

## **Fraseo Prosódico y Sintaxis en un Corpus Radiofónico**

### **PROSODIC PHRASING AND SYNTAX IN A RADIO CORPUS**

**Alejandro Carlos Renato**  
Instituto Genius de Tecnología  
Manaus, Brasil  
arenato@genius.org.br

#### **Abstract**

Speech segmentation in prosodic phrases has relation in first place with syntactic phrase structure, but it cannot be derived from this in a straightforward manner. In the past, several studies tried to relate both domains through rule systems. However, data driven stochastic systems show a better performance.

A corpus of 400 sentences, spoken by professional speakers was segmented in prosodic phrases in pauses and tonal changes. On another hand, a maximum entropy chunker was trained for segmentation of corpus in syntactic phrases -chunks-.

Scope and problems were analyzed in three dimensions: prosodic phrases characteristics, structural hierarchy and symmetry between syntactic components. Results point out that a simple chunker can partially account for the prediction of metric phrasing task, but further linguistic information is required.

**Keywords:** Metrical Theory, Chunker, Phrasing, Spanish, Sintaxis and Phonology.

#### **Resumen**

La segmentación del habla en frases prosódicas se relaciona en primer lugar con la estructura de frase sintáctica, pero no se deriva de la misma en una forma directa.. En el pasado, varios estudios intentaron relacionar los dos dominios mediante sistemas de reglas. Sin embargo, los sistemas probabilísticos entrenados en datos mostraron un desempeño superior.

Un corpus de 400 oraciones de 4 locutores profesionales de radio fue segmentado en frases prosódicas de acuerdo a pausas y cambios tonales. Por otra parte, fue entrenado un chunker basado en entropía máxima para la segmentación del corpus en frases sintácticas -chunks-.

Los alcances y problemas encontrados se analizaron en tres dimensiones: características de las frases prosódicas, jerarquía estructural y simetría entre componentes. Los resultados indican que un simple chunker puede cumplir con parte de la tarea de predicción de frases métricas, pero que se requiere más información lingüística.

**Palabras claves:** Teoría métrica, Chunker, Fraseo, Español, Sintaxis y Fonología.

## 1. INTRODUCCION

La información prosódica ha sido considerada de gran importancia a los fines de la comprensión del mensaje hablado, ya que por medio de ella el oyente puede extraer la estructura lingüística y su significado semántico y pragmático. Las pausas y los acentos de frontera dividen la señal continua del habla en frases prosódicas. Diversos estudios han demostrado que dichas frases prosódicas no son isomórficas con respecto a la estructura sintáctica de la emisiones [1], [2], [3] y [4].

Numerosos factores que inciden en el fraseo prosódico han sido indicados en la literatura: el equilibrio entre las estructuras de actuación y las prosódicas [1], [2], las relaciones sintácticas funcionales, como sujeto, verbo, objeto, o sujeto y predicado, la dinámica de la información nueva y dada, el largo de frases, la velocidad de habla, entre otros.

Básicamente, se distinguen dos tipos de aproximación al tema dentro de las tecnologías del habla. Por un lado, se encuentran aquellos que tratan de derivar la información prosódica a partir de un análisis sintáctico parcial [1]. Por el otro lado, los métodos probabilísticos, como la predicción de límites de frase basada en la secuencia de clases -POS- de palabras [4], o los límites oracionales predichos por modelos enegramáticos [11], han dado mejores resultados que las predicciones basadas en información sintáctica en aplicaciones reales de reconocimiento automático o síntesis.

Puede especularse que la mayor precisión de los métodos probabilísticos se debe a que están basados en información empírica, que atiende a las características del dominio y porque tanto la secuencia de palabras o la secuencia de clases de palabras reúne mayor información local y de contexto inmediato, en tanto que la información sintáctica distribuye dicha información en la estructura, sin darle a cada elemento un peso determinado. La relación local que existe entre una palabra y otra, por ejemplo en los métodos basados en enigramas, no es sólo sintáctica, como también semántica y de uso.

Los sistemas de reconocimiento del habla que requieren interpretación semántica y los sistemas de síntesis de habla irrestrictos (que no están sujetos a un dominio) se verían beneficiados por sistemas que posean información lingüística detallada. Esta es una buena motivación para el desarrollo de un modelo que prediga el fraseo métrico a partir de la información sintáctica.

Analizadores sintácticos basados en redes neuronales han demostrado un buen desempeño en la predicción de frases métricas [7] y [8], mientras que las secuencias de clases de palabras ha demostrado ser un buen predictor de la acentuación [7] y [9].

Los trabajos basados en la fonología métrica utilizan una metodología basada en pequeños corpus, juicios de gramaticalidad y contraejemplos, que han sido valiosos para descubrir heurísticas y nuevos problemas para la teoría. Sin embargo, en el caso del fraseo métrico las diferencias entre los corpus reales y las predicciones de fonología de lápiz y papel son grandes. La confluencia de diferentes métodos puede ser fructífera en un tema como el presente, el cual no se agota en una sola perspectiva. Abney [14] ha indicado este aspecto, no obstante reconociendo cierta impermeabilidad de la lingüística teórica para aceptar conceptos provenientes de la investigación psicolingüística y empírica.

## 2. ANTECEDENTES TEÓRICOS

La relación entre prosodia y sintaxis abarca una basta literatura, dentro de la cual sólo se mencionarán los trabajos que parten del planteo del isomorfismo entre estructuras prosódicas y estructuras de frase sintácticas. Dicha tendencia surge del encuentro de trabajos de psicolingüistas y de la teoría prosódica no lineal, siendo precursores en este sentido Selkirk [2] y Nespor y Vogel [3], en la teoría prosódica no lineal, y los trabajos de Gee y Grosjean [1], Bachenko y Fitzpatrick [4]

y Abney [13], [14] y [15] con respecto al esfuerzo de encontrar una síntesis entre teoría prosódica y sintaxis.

En el campo de la psicolingüística, la motivación estuvo relacionada con la búsqueda de evidencia psicológica de la existencia de estructuras gramaticales, paralela al surgimiento de la teoría generativa. Pueden mencionarse tres tipos de tareas utilizadas en dicha investigación: a) memorización –recall-, juicios de relación semántica -relatedness judgments [10] y estudios sobre pausa –pausing-. En ellos se esperaba que se revelase la estructura sintáctica utilizada en el procesamiento del habla.

Un estudio influyente en este sentido fue el de Martin [16], que investigó las diferencias encontradas entre el análisis de grupos naturales realizado por hablantes y las predichas por la teoría lingüística. Hablantes no expertos analizaban oraciones reuniendo en frases aquellas palabras que formarían grupos naturales, los cuales fueron analizados por medio de la técnica de clustering jerárquico [17]. Los resultados [16] muestran que los participantes tienden a agrupar de acuerdo al modelo (SV)O en oposición al modelo S(VO), predicho por la gramática.

La teoría fonológica no lineal [2] y [3] propone un mapeo entre constituyentes sintácticos y dominios prosódicos. Dos dominios fueron propuestos: frase fonológica y frase entonativa. Los dos dominios forman parte de una estructura arbórea jerárquica, la estructura prosódica oracional.

Las frases fonológicas requieren al menos de una palabra de contenido que se constituye en su núcleo. El núcleo prosódico juega un rol fundamental en la acentuación. El constituyente prosódico superior es la frase entonativa, construida por agrupación de dominios fonológicos adyacentes. Debe destacarse que los cortes sintácticos mayores son respetados: cada constituyente es vinculado o ligado a algún nodo sintagmático en la estructura superficial que establece un dominio entonacional separado. Por eso a) los constituyentes sintácticos desplazados, b) las cláusulas subordinadas y c) las oraciones parentéticas, son siempre dominios entonativos independientes. Los dominios prosódicos tienden a mantener su longitud. Debe señalarse que existen diferencias en la literatura sobre la cantidad de niveles existente en la jerarquía prosódica.

Gee y Grosjean [1] representa el primer intento, por medio del algoritmo PHI, de realizar una correspondencia entre las estructuras sintácticas y el componente prosódico de la gramática, dando inicio a las llamadas estructuras de actuación -Performance Structures-. El algoritmo propuesto, construido sobre la base de reglas, fue extendido bajo una sintaxis más abarcadora [4] y replanteado dentro del análisis de chunks y dependencias en Abney [13], [14] y [15].

Las estructuras de actuación poseen tres características. La primera, es que sus “unidades básicas” son pequeñas, y en algunos casos no se corresponden con constituyentes sintácticos. Sin embargo, una investigación más detallada indicó que dichas unidades tienen características en común y son susceptibles de ser analizadas en términos gramaticales. La segunda característica es que las unidades presentan una rica estructura jerárquica, hecho revelado por los experimentos basados en la longitud de las pausas. Y en tercer lugar, son más o menos simétricas, estableciendo que existirían condiciones prosódicas y sintácticas que harían que la pausa principal de una emisión se ubique en relación al verbo y el centro de la oración. Se postulan tres tipos básicos de frases fonológicas: a) las compuestas por determinantes, artículos y palabras gramaticales, más modificadores y un núcleo de frase, por ejemplo “la mesa grande”, que podría ser lexicalizado en lenguas más sintéticas, por ejemplo como equivalente a “mesota”, b) pronombres en función de sujeto u objeto más verbos auxiliares que se adjuntan a un verbo principal, por ejemplo: “hemos estado revisando esto”, y c) preposiciones más sus complementos, tales como “en la casa”.

En [4] critican la falta de una teoría sintáctica comprensiva en los primeros trabajos, el hecho de que sus reglas son limitadas y su sintaxis poco especificada para alcanzar una moderada cobertura.

El análisis [4] posee dos ideas fuertes: en primer lugar “describe un nivel de fraseo prosódico que es independiente de la semántica discursiva”. Segundo, “el fraseo es neutral al discurso, depende de una mezcla de factores sintácticos y no sintácticos: la constitución sintáctica, el orden de palabras de izquierda a derecha, y el largo de los constituyentes”. Al mismo tiempo “no es necesario llenar el espacio entre estructura sintáctica y fraseo, ya que el fraseo prosódico puede ignorar los límites sintácticos con el fin de satisfacer las restricciones del largo de frase”. Las reglas prosódicas utilizadas por los autores, intentan dar cuenta de dos aspectos del fraseo entonativo: la ubicación del límite y su prominencia.

Abney [12] introduce una nueva unidad de análisis, el chunk ( fragmento ) para explicar la relación entre sintaxis y fraseo prosódico. Cabe destacar que el chunker se ha convertido en un área de análisis por sí misma en el procesamiento del lenguaje natural, también denominado “shallow parsing” o “partial parsing”. El análisis en chunks permite una buena articulación de las estructuras prosódicas y las estructuras de actuación descritas en la literatura. Tiene como antecedente a las estructuras de actuación [1] y también puede emparentarse con el tradicional análisis de constituyentes inmediatos. El chunk es definido como “los fragmentos de un árbol que son dejados intactos después que hemos desligado los elementos problemáticos”, como por ejemplo, en el caso de la vinculación de una frase preposicional con uno entre varios núcleos nominales, o la ambigüedad de un vínculo como argumento o como modificador.

Klein [18] explora cómo la constitución prosódica podría articularse dentro de teoría HPSG, previa revisión de la teoría fonológica [2] y [3]. El estudio pone su atención en construcciones sintácticas fundamentales, que incluye frases verbales transitivas e intransitivas, modales, complementos nominales, sujetos pronominales, entre otros.

Quené y Kager [22] también parten de las proposiciones de [1] y [3] en el sentido que la prosodia abstracta de la oración no depende directamente de la superficie sintáctica, sino que tiene relación con la estructura prosódica de la oración. Además de este hecho, postula que la acentuación abstracta es afectada por la estructura temática, la cual especifica si los constituyentes funcionan como predicados, argumentos o modificadores, tanto como por la estructura de foco, la cual indica cuáles constituyentes sintácticos deben ser enfatizados con acentos específicos. Dicho trabajo se propone, -ver también [4]- para generar la prosodia de un sintetizador de habla. La derivación de una “prosodia abstracta” es realizada en forma automática, de la cual se deriva tanto la acentuación como el fraseo prosódico.

De Pijper y Sanderman [19] introducen el concepto de una medida de peso de los límites entre constituyentes ( PBS o perceptual boundary strength ) con base perceptiva, independiente de la existencia de eventos fonéticos, o estructuras sintácticas, fonológicas o prosódicas. Demuestran que es posible obtener una medición de los límites entre frases prosódicas por medio del juicio de hablantes no expertos, superior en este aspecto a la obtenida por otros métodos, como el ToBI, que requiere el juicio de expertos entrenados en transcripción. Dicho abordaje puramente perceptivo del grado de juntura entre palabras y constituyentes, no es común en la literatura, donde el problema estuvo siempre asociado a la sintaxis, la semántica o la pragmática, como se demuestra en esta reseña.

A modo de conclusión de los trabajos encontrados en la literatura, podríamos mencionar los siguientes problemas que enfrentan los sistemas revisados: restricciones semánticas de diversa índole, problemas sintácticos derivados del análisis parcial, como el tratamiento de las subordinadas, dificultad para predecir y tratar elementos pragmáticos, entre otros.

Un problema no tratado por la mayoría de los autores reseñados es la restricción semántica que opera sobre las frases prosódicas, las cuales deben poseer “unidad de sentido” [1]. Quené y Kager [22] señalan expresamente los problemas que también se encuentran en otros trabajos basados en el

análisis de la estructura prosódica. La mayoría de los problemas de acentuación (69 %) son causados por factores semánticos, pragmáticos y contextuales. Los autores no encontraron un grado elevado de correspondencia entre información nueva y dada y mayor o menor acentuación. Los autores piensan que la mayoría de los errores de este tipo no pueden solucionarse sin un mejor entendimiento de la relación entre acentuación y el conocimiento del mundo real.

La información dada y nueva representa un problema para los modelos de procesamiento automático del lenguaje natural. En [1] se asume que la información nueva tiene más probabilidad de aparición hacia el final de la emisión, por lo cual se asigna un índice de corte mayor, al mismo tiempo que la ramificación a la izquierda favorece el fraseo de la información nueva. Tratar el problema en forma directa, requeriría de un paso adicional, porque debe establecerse una memoria que almacene el contexto oracional, análisis de la correferencia más allá de los límites oracionales, como también paráfrasis, sinonimia, etc. El problema permanece con los sinónimos y paráfrasis inacentuados, y también con los acentos contrastivos, aunque se ha demostrado que poseen baja frecuencia [20].

### **3. METODOLOGIA**

La metodología utilizada en la literatura tuvo como fin relacionar por un lado la jerarquía prosódica, la cual se revela por los grados o niveles de juntura entre las frases, expresados con algún indicador numérico, y por el otro, el análisis sintáctico parcial. Los niveles de la jerarquía prosódica parten fundamentalmente de los establecidos en [1], [2] y [3], que los articula con un análisis sintáctico simple.

La unidad postulada es equivalente en términos generales: las estructuras de actuación [1], como las frases fonológicas [4], y [22], guardan una fuerte relación con el chunk [13], [14] y [15]. En [4] se incluye a los cambios tonales como indicadores de límites de frases. Tanto [13], [4] como [22] por otro lado, no asignan un límite entre todas las palabras, sino entre límites de frase.

#### **3.1. Corpus**

El corpus elegido para este trabajo corresponde a noticieros radiofónicos. Los 4 locutores, dos femeninos y dos masculinos, son profesionales pertenecientes a radios internacionales, como Radio de las Naciones Unidas, Radio Voz de América, y Radio Nederlands, en sus emisiones en español. Los corpora fueron recolectados, transcritos y etiquetados en parte en anteriores trabajos (Renato 2004). Los locutores elegidos, pertenecen a distintos dialectos del español, siendo todos ellos nativos.

La cantidad de emisiones por hablante promedia las 80, perteneciendo ellas a dos o más programas enteros, sin que se haya realizado una selección del material. Las oraciones descartadas fueron aquellas que presentaron problemas como ruidos de los locutores o de estudio, superposición con cortinas musicales o entre hablantes o llamadas telefónicas de oyentes ( que poseen la resolución frecuencial propia del canal telefónico, en 8 KHz). Las emisiones en su mayoría son oraciones complejas, por subordinación o coordinación. La división en emisiones se realizó teniendo en cuenta la estructura sintáctica y el final de la emisión indicado por un tono bajo, como es tradicional. Las emisiones presentan comúnmente la estructura de tonos de continuación entre frases ( generalmente ascendentes o medios ) y un tono bajo de finalidad.

El total de oraciones alcanza las 409, con un total de palabras de 9038, y un promedio de 22 palabras por oración, sumando en total 31 minutos de habla.

#### **3.2. Análisis prosódico**

Las frases prosódicas fueron definidas como aquellas delimitadas por límites prosódicos: cambios tonales a final de frase y pausas. A cada correlato acústico se estableció un valor numérico, siendo que la principal pausa encontrada en una frase recibe el valor de 3, otras pausas, el valor de 2, cambios tonales indicativos de fin de frase 1, y al resto no se le asignó ningún valor. En el corpus fue etiquetado estableciendo 3 es igual a “:”, 2 es igual a “;” y 1 es igual a “,”. El nivel 0 no fue marcado.

Las pausas utilizadas en el presente trabajo coinciden con las identificadas por un sistema de reconocimiento automático de habla en el momento de asignar una pausa intermedia, que es indicado con el modelo de silencio <sil>, según la convención. Existen al menos dos tipos de silencio o pausa en el habla continua. El primero aparece justo antes del momento de explosión de una consonante oclusiva, como /p,t,k,b,g,d/, y el segundo, como pausa entre palabras. Para fines de control del sistema, se realizó un promedio entre las pausas anteriores a oclusivas internas a palabras, y las que son externas a palabras, tomándose la diferencia como referencia. La diferencia entre las pausas reconocidas por el sistema y las agregadas o suprimidas por la revisión manual fue del orden del 3%. El alineamiento forzado fue corregido en un editor de habla, ya que los sistemas de reconocimiento automático no tienen una resolución temporal óptima.

El cambio tonal fue identificado por un algoritmo utilizado en sistemas de reconocimiento automático de habla para delimitar en frases. El grado de precisión alcanzado con respecto a las correcciones manuales hechas por el autor son del 96 %.

La elección de un sistema de reconocimiento automático está basada en varios motivos. En primer lugar, debido a que no se está realizando una medición relativa al acuerdo entre expertos. Las mediciones subjetivas mejor establecidas, son aquellas realizadas por oyentes “no” expertos, ya que no eligen y desechan frases de acuerdo a los resultados parciales, o no están condicionados por la teoría. Este problema, el hecho que sea necesaria una intervención de expertos, ha sido motivo de numerosas críticas en los últimos años con respecto a la utilización del sistema ToBI [21].

Las pausas dubitativas, debidas a problemas de lectura o disfluencias de los mismos locutores no forman parte del conjunto de mediciones del presente estudio.

Con respecto al análisis sintáctico, se entrenó un chunker probabilístico, basado en entropía máxima, presente en el conjunto de herramientas OpenNLP [22]. El chunker fue entrenado para frases nominales, verbales, preposicionales, adverbiales, y elementos como coordinantes y subordinantes. Los criterios para la transcripción en chunks pueden encontrarse en (). El corpus de entrenamiento está compuesto por 184757 palabras, extraídas del corpus utilizado en [23].

### **3.3. Análisis del corpus**

Gee y Grosjean [1] señalaron tres características de las estructuras de actuación: unidades más o menos pequeñas -relativa a chunks-, simetría, y una rica estructura jerárquica. Con el fin de ordenar la exposición hemos agrupado la evaluación del método de análisis escogido, teniendo en consideración dichas características. También hemos considerado importante ver la frecuencia y tipo de estructuras demarcadas por los tipos de límites estudiados, y si las estructuras prosódicas reflejan las tendencias tipológicas a agrupar S(VO) o (SV)O. Dicho análisis tiene como motivación encontrar las regularidades necesarias sobre las cuales se puedan construir gramáticas que realicen un parsing prosódico.

#### *3.3.1. Unidades prosódicas*

Las unidades delimitadas por límites o cortes, ya sean pausas o cambios tonales, son de variado tipo y longitud, aunque se corrobora una tendencia a que se delimite por frases prosódicas o chunks. La

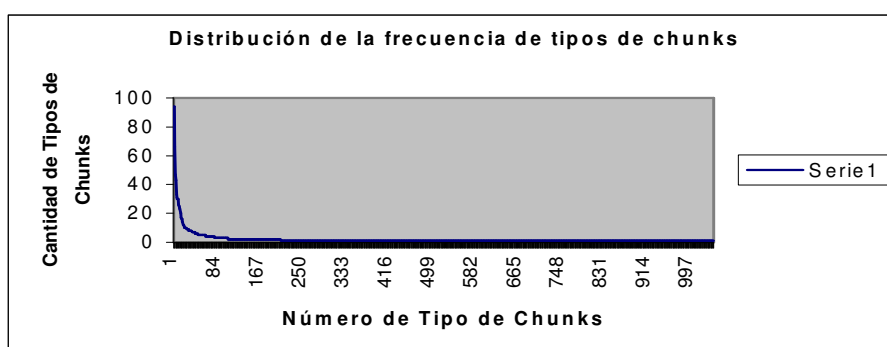
inclusión de frases preposicionales dentro de una frase nominal o formando una frase preposicional aparte, o la ambigüedad de la ligazón de las mismas ( preposition attachment), son los tipos de problemas más frecuentes, para los cuales debiera establecerse algún criterio semántico [1].

En las 409 oraciones se encontraron 2153 chunks delimitados por pausas o cambios tonales, y una cantidad de 1043 tipos de chunks diferentes de acuerdo a la secuencia de clases de palabras que los componen. En la tabla 1, se muestran los tipos de frases más comunes encontrados clasificados por clase de palabra:

Tabla 1. Frecuencia de ocurrencia de chunks clasificados por secuencias de clases de palabras. ( Primeros 60 tipos )

Freq.	Secuencia	Freq.	Secuencia	Freq.	Secuencia
94	NP NP	11	CC D NC	7	D NC SPS NP NP
94	NP	10	V CS D NC	6	V SPS NP
66	SPS D NC	10	SPS NC AQ	6	V SPS D NC
47	D NC	10	CC NP NP	6	SPS D NC SPC NC
43	SPS NP	10	CC NP	6	D NC SPS D NC
33	V	9	RG	6	CS
31	NC	9	P V	6	CC D NC AQ
30	D NC AQ	9	NC SPS NC	5	V SPS VN
30	AQ	9	CC NC	5	V SPS RG NC
26	SPS D NP SPS NP	8	V D NC AQ	5	V SPS NC
25	SPS NC	8	SPS D NC SPS NC	5	V NC
24	NP NP NP	8	SPS D NC SPS D NC	5	VN
23	SPS D NC AQ	8	SPS	5	V CS NP
20	V D NC	8	D W	5	SPS VN NC
17	D NC SPS NC	8	D	5	SPS VN D NC
16	SPS NP NP	7	VIF	5	SPS D AQ NC
16	NC AQ	7	SPS VN D NC AQ	5	RG AQ NC NP
14	D NC NP	7	SPS D NC SPS NP	5	PR V
12	SPS VN	7	SPC NC	5	NP CC NP
12	D NC SPS NP	7	D NP	5	CS D NC

La cantidad de tipos de chunks por clases de palabras ( 1053 ) llevaría a pensar que en realidad no existe una tendencia principal en la constitución sintáctica. Sin embargo, analizando mejor los datos se encuentra que los principales tipos más frecuentes se agrupan en un conjunto más homogéneo y que los tipos menos frecuentes, por ejemplo que aparecen 1 o 2 veces, constituye una larga lista de ejemplos susceptibles de ser agrupados en conjuntos más representativos. El gráfico 1 muestra la distribución de los tipos de chunks de acuerdo a su frecuencia de aparición:



En todo caso vemos que 836 tipos secuencias aparecen solo 1 vez, por ejemplo, ( VP RG ) – participio y adverbio-, 102 tipos de secuencias aparecen solo 2 veces, por ejemplo, ( V RG ) –verbo conjugado y adverbio-, y 26 tipos de secuencias aparecen solo 3 veces. De las 2153 frases, 1112 aparecen en secuencias que aparecen en el corpus 3 o más veces. Podemos ver por otra parte, que de los 102 tipos de secuencias que se presentan dos veces, 8 empiezan con ( SPS D NC ... ) – preposición, determinante, nombre común-, 13 con ( D NC ... ) –determinante, nombre común-, 11 son encabezados por coordinante y u o, y 14 son encabezados por un verbo conjugado. En cierto sentido un análisis más detallado muestra una cierta tendencia a formar chunks que son frases verbales, frases prepositivas o frases nominales o algún miembro en una coordinación.

De todas maneras, teniendo en cuenta que el conjunto de etiquetas POS utilizado es reducido, ( se unificaron todos los tipos de verbos conjugados y determinantes en dos categorías, para reducir la variación ) vemos que las clases constituidas están dispersas, siendo que la fuente de variación está más allá de la sintaxis, sino los tipos debieran coincidir en un conjunto más o menos reducido.

Una de las cuestiones que surgió en el presente estudio, en el que se eligió un corpus periodístico y una cantidad de oraciones superior a la que aparece normalmente en la literatura, es que la frase prosódica más común está constituida por uno a tres nombres propios, siendo a veces encabezados por preposición o coordinante. Forman una frase prosódica, en 94 oportunidades dos NP, en 94 una NP, en 24 casos 3 NPs, y en 10 casos SPS NP, ( preposición y nombre común ). Dicho tipo de frase es en muchos aspectos el tipo de frase prosódica o chunk más común de este corpus, que no se encontró en los estudios clásicos del tema. En la mayoría de los casos, se trata de aposiciones, y posee un nivel de pausa alto antes de cada secuencia.

Por último cabe mencionar que muchas veces las frases prosódicas constituidas son menores a chunks o unidades sintácticas ( por ejemplo, frases terminadas en preposiciones ) o están compuestas por una cantidad de chunks que sería difícil predecir por un conjunto de reglas. De las 2153 emisiones, 450 oraciones ( 20 % ), o son menores a chunks ( 163 – 7 % ) o están compuestas por una cantidad de chunks no alcanzados por las reglas de combinación ( 287 – 13 % ). Como ejemplo de una oración que tiene unidades menores a chunk podemos ver la siguiente -en este caso se trata de elementos focalizados-:

(McCormack/NP) 2 (manifestó/V a/SPS los/D periodistas/NC que/PR sin\_embargo/CA) 1 (Estados/NP Unidos/NP) 2 (y/CC citamos/VIP sus/DP palabras/NC) 2 (no/RN) 2 (tratará/VIF) 2 (de/SPS dictar/VN un/DI curso/NC particular/AQ de/SPS acción/NC) 2 (al/SPC gobierno/NC) 3 (boliviano/AQ)

Como ejemplo de una oración que posee chunks superiores a las reglas de combinación generalmente propuestas podemos ver el siguiente ejemplo:

(El/DA aumento/NC de/SPS los/DA salarios/NC podría/VIC aumentar/VN) 1 (las/DA inquietudes/NC sobre/SPS la/DA inflación/NC) 3 (y/CC es/VIP posible/AQ que/CS sea/VSP uno/PI de/SPS los/DA temas/NC a/SPS tratar/VN durante/SPS la/DA próxima/AQ reunión/NC de/SPS las/DA autoridades/NC del/SPC Banco/NP Central/NP estadounidense/AQ) 2 (la/DA semana/NC próxima/AQ) 2 (cuando/CS considerarían/VIP un/DI posible/AQ aumento/NC en/SPS los/DA tasas/NC de/SPS interés/NC)

Esto da cuenta que el método de chunker propuesto aquí, más las reglas de combinación, tiene una predicción cercana al 80 % del fraseo, que es una medida bastante inferior a la alcanzada por los sistemas probabilísticos, y con reglas ad hoc, cerca del 86 %. Uno de los casos problemáticos encontrados se relaciona con la necesidad de conocer con precedencia ( tener reglas específicas ) para casos como el siguiente:

(la/DA convención/NC internacional/AQ para/SPS la/DA represión/NC del/SPC financiamiento/NC al/SPC terrorismo/NC)



donde se debe conocer de antemano que es el “nombre” de una convención, es decir, de un solo referente del mundo real. En otros casos, dentro de una misma frase aparecen secuencias que pueden ser pensadas como nominalizaciones, pero que no tienen un único referente:

( la/DA perspectiva/NC colombiana/AQ sobre/SPS el/DA origen/NC del/SPC narcotráfico/NC en/SPS ese/DD país/NC latinoamericano/AQ )

### 3.3.2. Simetría

Con respecto a la simetría se realizó un análisis de las pausas y se analizó si la pausa de mayor longitud tiende a ubicarse en el centro de la emisión, de manera de favorecer un análisis biseccional de las emisiones. Con este fin se ubicó la pausa principal de la emisión ( la de mayor longitud ) y se evaluó su ubicación con respecto a la longitud total de la emisión en milisegundos. Los resultados muestran que existe una tendencia marcada a que la pausa se ubique de la mitad hacia el final de la emisión, siendo que esto ocurre con 328 oraciones, el 80 % del corpus. En el último cuarto, se encuentran el 60 % de las pausas principales, contabilizando 249 casos. Solamente 45, el 11 %, se ubica en el primer tercio de la emisión, correspondiendo en la mayoría de los casos a elementos desplazados o a otros elementos focalizados. El 20 % de la emisiones, 84 casos, se ubica en el segundo tercio, mientras que 280, el 68 %, en el último tercio. En este caso, no se obtiene una correspondencia con las tendencias que ubicarían a la pausa principal hacia el centro de la emisión. Estos datos deberían ser nuevamente analizados tomando como factor el largo de los constituyentes, principalmente de sujeto y objeto. La hipótesis que surge en este trabajo, es que la principal pausa se coloca antes del último chunk y anticipa el final de la emisión, sin que aparentemente posea un significado específico.

### 3.3.3. Jerarquía

Con respecto a la jerarquía, se evaluó si la marcación de pausas y cambios tonales refleja una jerarquía de la cual emerjan las estructuras prosódicas o las estructuras de actuación. Los resultados no fueron positivos en este sentido, si se toma en consideración una jerarquía compuesta por: a) pausa más larga, b) demás pausas y c) cambios tonales. En muchas oraciones, las pausas más largas no reflejan una división entre frases entonativas y prosódicas, sino de componentes interiores a las mismas, llegando hasta establecerse límites entre unidades altamente resistentes a la pausa. Veamos unos ejemplos:

1. ( Destaca/VIP que/CS los/DA países/NC latinoamericanos/AQ ) 1 ( han/VIP hecho/VP progresos/NC ) 1 ( para/SPS evitar/VN la/DA mortalidad/NC infantil/AQ ) 3 ( y/CC materna/AQ ) .FP

2. ( En/SPS un/DI mensaje/NC leído/VP por/SPS Rubens/NP Ricúpero/NP ) 1 ( secretario/NC general/AQ de/SPS la/DA UNTAG/NP ) 2 ( Koffi/NP Annan/NP destacó/VIP la/DA importancia/NC ) 1 ( de/SPS evitar/VN medidas/NC proteccionistas/AQ ) 2 ( y/CC de/SPS recurrir/VN 1 a/SPS soluciones/NC ) 3 ( multilaterales/AQ ) .FP

3. ( Desde/SPS el/DA lanzamiento/NC del/SPC pacto/NC ) 1 ( la/DA alianza/NC ) 3 ( ha/VIP trabajado/VP con/SPS éxito/NC ) 1 ( en/SPS programas/NC de/SPS lucha/NC contra/SPS el/DA VIH/NC sida/NC ) 2 ( la/DA eliminación/NC del/SPC trabajo/NC infantil/AQ ) 2 ( la/DA alfabetización/NC de/SPS adultos/NC ) 2 ( y/CC la/DA protección/NC de/SPS recursos/NC ambientales/AQ ) 2 ( percederos/AQ ) .FP

La pausa principal se ubica en 1 entre dos adjetivos coordinados dentro de una misma frase nominal, en 2, entre un nombre y un adjetivo de la misma frase. En el caso 3, si bien no se trata de la pausa principal, vemos un nivel de pausa 2 entre dos adjetivos que pertenecen a la misma frase prosódica. Los niveles y las frases en estos casos, no se relaciona con información dada y nueva en

aparición, y tendría que ver, posiblemente, con un refuerzo en el señalamiento de la pausa final, como se indica en el apartado sobre simetría. Las interpretaciones ad hoc no son fáciles de establecer, porque los casos de excepción son numerosos y diversos. Por ejemplo veamos los siguientes:

4. ( El/DA presidente/NC de/SPS Irán/NP ) 1 ( calificó/VIP a/SPS los/DA autores/NC de/SPS los/DA atentados/NC del/SPC **pasado/AQ** ) 3 ( **once\_de\_septiembre/ZP** ) 2 ( de/SPS ) 2 ( culto/NC de/SPS fanáticos/NC ) 1 ( que/PR se/P han/VIP ) 2 ( automutilado/VP lenguas/NC y/CC orejas/NC ) 1 ( y/CC sólo/RG pueden/VIP comunicarse/VN con/SPS quienes/PR consideran/VIP oponentes/AQ ) 2 ( mediante/SPS la/DA carnicería/NC y/CC la/DA devastación/NC ) .FP

En este caso, la ubicación de la pausa principal puede tener como objeto focalizar la fecha "once de septiembre", colocando una pausa entre ésta y el adjetivo "pasado". En otros casos, se puede ver que la pausa se ubica entre nombre y apellido:

5. ( Dato/NP Param/NP Cumaraswamy/NP ) 1 ( envió/VIP un/DI mensaje/NC al/SPC presidente/NC **George/NP** ) 2 ( **Bush/NP** ) 2 ( en/SPS el\_que/PR le/PP señaló/VIP ) 1 ( que/PR esos/DD tribunales/NC ) 1 ( no/RN garantizan/VIP el/DA derecho/NC ) 1 ( a/SPS una/DI representación/NC legal/AQ ) 2 ( y/CC a/SPS asesoramiento/NC ) 3 ( durante/SPS la/DA detención/NC ) .FP

Cabe destacar que a juicio del autor ( y en experimentos perceptivos que se llevan a cabo paralelamente en otros trabajos ), no se encuentra que éstas emisiones presenten alguna pausa dubitativa o anomalía prosódica.

La teoría fonológica no lineal señalaba que las subordinadas, aposiciones, elementos desplazados, entre otros, constituían dominios en sí mismos. Si bien las estructuras de pausas reflejan una separación en chunks o constituyentes sintácticos mínimos, la jerarquía no se encuentra bien reflejada en este sentido, porque con frecuencia las pausas principales se establecen dentro de las subordinadas, no habiendo un corte entre el elemento subordinante y la principal. Se ha encontrado en numerosas ocurrencias de verbos de comunicar o decir, por citar un ejemplo, que la pausa se ubica entre el sujeto y el predicado de la subordinada, no existiendo un corte antes del verbo principal.

(Se/P espera/VIP que/CS el/DA tema/NC del/SPC **terrorismo/NC**) 3 (ocupe/VSP el/DA papel/NC primordial/AQ de/SPS los/DA discursos/NC)

donde la primera frase prosódica, está conformada por parte de la oración principal y parte de la subordinada.

### 3.3.4. S(VO) vs. (SV)O

La tendencia general es que el verbo opere como un atractor tanto del sujeto como de sus complementos. En cuanto a si el sujeto forma una unidad prosódica con el sujeto (SV)O o con el objeto S(VO), la tendencia fue mayoritariamente a favor del objeto. Las tendencias fueron: a) 0.65 S(VO), b) 0.17 (SV)O, c) 0.098 (SVO), y d) (S) (V) (O) , 0.045. Los últimos casos generalmente no se consideran en los trabajos sobre el tema, pero dado que son casos significativos se incluyen aquí: c) es el caso donde SVO forman una unidad prosódica, y S V O donde cada componente forma una unidad prosódica. Es común que las oraciones breves pertenezcan a un mismo dominio entonativo. Por ejemplo:

( Éstos/PD son/VIP los/DA titulares/NC )

La tendencia no sólo ocurre con los complementos sino también con modificadores. Los predicativos, por ejemplo, forman casi en su totalidad una unidad con los verbos copulativos. La

tendencia general indica que antes del verbo tiende a existir un corte, sea un cambio tonal, una pausa breve o una pausa larga. En general, ésta no suele ser la pausa más larga que se ubica hacia el final de la emisión.

#### 4. CONCLUSIÓN

Las tendencias encontradas en trabajos realizados en su mayoría en inglés y en corpus de pequeñas dimensiones, no se comprueban en su totalidad para el presente trabajo. La tendencia a establecer una pausa prominente en el centro de la emisión, o la tendencia a que las subordinadas se encuentren en un dominio prosódico autónomo, no fue comprobado, así como tampoco se vislumbra una jerarquía prosódica a partir de los datos provenientes del análisis de pausas y cambios tonales. Como en otros trabajos, el chunk aparece como el tipo de unidad que actúe como basamento del análisis en frases prosódicas. Como datos presentes en la literatura, se comprueba una tendencia a ubicar un corte antes del verbo, y de estructuras que forman un dominio prosódico propio: condicionales, concesivas, causales, temporales con mientras y cuando, parentéticas, completivas con quién y cuyo, elementos desplazados a la izquierda, y en este corpus, las aposiciones. El comportamiento de los adverbios en mente merecen un estudio aparte. Sin embargo se requiere tener un corpus más grande, ya que existe una gran variedad de factores intervinientes para ser modelados que no son alcanzados por el tipo de análisis basado en reglas.

#### Bibliografía

- [1] Gee, James y Grosjean, F. (1983) *Performance Structures: A Psycholinguistic and Linguistic Appraisal*. *Cognitive Psychology* 15, 411-458.
- [2] Elisabeth O. Selkirk. (1984) *Phonology and Syntax: The Relation between Sound and Structure*. The MIT Press, Cambridge, MA.
- [3] Nespor, Marina y Vogel, Irene (1986). "Prosodic Phonology". Number 28 en *Studies in Generative Grammar*. Foris Publications. Dordrecht.
- [4] Bachenko, J. y Fitzpatrick, E. (1990) *A Computational Grammar of Discourse-Neutral Prosodic Phrasing in English*. *Computational Linguistics* 16(3), 155-170. 1990.
- [5] Black, A. y Taylor, P. (1994) "Assigning intonation elements and prosodic phrasing for English speech synthesis from high level linguistic input," en *Proceedings ICSLP*, pág. 715–718.
- [6] Black, A. y Taylor, P. (1997) "Assigning phrase breaks from part-of-speech sequences," en *Proc. Eurospeech '97, Rhodes, Greece*, pág. 995–998.
- [7] Cohen, A. S. (2004) "A survey of machine learning methods for predicting prosody in radio speech". University of Illinois at Urbana-Champaign. PhD Doctoral Thesis.
- [8] Wang, M. y Hirschberg, J. (1992) "Automatic classification of intonational phrase boundaries," *Computer Speech and Language*, vol. 6, pp. 175–196.
- [9] Hirschberg, J. (2001) "Training accent and phrasing assignment on large corpora," en *Data-Driven Techniques in Speech Synthesis*, R. I. Dampier, Ed. Boston: Kluwer Academic Publishers, pág. 239–273.
- [10] Levelt, W. (1970). "Hierarchical chunking in sentence processing". *Perception & Psychophysics* 8(2), 99-103.

- [11] A. Stolcke, E. Shriberg, R. Bates, M. Ostendorf, D. Hakkani, M. Plauche, G. Tur, & Y. Lu (1998). "Automatic Detection of Sentence Boundaries and Disfluencies based on Recognized Words". Proc. Intl. Conf. on Spoken Language Processing, vol. 5, pág. 2247-2250, Sydney, Australia.
- [12] Abney, Steven (1991). Chunks and Dependencies: Bringing Processing Evidence to Bear on Syntax. Paper presented at the Workshop on Linguistics and Computation, University of Illinois, Urbana/Champaign.
- [13] Abney, Steven. (1991). "Parsing By Chunks". En: Robert Berwick, Steven Abney and Carol Tenny (eds.), Principle-Based Parsing. Kluwer Academic Publishers, Dordrecht.
- [14] Abney, Steven (1992) "Prosodic Structure, Performance Structure and Phrase Structure". In: Proceedings, Speech and Natural Language Workshop, pp.425-428. Morgan Kaufmann Publishers, San Mateo, CA. 1992.
- [15] Abney, Steven (1995) "Chunks and Dependencies: Bringing Processing Evidence to Bear on Syntax". In: Computational Linguistics and the Foundations of Linguistic Theory. CSLI.
- [16] Martin, E. (1970). "Toward an analysis of subjective phrase structure". Psychological Bulletin 74(3), 153-166.
- [17] Johnson, S.C. (1967). "Hierarchical clustering schemes", Psychometrica 32.  
Jean-Yves Dommergues & François Grosjean. Performance structures in the recall of sentences. Memory & Cognition 9(5), 478-486. 1981.
- [18] Klein, Ewan. 2000. "Prosodic Constituency in HPSG". Grammatical Interfaces in HPSG. Ronnie Cann, Claire Grover and Philip Miller. 2000. Stanford University.
- [19] de Pijper y Sanderman (1994) "On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues". JASA 96 (4) October.  
Sluijter, A. y Terken, J. (1994). "Beyond sentence prosody: Paragraph intonation in Dutch". Phonetica 50, 180-188.
- [20] Christiansen, Morten y Chater, Nick. (2001) "Connectionist Psycholinguistic in Perspective". En: Christiansen y Chater (2001) "Connectionist Psycholinguistic". Ablex Publishing Corporation.
- [21] Kochanski, Greg. (2006). "Prosody beyond Fundamental Frequency". En: Sudhoff, S. y otros (2006) "Methods in Empirical Prosody Research". Language Context and Cognition. Walter de Gruyter. Berlin.
- [22] Quené, Hugo y Kager, René (1993), "Prosodic sentence analysis without parsing". En V.J. van Heuven & L.C.W. Pols (eds.), Analysis and Synthesis of Speech: Strategic Research towards High-Quality Text-to-Speech Generation, 115-130. [Speech Research 11.] : Mouton-De Gruyter. En: <http://www.let.uu.nl/~Rene.Kager/personal/publications.htm>.
- [23] OpenNLP. En: <https://sourceforge.net/projects/opennlp/>
- [24] Renato, Alejandro y Álvarez, José (2004). "Corpora of Latin American Spanish for Research in Prosody and Synthesis". 5th ISCA Speech Synthesis Workshop June 14-16th 2004, Carnegie Mellon University. Alejandro Renato, Jose Alvarez. En: <http://www.ssw5.org/program.html>.