

## Capítulo 7

### La prosodia en la documentación lingüística

*Nikolaus P. Himmelmann*

#### Introducción

Los aspectos prosódicos de un mensaje lingüístico, como la entonación y el acento léxico, son elementos esenciales de su estructura formal. A la fecha, lo esencial del análisis de los rasgos prosódicos todavía no se ha convertido en parte integral de la capacitación en trabajo de campo lingüístico y, por consiguiente, una documentación y descripción razonablemente detallada y amplia de los mismos todavía no forma parte de las prácticas estándares en el trabajo de campo lingüístico. Este capítulo aborda específicamente la *documentación* de rasgos prosódicos, esto es, la cuestión de qué clase de datos tiene que contener una documentación lingüística para que se pueda realizar un análisis prosódico lo más completo posible. Para aprovechar las sugerencias analizadas en este capítulo, es necesario un entendimiento básico de las unidades y procedimientos usados en el análisis prosódico. Si se desea revisar una introducción más amplia al trabajo de campo prosódico centrado en cuestiones de análisis y descripción, véase Himmelmann y Ladd (en prensa).

Dado que una documentación lingüística incluye un gran corpus de grabaciones de eventos comunicativos de diferentes tipos, bien puede cuestionarse si existe alguna necesidad de prestar atención especial a la prosodia cuando se compila dicha documentación. Siempre que las grabaciones sean de calidad aceptable,<sup>1</sup> no hay duda de que un corpus semejante puede usarse para analizar la prosodia incluso cuando no se haya prestado atención particular a los rasgos prosódicos al compilar el corpus.<sup>2</sup>

---

<sup>1</sup>Los rasgos que definen una buena grabación se enlistan en el apartado 2.1 del capítulo 4.

<sup>2</sup>Ejemplos de lo que puede hacerse—y de lo que no puede hacerse—en términos de análisis prosódico sólo a partir de un corpus de grabaciones se encuentran en las tesis sobre la entonación del dyirbal y del bininj gun-wok de King (1994) y Bishop (2002),

Sin embargo, existen esencialmente tres razones para prestar cierta atención especial a los rasgos prosódicos cuando se compila un corpus de datos primarios, de modo que resulte realmente útil para los fines prosódicos:

- 1) Los fenómenos prosódicos son bastante variables y sensibles a las influencias contextuales. Esto dificulta el reconocimiento de patrones distintivos básicos, pero se puede compensar si se tiene la misma enunciación pronunciada por varios hablantes distintos (o al menos múltiples versiones de la misma enunciación). Véase además el apartado 2.
- 2) Las palabras pronunciadas aisladamente son enunciaciones mínimas que materializan elementos tanto léxicos como de nivel enunciativo (post-léxicos). Por consiguiente, la difundida práctica de grabar palabras aisladas (listas de palabras) es de uso limitado para propósitos prosódicos. Véase además el apartado 3.
- 3) Los datos acústicos y auditivos (es decir, las grabaciones de enunciaciones espontáneas o elicidadas) no proporcionan evidencia directa en relación con la percepción de los hablantes nativos, es decir, lo que éstos de hecho perciben como contrastes prosódicos relevantes (aunque el material conversacional puede proporcionar evidencia indirecta; véase más adelante). La manera más directa de obtener datos de percepción es realizar experimentos de percepción, como se analiza de forma más amplia en el apartado 5.

Antes de desarrollar estos puntos, el apartado 1 explica con un poco más de detalle a qué se refiere aquí el término *prosodia*. Además, cuando se analicen los puntos (1) y (2) se sugerirá repetidamente que los datos elicitados pueden proporcionar materiales útiles para complementar los datos de grabaciones espontáneas. Sin embargo, la obtención de datos prosódicos no es una tarea fácil, como se analiza en el apartado 4.

---

respectivamente. La tesis de King se basa exclusivamente en grabaciones en cinta de textos narrativos y de procedimientos realizadas por R. M. W. Dixon en las décadas de 1960 y 1970. Con base en este material, King puede proponer algunos rasgos clave de la entonación del dyirbal. Sin embargo, en diversos puntos tiene que tomar nota del hecho de que los géneros disponibles (en su mayoría narrativos) limitan severamente el alcance de su análisis. Además, advierte que mucho de su análisis será especulativo mientras no sea posible evaluar si los hablantes de dyirbal también perciben como significativas las distinciones prosódicas que él establece a partir de datos acústicos. Sobre la percepción, véase también el apartado 5.

## 1. Los fenómenos prosódicos

La Tabla 1 enlista los principales fenómenos prosódicos de acuerdo con los diferentes dominios en que se manifiestan, es decir, la onda sonora grabable (dominio acústico), la impresión perceptiva (dominio auditivo) y la prosodia como componente del sistema lingüístico (dominio de la categoría fonológica). La columna del extremo derecho enlista las funciones más conocidas que se transmiten mediante rasgos prosódicos (pero por supuesto también a través de medios no prosódicos).

Al analizar la prosodia, es importante mantener los diferentes dominios separados y estar consciente del hecho de que no hay correspondencias inequívocas entre los rasgos en diferentes dominios. Si se toma sólo el tono como ejemplo, existen correspondencias regulares entre los cambios de *frecuencia fundamental* (F0) observadas en la señal acústica, cambios en el tono percibidos por el oído humano y distinciones tonales y de entonación. Sin embargo, estas correspondencias no consisten en relaciones sencillas y directas entre los dominios. Por ejemplo, existen cambios en la frecuencia fundamental que por lo general no se perciben como tales por el oído humano. Éstos se conocen como *perturbaciones microprosódicas* e incluyen fenómenos como la disminución de la F0 a menudo inducida por consonantes sonoras.<sup>3</sup> Más aún, si bien es cierto que las categorías tonales y de entonación están marcadas principalmente por cambios en el tono, otros parámetros auditivos como la longitud, el volumen y la calidad de la voz también funcionan como marcas de estas categorías.

---

<sup>3</sup>La Figura 3 contiene un ejemplo muy claro de este efecto en el que la /b/ del *baa* ocasiona un “decaimiento” notable en la curva de F0. Laver (1994: 452-456) proporciona un análisis más completo de las perturbaciones microprosódicas.

*Tabla 1.* Fenómenos prosódicos de acuerdo con el dominio en que se manifiestan

Acústico	Auditivo	Categoría fonológica	Significado/función
<ul style="list-style-type: none"> <li>– frecuencia fundamental</li> <li>– duración</li> <li>– intensidad</li> <li>– características espectrales<sup>4</sup></li> <li>– pausas/silencios</li> </ul>	<ul style="list-style-type: none"> <li>– tono</li> <li>– longitud</li> <li>– volumen</li> <li>– intensidad/prominencia</li> <li>– ritmo/tempo</li> <li>– agrupación</li> <li>– calidad de la voz (chirriante, etc.)</li> </ul>	<ul style="list-style-type: none"> <li>– tono</li> <li>– cantidad</li> <li>– acento (léxico)</li> <li>– entonación</li> <li>– niveles en la jerarquía prosódica (sílaba, pie, etc.)</li> </ul>	<ul style="list-style-type: none"> <li>– delimitar unidades</li> <li>– distinguir unidades léxicas</li> <li>– categorías gramaticales</li> <li>– actitud del hablante</li> <li>– modalidad de la oración</li> <li>– estructura de la información</li> <li>– tareas interactivas</li> </ul>

En el presente capítulo, las distinciones mencionadas arriba y la terminología correspondiente se observarán de forma bastante estricta. Muchos de los términos se utilizan ampliamente en la literatura en el sentido en que se emplean aquí; sin embargo, puede valer la pena señalar que en rigor la distinción también se aplica a los términos *intensidad* (*stress*) y *acento léxico* (*lexical accent*), que se utilizan en la literatura especializada en muchas formas distintas y a menudo un tanto confusas. Ambos términos se refieren al fenómeno de que una sílaba determinada es en cierto sentido más prominente que las sílabas vecinas; sin embargo, el *acento léxico* aquí designa que esta propiedad se refiere a la estructura fonológica de los elementos léxicos (es decir, como una categoría fonológica), mientras que *intensidad* se refiere a una impresión auditiva (que puede o no tener correlatos fonológicos o acústicos definidos). Entendido así, el acento léxico puede realizarse de diferentes maneras, incluida la “intensidad” o un cambio fijo en el tono (el llamado acento

<sup>4</sup>Las “características espectrales” aquí se refieren a aquellos aspectos de la estructura del formante de los sonidos del habla que reflejan rasgos prosódicos; por ejemplo, la distribución de la energía a lo largo del espectro de frecuencia puede ser un correlato acústico del acento.

melódico o acento tonal, que se encuentra, por ejemplo, en japonés; cf. un análisis más amplio en Beckman 1986 y Gussenhoven 2004).

Aquí no tenemos el espacio suficiente ni la necesidad de abordar con detalle todas las funciones y fenómenos prosódicos enlistados en la Tabla 1. La finalidad principal de esta tabla es dar una definición extensional de la gama de fenómenos a los que se refiere el término *prosodia* en este capítulo. Se puede encontrar una introducción detallada a la fonética (acústica y auditiva) de rasgos prosódicos en Laver (1994: 431-546; véase también Ladefoged 2003: 75-103). Las categorías fonológicas principales se analizan en Ladd (1996), Cruttenden (1997), Hirst y di Cristo (1998), Hyman (2001), Yip (2002), Gussenhoven (2004) y Jun (2005), entre otros. Estas obras también proporcionan información útil en relación con la variabilidad translingüística de los rasgos prosódicos.

En principio, el análisis de este capítulo se aplica a todos los rasgos prosódicos enlistados en la Tabla 1. Sin embargo, la entonación y el acento léxico se mencionarán como los ejemplos principales y con frecuencia se destacarán para hacer comentarios adicionales, porque son dos categorías que se han obviado ampliamente en el trabajo de campo lingüístico, en oposición al tono, por ejemplo, que ya es un tema estándar.

## **2. La necesidad de trabajar con varios hablantes**

El trabajo de campo lingüístico a menudo consiste en una cooperación cercana con sólo uno o tal vez dos hablantes nativos que son los principales colaboradores o “informantes”, en el sentido de que: a) proporcionan mucha de la información elicitada de la lengua (los textos con frecuencia se graban con una gama más amplia de hablantes); y b) cualquier dato proporcionado por otros hablantes se procesa y verifica con ellos. Este procedimiento se basa en el hecho de que respecto a los elementos gramaticales nucleares la información brindada por diferentes hablantes tiende a no diferir (o a diferir sólo de forma mínima). Así, por ejemplo, si un hablante expresa que el artículo definido tiene que anteceder al sustantivo y no puede ser pospuesto, con toda probabilidad esto será confirmado por todos los demás hablantes en la comunidad.

Mientras que este arreglo funciona razonablemente bien para los elementos estructurales más básicos de una lengua, se vuelve más y más problemático cuando se están investigando otros elementos lingüísticos más complejos y variables. La fonética de los rasgos prosódicos es altamente variable y depende de una serie de factores complejos, incluyendo variables del hablante y el contexto. Hay muy pocos valores absolutos, si es que

alguno. Lo que es un tono alto para un hablante, puede ser un tono bajo para otro; lo que es fuerte en un contexto, sólo es normal en otro, etc. Además, la percepción de los rasgos prosódicos por parte del investigador tiende a estar influida en gran medida por su propio sistema prosódico, que distorsiona todavía más los datos y complica el análisis. Por consiguiente, en las primeras etapas de la investigación prosódica de una lengua tiende a ser extremadamente difícil reconocer un patrón básico en los datos grabados. Este problema es particularmente apremiante en el caso de la entonación, la cual por esta razón sirve como el ejemplo principal en esta sección, aunque también puede presentarse con respecto al acento léxico o al tono.

La forma más sencilla de resolver el problema del reconocimiento de patrones es que varios hablantes “hagan lo mismo”, esto es, que produzcan la misma enunciación en el mismo contexto y con la misma intención. La Figura 1 ilustra el problema y la solución sugerida. Muestra las gráficas originales y simplificadas de la frecuencia fundamental (F0) correspondiente al segmento (*was für große Ohren du hast* “(qué grandes) orejas tienes”, tomado de grabaciones del cuento popular “Caperucita Roja” narrado por cinco hablantes de alemán.<sup>5</sup> Todos los hablantes producen una elevación en la sílaba acentuada inicial *Oh* y después un descenso continuo hasta el final de la enunciación. Adviértase qué tan variable es la elevación inicial (área sombreada en las gráficas simplificadas). Para el hablante JH es bastante larga y comienza muy pronunciada, pero después se vuelve más plana, mientras que para el hablante NF es pronunciada y breve. Ciertamente, la elevación inicial del hablante JN es muy menor y podría argumentarse que no existe en absoluto en esta sílaba. Sin embargo, como los cinco hablantes están “haciendo lo mismo”, es decir, produciendo la misma enunciación en el mismo contexto (la lectura de un cuento en voz alta) y con la misma intención (expresar sorpresa por los cambios radicales en el aspecto de la abuela), también es válido suponer que las diferentes elevaciones y descensos en las F0 constatadas en estas gráficas de hecho son realizaciones de la misma categoría, es decir, el decaimiento nuclear del alemán estándar del norte (simbolizado con H\*+L en la notación ToBI). O, considerado desde el punto de vista de quien intenta detectar un patrón básico, el hecho de que se pueda suponer razonablemente que las cinco ejecuciones son “lo mismo” al nivel del sistema de la lengua permite

---

<sup>5</sup>Las gráficas se presentan en dos versiones: la columna derecha presenta los registros originales de F0 y la columna izquierda presenta una versión simplificada de los mismos. Véase Grabe (1998: capítulo 2) para encontrar información adicional sobre los procedimientos usados para reunir y procesar los datos. Su tesis está disponible en <http://www.phon.ox.ac.uk/~esther/thesis.html>.

reconocer un patrón común, que es la elevación en la sílaba acentuada más el descenso continuo hasta el final de la unidad de entonación.<sup>6</sup>

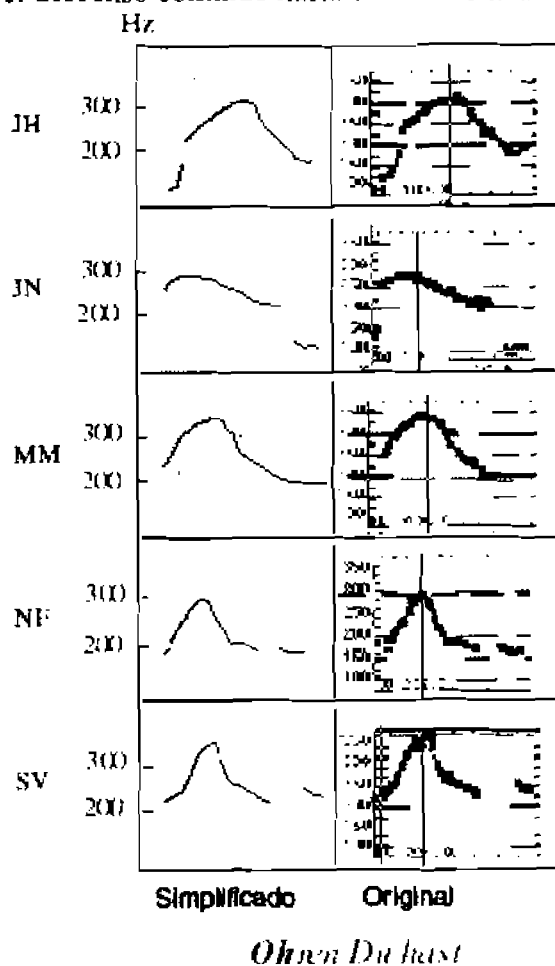


Figura 1. Ejecuciones múltiples de la misma expresión en alemán, *Ohren Du hast* “¿(qué grandes) orejas tienes!”, del cuento “Caperucita Roja”. Las gráficas aparecen en su formato original (derecha) y en un formato simplificado (izquierda). (Tomado de Grabe 1998: 245, Apéndice C.)

<sup>6</sup>Los detalles precisos del análisis no nos ocupan aquí. Nótese que Grabe (1998: capítulo 3, sección 2) propone que el decaimiento nuclear en el alemán estándar del norte permite dos realizaciones alternativas principales, una con una clara elevación en la sílaba acentuada y otra en la que el tono está más o menos a nivel en la sílaba acentuada (como con el hablante JN en la Figura 1). Se argumenta que la distinción entre estas dos realizaciones (fonéticas) de la misma categoría fonológica es gradual.

Aquí “hacer lo mismo” implica tres cuestiones. Primero, las enunciaciones tienen que ser idénticas (o al menos muy similares) en términos segmentales (deben corresponder al mismo segmento), porque si son diferentes tienen efectos microprosódicos diversos y no es una tarea sencilla filtrarlos para reconocer un patrón básico. Segundo, las enunciaciones tienen que transmitir el mismo sentido y, lo que es más importante, tienen que ejecutarse con la intención de lograr el mismo acto ilocutivo. Es consabido que la misma enunciación—el mismo segmento—puede emplearse para hacer una pregunta, dar una orden, hacer un comentario irónico, expresar sorpresa, etc. Todas estas funciones distintas afectan la configuración prosódica y por consiguiente tienen que ser controladas si se buscan enunciaciones idénticas en términos prosódicos. Tercero, las enunciaciones tienen que producirse en circunstancias idénticas (o muy similares), por ejemplo, como comentarios espontáneos entre adolescentes, en un ambiente de trabajo entre personas de condición social diferente, etcétera.

Con respecto al número de enunciaciones “idénticas” requeridas para un análisis prosódico detallado, existen algunas guías generales. El mínimo absoluto para reconocer un patrón con cierto grado de confiabilidad es de tres ejecuciones, porque con sólo dos versiones de la misma enunciación es difícil, cuando no imposible, decidir cuáles aspectos divergentes son distintivos y cuáles son azarosos. Un buen comienzo para un análisis detallado es contar con cuatro versiones de la misma enunciación, idealmente dos de varones y dos de mujeres. Si se tienen ocho diferentes versiones, los análisis estadísticos se vuelven más viables y útiles. Con 10-12 hablantes, el tamaño de la muestra se aproxima al de muchos trabajos sobre lenguas bien documentadas, como el inglés o el japonés.

No hay un límite superior establecido para el tamaño de la muestra y, dependiendo de los fenómenos que estén siendo investigados, pueden necesitarse muestras más grandes que también tomen en cuenta variables como la edad, el registro y el dialecto local. Reiterando, en el típico entorno de una lengua hasta la fecha no documentada, hablada por una pequeña cantidad de hablantes, muestras de cuatro a diez versiones de la misma enunciación proporcionarán una buena base para un análisis prosódico detallado y por consiguiente mejorarán en gran medida la base de datos para la investigación prosódica.

Adviértase también que, si bien es preferible, no es absolutamente necesario que las diferentes versiones sean producidas por hablantes distintos. También podrían haber sido producidas por el mismo hablante o hablantes en diferentes ocasiones. Lo que es más importante, las



repeticiones más o menos inmediatas de la misma enunciación (como cuando se pide al hablante que repita algo que acaba de decir o que diga algo dos veces) por lo general no producen distintas versiones de la misma enunciación, porque la repetición usualmente influye en la prosodia.

Debe resultar obvio que incluso en un corpus muy grande de grabaciones de habla más o menos espontánea resultará difícil encontrar una serie de cuatro a diez versiones de la misma enunciación con la intencionalidad propuesta. Por ejemplo, puede haber cientos o incluso miles de enunciaciones que se identifiquen razonablemente y sin temor a equivocarse como preguntas polares (es decir, preguntas que se responden con las opciones *sí* o *no*, como “¿Vendrá hoy?”). Sin embargo, ¿cuántas de éstas serán idénticas o al menos muy similares en términos segmentales? Es más, las circunstancias en que se hace la pregunta tal vez no sean en realidad comparables, todo lo cual dificulta determinar aquellos aspectos en la configuración prosódica que se relacionan con distinciones de categoría. Para estar seguros, en el caso de las preguntas polares, quizás sea posible determinar estos aspectos con un grado de certidumbre razonable a partir de una muestra suficientemente grande de habla espontánea. Sin embargo, resulta más engorroso hacer esto basándose sólo en una muestra de este tipo y puede resultar más y más difícil cuando se investigan cuestiones más complejas. En particular, cuando se estudian problemas de la configuración prosódica de la estructura de la información (foco, contraste, neutralización, etc.), el número de variables que se deben controlar y explicitar puede llegar a ser tan alto que los resultados no dejarán de ser especulativos.

Idealmente, pues, una documentación lingüística exhaustiva debe contener series de diferentes versiones de la misma enunciación y cada serie debe representar una función en la que la prosodia pueda ser relevante (esto es, una serie para preguntas polares, otra para enunciaciones nuevas, otra para peticiones corteses y así sucesivamente). Mientras que semejantes series pueden llegar a ocurrir en un corpus suficientemente grande de grabaciones espontáneas sin que se haya prestado atención particular al tema del análisis prosódico, existen tres formas de asegurar que las series queden representadas en la documentación.

Primero, trabajar con instrumentos de elicitación, como video clips, cuentos con dibujos o juegos en los que un hablante instruye a otro para que identifique un objeto dentro de un conjunto o para que encuentre una ruta a través de un paisaje imaginario (la llamada “tarea del mapa”). Esto producirá enunciaciones similares, si no es que verdaderamente idénticas.<sup>7</sup>

<sup>7</sup>En el capítulo 6 y en el sitio de Internet del libro pueden encontrarse referencias adicionales y ligas a estímulos detonadores.

Los juegos en que los hablantes realizan diferentes tipos de actos de habla (por ejemplo, hacer una pregunta, dar direcciones, confirmar una solución propuesta) son particularmente útiles, siempre que la estructura del juego fuerce a los hablantes a hablar acerca del mismo “mundo” (es decir, a usar los mismos elementos léxicos), de forma que las enunciaciones lleguen a ser razonablemente similares con respecto a su estructura segmental.

El segundo método para producir conjuntos de datos relevantes es intentar la elicitación directa de datos pidiendo a los hablantes que produzcan enunciaciones o, más bien, pequeñas narraciones preparadas con anticipación. El principal problema aquí es cómo presentar las enunciaciones deseadas de forma que la prosodia no sea influida por el estímulo. Observaremos más de cerca el problema de la elicitación en el apartado 4. Aquí se presentan algunos ejemplos de la clase de enunciaciones que uno puede elicitar con una indicación de la función prosódica que se desea obtener, la cual se señala entre corchetes:

1. ¿Ya llegó X? No, no lo/la he visto todavía.  
[par adyacente de pregunta polar + respuesta]
2. (En el mercado) ¿Qué busca? (Estoy buscando) las verduras.  
[par adyacente de pregunta tipo qué/quién/dónde/cuándo + respuesta]
3. ¡Siéntese, por favor!  
[petición cortés]
4. (Grupo de personas paradas junto al camino, obviamente agitadas.  
¿Qué pasó?):  
¡Se volcó un camión! o ¡El perro mató a un cerdo!  
[enunciaciones nuevas]
5. Me gusta la camiseta *azul*, no la *roja*.  
[foco contrastivo]

6. ¿Alguna vez has comido un gusano de maguey? No, no como gusanos.  
[neutralización de la presuposición]<sup>8</sup>
7. (Sorpresa:) ¡Qué grande estás ya!  
[actitud afectiva del hablante]

Esta lista de ejemplos no está completa y debe ampliarse y adaptarse de acuerdo con el contexto y composición del proyecto. Sin embargo, como obtener tales ejemplos por lo general no resultará tarea fácil y no será algo que los hablantes nativos harán con entusiasmo, se debe invertir tiempo considerable en el bosquejo del conjunto adecuado de ejemplos y probarlos todos con un colaborador cercano antes de acercarse a un número más grande de hablantes para hacer una grabación.

Un factor que debe considerarse al hacer un borrador de los ejemplos es la estructura segmental. Los ejemplos deben incluir las menos fricativas posible y, en general, deben evitar las consonantes sordas en todos los modos de articulación. De hecho, el ejemplo ideal consta sólo de vocales y nasales, lo cual por supuesto es un ideal que difícilmente se logrará si se busca elaborar ejemplos que tengan sentido y sean culturalmente apropiados. Tener estímulos bien formados en términos semánticos y pragmáticos y que sean culturalmente apropiados será en general la preocupación más importante, ya que de otro modo la elicitación no funcionará en absoluto.

La tercera forma de obtener conjuntos de datos comparables para el análisis prosódico es asegurarse de que el corpus de grabaciones contenga un número suficiente de enunciaciones con *entonación estilizada*. Un ejemplo típico y muy cotidiano de la entonación estilizada es la “curva vocativa” que se forma cuando una persona llama a otra (Ladd 1996: 88, 136 y ss.). Puede haber diferentes curvas vocativas: una para el llamado común (“¡Pedro!”), otra para los pregones en el mercado, otra para los vendedores ambulantes que anuncian su mercancía, etc. En muchas lenguas, producir listas (decir, por ejemplo, “ellos tenían muchas vacas, cabras, pollos y perros”) también implica una entonación especial, en cierta forma

---

<sup>8</sup>Los ejemplos (4) a (6) apuntan a distinciones en la estructura de la información, un tema bastante complejo que no puede abordarse de forma adecuada aquí. Véanse Lambrecht (1994) y Jacobs (2001) para un análisis minucioso de algunas de las distinciones básicas y problemas relacionados, Ladd (1996) para la función que puede desempeñar la prosodia en el marcado de la estructura de la información, Drubig (2003) para un estudio tipológico y Dinroth (2002) para una tarea de elicitación centrada en la estructura de la información.

estilizada (“entonación de lista”; véase también la siguiente sección). Además, la entonación estilizada es un rasgo común de muchas formas de habla ritual, en particular del “habla cantada”.

Para los propósitos del análisis prosódico, la principal ventaja de las curvas de entonación estilizada es el hecho de que por su naturaleza misma, los contrastes de entonación son más estables y más marcados que en la entonación no estilizada. Por consiguiente, los patrones suelen reconocerse con mucha mayor facilidad. De hecho, mientras que los hablantes nativos con frecuencia no tienen intuiciones muy claras sobre los patrones de entonación no estilizados, a menudo conocen los patrones estilizados y fácilmente pueden imitarlos.

Es obvio que los patrones empleados en entonaciones estilizadas difieren de aquéllos empleados en las no estilizadas; de forma similar, puede darse el caso de que los patrones de entonación elicitados difieran muy claramente de aquéllos encontrados en el habla espontánea (compárese el fenómeno de la “entonación de lectura” encontrado en muchas lenguas europeas). A este respecto, debe entenderse claramente que los conjuntos de datos elicitados y las enunciaciones estilizadas tienen la función de permitir que el investigador comience el análisis prosódico y, en específico, el análisis de la entonación, pues le dan una idea básica de qué clase de contrastes se presentan en la lengua y le permiten desarrollar hipótesis que tendrá que corroborar con el material espontáneo. Por supuesto, un análisis prosódico exhaustivo tiene que describir la gama completa de fenómenos que se encuentran en un corpus de grabaciones espontáneas.

### **3. La grabación de palabras**

Una práctica muy extendida en la lingüística es la grabación de listas de palabras elicitadas para poder verificar las propias transcripciones y documentar la estructura sonora básica de los elementos léxicos. El formato estándar en tales grabaciones es pronunciar primero la traducción de la palabra en la lengua de contacto empleada (o el número de la palabra en una lista de palabras) y luego la palabra en la lengua documentada, a menudo repetida una o dos veces más. De esta forma, las palabras se graban “aisladas”, lo cual a menudo significa “en su forma más básica, libre de cualquier influencia contextual ‘contaminante’”. Sin embargo, esto es una concepción equivocada, ya que emitir una palabra aislada ya constituye una enunciación, aunque mínima, y esto es lo determinante para la prosodia. Sobre todo, las “palabras aisladas” no sólo revelan sus elementos prosódicos de tipo léxico (el tono o el acento léxicos), sino también los

elementos de la prosodia enunciativa (comúnmente declarativa). Éste puede parecer un punto bastante trivial pero incluso en la literatura especializada no se había hecho hasta hace muy poco esta distinción de forma consistente.<sup>9</sup>

Como ejemplo, compárese las Figuras 2 y 3. La Figura 2 muestra la forma de onda y la gráfica de F0 para una palabra del waima'a, *kaluha* “nube”, grabada de forma aislada. La Figura 3 muestra la forma de onda y la gráfica de F0 para una enunciación breve en waima'a, *kii baa ini* “hay gente peleando” (literalmente, “la gente se golpea”, una respuesta nueva a una pregunta tipo “¿qué pasa?”). Nótese que la gráfica de F0 es esencialmente idéntica en las dos figuras: comienza plana a la mitad del rango tonal,<sup>10</sup> se eleva, comienza a descender de nuevo en la penúltima sílaba y continúa cayendo en la última sílaba. Por consiguiente, surge la pregunta de si la elevación en la penúltima sílaba en *kaluha* es parte de la configuración léxica de este elemento y refleja al menos en parte un acento léxico regular en la penúltima sílaba. Una alternativa es que esta elevación y descenso en las dos últimas sílabas —un patrón observado prácticamente en todos los elementos léxicos del waima'a pronunciados en aislamiento— se debe exclusivamente al hecho de que emitir una palabra aislada en waima'a también implica elementos a nivel de una enunciación declarativa estándar en esta lengua. (Mientras redacto este capítulo, creo que la última opción es correcta, pero esto requiere más investigación y evaluación. Para los propósitos actuales, no viene al caso decidir cuál de las dos opciones resulta correcta. El punto que debe quedar claro es que las palabras aisladas siempre y necesariamente despliegan rasgos de entonación propios del nivel enunciativo.)

---

<sup>9</sup>Bruce (1977) es considerada por muchos la primera obra moderna donde la distinción se aplica de forma completa y consistente.

<sup>10</sup>Como ya se mencionó, el “decaimiento” inicial en la Figura 3 es una microperturbación ocasionada por la /b/ en *baa*. La enunciación de la Figura 2 es de un hablante masculino, la de la Figura 3 es de una hablante femenina, y es por tanto en general bastante más alta. Los archivos de ondas con las enunciaciones de las Figuras 2-5 están disponibles en el sitio web del libro.

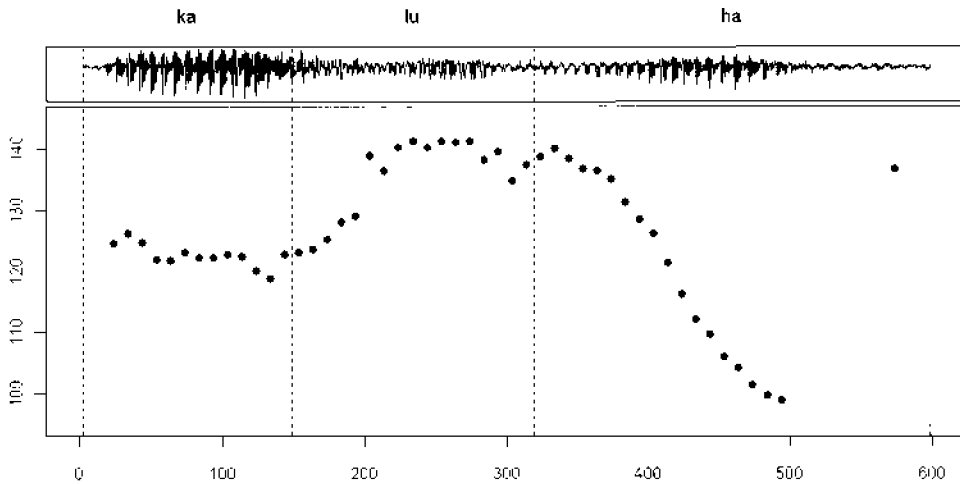


Figura 2. Forma de onda y frecuencia fundamental de una palabra aislada del waima'a (*kaluha* "nube")

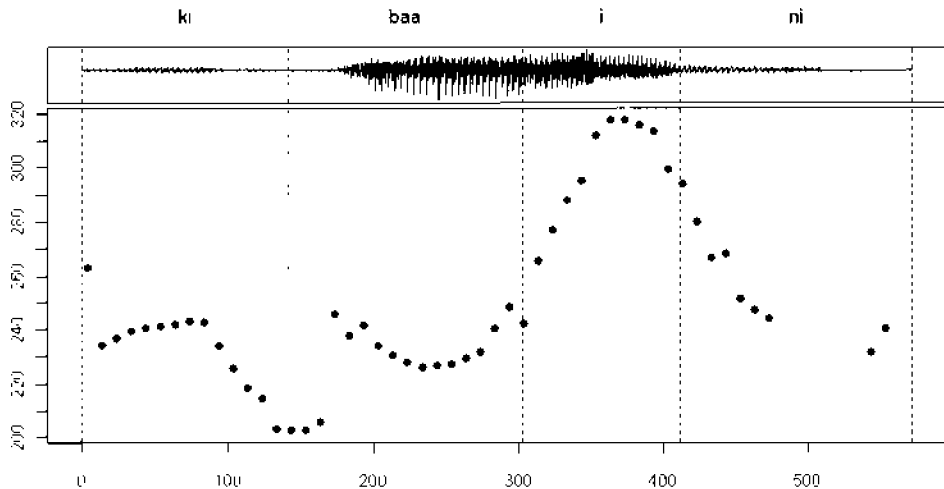


Figura 3. Forma de onda y frecuencia fundamental de una enunciación breve del waima'a (*kii baa ini* "hay gente peleando")

Para poder separar claramente los elementos prosódicos léxicos y post-léxicos (al nivel de la enunciación), ahora se ha vuelto una práctica común en la investigación prosódica (pero también en muchos estudios fonéticos segmentales) no grabar palabras aisladas, ni siquiera cuando el léxico sea la inquietud principal. En cambio, lo ideal es grabar la o las palabras meta en diferentes posiciones en una *frase portadora*, como en los siguientes ejemplos:

8. La palabra meta, “América”, en diferentes posiciones en una frase portadora
- a. “América” es una palabra que conozco. [posición inicial]
  - b. Dije “América” una vez. [posición interna]
  - c. Ella dijo “América”. [posición final]

Como se observa en estos ejemplos, las diferencias de posición por lo general tendrán implicaciones para la estructura de la información, que pueden, pero no tienen que, correlacionarse con distinciones prosódicas post-léxicas. Más aún, como el propósito de estas grabaciones es comparar características de diferentes elementos léxicos, la frase portadora comúnmente lleva elementos muy generales, en particular verbos como “decir”, “oír” o “conocer” que en principio son compatibles con todos los elementos léxicos.

Las Figuras 4 y 5 muestran el efecto de la posición dentro de la frase portadora con otro ejemplo del waima’a. Aquí, la palabra meta *aboo* “abuelo o persona mayor respetada” se presenta al final de una frase portadora (*ne ehe aboo* “ella dijo *aboo*”) y al comienzo de otra (*aboo aku de nau* “*aboo* es una palabra que conozco”, literalmente “*aboo* la sé”).<sup>11</sup> Nótese cómo el cambio en la posición se correlaciona con un evidente cambio en el tono (elevación y descenso en *boo* en posición final, elevación tardía en *boo* en posición inicial). Sin embargo, debe advertirse también lo que permanece constante en las dos posiciones. Lo más importante es que en ambas realizaciones *boo* es más o menos el doble de largo de la sílaba inicial *a*. Por consiguiente, podemos suponer que *boo* contiene una vocal larga como parte de su configuración léxica y que la longitud silábica en las dos grabaciones no se debe a un efecto de nivel enunciativo.

---

<sup>11</sup>El hablante, por supuesto, conoce la palabra *aboo*, pero colocarla en posición inicial y no usar una negación (es decir, utilizar el equivalente de “*aboo* la sé” como estímulo) no le pareció apropiado.

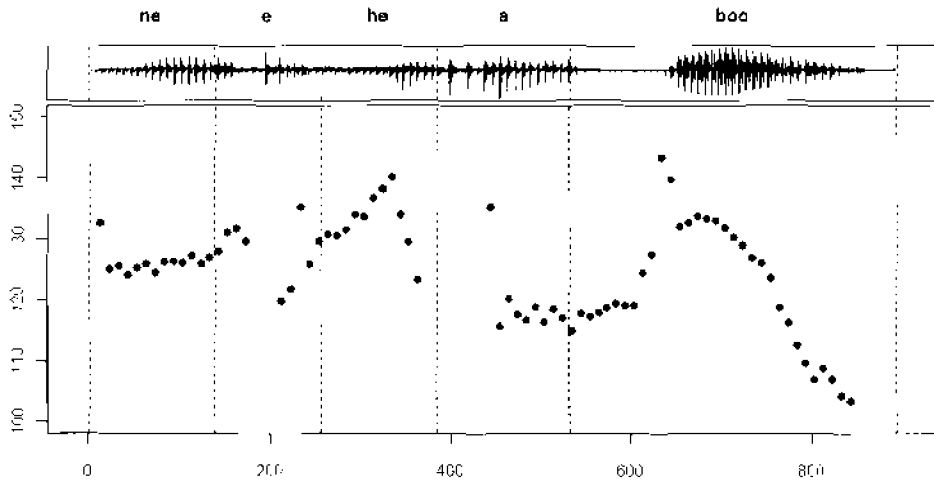


Figura 4. Frase portadora del waima'a con la palabra meta al final (*ne e he boo* "él/ella dijo *boo*")

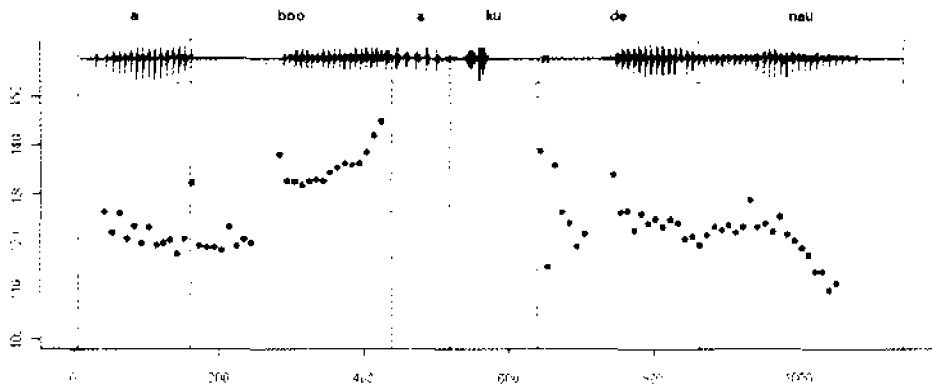


Figura 5. Frase portadora del waima'a con la palabra meta al inicio (*boo aku nau* "aboo no la sé")

Si trabajar con frases portadoras resulta ser demasiado incómodo o no funciona por alguna otra razón (véase la siguiente sección), se puede intentar grabar palabras en mini listas de tres o cuatro elementos, alternando la posición de las palabras incluidas en la lista, como en el ejemplo 9.

#### 9. Mini listas con órdenes alternativos

- a. *América, África, Antártida*
- b. *África, Antártida, América*
- c. *Antártida, América, África*
- etc.



Aunque no es tan útil como grabar frases portadoras, estas mini listas a menudo permiten que se haga al menos una distinción entre la prosodia enunciativa final y no final, siempre y cuando los hablantes utilicen la entonación de lista y no únicamente produzcan tres enunciaciones aisladas en una secuencia rápida. Como en el ejemplo de la frase portadora, se puede formular la hipótesis de que los elementos prosódicos que permanecen idénticos a lo largo de posiciones diferentes en la lista pertenecen más al léxico que al nivel enunciativo.

#### 4. El problema de la elicitación

La mayoría de los procedimientos descritos en los dos apartados anteriores consisten en elicitación de datos prosódicos al pedirles a los hablantes que produzcan diversas clases de enunciaciones o mini narraciones. Cualquier tipo de elicitación implica casi siempre problemas relacionados con la naturalidad y confiabilidad de los datos obtenidos, pero la elicitación de datos prosódicos es particularmente propensa a causar distorsiones mayores, ya que los rasgos prosódicos son altamente sensibles a las influencias contextuales. Por lo tanto, no tiene mucho caso simplemente pedir a los hablantes nativos que repitan lo que dice el investigador o alguno de sus colaboradores. En casi todas las circunstancias, esto producirá enunciaciones bastante distorsionadas que más bien imitarán los rasgos de la enunciación presentada o desplegarán los rasgos prosódicos propios de la repetición.

El procedimiento más utilizado en la investigación prosódica con lenguas que poseen una tradición de escritura bien establecida es hacer que los hablantes *lean* las enunciaciones meta. Si bien no influye directamente en la prosodia, como los modelos de imitación, este procedimiento puede enfrentar otros problemas. Lo más importante es que las tareas de lectura requieren que los hablantes de hecho *actúen* el tipo de enunciación que se quiere lograr. Obviamente, no tiene mucho caso hacer que alguien lea una pregunta o una exclamación de sorpresa con una voz monótona, no comprometida y más bien plana. No todos los hablantes son capaces o están dispuestos a participar en una actuación de este tipo. Los estímulos de lectura exitosos también suponen que los hablantes sean lectores razonablemente competentes. Con frecuencia, éste no será el caso ni siquiera en comunidades con hablantes alfabetizados en una lengua dominante, pero no acostumbrados a ver su lengua escrita (leer en esas circunstancias será un proceso lento y tendrá un estilo de “palabra por

palabra”). Otra complicación puede surgir del hecho de que la entonación de lectura difiere en forma significativa de la entonación conversacional.

En sociedades analfabetas, los estímulos escritos obviamente no funcionarán en absoluto. La alternativa principal aquí es intentar diversas clases de dramatizaciones o tareas experimentales que incluyan video clips y otros estímulos (mencionados en el apartado 2). Las dramatizaciones pueden funcionar cuando se preparan cuidadosamente con un miembro del equipo local. Los hablantes simulan que se encuentran en una situación determinada y reaccionan con una enunciación breve apropiada y ensayada con anticipación. Así, por ejemplo, se puede pedir que un par de hablantes simulen toparse en el mercado, para que luego uno de ellos pregunte “¿qué pasa aquí?” y el otro responda con la enunciación meta “la gente se está peleando”. En el mejor de los casos, los hablantes que participan en esta dramatización se enfrascarán de hecho en una pequeña conversación y después de este par pregunta-respuesta imaginario continuarán con una breve secuencia de enunciaciones adicionales. No siempre será posible lograr que usen exactamente la enunciación meta preparada con anticipación, pero las variaciones menores en esta composición de segmentos en general no causarán mayor problema para efectos comparativos. Entre más realista sea la dramatización, mejor será la calidad de los datos prosódicos producidos de esta forma.

Al preparar las escenificaciones y las tareas experimentales debe tenerse en cuenta que éstas con toda probabilidad serán eventos comunicativos muy extraños para los hablantes nativos que no estén familiarizados con la idea básica de las dramatizaciones, experimentos o entrevistas. Por consiguiente, el investigador debe estar preparado para enfrentar muchos obstáculos cuando intente recolectar datos de esta forma. Las carcajadas continuas o las risitas nerviosas debido al carácter inusual y artificial de la situación son un problema bastante común. Los hablantes también pueden cambiar el acto de habla; es decir, en vez de responder con una afirmación (“Se fue al mercado”) quizás produzcan una orden (“¡Ve al mercado!”). Además, no es raro que cuando se pide a los hablantes que narren una secuencia breve vista en un video clip, hagan en cambio un comentario sobre el tipo de vestimenta que la gente estaba usando o el color del cielo que se aprecia en el video, en vez de acometer la tarea solicitada. Por consiguiente, puede requerirse considerable tiempo e ingenio al diseñar estímulos apropiados para lograr que las tareas experimentales funcionen o para diseñar formas útiles de dramatización en una comunidad determinada. Sin embargo, este esfuerzo valdrá la pena porque los datos así generados

resultarán muy útiles no sólo para los análisis prosódicos, sino a menudo también para otros tipos de análisis.

## 5. Los experimentos de percepción

Todos los procedimientos presentados hasta ahora en este capítulo se enfocan en la producción de datos, es decir, conjuntos de enunciaciones que pueden analizarse acústica y auditivamente. Sin embargo, la producción de datos no proporciona ninguna base para determinar qué componentes de la señal compleja perciben los propios hablantes como rasgos prosódicos distintivos. A partir de las investigaciones sobre las lenguas europeas, se sabe que no toda la información relevante en términos acústicos es percibida como tal por los hablantes nativos. Por consiguiente, es necesario que los datos respondan a preguntas como: ¿los hablantes nativos perciben este rasgo claramente observable (por ejemplo, un cambio en la dirección del tono o un aumento de volumen o de duración)?, ¿lo perciben en el mismo punto donde se observa en la señal?, ¿cuál de los principales fenómenos observados para los acentos léxicos perciben como distintivo: el tono, la duración, la longitud o el timbre? La forma más directa de responder estas preguntas es llevar a cabo experimentos de percepción en los que se alteran los parámetros prosódicos observados en un conjunto de enunciaciones, para que los hablantes evalúen las enunciaciones modificadas o las comparen con otras no modificadas. Por ejemplo, puede disminuirse el volumen de una sílaba con acento léxico para probar si la sílaba todavía se percibe como prominente. O bien, se puede disminuir la elevación final en un enunciado interrogativo o desplazarla a una sílaba anterior para determinar si se sigue percibiendo como pregunta.

No es tarea fácil preparar y realizar experimentos de percepción de este tipo y a la fecha se ha informado sobre muy pocos experimentos de percepción para lenguas fuera de Europa y Japón.<sup>12</sup> En cierto modo, la parte más sencilla es la preparación de los estímulos, ya que las herramientas de

---

<sup>12</sup>Mucho del trabajo reciente en este campo lo han realizado investigadores del Centro de Lingüística de la Universidad de Leiden, Países Bajos, en el Laboratorio de Fonética, y han trabajado sobre todo con lenguas de Indonesia, en particular el malayo. Véanse Ebing (1997), Odé (1997, 2002), van Zanten *et al.* (2003) y Stoel (2005: 108-208) para ejemplos y referencias. Estos trabajos también proporcionan un análisis detallado sobre la forma en que pueden diseñarse y aplicarse experimentos prosódicos. Otro grupo de investigadores, asociado con el *Linguistic Lab* de Bochum, Alemania, está realizando un extenso trabajo sobre la percepción de diferencias prosódicas entre dialectos del ruso; véase la dirección electrónica (<http://www.ruhr-uni-bochum.de/lilab/Index.htm>).

análisis del habla, como EMU, PRAAT, Wave Surfer o Speech Analyzer, permiten una modificación relativamente sencilla y directa del tono y otros parámetros prosódicos en enunciaciones digitalizadas. La parte más difícil es encontrar una forma de realizar las pruebas, en especial en sociedades que tienen poco o ningún contacto con la experimentación.

En otras palabras, los experimentos de percepción enfrentan los mismos problemas que la elicitación. Los problemas pueden aparecer desde el comienzo si, por ejemplo, los hablantes se rehúsan a utilizar los audífonos (que son la mejor forma de asegurar que escuchen cuidadosamente los estímulos). Sin embargo, el reto principal consiste en definir una tarea practicable. Por lo general, no será posible pedir directamente la identificación de propiedades prosódicas con preguntas como “¿cuál es el rasgo más destacado?”, “¿es X más alta que Y?”; etc. En cambio, lo que puede funcionar son tareas que impliquen cierta clase de comparación y clasificación de elementos distintos, incluyendo preguntas como “¿cuál de estos dos elementos es más natural o más apropiado?”, “¿cuál se escucha con mayor frecuencia?” o “¿qué palabra elegirías para hablar en público?”, entre otras. Adicionalmente, los comentarios generales sobre los estímulos (“esto suena bastante extraño”, “así habla la gente del sur”, etc.) pueden proporcionar señales importantes, aunque contribuirán a que el conjunto de datos sea muy heterogéneo y difícil de cuantificar.

En relación con lo anterior, debe advertirse que los datos conversacionales no experimentales también pueden proporcionar pistas importantes sobre qué rasgos prosódicos se perciben como relevantes en una comunidad de habla determinada. Un ejemplo un tanto trivial, pero aun así relevante, es que en un corpus conversacional se pueden detectar ejemplos de enunciaciones que los participantes tratan como preguntas y que se pueden comparar con enunciaciones similares en términos prosódicos, pero que el oyente no toma como preguntas. Son todavía más complejos los ejemplos en los que un énfasis mal colocado o una curva de entonación equivocada producen un malentendido, que provoca a su vez una secuencia de corrección (véanse las contribuciones en Couper-Kuhlen y Selting 1996, y Couper-Kuhlen y Ford 2004, para observaciones y ejemplos relevantes).

## **Agradecimientos**

Me gustaría agradecer a Bruce Birch y Bob Ladd por haber analizado conmigo ampliamente las cuestiones e ideas presentadas en este capítulo. Ellos, por supuesto, no son responsables del uso que he hecho de sus

conocimientos. Agradezco también a Jan Strunk por su ayuda con el diseño gráfico de las figuras.

Agradezco de forma muy especial a Mauricio C. A. Belo, mi colaborador waima'a, quien pacientemente pasó por varias pruebas tanto de los procedimientos analizados aquí como de otros que no resultaron ser productivos. Puede encontrarse información adicional y agradecimientos completos al proyecto waima'a en la dirección:

[http://www.mpi.nl/dobes/WebpageDobes1/SubpagesTeams/  
SubpageWaimaa/Frameset.htm](http://www.mpi.nl/dobes/WebpageDobes1/SubpagesTeams/SubpageWaimaa/Frameset.htm)

Este trabajo se hizo posible gracias a la beca de investigación otorgada por la Fundación Volkswagen y estoy muy agradecido por tan generoso apoyo.