesta de formalización para desambiguarlos automática o informáticamente <i>Hernán Martínez, Lino Urdaneta, Carmen Luisa Domínguez</i> | 59 |
| Características del discurso oral pedagógico <i>Teresa Molina</i> | 80 |
| La necesidad de planificar una norma lingüística en lenguas de señas para usos académicos <i>Lionel Antonio Tovar</i> | 97 |
| Normas | 135 |

Estudio fonético-sintáctico de algunos marcadores discursivos: propuesta de formalización para desambiguarlos automática o informáticamente.

Hernán Martínez

Lino Urdaneta

Carmen Luisa Domínguez

Universidad de Los Andes

Resumen

El estudio que se presenta a continuación se propone postular una serie de presupuestos que sirvan como base a una posterior formalización morfofonológica y sintáctica de los marcadores discursivos, de manera que éstos puedan ser reconocidos por sistemas informáticos y diferenciarlos de las formas homófonas correspondientes (cuando las haya). Para ello, se estudiaron fonética y sintácticamente cuatro marcadores: bueno, ¿no?, ahora y este. Los resultados del estudio parecen mostrar que mecanismos sintácticos y fonéticos pueden entrecruzarse y formalizarse para llegar a conformar una herramienta de tratamiento automático de textos de cierta eficacia para desambiguar parcialmente los marcadores discursivos de sus respectivos homófonos.

Abstract

The following study postulates a series of considerations to serve as a basis for a posterior morpho-phonological and syntactic formalization of discourse markers, so that they can be recognized by data processing systems and be differentiated from eventual homophone corresponding forms. For this purpose, we studied four Spanish markers: *bueno*, *¿no?*, *ahora* and *este*. The results seem to show that syntactic and phonetic mechanisms can be intercepted and formalized in order to constitute an automatic text treatment tool, capable, to a certain extent, to render – rather effectively– partially discourse markers unambiguous from homophone forms.

Introducción

En este estudio se propone un principio de formalización de los marcadores discursivos basada en criterios fonéticos y sintácticos que pueda ser utilizada para desambiguarlos automáticamente de sus respectivas formas homófonas. El modelo propuesto plantea que los marcadores discursivos pueden identificarse durante el reconocimiento del habla por medio de criterios morfofonológicos, tanto por parte de los participantes de la comunicación como por sistemas informáticos de tratamiento lingüístico automático. El modelo planteado refuerza también la tesis de que los mecanismos sintácticos y fonéticos se entrecruzan para contribuir a la identificación de estas partículas (Heeman, 1999).

Uno de los principales motivos que nos han guiado a realizar esta investigación es la falta absoluta de una formalización sintáctico-fonética que le permita a un sistema informático distinguir, de una manera fiable, a un marcador discursivo de su correspondiente homófono no-marcador en el caso del español de Venezuela. Así pues, esta investigación pretende ir encaminada hacia una propuesta de formalización que posteriormente permita utilizar autómatas finitos deterministas y algoritmos de intersección entre ellos para desambiguar automáticamente dichos marcadores.

La formalización informática de los marcadores discursivos permitirá, entre otras cosas, procesar de manera más adecuada textos escritos y orales; aplicaciones como traductores automáticos, sistemas de diálogo, correctores automáticos y sistemas de búsqueda podrían mejorar su desempeño si pudiesen distinguir los marcadores discursivos de sus respectivos homófonos, ya que podrían valerse de dicha información para

procesarlos o, simplemente, filtrarlos de manera que no entorpezcan la búsqueda o la traducción.

En esta investigación se han estudiado los marcadores discursivos *bueno*, *¿no?*, *ahora* y *este*. Se han escogido éstos y no otros porque cada uno presenta homófonos de una clase funcional distinta: *bueno* (adjetivo), *¿no?* (adverbio de negación), *ahora* (adverbio de tiempo) y *este* (pronombre demostrativo, adjetivo demostrativo y sustantivo). Esto permitió comprobar si la formalización podía distinguir los marcadores de homófonos con función y distribución distinta.

Marco Teórico

Como se señaló arriba, esta investigación pretende ir encaminada hacia una propuesta de formalización sintáctico-fonética que permita distinguir los marcadores discursivos que comparten una misma forma morfológica con otra(s) palabra(s). Específicamente, se estudiaron los marcadores *bueno*, *ahora*, *este* y *¿no?*.

La entonación y los marcadores discursivos

La oralidad presenta mecanismos prosódicos que la caracterizan y la hacen única dentro de las manifestaciones del lenguaje; son esos mecanismos de naturaleza suprasegmental los que organizan el hilo fónico. Sin ellos sería muy difícil hacer de lo oral algo inteligible; uno de los elementos prosódicos básicos que organizan el habla es pues la *entonación*.

Es justamente a través de la entonación que el hablante organiza su discurso y transmite ciertas informaciones que no están en los elementos gramaticales que conforman el enunciado (Obediente, 1998:204). La entonación es un fenómeno que sólo puede observarse en la oralidad y mediante el cual adquieren una expresión concreta tanto el aspecto comunicativo de lo hablado como su significado y orden sintáctico. Es decir, que la entonación nos sirve para estructurar el sentido de la frase, convirtiéndose de hecho en uno de los medios **más seguros** para lograr una comunicación más eficaz. La entonación se percibe sujeta, ligada, unida a la estructura léxico-sintáctica y a las particularidades estilísticas de la expresión hablada. De ahí que digamos que la entonación contribuye a organizar el discurso oral, ya que ayuda a distinguir segmentos y palabras dentro del discurso oral.

En este trabajo se pretende observar si existen parámetros entonativos que nos permitan distinguir un marcador discursivo de su homófono en la lengua oral. Creemos que los marcadores discursivos presentan ciertas peculiaridades entonativas que nos permiten diferenciarlos de partículas morfofonológicamente idénticas y que los hacen distintos. Una de esas características particulares es que los marcadores discursivos “(...) se aglutinan frecuentemente alrededor de pausas (reales o virtuales), como las plaquetas en una herida detienen el flujo de la sangre, para contribuir en la delimitación de las unidades de información (...)” (Álvarez, 2000). La entonación, el ritmo, las pausas y el alargamiento al final de la palabra sirven para delimitar las unidades de información. Por ejemplo, y adelantándonos a los resultados, el marcador *bueno* se agrupa alrededor de las pausas que separan dichas unidades. Los marcadores pueden llegar incluso a funcionar por sí mismos como pausas, construyendo así la conclusión de un periodo explicativo.

A nivel del enunciado, los marcadores, cuya configuración fonética creemos particular, cumplen ciertas funciones sintácticas.

Sintaxis y formalización automática de los marcadores

El primer paso es intentar establecer qué funciones sintácticas podrían atribuírsele al marcador o intentar ubicar su lugar dentro del proceso sintáctico discursivo. Como se podrá observar en los antecedentes, la totalidad de los autores revisados coinciden en señalar que un marcador discursivo está fuera de la predicación oracional; asimismo, la mayoría coincide en que se trata de partículas sintácticas cuya principal tarea es la de guiar la descodificación que el receptor hace del discurso sin formar parte de su contenido proposicional. Si bien, en términos generales, estamos de acuerdo con ambas proposiciones, nos parece necesario matizarlas y someterlas a examen. Aunque parece cierto que los marcadores discursivos no forman parte de la predicación oracional, también nos resulta absolutamente necesario aclarar que suponemos que sí forman parte de un sistema predicativo aunque éste no sea de carácter oracional. También creemos necesario comentar que, si la función principal de un marcador es la de guiar al interlocutor, no por ello dichos marcadores dejan de formar parte del texto (e incluso de su *dictum*). Ambas aclaratorias se presentan como absolutamente pertinentes porque ambas proposiciones parecen partir del supuesto de que el marcador entra en una relación por completo ajena al proceso sintáctico que relaciona al resto de las unidades de la línea. Es

verdad que un marcador discursivo no admite su clasificación en ninguna de las tradicionales partes del discurso, pero ello no quiere decir que forme sólo parte marginal de éste; creemos que es posible dar interpretaciones paralelas que den cuenta de que, en efecto, es posible utilizar un mismo procedimiento sintáctico para relacionar adjetivos, nombres, verbos, locuciones, preposiciones, adverbios, conjunciones, pronombres, oraciones, sintagmas, y **marcadores discursivos**, teniendo en cuenta que el paradigma al que pertenecen estos últimos es bastante complejo.

Para llevar a cabo esta tarea, nos basamos en modelos dependenciales y/o funcionales y no en modelos constitutivos (que, al parecer, han contribuido a ocultar la relación del marcador con el proceso textual). Prácticamente, cualquier modelo dependencial (Tesnière, Harris, Hjelmslev, entre otros) podría dar más o menos cuenta de la situación sintáctica del marcador. Así pues, si se sostiene que ciertos elementos pueden o no depender de otros (que también pueden depender de otros, y así sucesivamente) de manera diversa (dicha dependencia puede darse de manera recíproca o unilateral) y que las funciones dependenciales que contraen entre sí explicarían en su totalidad las relaciones entre el texto y cada una de sus partes (en todos los niveles) entonces, necesariamente, las partículas que nos ocupan deberían también entrar en funciones dependenciales con el texto y sus partes.

Es por ello que en la presente investigación nos hemos valido del modelo dependencial de la sintaxis léxica (Subirats, 2001) –basada en la **teoría de la predicación** que desarrolló Zellig Harris (Harris, 1991)– para llevar a cabo la primera parte de nuestra tarea. Este modelo tiene como premisa fundamental el que las relaciones predicativas (entre predicado y argumentos) son redundantes (sus posibilidades combinatorias se encuentran restringidas) y son las encargadas de transmitir información (del tipo que Halliday denominaría *ideativo*) en las lenguas naturales. Según este principio, las relaciones predicativas son de tipo dependencial; hay elementos que requieren la aparición de otros para poder formar parte de cierta estructura; otros elementos no requieren nada; algunos elementos tienen una dependencia recíproca; por último, las unidades mayores en las que puede aparecer un predicado están de antemano determinadas por éste, es decir, cada predicado, además de depender de cierto número de elementos, es concebido como un conjunto de propiedades sintácticas que permiten o restringen el tipo de proyección oracional de la cual pueden formar parte. Verbigracia, un adjetivo que requiera la presencia de un nombre es visto como un predicado, y lo mismo ocurre con adverbios,

preposiciones, etc. Una preposición *en* requiere de dos argumentos para poder funcionar ((Francisco comió) *en* (la casa)). Aquí, la preposición *en* requiere necesariamente de los argumentos *Francisco comió* y *la casa* para poder entrar en relación textual; *en* no se puede presentar si faltan ambos o uno de los argumentos: **Francisco comió en* o ?*en la casa* (*esta última*, aunque gramatical, es inaceptable si se escucha aislada, sin que se pueda recuperar el primer argumento).

Este criterio que se ha esbozado, de manera bastante sucinta, funciona para dar cuenta del proceso de formación oracional; en este estudio nos servimos de él para insertar los marcadores discursivos en el mismo proceso textual; en otras palabras, para explicar ciertas características de los marcadores discursivos se utilizaron las mismas herramientas analíticas con las cuales se explican ciertas características del resto de elementos léxicos que concurren en un enunciado. Este criterio no pretende imponerse como “verdad” lingüística sino como herramienta operativa que permitirá utilizar autómatas finitos deterministas como modelos de formalización para el procesamiento informático.

Del nuevo enfoque que este tipo de teoría le da a estas partículas se desprenden las siguientes consecuencias:

- i. Todos los marcadores pueden ser tratados como predicados: todo marcador para poder funcionar **requiere** de la presencia de ciertos elementos.
- ii. Los marcadores, por el hecho de estar fuera de la predicación oracional, son predicados de segundo nivel (P_p): todo marcador se puede insertar “después” de haber sido insertado el último predicado que restringe el tipo de estructura oracional en la que puede aparecer, motivo por el cual cualquier partícula de este tipo podría llegar a tener como argumento otro predicado léxico (es por ello que se le considera el predicado de un predicado).
- iii. De manera general, los marcadores pueden pertenecer a dos tipos de predicados: a los P_p (predicados de un argumento) o a P_{pp} (predicados de dos argumentos).

De estas condiciones se deduce que un marcador discursivo como *¿no?* o *ahí* requiere necesariamente de la presencia de un (y sólo uno) elemento para poder pasar a formar parte de un texto: ((Ella fue a mi

casa) *¿no?*) o ((Dame un cafecito) *ahí*). Un concatenador, conector o reformulador requiere, a su vez, de dos argumentos (siempre dos, aunque puede darse el caso de que se den en catálisis): ((el cambio de clima) *bueno...* (una parte forestación)). Esta clasificación, aunque simple, nos permite equiparar estas unidades discursivas con el resto de los elementos léxicos que conforman la enunciación.

El requerimiento argumental, aunado a las restricciones que las formas homófonas a los marcadores imponen sobre el tipo de construcción oracional en la que pueden aparecer (construcciones que este modelo relaciona de manera transformacional) permite, en principio, distinguir sintácticamente los marcadores de dichas formas morfofonológicamente idénticas. Estos principios nos permitieron, tanto relacionar los marcadores con resto de las unidades textuales, como distinguirlos de sus homófonos siguiendo un procedimiento mecánico.

Otra de las directrices que configuraron el modelo de nuestra investigación apuntó hacia una formalización computacional de los marcadores discursivos. Los principios teóricos que acabamos de exponer nos permitirán incluir los resultados de esta investigación dentro de un proyecto de tratamiento automático de textos. Los resultados obtenidos pueden ser formalizados, y parcialmente lo han sido, en forma de autómatas finitos deterministas que reconozcan marcadores discursivos; esto permite etiquetar un texto, en forma de autómata también, y desambiguar los marcadores parcialmente interceptando el texto etiquetado con transductores finitos deterministas que formalicen las restricciones sintáctico-formales de dichas partículas y de los respectivos homófonos. Esto contribuye a identificar las relaciones predicativas que un marcador sostiene con distintos niveles textuales. Como se verá más adelante, esta propuesta permite, tanto incluir a los marcadores dentro de este tipo de procesamiento electrónico, como distinguirlos sintácticamente y fonéticamente de sus homófonos.

De todo lo expuesto hasta aquí resulta obvio el sostener que hay un vínculo estrecho entre la fonética y la sintaxis: en los marcadores resalta una colaboración entre la entonación y la función sintáctica que permite distinguir sus funciones de las de sus formas homófonas (cuando las hay). Estos presupuestos se pondrán a prueba en el análisis que sigue más adelante.

Antecedentes

Los marcadores discursivos se han convertido en objeto de no pocos estudios en diversas lenguas. Dichos estudios, y sobre todo en lengua española, se han acrecentado de una manera impresionante en los últimos años, centrándose sobre todo en el nivel de análisis sintáctico, algunos de los cuales han arrojado como resultado taxonomías que han servido de base a otros trabajos descriptivos.

Quizá el estudio más representativo para el español es el de María Martín Zorraquino y José Portolés, quienes desde una perspectiva pragmática sostienen que los marcadores del discurso

“...son unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional -son, pues, elementos marginales- y poseen un cometido coincidente en el discurso: el de guiar, de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas, las inferencias que se realizan en la comunicación” (Martín Zorraquino y Portolés, 1999:4057)

Dichos autores, en el mismo texto, restringen unas líneas más adelante el concepto diciendo que “el significado de los marcadores contribuye al procesamiento de lo que se comunica y no a la representación de la realidad comunicada” (idem). Martín Zorraquino y Portolés conciben, entonces, estas partículas como orientadores del discurso.

A diferencia de los estudios sintácticos sobre marcadores discursivos, los prosódicos parecen representar el terreno menos explorado. Martín Zorraquino y Montolío han dicho que “otro ámbito de estudio pendiente de una investigación más profunda es el que afecta a los rasgos suprasegmentales (como la entonación, la cantidad silábica, la posible autonomía tonal de los marcadores –su delimitación por pausas–, etc.)” (Martín Zorraquino y Montolío Durán 1998:13-14).

Dentro del ámbito prosódico pues, y más precisamente desde un punto de vista similar al nuestro, Peter Heeman *et al* en 1999, en un trabajo titulado *Identifying discourse markers in spoken dialog*, presenta un modelo para el reconocimiento o identificación de los marcadores del discurso en el habla basado en el *POS-Based Language Model*. Dicho modelo plantea que los marcadores discursivos pueden identificarse durante el reconocimiento del habla, tanto por parte de los oyentes como de sistemas informáticos; además, sostienen que los marcadores se pueden identificar en el continuum con la implementación del modelo *POS tagging* (Parts

Of Speech [etiquetado de partes del discurso]): las partes del discurso (lemas) tienen una cierta configuración sintáctica y unas características fonéticas que le “indican” al oyente cuál es su función dentro de la línea textual. El oyente reconoce los marcadores en el discurso por: 1) la posición sintáctica que ocupa la partícula en la oración discursiva; y 2) por ciertas características prosódicas o fonéticas en la señal acústica (por lo general los marcadores presentan una entonación particular y están flanqueados por silencios o pausas). Por otra parte, en este trabajo se sugiere que los marcadores pueden ser usados para ayudar al oyente a saber cuál es el papel que juega en el diálogo: se enfoca entonces la relación interlocutiva.

Por su parte, y dentro del ámbito de la lengua española, Pedro Martín Butragueño, en un trabajo titulado *Hacia una descripción prosódica de los marcadores discursivos. Datos del español de México*, y en base a la observación de 214 datos procedentes de la oralidad y de tres estilos diferentes (I, conversación grabada; II, cuestionario leído; III, texto leído), y clasificados según la taxonomía propuesta por Martín Zorraquino y Portolés en 1999, llega a proponer que los marcadores discursivos muestran “un aire de familia prosódica. Con todo, son tan maleables, en especial algunos de ellos, que la línea melódica sigue con relativa docilidad las sutiles diferencias de sentido a que se adscriben” (Martín Butragueño, 2000). También, propone el autor al final de su artículo que cuanto más informal sea el estilo de habla menos específica es la marca prosódica que caracteriza al marcador discursivo.

Hipótesis: Los marcadores discursivos presentan rasgos fonéticos y sintácticos que confluyen para contribuir a su distinción con respecto a sus formas homófonas.

Objetivos: Determinar cómo funcionan los marcadores del discurso sintáctica y fonéticamente para diferenciarse de sus respectivos homófonos; con esto se pretende establecer una sistematización y formalización de los marcadores discursivos que pueda llegar a utilizarse por desambiguadores automáticos.

Metodología

El Corpus

Este estudio se ha basado en una grabación de la variedad del español hablado en la ciudad de Mérida, Venezuela, transcrita y publicada en *El Habla de Mérida* por Domínguez y Mora en 1998. Se escogió al hablante con el número 23, cuyo código es MDD5MA (índice 7.62). Se trata de un hombre de 66 años, chofer, que no culminó sus estudios de primaria. Vive en vivienda propia, de padre obrero y madre doméstica.

En este corpus se buscaron las apariciones de los marcadores discursivos *bueno*, *¿no?*, *ahora* y *este*.

Para el análisis acústico-sintáctico

Para el análisis acústico se utilizó el programa *Speech Analyzer* de *SIL International* en su versión 1.5. Con este programa se segmentó el habla y posteriormente se extrajo la curva entonativa de los marcadores y sus homófonos. Se midió el ascenso o descenso de la curva melódica y la duración de las pausas para determinar las características acústicas particulares de cada una de dichas partículas.

La descripción sintáctica se realizó para permitir una posterior formalización en autómatas que pueden ser procesados en aplicaciones como INTEX o UNITEX.

Análisis

El análisis se llevó a cabo de la siguiente manera:

- i. Utilizando conjuntamente la grabación del hablante 23 y su transcripción, se identificaron los marcadores que se decidió estudiar: *bueno*, *ahora*, *¿no?* y *este*.
- ii. Cada uno de dichos marcadores fue estudiado y clasificado según la taxonomía de Martín Zorraquino y Portolés (1999).

Análisis acústico

Como señalábamos anteriormente, para el análisis acústico utilizamos el programa *Speech Analyzer* para determinar la configuración acústica de los marcadores discursivos. En el programa revisamos el ascenso o descenso¹ y la duración en el tiempo de la onda entonativa. De dicho análisis se desprende lo siguiente para los cuatro marcadores estudiados:

Bueno

La mayoría de los *bueno* que funcionan como marcadores de inicio (67.7%) presentan una curva entonativa ascendente (↓) y en promedio el ascenso es de 13.86 Hz, el resto (33.3%) presenta curva entonativa descendente (↑). Los que funcionan como reformuladores, de modalidad deóntica y epistémica (58.3%), por el contrario, presentan una curva que en promedio desciende 5.86 Hz.; también encontramos algunos (41.7%) de este tipo de marcadores cuya curva asciende, todos éstos acompañados por lo general de un alargamiento vocálico de la primera sílaba. Los *bueno* marcadores de inicio están, generalmente, flanqueados por pausas largas.

Ahora

La curva entonativa en estos marcadores que están ubicados dentro de la clasificación de marcadores como conectores contraargumentativos presentan en el 100% de los casos de nuestro corpus un ascenso promedio bastante brusco de la curva entonativa: 31.4 Hz.

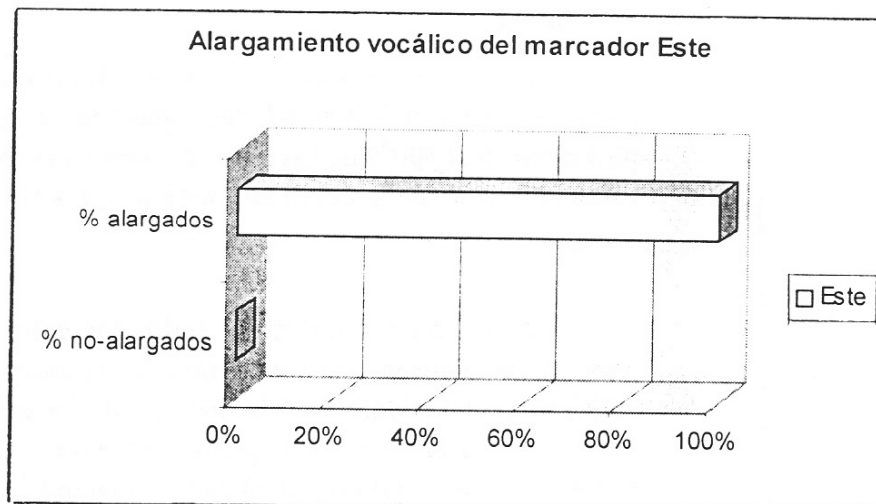
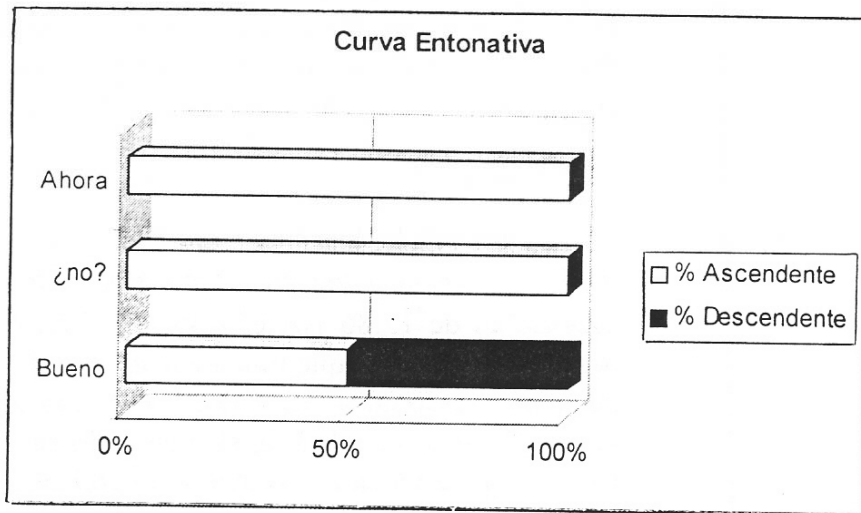
¿No?

Este marcador muestra en todos los casos de nuestro corpus una curva ascendente al final; en promedio el ascenso es de 32.96 Hz. En contraste, el *no* adverbio en la mayoría de los casos (70%) presenta una curva descendente que en promedio baja un 4.5 Hz. Sin embargo, encontramos algunos *no* (30%) que al igual que los *¿no?* presentan curva ↑, pero el aumento es insignificante en comparación con los *¿no?*: 5.2 Hz.

Este

Este marcador, en promedio, presenta un descenso en la curva del fundamental de 9.8 Hz. en el 100% de los casos. Dicho descenso siempre va acompañado por un alargamiento vocálico de la última sílaba el cual, en promedio, dura alrededor de 0.46 seg, duración que hace que el marcador

se diferencia de los homófonos sustantivo, adjetivo demostrativo y pronombre.



Análisis sintáctico

En principio, cada marcador y cada una de sus funciones se clasificó sintácticamente según su requerimiento argumental en P_p o en P_{pp} . Una

vez clasificado cada marcador según su requerimiento argumental, se compararon las propiedades de éste con las propiedades que presentaban las formas morfológicamente análogas que presentaban el mismo requerimiento argumental (sus homófonos). El llevar a cabo esto permitió distinguir sintácticamente los marcadores de sus respectivas formas homófonas. A continuación se detallarán las propiedades sintácticas de cada uno de los marcadores estudiados.

Bueno

Encontramos, básicamente, en el discurso del hablante 23 cinco funciones de *bueno*: 1) Inicio de turno; 2) reformulador; 3) conector; 4) de modalidad epistémica; 5) de modalidad deóntica. Los marcadores *bueno* con función conectiva o reformuladora fueron clasificados como P_{pp} . El resto de las funciones fueron clasificadas como P_p . Esta primera partición obliga a generalizar dos tipos de marcador *bueno* según lo que acabamos de exponer. El *bueno* con función reformuladora o conectiva se distingue del adjetivo correspondiente en esta primera partición: mientras el segundo es un P_p , el primero es un P_{pp} . Ya esto es suficiente para que una gramática electrónica pueda distinguirlos sintácticamente (siempre y cuando los dos argumentos de los P_{pp} estén presentes).

Como el *bueno* con las demás funciones (inicio de turno, epistémico y deóntico) es un P_p al igual que su homófono adjetival, hay que distinguirlo por sus propiedades sintácticas (las restricciones que cada uno impone) o por su relación distribucional (superficial, lo que le precede o le sigue). Una matriz de propiedades sintácticas que define al adjetivo predicativo puede ser representada como sigue:

| | |
|---|---|
| Adjetivo Predicativo <i>bueno</i> | |
| Proyección oracional: P_p Arg. ser <i>bueno</i> | |
| Relativización | + |
| Reducción de Oración de Relativo | + |
| Nominalización | + |
| Reducción de Verbo de Soporte | + |
| Concordancia | + |
| Verbalización | - |
| Selección de Subjuntivo | + |
| Adverbialización (-mente) | + |
| Oración especificativa | + |
| Oración explicativa | + |

Tabla 1. Propiedades sintácticas del adjetivo predicativo bueno

Cada símbolo + especifica que el adjetivo predicativo *bueno* acepta dicha propiedad o transformación, razón por la cual puede encontrarse en cada una de dichas construcciones sintácticas (Subirats: 2001). Así, la Tabla 1 muestra que el adjetivo *bueno* puede formar parte de una oración relativa explicativa o especificativa; por lo tanto, dicho adjetivo puede ser relativizado y esta construcción puede aceptar una **Reducción de Oración de Relativo** como la siguiente (Subirats: 2001):

El niño; el niño es bueno

[Rel] → *El niño, que es bueno*

[RedORel] → *El niño bueno*

También acepta una nominalización con el artículo lo: lo bueno de ir al fútbol. El adjetivo bueno tiene que concordar con su argumento, se puede adverbializar, puede seleccionar el subjuntivo en sus completivas con *ser* y acepta una reducción del verbo de soporte (*ser*).

Así pues, el adjetivo predicativo presenta una serie de propiedades sintácticas y morfológicas que lo caracterizan. En la Tabla 2 se presentará la misma matriz de propiedades sintácticas, pero aplicada al marcador *bueno* P_p.

| | |
|---|---|
| Marcador <i>bueno</i> (epistémico, deóntico, inicio de turno) | |
| Proyección oracional: P _p Arg. <i>bueno</i> (deóntico) <i>bueno</i> Arg. (inicio de turno y epistémico) | |
| Relativización | - |
| Reducción de Oración de Relativo | - |
| Nominalización | - |
| Reducción de Verbo de Soporte | - |
| Concordancia | - |
| Verbalización | - |
| Selección de Subjuntivo | - |
| Adverbialización (<i>-mente</i>) | - |
| Oración especificativa | - |
| Oración explicativa | - |

Tabla 2. Propiedades sintácticas del marcador (epistémico, deóntico, inicio de turno) *bueno*

Como puede observarse, las propiedades sintácticas y morfológicas que el adjetivo *bueno* presentaba no pueden ser aceptadas o regidas por el marcador; el marcador *bueno* no puede adverbializarse, ni nominalizarse, ni construirse como atributo concordado, ni relativizarse, ni presentar un verbo

de soporte, etc. Esto establece una base de comparación que permite distinguir formalmente a ambos predicados P_p en la línea textual. Así, una oración como *bueno es una casa* no cumple con la propiedad de concordancia, razón por la cual un transductor descartaría la descripción estructural del adjetivo predicativo y le asignaría a *bueno* la etiqueta de **marcador**. Siguiendo estos parámetros distinguimos sintácticamente 12 de los 18 marcadores *bueno* que encontramos en el discurso del hablante 23.

Ahora

El mismo método empleado para diferenciar al marcador *bueno* sirve para hacerlo con el marcador conversacional *ahora*. Este, al igual que su homófono adverbial, es un P_p que puede tener un predicado como argumento. La matriz de propiedades de este adverbio predicativo se comparó con la matriz de este marcador conversacional, lo que permitió distinguir los casos ambiguos de los no ambiguos en el texto del hablante 23; así, de 5 apariciones de este marcador, la única ambigua fue: *Ahora, venga acá...este...todo eso le pasa a uno ¿no?* Del contexto se desprende que éste no es un adverbio de tiempo sino que funciona como un marcador; pero un autómata o transductor que formalice al P_p *ahora* como adverbio que requiere un verbo como argumento directo no podría diferenciar que este *ahora* es un marcador y no un adverbio porque su cercanía en la línea con el verbo es típica de los adverbios. En cambio, en un enunciado como *Ahora...después...vino...la conclusión de que las cosas pues...fueron aumentando...sí* puede ser parcialmente desambiguado porque el contexto sintáctico de *ahora* (a principio de oración, seguido de pausa y, luego seguido de otro adverbio temporal) suele ser el de un marcador y no el de un adverbio.

¿No?

La partícula negativa *no* puede tener (según la sintaxis léxica) dos orígenes sintácticos distintos. Puede ser un argumento (N) es decir, un elemento léxico que está dominado por un predicado y que no necesita nada para formar una oración. Por otro lado, cuando es un adverbio de negación, se trata, según Harris, de la huella de la aplicación del predicado metadiscursivo *negación*. Como se puede observar, esto lo distingue del marcador *¿no?* el cual es un P_p con una proyección oracional *Arg. ¿no?* De 46 apariciones, 19 (41,3%) de ellas presentan una ambigüedad estructural entre la huella de la reducción del metapredicado y la del

marcador discursivo (la desambiguación es dramáticamente alta si el autómata se programa para reconocer los signos de interrogación ¿? que suelen delimitar a este marcador en la escritura).

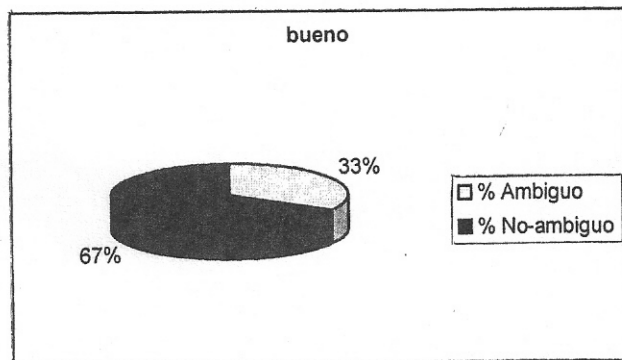
Este

Con respecto a este marcador ocurre algo bastante particular: no hemos podido incluirlo dentro de una estructura predicativa, es decir, no es ni un predicado ni un argumento de algún predicado; desde el punto de vista metodológico utilizado en esta investigación, este hecho sugiere que quizá no pueda ser tratado (a fines formales) como un marcador discursivo sino tan sólo como una pausa llena (son necesarios más estudios para poder llegar a una conclusión). Así, para desambiguarlo sintácticamente sólo podemos contar con criterios distributivos lineales (como p. ej. la adyacencia de un verbo). Además, su distribución parece incluso mucho más libre que la del resto de los marcadores: puede aparecer incluso a mitad de una “rama” de un árbol dependencial. El contexto lineal puede hacer que dicha pausa llena (o marcador) se confunda con el sustantivo (punto cardinal), el pronombre o con el adjetivo predicativo. Si bien los criterios dependenciales no funcionan con este marcador, las propiedades sintácticas y morfológicas que el adjetivo predicativo posee (concordancia, reducción de oración de relativo, reducción del verbo de soporte, etc.) lo distinguen del marcador; en cuanto al pronombre o al nombre, éstos se distinguen precisamente porque podrían llegar a presentar predicados que dependen de ellos, mientras que si se considera a *este* como pausa llena, no dependería sintácticamente de ningún otro elemento y, a su vez, ningún elemento dependería de su aparición. En todo caso, los criterios distributivos permitieron desambiguar 6 de 7 casos (85,7%). Ejemplo: *Y la edad es...este...nací en mil novecientos veinticuatro, a estas épocas...* En este ejemplo puede observarse que *este* no es un adjetivo demostrativo (no precede a un sustantivo) ni tampoco un pronombre (no hay concordancia con *edad*).

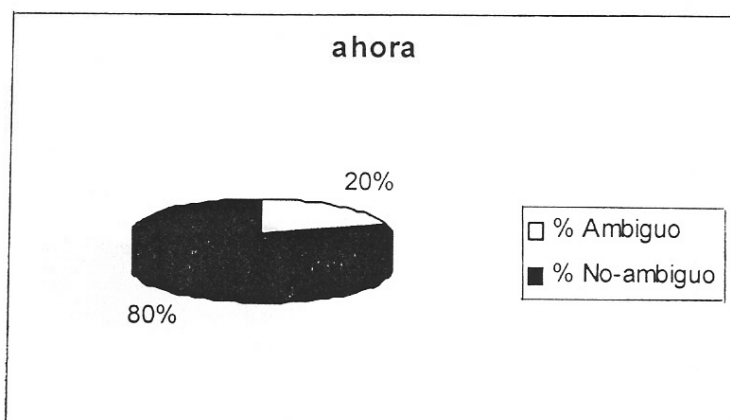
Resultados

Partiendo de todo lo expuesto *supra* se puede llegar a afirmar lo siguiente:

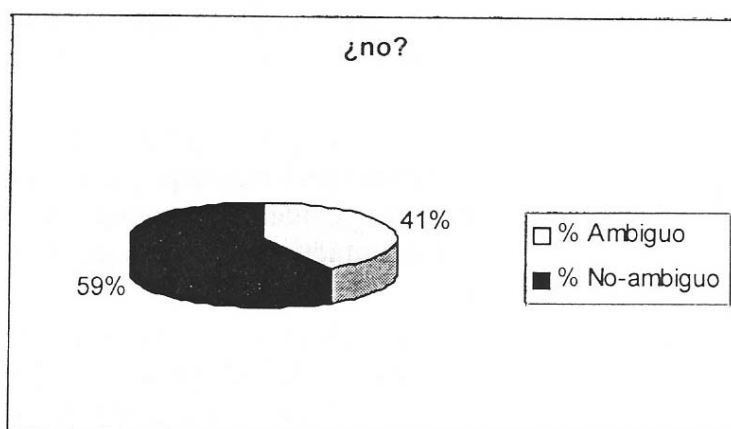
- **Bueno:** del total general, sintácticamente el 66,7% de este marcador se puede desambiguar: 50% del marcador *bueno* de inicio de turno y 75% de los que funcionan como reformulador, conector, deóntico y epistémico. Fonéticamente, los de inicio presentan en un 66,7% un ascenso de la curva; es este ascenso lo que nos permite desambiguar o distinguir algunos marcadores de sus homófonos. En los reformuladores, conectores, deónticos y epistémicos, en un 58,3% de los casos la curva es descendente. En total, el marcador *bueno* se puede desambiguar fonéticamente en el 50% de los casos; en conjunto, sintáctica y fonéticamente, sólo el 11% de dicho marcador persiste en su ambigüedad. De estos números se desprende que, en muchos de los casos, cuando este marcador es ambiguo sintácticamente, la entonación contribuye a establecer su función; por el contrario, en varios casos en los que el marcador no se confundía sintácticamente con sus homófonos, no se presentaba un ascenso en la curva entonativa: la sintaxis ayudando a la fonética y la fonética ayudando a la sintaxis.



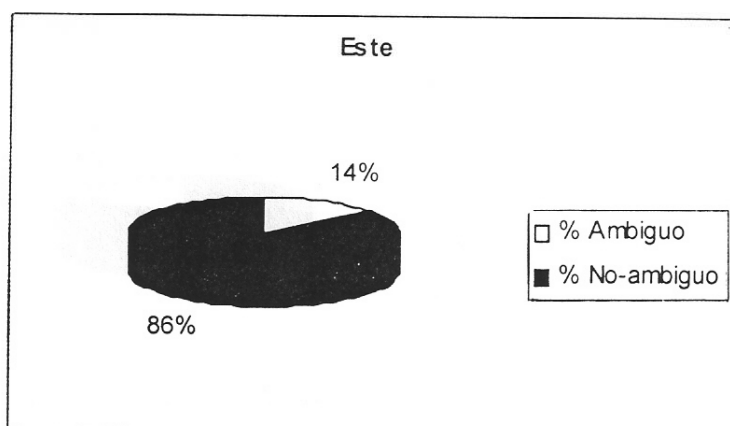
- *Ahora*: el 20% de los casos presenta ambigüedad a nivel sintáctico. Sin embargo, esta partícula puede llegar a distinguirse fonéticamente de su homónimo en un 100% de los casos: la curva entonativa que presenta el marcador es ascendente.



- *¿No?*: a nivel sintáctico, este marcador resulta imposible desambiguarlo en el 41,3% de los casos. Fonéticamente la ambigüedad de este marcador es del 0%.



- *Este*: de los casos estudiados, el 14,3% resultó ser sintácticamente ambiguo. Pero, como en los casos presentados arriba, la totalidad de las ocurrencias de este marcador se desambiguan fonéticamente, ya que el 100% de los casos presenta un alargamiento vocálico de la última sílaba y una curva entonativa descendente.



- De manera general puede sostenerse que la sintaxis y la fonética actúan conjunta y simultáneamente para disminuir en la medida de lo posible la ambigüedad estructural que puede encontrarse en la oralidad; esto ha sido visto, de forma palpable, en el caso de los marcadores discursivos que hemos estudiado en el hablante 23 del CSM. Estos resultados parecen apoyar la tesis de Heeman (1999): el oyente reconoce los marcadores discursivos por ciertas características fonéticas en la señal acústica y por la posición sintáctica que ocupan dentro del discurso.
- Creemos que la hipótesis de que los marcadores discursivos contraen funciones dependenciales (predicativas) tiene cierta base empírica y que esta concepción puede contribuir a esclarecer el papel sintáctico que éstos juegan en el texto.
- La inexistencia de relaciones predicativas que presenta *este* probablemente indique que sintácticamente –según el modelo que hemos aplicado– no funciona tanto como un marcador discursivo, sino más bien como una pausa llena o de hesitación (Blondet, Maria Alejandra, 1999), pero hacen falta estudios que, o bien corroboren esta hipótesis, o bien la rechacen.

Comentario final

En la actualidad estamos utilizando los datos sintácticos formales arrojados por esta investigación para escribir algoritmos de reconocimiento automático para el marcador discursivo *bueno*. En un corpus de prueba, (la transcripción electrónica de un hablante del *Corpus Sociolingüístico de Mérida*, con el código MDD5FB) dicha formalización ha demostrado un 100% de eficiencia para distinguir el marcador de su homófono adjetivo. Esto demuestra que la formalización de los marcadores discursivos que hemos propuesto puede ser utilizada de una forma eficaz por sistemas informáticos de tratamiento textual. Esta formalización la llevamos a cabo en forma de autómatas y transductores finitos deterministas escritos con un programa denominado INTEX. Hay que acotar, sin embargo, que para que pueda aprovecharse al máximo la formalización propuesta, es necesario que se utilicen programas que puedan relacionar voz y texto de manera simultánea, de forma que puedan también aprovecharse al máximo los aspectos acústicos de la formalización.

Referencias bibliográficas

- Álvarez Muro, Alexandra. 2000. *Poética del habla cotidiana*. Mérida, Venezuela: Universidad de los Andes.
- Blondet, María Alejandra. 1999. *Estudio acústico-prosódico de los fenómenos de hesitación: análisis contrastivo entre los dialectos andino y central*. Universidad de Los Andes: trabajo de Grado para la obtención del título Magister Scientiae en Lingüística.
- Harris, Zellig Sabettai. 1991. *A theory of language and information. A mathematical approach*. Oxford: Clarendon Press.
- Domínguez, Carmen Luisa y Mora, Elsa. *Corpus sociolingüístico de la ciudad de Mérida*. Mérida: ULA
- Heeman, Peter., et al. 1999. Identifying discourse markers in spoken dialog. Ponencia presentada en el AAAI Spring Symposium on Applying Learning and Discourse Processing. Stanford.
- Martín Butragueño, Pedro. 2000. *Hacia una descripción prosódica de los marcadores discursivos. Datos del español de México*. México. Disponible en: http://www.linguas.net/preseea/contenido/documentos_detalle.asp?cod=3&Doc=8
- Martín Zorraquino, María., Montolío Durán, E. (coords.). 1998. *Los marcadores del discurso. Teoría y análisis*. Madrid: Arco-Libros.

- Martín Zorraquino, María., Portolés, José. 1999. Los marcadores discursivos. En Ignacio Bosque y Violeta Demonte, coords. *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe.
- Obediente, Enrique. 1998. *Fonética y fonología*. Mérida: Universidad de Los Andes.
- Portolés, José. 1998. *Marcadores del discurso*. Barcelona, España: Ariel.
- Subirats, Carlos. 2001. *Introducción a la sintaxis léxica del español*. Madrid: Iberoamericana.

Notas

Cuando hablamos de ascenso o descenso nos referimos al movimiento hacia arriba o hacia abajo desde el punto de partida de la curva entonativa, el cual puede ubicarse en el espectro en varios puntos.